PG PHY S1-02 DSC

TWO YEAR MASTER'S DEGREE COURSE

(CBCS in Accordance with NEP 2020)

SEMESTER-I

MATHEMATICAL PHYSICS

PHYSICS

Hiranya Chandra Bhuyan School of Science and Technology

KRISHNA KANTA HANDIQUI STATE OPEN UNIVERSITY

Subject Experts

Prof. Anurup Gohain Baruah, Department of Physics, Gauhati University
Prof. Mrinal Kumar Das, Department of Physics, Tezpur University
Dr. Debajyoti Barooah, Department of Physics, Cotton University

Programme Coordinator : Dr. Antara Mahanta Barua, Discipline of Electronics, KKHSOU

Course Coordinator : Dr. Kabita Baruah, Discipline of Physics, KKHSOU

SLM Preparation Team

UNITS	CONTRIBUTORS
1, 4	Dr. Kabita Baruah, Discipline of Physics, KKHSOU
2, 3, 8	Dr. Kabita Baruah, Discipline of Physics, KKHSOU
	Dr. Nayan Moni Nath, Department of Physics, Barnagar College, Sorbhog
5, 6, 14	Dr. Sagar Singh Goyari, Department of Physics, Rajib Gandhi University
	Dr. Chandra Kamal Borah, Discipline of Physics, KKHSOU
7	Dr. Chandra Kamal Borah, Discipline of Physics, KKHSOU
9	Dr. Dhritiman Gogoi, Department of Physics, Pandu College
	Dr. Kabita Baruah, Discipline of Physics, KKHSOU
10	Dr. Debajyoti Dutta, Department of Physics, Bhattadev University
	Dr. Antara Mahanta Barua, Discipline of Electronics, KKHSOU
11, 12, 13	Prof. Sunanda Naik, Department of Applied Sciences, Gauhati University
	Dr. Kabita Baruah, Discipline of Physics, KKHSOU

Editorial Team

Content : **Dr. Prabal Jyoti Phukan,**
Department of Physics, Dibrugarh University
: **Dr. Pranab Jyoti Bhuyan,**
Department of Physics, Cotton University

Structure & Format & Graphics: Dr. Kabita Baruah, KKHSOU

July, 2026 **ISBN:978-93-7839-002-9**

This Self Learning Material (SLM) of the Krishna Kanta Handiqui State University is madeavailableunderaCreativeCommonsAttribution-NonCommercial-ShareAlike4.0License(International): http.//creativecommons.org/licenses/by-nc-sa/4.0

Printed and published by Registrar on behalf of the Krishna Kanta Handiqui State Open University.

Headquarters: Patgaon, Rani Gate, Guwahati-781017
CityOffice:ReshamNagar,Khanapara,Guwahati-781022Web:www.kkhsou.in

CONTENTS

		PAGE NO.
UNIT 1 :	**Review of vector calculus** Vector operators: Overview of Vector operators, Gradient of Scalar function, Divergence of a vector field, Curl of a vector field, Important Vector Identities, Physical meaning of operators; Numerical Exercises	**7-26**
UNIT 2 :	**Linear Vector Spaces** N-dimensional vector space; Basis and dimensions: orthogonal and orthonormal basis; Inner product; Hilbert space; Gram–Schmidt orthonormalisation; Outer product; Applications.	**27-44**
UNIT 3 :	**Linear Operators** Overview of Linear operators; Algebra of operators; Matrix representation of linear operators; eigenvalue problems (introductory).	**45-64**
UNIT 4 :	**Matrices** Special matrices: symmetric, antisymmetric, orthogonal, hermitian, anti-hermitian, unitary; Eigenvalues and eigenvectors; Cayley–Hamilton theorem; Diagonalization of Matrix.	**65-86**
UNIT 5 :	**Tensor Analysis** Basics of tensors; Covariant and contravariant tensors; Symmetric and Anti-symmetric Tensor; Four Vectors line element; Metric tensor; Associated tensors.	**87-100**
UNIT 6 :	**Tensor Calculus and Geometry** Christoffel symbols; Covariant derivative; Geodesics; Riemann-Christoffel curvature tensor; Bianchi identities.	**101-115**
UNIT 7 :	**Differential Equations** Order of a differential equation, Degree of a differential equation; First and second order linear differential equations.	**116-135**
UNIT 8 :	**Special Functions** Types of polynomials: Legendre, Bessel, Hermite and Laguerre polynomials; Rodrigues formula; Generating functions; Orthogonality; Applications.	**136-172**
UNIT 9 :	**Laplace Transform** Laplace transform of elementary functions: properties, shifting and scaling; Derivatives and integrals of Laplace transform; Convolution theorem; Inverse Laplace transformation; Solutions of differential equations.	**173-187**

UNIT 10 : Fourier Integrals **188-208**
Preliminaries of Special Functions: Gamma, Error, Bessel functions; Fourier Integral Representations; Proof of Fourier Integral Theorem.

UNIT 11 : Fourier Transform **209-222**
Fourier transformation and properties; Fourier transform of derivatives; Shift theorem, Change of scale; Modulation theorem; Convolution theorem; Parseval's Identity.

UNIT 12 : Complex Analysis-I **223-233**
Complex Functions; Cauchy–Riemann Equations; Analytic and Harmonic functions

UNIT 13 : Complex Analysis-II **234-252**
Complex integration, Cauchy theorem and integral formula; Taylor and Laurent series; Poles; Residues.

UNIT 14 : Group Theory **253-272**
Basics of Group Theory: Definition, Examples; Discrete Groups: Cyclic, Dihedral, Permutation groups; Subgroups; Lagrange theorem; Cosets and Conjugacy classes; Introduction to Group Representations; Rotational Groups: SO(2), SO(3), U(1), SU(2).

COURSE INTRODUCTION

Mathematical Physics course is the second Discipline Specific Course (DSC) of the first semester of Master of Science in Physics programme. The main objective of this course is to develop a clear understanding of mathematical methods used in physics and the mathematical tools required to analyse physical systems. This course covers fundamental mathematical structures in physics and advanced techniques for solving physical problems. It begins with a review of vector calculus and then discusses linear vector spaces, linear operators, and matrices. These topics provide the mathematical foundation for many physical theories. It also covers tensor analysis and tensor calculus, which are useful in describing geometrical properties of space and coordinate transformations. In addition, basic differential equations are introduced, which are commonly used to describe physical systems and dynamical processes. It also includes special functions such as Legendre, Bessel, Hermite, and Laguerre polynomials, which frequently appear in the solutions of physical equations. The block also discusses integral transform methods such as Laplace transforms and Fourier transforms, which are powerful tools for solving differential equations. Further, the course introduces complex analysis and the fundamentals of group theory, which are important for understanding symmetry and mathematical structures in modern physics. After completing this course, learners will be able to apply various mathematical methods to solve problems in physics and develop a deeper understanding of the mathematical framework used in advanced physical theories. It consists of 14 units.

The first unit covers the fundamental concepts of vector calculus, including important vector operators such as gradient, divergence, and curl, along with their physical significance. It also discusses their applications in describing physical fields like electric, magnetic, and fluid flow fields.

The second unit covers the concept of linear vector spaces and their mathematical foundations in physics. It includes basis, dimension, inner products, orthogonal and orthonormal bases, Hilbert space, Gram–Schmidt orthonormalisation, and outer products.

The third unit covers the theory of linear operators and their algebraic properties. It also introduces matrix representations of operators and the concept of eigenvalue problems, which form the basis of many physical theories.

The fourth unit covers matrices and their applications in physics, with emphasis on special matrices such as symmetric, Hermitian, unitary, and orthogonal matrices. It also discusses eigenvalues, eigenvectors, the Cayley–Hamilton theorem, and matrix diagonalization.

The fifth unit covers the basic concepts of tensor analysis, including covariant and contravariant tensors, line elements, and metric tensors. These concepts provide the foundation for studying physical laws in different coordinate systems.

The sixth unit covers tensor calculus and its geometrical applications. It introduces Christoffel symbols, covariant differentiation, geodesics, curvature tensors, and Bianchi identities, which are essential in differential geometry and relativity.

The seventh unit covers the fundamental theory of differential equations, including their order and degree. It also discusses first and second-order linear differential equations and their applications in physical systems.

The eighth unit covers special functions commonly used in mathematical physics, including Legendre, Bessel, Hermite, and Laguerre polynomials. It also explains generating functions, Rodrigues' formula, orthogonality relations, and applications of Bessel functions.

The ninth unit covers the Laplace transform and its important properties, including shifting, scaling, differentiation, integration, and convolution theorems. It also discusses inverse Laplace transforms and their applications in solving differential equations.

The tenth unit covers Fourier integral representations and their importance in mathematical physics. It also introduces special functions such as Gamma, Error, and Bessel functions, along with the Fourier integral theorem.

The eleventh unit covers the Fourier transform and its major properties, such as shift theorem, scaling property, modulation theorem, convolution theorem, and Parseval's identity. These tools are widely applied in wave analysis and signal processing.

The twelfth unit covers the fundamentals of complex analysis, including complex functions, Cauchy–Riemann equations, analytic functions, and harmonic functions. These concepts form the basis for advanced complex variable theory.

The thirteenth unit covers complex integration and important results such as Cauchy's theorem and Cauchy's integral formula. It also introduces Taylor series, Laurent series, poles, and residues for evaluating complex integrals.

The fourteenth unit covers the basic principles of group theory and its applications to symmetry in physics. It discusses groups, subgroups, cosets, conjugacy classes, group representations, and important rotational groups such as SO (2), SO(3), U(1), and SU(2).

While going through the units, the learner will notice some boxes containing additional information that help in understanding important or unfamiliar terms. Sections titled "ACTIVITY" provide opportunities for the learner to apply knowledge through questions related to the topic. Some useful concepts are also provided under the heading "LET US KNOW". At the end of each section, "CHECK YOUR PROGRESS" questions are included to help the learner assess understanding of the topic. Attempting these questions and comparing the responses with the "ANSWERS TO CHECK YOUR PROGRESS" given at the end of each unit helps the learner evaluate progress in the course.

UNIT 1: REVIEW OF VECTOR CALCULUS

UNIT STRUCTURE

1.1 Learning Objectives
1.2 Introduction
1.3 Vector Operators
    1.3.1 Overview of Vector Operators
    1.3.2 Gradient of Scalar Function
    1.3.3 Divergence of a Vector Field
    1.3.4 Curl of a Vector Field
    1.3.5 Important Vector Identities
    1.3.6 Physical Meaning of Operators
1.4 Numerical Exercises
1.5 Let Us Sum Up
1.6 Answers to Check Your Progress
1.7. Further Reading
1.8 Model Questions

1.1 LEARNING OBJECTIVES

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

- apply the basic concepts of vector algebra and represent vectors in component form
- distinguish between scalar and vector fields and explain their physical significance
- use the del operator (∇) in vector calculus
- compute the gradient, divergence, and curl of scalar and vector fields
- interpret the physical meaning of gradient, divergence, and curl in field theory
- implement vector calculus concepts to solve simple problems and relate them to integral theorems

1.2 INTRODUCTION

This is the first unit of the course on Mathematical Physics which provides a review of vector calculus. Vector calculus offers the mathematical framework to describe physical quantities that vary in space, such as scalar and vector fields encountered in physics. The learner is expected to be familiar from previous studies with basic vector algebra**,** including representation of vectors in Cartesian form, vector addition, scalar (dot) and vector (cross) products, and simple geometrical interpretation of vectors.

Building on this foundation, this unit introduces differential operators such as gradient, divergence, and curl, along with their physical significance. In this unit, the learner will study scalar and vector fields, the del operator, gradient, divergence, curl, and an introduction to integral concepts and theorems connecting local and global properties of fields.

1.3 VECTOR OPERATORS

In many physical situations, quantities such as temperature, pressure, velocity, and electric field vary continuously from one point to another in space. To understand how these quantities change, a systematic framework is needed to describe their spatial behavior. Vector operators provide such a framework by extending the concepts of calculus to scalar and vector fields. These operators involve differentiation with respect to spatial coordinates and enable the analysis of how physical quantities evolve in different directions and regions of space.

1.3.1 Overview of Vector Operators

The most fundamental of vector operator is the del (or nabla) operator, denoted by ∇, from which important operations such as gradient, divergence, and curl are derived. Each of these operations has a clear physical interpretation: the gradient describes the rate and direction of change of a scalar field, the divergence measures the source or sink nature of a vector field, and the curl characterizes its rotational behavior.

Del Operators

The del operator, denoted by the symbol ∇, is a fundamental vector differential operator used to describe spatial variations of physical quantities. In Cartesian coordinates, it is defined as:

$$\nabla = \hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z} \qquad (1.1)$$

This expression shows that the del operator consists of unit vectors along the coordinate axes multiplied by corresponding partial differential operators.

Here, $\hat{\imath}, \hat{\jmath}, \hat{k}$ are unit vectors along the x, y and z directions respectively, indicating that the operator has a directional character like a vector. The terms $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ represent partial differentiation with respect to the spatial coordinates, and they measure how a function changes along each coordinate direction. Thus, the del operator combines the features of both vector algebra and differential calculus, and hence it is called a vector differential operator. When the del operator operates on a scalar field, it produces a vector field, whereas when it operates on a vector field, it can produce either a scalar or another vector depending on the type of operation performed. These operations include:

- The gradient, which is obtained by applying ∇ to a scalar field, and results in a vector field representing the rate and direction of maximum change.
- The divergence**,** obtained through the dot product $\nabla \cdot \vec{A}$, which gives a scalar quantity indicating the source or sink nature of the vector field.
- The curl**,** obtained through the cross product $\nabla \times \vec{A}$, which yields a vector representing the rotational behavior of the field.

The operation of the del operator therefore depends on how it is combined with scalar or vector fields, and each type of operation has a distinct mathematical form and physical interpretation.

1.3.2 Gradient of a Scalar Field

In mathematical physics, many physical quantities such as temperature, pressure, electric potential, and gravitational potential vary from point to point in space. These quantities are described by scalar fields. To understand how rapidly such quantities change and in which direction the change is maximum, we introduce an important differential operator called the gradient**.** More precisely, the gradient maps a differentiable scalar field to its associated vector field of steepest ascent. The gradient is a differential operator that converts a scalar function into a vector quantity. It gives information about the rate of change of a scalar quantity, the direction in which this change is maximum.

Let, $\phi(x,y,z)$be a scalar function defined in space.

Then the gradient of the scalar function is defined as

$$\text{grad } \phi = \hat{\imath}\frac{\partial \phi}{\partial x} + \hat{\jmath}\frac{\partial \phi}{\partial y} + \hat{k}\frac{\partial \phi}{\partial z} \qquad (1.2)$$

It is commonly written using the del operator ∇ as

$$\boxed{\text{grad } \phi = \nabla\phi}$$

where $\nabla = \hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}$

A) Directional Derivative

The directional derivative measures the rate of change of a scalar field in a specified direction.

Let $\phi(x,y,z)$ be a scalar field and let $\hat{n}$ be a unit vector representing a direction. Then the directional derivative of ϕ along the direction $\hat{n}$ is defined by

$$\frac{d\phi}{ds} = \nabla\phi \cdot \hat{n}$$

This expression shows that the directional derivative is the component of the gradient along the specified direction.

Important observations:

1. It measures the rate of change of the scalar field in each direction.
2. Its maximum value occurs when the direction coincides with the gradient.
3. The maximum value equals the magnitude of the gradient.

Thus, $\left(\frac{d\phi}{ds}\right)_{\max} = |\nabla\phi|$

For example, let $\phi = x^2 + y^2 + z^2$. Find the directional derivative at (1,1,1) in the direction of $\vec{a} = \hat{\imath} + 2\hat{\jmath} + 2\hat{k}$

Solution: First, $\nabla\phi = 2x\hat{\imath} + 2y\hat{\jmath} + 2z\hat{k}$

At (1,1,1) $\nabla\phi = 2\hat{\imath} + 2\hat{\jmath} + 2\hat{k}$

Unit vector along $\vec{a}$:

$$\hat{n} = \frac{\hat{\imath} + 2\hat{\jmath} + 2\hat{k}}{3}$$

Therefore, $\frac{d\phi}{ds} = (2\hat{\imath} + 2\hat{\jmath} + 2\hat{k}) \cdot \frac{\hat{\imath}+2\hat{\jmath}+2\hat{k}}{3}$

$$= \frac{2+4+4}{3} = \frac{10}{3}$$

Hence, the directional derivative is $\frac{10}{3}$

1.3.3 Divergence of a Vector Field

The divergence of a vector point function $\vec{F}$is represented by div$\vec{F}$. It measures the rate at which a vector field spreads out or converges at a point.

Let the vector field be $\vec{F} = F_1\hat{\imath} + F_2\hat{\jmath} + F_3\hat{k}$.

Then, the divergence of $\vec{F}$is defined as the dot product of the del operator ∇ with the vector field:

$$\text{div}\,\vec{F} = \nabla \cdot \vec{F} = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \cdot \left(\hat{\imath}F_1 + \hat{\jmath}F_2 + \hat{k}F_3\right) \qquad (1.3)$$

On simplifying, we obtain $\text{div}\,\vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$.

Hence, the divergence of a vector field is a scalar function defined at every point in space.

1.3.4 Curl of a Vector Field

The curl of a vector field measures the rotational tendency of the field at a point. It indicates how much and in what direction the field circulates locally. The curl of a vector point function $\overrightarrow{F}$is defined as the cross product of the del operator ∇ with the vector field:

$$\text{curl}\,\vec{F} = \vec{\nabla} \times \vec{F} \qquad (1.4)$$

Here, $\vec{F} = F_1\hat{\imath} + F_2\hat{\jmath} + F_3\hat{k}$

$\therefore$ Curl $\vec{F} = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \times \left(F_1\hat{\imath} + F_2\hat{\jmath} + F_3\hat{k}\right)$

This cross product is conveniently evaluated using the determinant form:

$$\text{Curl } \vec{F} = \begin{vmatrix} \hat{\imath}\hat{\jmath}\hat{k} \\ \frac{\partial}{\partial x}\frac{\partial}{\partial y}\frac{\partial}{\partial z} \\ F_1 F_2 \ F_3 \end{vmatrix}$$

Curl $\vec{F}$ is a vector quantity.

Expanding the determinant, we obtain

$$\text{Curl } \vec{F} = \hat{\imath}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right) + \hat{\jmath}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right) + \hat{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)$$

Thus, curl of a vector can be physically interpreted as

- If curl $\vec{F} = 0$, the field is said to be irrotational.
- If curl $\vec{F} \neq 0$, the field has rotational character

1.3.5 Important Vector Identities

In vector calculus, the operator's gradient, divergence, and curl satisfy certain fundamental identities that play an important role in mathematical physics. These identities are useful in simplifying field equations and establishing important physical laws in electrostatics, fluid mechanics, and electromagnetic theory.
Two important identities are discussed below.

(i) Curl of the Gradient of a Scalar Function
For any scalar function $\phi(x, y, z)$,

$$\nabla \times (\nabla\phi) = 0$$

This identity states that the curl of the gradient of a scalar field is always zero.
Proof:
Let $\phi(x, y, z)$be a scalar function.
Its gradient is

$$\nabla\phi = \hat{\imath}\frac{\partial\phi}{\partial x} + \hat{\jmath}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}$$

Taking curl on both sides,

$$\nabla \times (\nabla\phi) = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\phi}{\partial x} & \frac{\partial\phi}{\partial y} & \frac{\partial\phi}{\partial z} \end{vmatrix}$$

Expanding the determinant,

$$\nabla \times (\nabla\phi) = \hat{\imath}\left(\frac{\partial^2\phi}{\partial y\,\partial z} - \frac{\partial^2\phi}{\partial z\,\partial y}\right) + \hat{\jmath}\left(\frac{\partial^2\phi}{\partial z\,\partial x} - \frac{\partial^2\phi}{\partial x\,\partial z}\right) + \hat{k}\left(\frac{\partial^2\phi}{\partial x\,\partial y} - \frac{\partial^2\phi}{\partial y\,\partial x}\right)$$

For a well-behaved scalar function, the mixed partial derivatives are equal:

$$\frac{\partial^2\phi}{\partial y\,\partial z} = \frac{\partial^2\phi}{\partial z\,\partial y}$$
$$\frac{\partial^2\phi}{\partial z\,\partial x} = \frac{\partial^2\phi}{\partial x\,\partial z}$$

$$\frac{\partial^2 \phi}{\partial x\, \partial y} = \frac{\partial^2 \phi}{\partial y\, \partial x}$$

Hence, $\nabla \times (\nabla \phi) = 0$

Thus proved.

Physical significance:

This shows that the gradient field of a scalar function is always irrotational**.**

(ii) Divergence of the Curl of a Vector Function

For any vector function $\vec{A}$,

$$\nabla \cdot (\nabla \times \vec{A}) = 0$$

This identity states that the divergence of the curl of a vector field is always zero.

Proof

Let, $\vec{A} = A_x \hat{\imath} + A_y \hat{\jmath} + A_z \hat{k}$

Then, $\nabla \times \vec{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\hat{\imath} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\hat{\jmath} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{k}$

Taking divergence, $\nabla \cdot (\nabla \times \vec{A}) = \frac{\partial}{\partial x}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)$

Expanding,

$$\nabla \cdot (\nabla \times \vec{A}) = \frac{\partial^2 A_z}{\partial x\, \partial y} - \frac{\partial^2 A_y}{\partial x\, \partial z} + \frac{\partial^2 A_x}{\partial y\, \partial z} - \frac{\partial^2 A_z}{\partial y\, \partial x} + \frac{\partial^2 A_y}{\partial z\, \partial x} - \frac{\partial^2 A_x}{\partial z\, \partial y}$$

Grouping similar terms,

$$\nabla \cdot (\nabla \times \vec{A}) = \left(\frac{\partial^2 A_z}{\partial x\, \partial y} - \frac{\partial^2 A_z}{\partial y\, \partial x}\right) + \left(\frac{\partial^2 A_y}{\partial z\, \partial x} - \frac{\partial^2 A_y}{\partial x\, \partial z}\right) + \left(\frac{\partial^2 A_x}{\partial y\, \partial z} - \frac{\partial^2 A_x}{\partial z\, \partial y}\right)$$

Using equality of mixed partial derivatives,

$$\nabla \cdot (\nabla \times \vec{A}) = 0$$

Thus proved.

Physical significance: This shows that the curl of any vector field is always solenoidal**.**

(iii) Curl of the Curl of a Vector Field

For any vector field $\vec{A}$,

$$\nabla \times (\nabla \times \vec{A}) = \nabla(\nabla \cdot \vec{A}) - \nabla^2 \vec{A}$$

This identity is useful in electromagnetic theory and wave equations. It relates the curl of the curl of a vector field to the gradient of divergence and the vector Laplacian.It plays an important role in deriving the electromagnetic wave equation from Maxwell's equations.

Applications of these identities

These identities have important applications in physics:

1. In electrostatics, since electric field is derived from scalar potential, its curl is zero.

2. In magnetostatics, magnetic field is often expressed as the curl of a vector potential, and hence its divergence vanishes automatically.
3. In fluid mechanics, these identities help in analyzing irrotational and incompressible flows.

The three fundamental vector operators—gradient, divergence, and curl—provide mathematical tools for describing the local behaviour of scalar and vector fields. The gradient measures the rate of change of a scalar field, divergence measures the source or sink strength of a vector field, and curl measures its rotational tendency. Having defined these operators and established their important identities, we now discuss their physical significance in various physical systems.

1.3.6 Physical Meaning of Operators

Operators acquire clear physical meaning when applied to scalar and vector fields, as they describe how physical quantities vary in space and relate to observable phenomena.

A) Physical Interpretation of Gradient of a Scalar

The gradient of a scalar field has a clear geometrical and physical interpretation, which helps in understanding how a physical quantity varies in space:

1. The gradient vector at any point is directed along the direction in which the scalar function increases most rapidly. This means that if one moves in the direction of the gradient, the value of the scalar field rises at the highest possible rate.
2. The magnitude of the gradient represents the maximum rate of increase of the scalar function per unit distance at that point. A larger magnitude indicates a steeper change in the scalar quantity, whereas a smaller magnitude indicates a more gradual variation.
3. The gradient is always perpendicular (normal) to surfaces of constant scalar value, known as level surfaces or equipotential surfaces. This property is important in visualizing scalar fields, as it shows that there is no change in the scalar quantity along the surface itself, but only in the direction normal to it.

For example:

1.In a temperature distribution, the gradient indicates the direction of maximum temperature rise.

2. In electrostatics, the electric field is related to the gradient of electric potential b

$$\vec{E} = -\nabla V$$

This implies that the electric field points in the direction of the maximum decrease of potential**,** and its magnitude indicates how rapidly the potential changes in space.

B) Physical Interpretation of Divergence of a vector

To understand the physical significance of divergence, let us consider the flow of a fluid.

Suppose a very small rectangular parallelepiped is situated at a point $(x, y, z)$with dimensions $dx, dy,$and dzalong the x-, y-, and z-axes respectively. Let the velocity of the fluid at that point be represented by the vector field

$$\vec{V} = V_x\hat{\imath} + V_y\hat{\jmath} + V_z\hat{k}$$

Flow along the x-direction

The amount of fluid entering the element through the face perpendicular to the x-axis per unit time is proportional to

$$\rho V_x \, dy \, dz$$

where ρ is the density of the fluid.

As the fluid passes through the element, the velocity and density may vary with position. Therefore, the amount of fluid leaving through the opposite face per unit time is

$$\left(\rho V_x + \frac{\partial(\rho V_x)}{\partial x} dx\right) dy \, dz$$

Hence, the net outward flow along the x-direction is

$$\frac{\partial(\rho V_x)}{\partial x} dx \, dy \, dz$$

Similarly, along the y- and z-directions, the outward flows are

$$\frac{\partial(\rho V_y)}{\partial y} dx \, dy \, dz$$

and

$$\frac{\partial(\rho V_z)}{\partial z} dx \, dy \, dz$$

respectively.

Adding all three contributions, the total outward mass flow per unit time becomes

$$\left[\frac{\partial(\rho V_x)}{\partial x} + \frac{\partial(\rho V_y)}{\partial y} + \frac{\partial(\rho V_z)}{\partial z}\right] dx \, dy \, dz$$

or $\nabla \cdot (\rho\vec{V}) \, d\tau$

where $d\tau = dx \, dy \, dz$ is the volume element.

From the principle of conservation of mass, the rate of decrease of mass inside the volume element must equal the net outward flow of mass. Therefore,

$$\frac{\partial \rho}{\partial t} d\tau + \nabla \cdot (\rho\vec{V}) d\tau = 0$$

Dividing throughout by $d\tau$, we obtain

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\vec{V}) = 0$$

This is called the general equation of continuity. It expresses the fundamental law of conservation of mass in fluid motion.

For an incompressible fluid, the density remains constant, that is,

$$\frac{\partial \rho}{\partial t} = 0, \rho = \text{constant}$$

Hence, the continuity equation reduces to

$$\nabla \cdot \vec{V} = 0$$

This condition shows that the velocity field has zero divergence and is therefore called a solenoidal vector field.

Thus, the condition

$$\nabla \cdot \vec{V} = 0$$

is the incompressibility condition, which represents a special case of the continuity equation for constant density.

Physical Meaning

The divergence of a velocity field represents the net outward flow of fluid per unit volume at a point:

1. Positive divergence $(\nabla \cdot \vec{V} > 0)$: fluid is spreading outward (source).
2. Negative divergence $(\nabla \cdot \vec{V} < 0)$: fluid is converging inward (sink).
3. Zero divergence $(\nabla \cdot \vec{V} = 0)$: no net gain or loss of fluid; the flow is incompressible (solenoidal).

Hence, divergence provides a quantitative measure of the source or sink character of a vector field and plays a central role in fluid dynamics and field theory.

C) Physical interpretation of Curl of a vector

To understand the physical significance of curl, consider the motion of a fluid.

In fluid mechanics, the curl of the velocity field is directly related to the vorticity of the fluid, which measures the local angular rotation of fluid elements. If the fluid is flowing in such a way that neighbouring fluid layers move with different velocities, velocity gradients are produced. These velocity gradients exert unequal forces on opposite sides of a small paddle wheel placed in the flow, producing a net torque and causing the wheel to rotate.

In such a case,

$$\nabla \times \vec{V} \neq 0$$

and the flow is said to be **rotational**.

On the other hand, if the velocity gradients are arranged such that the torque on one side of the paddle wheel is exactly balanced by the torque on the opposite side, there is no resultant turning effect.

In that case,

$$\nabla \times \vec{V} = 0$$

and the flow is said to be **irrotational**.

Thus, curl provides a quantitative measure of the local rotational tendency of the fluid. A larger value of curl indicates stronger local rotation, while zero curl indicates absence of local angular rotation. It is important to note that curl measures local rotation (vorticity), not the overall shape of streamlines**.** A flow may have curved or circular streamlines and still have zero curl if there is no local angular rotation of fluid elements.

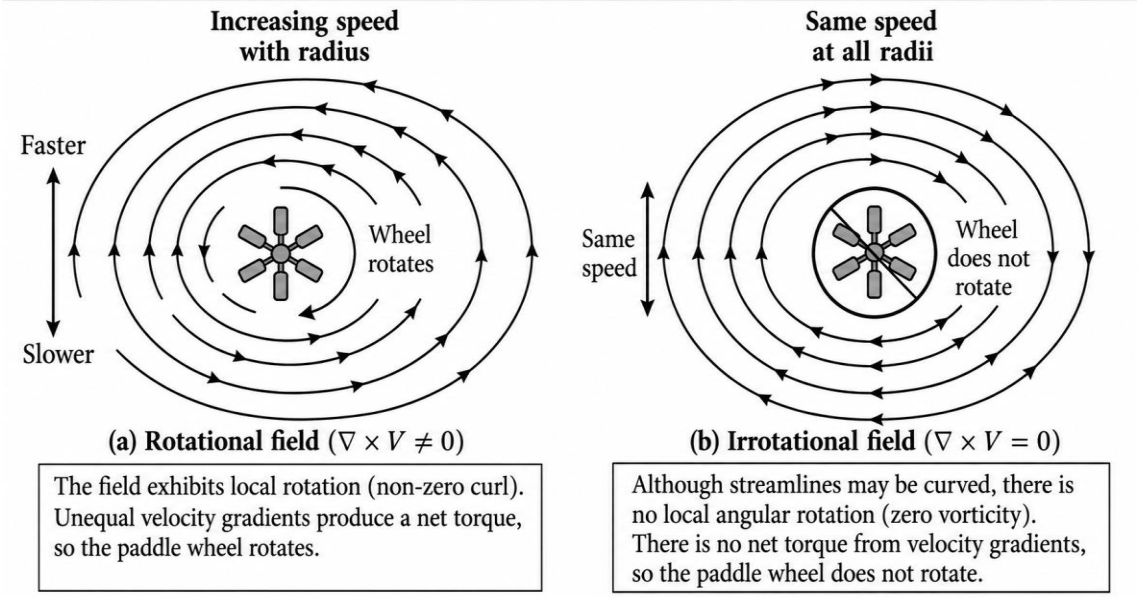


Figure 1.1(a) In a rotational flow field, the fluid exhibits local rotation (non-zero curl)**.** Due to unequal velocity gradients across the fluid element, a small paddle wheel experiences a net torque and rotates. **(b)** In an irrotational flow field, although the streamlines may be curved, there is no local angular rotation (zero vorticity)**.** The velocity gradients produce no net torque, and hence the paddle wheel does not rotate.

Electromagnetism:

In electromagnetism, the concept of curl plays a fundamental role in describing how electric and magnetic fields vary in space and time. One of the key relations involving curl is given by Maxwell's equation*:*

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

This equation states that a time-varying magnetic field gives rise to a circulating electric field. In other words, the electric field forms closed loops around regions where the magnetic field changes with time. The presence of a non-zero curl of the electric field indicates that the field has rotational character even in the absence of electric charges.

This behaviour is analogous to the magnetic field produced around a current-carrying conductor, where the magnetic field is illustrated infigure 1.2**.**

The direction of the curl vector is determined by the right-hand rule and represents the axis about which the field circulates. The magnitude of the curl indicates the strength of this local circulation.

Thus, curl provides a clear physical picture of how changing magnetic fields generate rotational electric fields and forms the basis of electromagnetic induction.

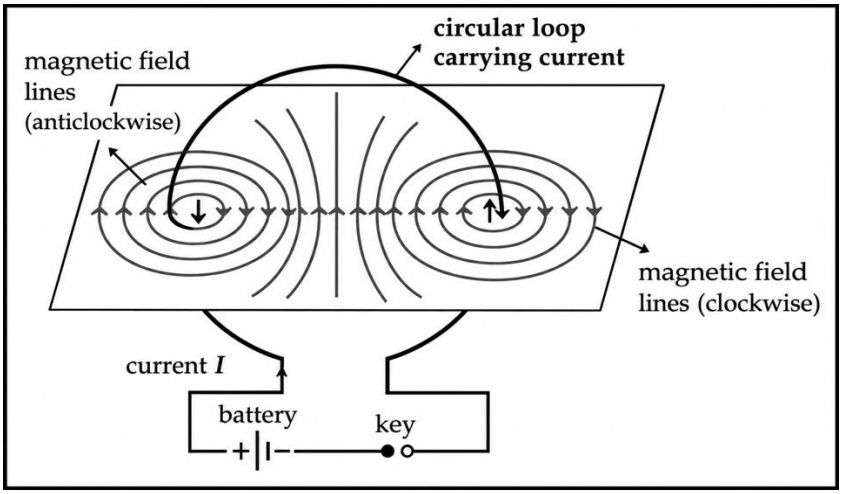


Figure 1.2: Magnetic field lines around a current-carrying circular loop, forming closed curves that illustrate the rotational nature (non-zero curl) of the magnetic field.

Thus, curl provides a clear physical picture of how changing magnetic fields generate rotational electric fields, forming the basis of important phenomena such as electromagnetic induction**.**

Rotational and Irrotational Vector

Consider a rigid body rotating with angular velocity $\vec{\omega}$.

If $\vec{r}$denotes the position vector of a point on the body, then the linear velocity of that point is given by

$$\vec{V} = \vec{\omega} \times \vec{r},$$

Here. $\vec{\omega} = \omega_1\hat{\imath} + \omega_2\hat{\jmath} + \omega_3\hat{k}$ and $\vec{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$

$$\therefore \text{Curl } \vec{V} = \vec{\nabla} \times (\vec{\omega} \times \vec{r}) = \vec{\nabla} \times \left[(\omega_1\hat{\imath} + \omega_2\hat{\jmath} + \omega_3\hat{k}) \times \left(x\hat{\imath} + y\hat{\jmath} + z\hat{k}\right)\right]$$

where, $\vec{\omega}$ is the angular velocity vector, $\vec{V}$ is the linear velocity, $\vec{r}$ is the position vector.

$$\therefore \text{Curl } \vec{V} = \vec{\nabla} \times \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \omega_1 & \omega_2 & \omega_3 \\ x & y & z \end{vmatrix} = \vec{\nabla} \times \left[(\omega_2 z - \omega_3 y)\hat{\imath} + (\omega_3 x - \omega_1 z)\hat{\jmath} + (\omega_1 y - \omega_2 x)\hat{k}\right]$$

$$= \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \times \left[(\omega_2 z - \omega_3 y)\hat{\imath} + (\omega_3 x - \omega_1 z)\hat{\jmath} + (\omega_1 y - \omega_2 x)\hat{k}\right]$$

$$= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (\omega_2 z - \omega_3 y) & (\omega_3 x - \omega_1 z) & (\omega_1 y - \omega_2 x) \end{vmatrix}$$

$$= (\omega_1 + \omega_1)\hat{\imath} - (-\omega_2 - \omega_2)\hat{\jmath} + (\omega_3 + \omega_3)\hat{k}$$

$$= 2(\omega_1\hat{\imath} + \omega_2\hat{\jmath} + \omega_3\hat{k})$$

$$= 2\omega$$

This result shows that curl measures the rotational character of motion. For a rigidly rotating body, the curl of velocity is directly related to angular velocity, indicating that curl represents twice the local angular rotation of the medium

If curl $\vec{F}$ =0, the field $\vec{F}$ is termed as irrotational.

Let us illustrate with an example:

Consider a fluid undergoing rigid-body rotation about the z-axis. The velocity field of the fluid is given by

$$\vec{V} = -\omega y\,\hat{\imath} + \omega x\,\hat{\jmath}$$

where ω is a constant representing the angular speed of rotation. Find the curl of the velocity field and interpret its physical meaning.

Solution: The curl of the $\vec{V}$ is given by

$$\nabla \times \vec{V} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix}$$

The velocity componentsare given by

$$V_x = -\omega y,\ V_y = \omega x,\ \ V_z = 0$$

Evaluating the determinant, we have

$$\nabla \times \vec{V} = \hat{\imath}\left(\frac{\partial 0}{\partial y} - \frac{\partial(\omega x)}{\partial z}\right) - \hat{\jmath}\left(\frac{\partial 0}{\partial x} - \frac{\partial(-\omega y)}{\partial z}\right) + \hat{k}\left(\frac{\partial(\omega x)}{\partial x} - \frac{\partial(-\omega y)}{\partial y}\right)$$

Computing the derivatives, we have

$$\nabla \times \vec{V} = \hat{\imath}(0-0) - \hat{\jmath}(0-0) + \hat{k}(\omega - (-\omega)) = 2\omega\hat{k}$$

The fluid particles move in circular paths like water in a whirlpool. The curl vector points along the axis of rotation (here, the z-axis). The magnitude of curl equals twice the angular velocity of the rotating fluid.

The concepts of gradient, divergence, and curl are further connected to integral theorems such as Green's theorem, Stokes' theorem, and Gauss's divergence theorem, which are studied in advanced vector calculus."

Integral Theorems of Vector Calculus

The differential operator's gradient, divergence, and curl are closely connected to integral theorems that relate local and global properties of fields.

(i) Gauss's Divergence Theorem

It states that $\iiint_V (\nabla \cdot \vec{A})dV = \iint_S \vec{A} \cdot d\vec{S}$

Physical meaning: The total outward flux through a closed surface equals the total divergence inside the volume.

Application: Electric flux in electrostatics.

(ii) Stokes' Theorem

It states that $\oint_C \vec{A} \cdot d\vec{r} = \iint_S (\nabla \times \vec{A}) \cdot d\vec{S}$

Physical meaning: Circulation around a closed path equals total rotation through the surface.

Application: Magnetic circulation.

(iii) Green's Theorem

It states that $\oint_C (Pdx + Qdy) = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$

Physical meaning: Relates line integral around a closed curve to area integral over the enclosed region.
Application: Planar fluid flow.

1.4 NUMERICAL EXERCISES

To develop a clear understanding of gradient, divergence, and curl, let us consider the following numerical exercises:
Q1: Consider the scalar function

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

Find the gradient of the scalar function and determine its value at the point $(1,2,2)$.
Solution: The gradient of a scalar function $\phi(x, y, z)$is defined as

$$\nabla\phi = \hat{\imath}\frac{\partial\phi}{\partial x} + \hat{\jmath}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}$$

Calculate partial derivatives

$$\frac{\partial\phi}{\partial x} = 2x, \frac{\partial\phi}{\partial y} = 2y, \frac{\partial\phi}{\partial z} = 2z$$

$$\therefore \nabla\phi = 2x\,\hat{\imath} + 2y\,\hat{\jmath} + 2z\,\hat{k}$$

At $(1,2,2)$ $\nabla\phi = 2(1)\hat{\imath} + 2(2)\hat{\jmath} + 2(2)\hat{\mathrm{k}} = 2\hat{\imath} + 4\hat{\jmath} + 4\hat{\mathrm{k}}$

Thus, $\boxed{\nabla\phi = 2\hat{\imath} + 4\hat{\jmath} + 4\hat{k}}$

Q2: Given a vector field

$$\vec{F} = x^2\hat{\imath} + y^2\hat{\jmath} + z^2\hat{k}$$

find the divergence of $\vec{F}$and evaluate it at the point (1,2,3)

Solution: The divergence of a vector field

$$\vec{F} = F_x\hat{\imath} + F_y\hat{\jmath} + F_z\hat{k}$$

is defined as $\nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$

Here, $F_x = x^2, F_y = y^2, F_z = z^2$

By partial differentiation,

$$\frac{\partial F_x}{\partial x} = 2x, \frac{\partial F_y}{\partial y} = 2y, \frac{\partial F_z}{\partial z} = 2z$$

Therefore,

$$\nabla \cdot \vec{F} = 2x + 2y + 2z \quad (1.5)$$

At (1, 2, 3), equation (1.5) becomes

$$\nabla \cdot \vec{F} = 2(1) + 2(2) + 2(3) = 2 + 4 + 6 = 12$$

Thus, $\nabla \cdot \vec{F} = 12$

Q3: The velocity field of a fluid is given by

$$\vec{V} = (2x)\hat{\imath} + (3y)\hat{\jmath} + (4z)\hat{k}$$

1. Find the divergence of the velocity field.
2. Determine whether the fluid behaves like a source or an incompressible flow at the point (1,1,1)

Solution: The divergence of a velocity field is

$$\nabla \cdot \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$$

The velocity components are given by

$$V_x = 2x, V_y = 3y, V_z = 4z$$

Computing the partial derivatives of the components, we obtain

$$\frac{\partial V_x}{\partial x} = 2x, \frac{\partial V_y}{\partial y} = 2y, \frac{\partial V_z}{\partial z} = 2z$$

Therefore, $\nabla \cdot \vec{V} = 2x + 2y + 2z$

At the point (1, 1, 1)$\nabla \cdot \vec{V} = 2(1) + 2(1) + 2(1) = 6$

Since the divergence is positive, there is a net outward flux of fluid from an infinitesimal volume around the point. Hence, the point behaves like a source.

$$\nabla \cdot \vec{V} \neq 0$$

the flow is not incompressible.

Q4: Given the vector field

$$\vec{F} = -y\,\hat{\imath} + x\,\hat{\jmath} + 0\,\hat{k},$$

find the curl of $\vec{F}$.

Answer: The curl of a vector field is defined as

$$\nabla \times \vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}.$$

Here, the components are

$$F_x = -\mathrm{y}\,, F_y = x, F_z = 0$$

Evaluating the determinant, we have

$$\nabla \times \vec{F} = \hat{\imath}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) - \hat{\jmath}\left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right) + \hat{k}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)$$

$$\nabla \times \vec{F} = \hat{\imath}\left(\frac{\partial 0}{\partial y} - \frac{\partial \mathrm{x}}{\partial z}\right) - \hat{\jmath}\left(\frac{\partial 0}{\partial x} - \frac{\partial(-\mathrm{y})}{\partial z}\right) + \hat{k}\left(\frac{\partial \mathrm{x}}{\partial x} - \frac{\partial(-\mathrm{y})}{\partial y}\right)$$

Computing the derivates, we have

$$\nabla \times \vec{F} = \hat{\imath}(0-0) - \hat{\jmath}(0-0) + \hat{k}(1-(-1))$$
$$\nabla \times \vec{F} = 2\hat{k}$$

This vector field represents rotational motion in the xy-plane.

Q5. Given the scalar field

$$\phi = x^2y + y^2z + z^2x$$

find:(i) the gradient of ϕ

(ii) the directional derivative at the point $(1,1,1)$in the direction of the vector

$$\vec{a} = \hat{\imath} + 2\hat{\jmath} + 2\hat{k}$$

Solution: The gradient is

$$\nabla\phi = \frac{\partial\phi}{\partial x}\hat{\imath} + \frac{\partial\phi}{\partial y}\hat{\jmath} + \frac{\partial\phi}{\partial z}\hat{k}$$

Computing the partial derivatives,

$$\frac{\partial\phi}{\partial x} = 2xy + z^2$$
$$\frac{\partial\phi}{\partial y} = x^2 + 2yz$$
$$\frac{\partial\phi}{\partial z} = y^2 + 2zx$$

Thus, $\nabla\phi = (2xy + z^2)\hat{\imath} + (x^2 + 2yz)\hat{\jmath} + (y^2 + 2zx)\hat{k}$

At $(1,1,1)$, $\nabla\phi = 3\hat{\imath} + 3\hat{\jmath} + 3\hat{k}$

Unit vector: $\hat{n} = \frac{\hat{\imath}+2\hat{\jmath}+2\hat{k}}{3}$

Directional derivative:$\nabla\phi \cdot \hat{n}$

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

Hence, directional derivative = 5

Q6: Given the vector field

$$\vec{A} = x^2y\hat{\imath} + y^2z\hat{\jmath} + z^2x\hat{k}$$

find its divergence and determine whether the field is solenoidal.

Solution:

$$\nabla \cdot \vec{A} = \frac{\partial}{\partial x}(x^2y) + \frac{\partial}{\partial y}(y^2z) + \frac{\partial}{\partial z}(z^2x)$$
$$= 2xy + 2yz + 2zx$$

Since$\nabla \cdot \vec{A} \neq 0$

the field is not solenoidal.

Q7.Given the vector field

$$\vec{A} = yz\hat{\imath} + zx\hat{\jmath} + xy\hat{k}$$

find the curl and determine whether the field is rotational or irrotational.

Solution

$$\nabla \times \vec{A} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ yz & zx & xy \end{vmatrix}$$

Evaluating,$\nabla \times \vec{A} = (x - x)\hat{\imath} + (y - y)\hat{\jmath} + (z - z)\hat{k} = 0$
Hence, the field is irrotational**.**

CHECK YOUR PROGRESS

Q1. Find the directional derivative of

$$\phi = x^2y + yz$$

at point (1, 1, 1) in the direction of vector $\vec{A} = \hat{\imath} + 2\hat{\jmath} + 2\hat{k}$

Q2. Find the gradient of $\phi = \ln(x^2 + y^2 + z^2)$

Q3.Evaluate $\nabla \cdot \vec{A}$at $(1, 2, 3)$for $\vec{A} = xy\hat{\imath} + yz\hat{\jmath} + zx\hat{k}$

Q4. Show that the vector field $\vec{A} = r^n\vec{r}$ has divergence

$$\nabla \cdot \vec{A} = (n + 3)r^n$$

Q5. Evaluate $\nabla \times \vec{A}$at point (1,1,1) for

$$\vec{A} = x^2\hat{\imath} + y^2\hat{\jmath} + z^2\hat{k}$$

Q6. Verify for the scalar function $\phi = x^2y + yz^2$ that $\nabla \times (\nabla\phi) = 0$.

Q7. Verify for the vector field $\vec{A} = xy\hat{\imath} + yz\hat{\jmath} + zx\hat{k}$ that $\nabla \cdot (\nabla \times \vec{A}) = 0$.

Q8. Determine whether the vector field. $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$ is rotational or irrotational.

1.5 LET US SUM UP

- Scalar and vector quantities form the basis of physical description, where scalars are characterized by magnitude alone, while vectors possess both magnitude and direction and obey the laws of vector algebra.

- Representation of vectors in cartesian coordinates using unit vectors $\hat{\imath}, \hat{\jmath}, \hat{k}$ provides a systematic approach for mathematical manipulation and physical interpretation.
- Fundamental vector operations such as dot product and cross product are essential in analyzing physical quantities like work, flux, and rotational effects.
- A scalar field and a vector field describe spatial distributions of physical quantities, forming the foundation for field theory in physics.
- The del operator (∇) acts as a vector differential operator and serves as the fundamental tool for defining vector calculus operations.
- The gradient of a scalar field produces a vector field that indicates the direction of maximum increase of the scalar quantity and its rate of variation.
- The divergence of a vector field yields a scalar function representing the net outward flux per unit volume, thereby identifying sources, sinks, and solenoidal fields.
- The curl of a vector field produces a vector that quantifies the local rotational behavior of the field and distinguishes between rotational and irrotational fields.
- The physical interpretations of gradient, divergence, and curl play a crucial role in understanding phenomena in fluid dynamics and electromagnetism.
- Integral concepts such as line integrals, surface integrals, and volume integrals extend the analysis from local behavior to global properties of fields.
- The fundamental theorems of vector calculus—Green's theorem, Stokes' theorem, and Gauss's divergence theorem—establish the connection between differential and integral formulations of physical laws.
- Overall, vector calculus provides a unified mathematical framework essential for advanced studies in classical mechanics, electromagnetism, fluid mechanics, and quantum physics**.**

✓1.6 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: Given,

$$\phi = x^2y + yz$$

Gradient ofϕ, $\nabla\phi = (2xy)\hat{\imath} + (x^2+z)\hat{\jmath} + y\hat{k}$

At (1,1,1) $\phi = 2\hat{\imath} + 2\hat{\jmath} + \hat{k}$

Now, the unit vector $\vec{A} = \hat{\imath} + 2\hat{\jmath} + 2\hat{k}$, $|\vec{A}| = 3$

$$\hat{A} = \frac{1}{3}(1,2,2)$$

Now, directional derivative is given by

$$\nabla\phi \cdot \hat{A} = \frac{1}{3}(2 + 4 + 2) = \frac{8}{3}$$

Answer to Q2: Given,

$$\phi = \ln(x^2 + y^2 + z^2)$$

$$\Rightarrow \nabla\phi = \frac{1}{x^2 + y^2 + z^2}(2x\hat{\imath} + 2y\hat{\jmath} + 2z\hat{k})$$

$$\Rightarrow \boxed{\nabla\phi = \frac{2(x\hat{\imath} + y\hat{\jmath} + z\hat{k})}{x^2 + y^2 + z^2}}$$

Answer to Q3: Given,

$$\nabla \cdot \vec{A} = x + y + z$$

At (1,2,3) $\nabla \cdot \vec{A} = 6$

Answer to Q4: Given, $\vec{A} = r^n\vec{r}$

Using identity: $\nabla \cdot (f\vec{r}) = f(\nabla \cdot \vec{r}) + \vec{r} \cdot \nabla f$

$$= r^n(3) + r \cdot (nr^{n-1}\hat{r})$$
$$= 3r^n + nr^n$$

Answer to Q5: Given, $\vec{A} = x^2\hat{\imath} + y^2\hat{\jmath} + z^2\hat{k}$

Now, $\nabla \times \vec{A} = \nabla \times \left(x^2\hat{\imath} + y^2\hat{\jmath} + z^2\hat{k}\right)$

Since, all cross partial derivatives = 0

$$\nabla \times \vec{A} = 0$$

Answer to Q6: Given the scalar function

$$\phi = x^2y + yz^2$$

We must verify that. $\nabla \times (\nabla\phi) = 0$

First, find the gradient of ϕ:

$$\nabla\phi = \hat{\imath}\frac{\partial\phi}{\partial x} + \hat{\jmath}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}$$

Now, $\frac{\partial\phi}{\partial x} = 2xy$, $\frac{\partial\phi}{\partial y} = x^2 + z^2$, $\frac{\partial\phi}{\partial z} = 2yz$

Therefore, $\nabla\phi = 2xy\hat{\imath} + (x^2 + z^2)\hat{\jmath} + 2yz\hat{k}$

Now taking curl,

$$\nabla \times (\nabla\phi) = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2xy & x^2 + z^2 & 2yz \end{vmatrix}$$

Expanding, $\nabla \times (\nabla\phi) = \hat{\imath}\left(\frac{\partial}{\partial y}(2yz) - \frac{\partial}{\partial z}(x^2 + z^2)\right) - \hat{\jmath}\left(\frac{\partial}{\partial x}(2yz) - \frac{\partial}{\partial z}(2xy)\right)$
$+ \hat{k}\left(\frac{\partial}{\partial x}(x^2 + z^2) - \frac{\partial}{\partial y}(2xy)\right)$

Now, $\frac{\partial}{\partial y}(2yz) = 2z$, $\frac{\partial}{\partial z}(x^2 + z^2) = 2z$, $\frac{\partial}{\partial x}(2yz) = 0$, $\frac{\partial}{\partial z}(2xy) = 0$, $\frac{\partial}{\partial x}(x^2 + z^2) = 2x$,
$\frac{\partial}{\partial y}(2xy) = 2x$

Substituting, $\nabla \times (\nabla\phi) = \hat{\imath}(2z - 2z) - \hat{\jmath}(0 - 0) + \hat{k}(2x - 2x) = 0$

Hence, $\nabla \times (\nabla\phi) = 0$

Thus verified.

Answer to Q7: Given the vector field

$$\vec{A} = xy\hat{\imath} + yz\hat{\jmath} + zx\hat{k}$$

We haveto verify that

$$\nabla \cdot (\nabla \times \vec{A}) = 0$$

First, calculate the curl of $\vec{A}$:

$$\nabla \times \vec{A} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy & yz & zx \end{vmatrix}$$

Expanding,

$$\nabla \times \vec{A} = \hat{\imath}\left(\frac{\partial}{\partial y}(zx) - \frac{\partial}{\partial z}(yz)\right) - \hat{\jmath}\left(\frac{\partial}{\partial x}(zx) - \frac{\partial}{\partial z}(xy)\right) + \hat{k}\left(\frac{\partial}{\partial x}(yz) - \frac{\partial}{\partial y}(xy)\right)$$

Now,

$\frac{\partial}{\partial y}(zx) = 0, \frac{\partial}{\partial z}(yz) = y, \frac{\partial}{\partial x}(zx) = z, \frac{\partial}{\partial z}(xy) = 0, \frac{\partial}{\partial x}(yz) = 0, \ \frac{\partial}{\partial y}(xy) = x$

So, $\nabla \times \vec{A} = -y\hat{\imath} - z\hat{\jmath} - x\hat{k}$

Now take divergence:

$$\nabla \cdot (\nabla \times \vec{A}) = \frac{\partial(-y)}{\partial x} + \frac{\partial(-z)}{\partial y} + \frac{\partial(-x)}{\partial z}$$

Since,

$$\frac{\partial(-y)}{\partial x} = 0, \frac{\partial(-z)}{\partial y} = 0, \frac{\partial(-x)}{\partial z} = 0$$

Therefore,

$$\nabla \cdot (\nabla \times \vec{A}) = 0$$

Hence verified.

Answer to Q8: Given, $\vec{\boldsymbol{F}} = -\boldsymbol{y\hat{\imath}} + \boldsymbol{x\hat{\jmath}}$

$$\boldsymbol{\nabla \times \vec{F}} = \left(\frac{\boldsymbol{\partial x}}{\boldsymbol{\partial x}} - \frac{\boldsymbol{\partial(-y)}}{\boldsymbol{\partial y}}\right)\boldsymbol{\hat{k}} = (\mathbf{1} - (-\mathbf{1}))\boldsymbol{\hat{k}} = \mathbf{2}\boldsymbol{\hat{k}}$$

As the curl is non-zero, so the field is rotational.

1.7 FURTHER READING

- Dass H. K.and Verma, Rama (2022), *Mathematical Physics*, S.Chand Publication, New Delhi.
- Arfken, G. B., Weber, H. J., and Harris, F. E. (2013),*Mathematical Methods for Physicists* :A Comprehensive Guide, 7th Edition, Academic Press, Amsterdam.
- Schey, H. M. (2005),*Div, Grad, Curl and All That* : An informal Text on Vector Calculus, 4th Edition, W. W. Norton & Company, New York.
- Boas, M. L. (2006), *Mathematical Methods in the Physical Sciences*, 3rd Edition, Wiley, New Delhi.
- Spiegel, M. R. (1968), *Schaum's Outline of Vector Analysis*, McGraw-Hill, New York.
- Marsden, J. E. and Tromba, A. (2012), *Vector Calculus*, 6th Edition, W. H. Freeman, New York.

⍰1.8 MODEL QUESTIONS

1. Define the del (∇) operator. Explain its role in vector calculus.
2. Explain the physical meaning of gradient, divergence, and curl with suitable examples.
3. What is a solenoidal vector field? Give its mathematical condition and physical significance.
4. Distinguish between rotational and irrotational vector fields with examples.

5. Show that the directional derivative of a scalar function is maximum in the direction of its gradient.
6. Prove that $\nabla \times (\nabla\phi) = 0$for any scalar function ϕ.

7. Verify that $\nabla \cdot (\nabla \times \vec{A}) = 0$for a well-behaved vector field $\vec{A}$.

8. Find the gradient of $\phi = x^2 + y^2 + z^2$ and evaluate it at the point (1, 2, 2).

9. Determine whether the vector field
$$\vec{F} = (x^2 - yz)\hat{\imath} + (y^2 - zx)\hat{\jmath} + (z^2 - xy)\hat{k}$$
is solenoidal.

10. A velocity field is given by $\vec{V} = (2x - y)\hat{\imath} + (x + 3y)\hat{\jmath} + 4z\hat{k}$.Find its divergence and comment on whether the flow is compressible or incompressible.

*** *** ***

UNIT 2: LINEAR VECTOR SPACES

UNIT STRUCTURE

2.1 Learning Objectives
2.2 Introduction
2.3 Introduction to Linear Vector Spaces
2.4 N-Dimensional Vector Space
2.5 Basis and dimensions
2.5.1 Orthogonal and Orthonormal Basis
2.6 Inner Product
2.7 Hilbert Space
2.8 Gram–Schmidt Orthonormalization
2.9 Outer Product
2.10 Applications
2.11 Let Us Sum Up
2.12 Answers to Check Your Progress
2.13 Further Reading
2.14 Model Questions

2.1 LEARNING OBJECTIVES

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

- explain the concept and properties of vector spaces
- describe the basis and determine the dimension of a vector space
- analyse the concepts of orthogonality and orthonormality in vector spaces
- explain the concept of Hilbert space and complex vector spaces
- apply the Gram–Schmidt orthonormalization process.

2.2 INTRODUCTION

In the previous unit, you studied the fundamental concepts of vectors, scalar and vector fields, and the vector differential operators such as gradient, divergence, and curl. These concepts provided the mathematical tools required to describe physical quantities and their variations in space. The study of vector calculus is largely based on vectors in two- and three-dimensional Euclidean space and their applications in various branches of physics. While ordinary vectors are useful for describing physical quantities such as displacement, velocity, force, and electric fields, many problems in mathematics and physics involve more general objects, including matrices, tensors, functions, polynomials, and solutions of differential equations. Although these objects may appear different in nature, they possess common algebraic properties that allow them to be treated in a unified manner.In this unit, you will learn the fundamental concepts of linear vector spaces, including vector space axioms, linear dependence and independence, basis sets, dimensionality, inner products, dual spaces, orthogonality, orthonormal bases, Gram–Schmidt orthonormalization, outer products, and Hilbert spaces. These concepts form the mathematical foundation for advanced

topics in mathematical physics and are essential for the study of operators, eigenvalue problems and function spaces.

2.3 INTRODUCTION TO LINEAR VECTOR SPACES

A **linear vector space (LVS),** denoted by V, is a collection of elements $|\psi\rangle, |\phi\rangle, |\chi\rangle, ...$, called **vectors,** together with an operation known as **addition**, which satisfies the following properties:
(i) $|\phi\rangle + |\psi\rangle = |\psi\rangle + |\phi\rangle \in V$ for every $|\phi\rangle, |\psi\rangle \in V$.
(ii) $(|\phi\rangle + |\psi\rangle) + |\chi\rangle = |\phi\rangle + (|\psi\rangle + |\chi\rangle)$ for all $|\phi\rangle, |\psi\rangle, |\chi\rangle \in V$.
(iii) There exists a unique null vector $|\Omega\rangle \in V$ such that $|\phi\rangle + |\Omega\rangle = |\phi\rangle$ for every $|\phi\rangle \in V$.
(iv) For every $|\phi\rangle$, there exists a unique vector $-|\phi\rangle \in V$ satisfying $|\phi\rangle + (-|\phi\rangle) = |\Omega\rangle$.

The null vector $|\Omega\rangle$ should not be confused with the symbol $|0\rangle$, which is often used to denote a specific non-null vector, such as the ground state of a quantum mechanical system.

In addition to vector addition, an LVS is also equipped with an operation called scalar multiplication**,** where the scalars $a, b, ...$ belong to a field. The term field is used here in its mathematical sense. For our purposes, it is sufficient to consider the fields of real numbers (R) and complex numbers $(\mathbb{C})$, whose properties are already familiar. Scalar multiplication is required to satisfy the following properties:

(v) $a\,|\phi\rangle \in V,\ \forall a \in R$ or C and $|\phi\rangle \in V$.

(vi) $a(b\,|\psi\rangle) = (ab)\,|\psi\rangle$.

(vii) $a(|\psi\rangle + |\phi\rangle) = a\,|\psi\rangle + a\,|\phi\rangle$.

(viii) $(a + b)\,|\psi\rangle = a\,|\psi\rangle + b\,|\psi\rangle$.

(ix) $1\,|\psi\rangle = |\psi\rangle$.

(x) $0\,|\psi\rangle = 0$.

These properties ensure that scalar multiplication is well defined and is compatible with the addition operation in the linear vector space.Because of property given by (x), it is common practice to use the symbol 0 both for the scalar zero and for the null vector $|\Omega\rangle$. Consequently, one often encounters equations in which the left-hand side is a vector while the right-hand side is written simply as 0, even though strict notation would require the right-hand side to be the null vector $|\Omega\rangle$.

A linear vector space is also commonly referred to as a linear space or simply a vector space**.** The vectors $|\phi\rangle$, $|\psi\rangle$, and similar elements are often called ket vectors. When the scalars $a, b, ...$ are restricted to real numbers, the space is known as a **r**eal linear space (LVS**).** If the scalars are allowed to be complex numbers, the space is called a

complex linear space. Throughout the discussion, an asterisk ($*$) will be used to denote the complex conjugate of a complex number.

Dual Vector Space

Associated with every linear vector space $V$is another linear vector space, called its dual space, and denoted by V^*. The elements of the dual space are represented by

$$\langle \phi \mid, \langle \psi \mid, \ldots$$

and are called bra vectors. The corresponding elements of V,

$$\mid \phi\rangle, \ \mid \psi\rangle, \ \ldots,$$

are called ket vectors.

An important property of dual spaces is that the dual of the dual space is the original vector space itself. Symbolically,

$$(V^*)^* \equiv V$$

Thus, taking the dual space twice brings us back to the original linear vector space.

For a finite-dimensional vector space, the concept of a dual space can be understood through the familiar example of Euclidean space. If the vectors of an n-dimensional vector space are represented by $(n \times 1)$column matrices,

$$\mid \psi\rangle = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

then the corresponding elements of the dual space are represented by $(1 \times n)$row matrices,

$$\langle \psi \mid = (x_1^* \quad x_2^* \quad \cdots \quad x_n^*)$$

where the superscript ($*$) denotes the complex conjugate of the corresponding component. From a mathematical standpoint, the dual space (V^ {*}) is defined as the collection of all linear functionals on (V). A linear functional is a mapping that associates each vector in (V) with a scalar while preserving the operations of vector addition and scalar multiplication. In the Dirac notation used throughout quantum mechanics, these linear functionals are represented by bra vectors.

Thus, the dual space may be regarded as the space of all bra vectors corresponding to the ket vectors of (V). The interaction between bras and kets gives rise to the inner product, which plays a fundamental role in defining orthogonality, norms, and many other important concepts in linear vector spaces.

2.4 N-DIMENSIONAL VECTOR SPACE

The concept of dimension plays a fundamental role in the study of linear vector spaces. Intuitively, the dimension of a vector space indicates the number of independent directions required to describe all vectors in the space.

A linear vector space is said to be n-dimensional if every vector in the space can be uniquely represented using (n) independent components. For example, ordinary vectors in a plane require two components and therefore form a two-dimensional space, while vectors in ordinary three-dimensional space require three components and form a three-dimensional space.

In general, a vector space may have any finite dimension (n). Although it is not possible to visualize spaces of dimension greater than three, the mathematical principles governing them are direct extensions of those familiar from two- and three-dimensional geometry.

An important result in linear algebra states that every finite-dimensional LVS of dimension (n) is isomorphic to (R^{n}), the set of ordered (n)-tuples of real numbers. This means that the vectors in such a space can be represented in a form analogous to ordinary vectors in (n)-dimensional Euclidean space, while preserving the rules of vector addition and scalar multiplication.

Therefore, many properties of finite-dimensional vector spaces can be understood by extending the geometric intuition developed from two- and three-dimensional vectors. This correspondence also provides the foundation for the study of matrices, linear transformations, and quantum-mechanical state vectors.

To develop a precise understanding of dimension and vector representation, it is necessary to introduce the concept of a basis, which will be discussed in the next section.

2.5 BASIS AND DIMENSIONS

A basis of an LVS (V) is a set of vectors$\{|\,\psi_k\rangle,\ k = 1,2,\ldots\}$ that satisfies two essential conditions.

(a)Linear Independence

The vectors of the basis must be linearly independent. This means that no vector in the set can be represented as a linear combination of the remaining vectors. Equivalently, the equation of the form

$$\sum_k a_k\,|\,\psi_k\rangle = 0$$

has only the trivial solution, $a_k = 0$ for all k. In other words, the null vector can be obtained as a linear combination of the vectors $|\,\psi_k\rangle$only when all the coefficients $a_k$are zero.

b) Spanning the Space

The vectors must also span the entire space (V). A set of vectors $\{|\,\psi_k\}$is said to span the vector space V if every vector $|\,\chi\rangle \in V$can be represented as

$$|\,\chi\rangle = \sum_k b_k\,|\,\psi_k\rangle,$$

for appropriate values of the coefficients b_k. In other words, every vector in the space can be generated from linear combinations of the vectors in the set.

The notion of span can be illustrated using ordinary three-dimensional Euclidean space. The vector (e_x) spans the (x)-axis, while the sets (e_x,e_y) and ($e_x + e_y$,$2e_x - 3e_y$) span the entire (xy)-plane. More generally, a set of vectors spans a space if every vector in that space can be generated from suitable linear combinations of the vectors in the set.

It is important to note that linear independence and spanning are separate requirements. A set of vectors may satisfy one condition without satisfying the other. For example, in three-dimensional Euclidean space, the vectors (e_x,e_y) are linearly independent, but they do not span the whole space. Conversely, the set (e_x, e_y,e_z, $e_x + e_y + e_z$) spans the space, yet it is not linearly independent because one of the four vectors can be expressed as a linear combination of the other three.

Dimensionality of an LVS

Although a vector space may possess many different bases, all bases of a given LVS contain the same number of vectors. This invariant number is called the **dimension** or **dimensionality** of the vector space.

The following properties characterize an (n)-dimensional LVS:

- Every basis of the space contains exactly (n) vectors.
- No set of (r) vectors can be linearly independent if (r > n).
- No set of (r) vectors can span the entire space if (r < n).

Thus, the dimension of a vector space specifies the exact number of independent vectors required to span the space completely.

Infinite-Dimensional Linear Vector Spaces

An LVS is said to be **infinite-dimensional** if it is not possible to find a finite set of vectors that forms a basis for the space. In other words, there exists no basis consisting of a finite number (n) of vectors capable of spanning the entire space.

A typical example of an infinite-dimensional vector space is the set of all polynomials in the complex variable z, represented by

$$P(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots ,$$

where the coefficients $a_0, a_1, a_2,$ ...are complex numbers. Since polynomials of arbitrarily high degree are included, no finite set of basis vectors can span the entire space. A basis for this vector space therefore contains infinitely many elements, namely $\{1, z, z^2, z^3, \ldots\}$.Consequently, the space is infinite-dimensional.

Many vector spaces encountered in quantum mechanics, differential equations, and functional analysis are infinite-dimensional. Although several concepts developed for finite-dimensional vector spaces—such as linear independence, spanning sets, and basis vectors—remain applicable, certain properties and theorems valid in finite

dimensions may not hold automatically in the infinite-dimensional case. Consequently, the study of infinite-dimensional vector spaces often requires additional mathematical tools and careful treatment.

2.5.1 Orthogonal and Orthonormal Basis

A basis $\{| \phi_k\}$of a linear vector space is said to be an orthogonal basis if every pair of distinct basis vectors is orthogonal, that is,

$$\langle \phi_j \mid \phi_k \rangle = 0, j \neq k$$

In an orthogonal basis, the basis vectors are mutually orthogonal, but they are not necessarily normalized. Consequently, the inner product of a basis vector with itself need not be equal to unity. Thus,

$$\langle \phi_k \mid \phi_k \rangle \neq 1,\text{in general.}$$

Therefore, an orthogonal basis becomes an orthonormal basis only when each basis vector is normalized to have unit magnitude. Under this additional condition,

$$\langle \phi_k \mid \phi_k \rangle = 1,\text{for every } k$$

Combining the orthogonality and normalization conditions, the orthonormality relation can be expressed compactly as

$$\langle \phi_j \mid \phi_k \rangle = \delta_{jk},$$

where δ_{jk} is the Kronecker delta, defined by

$$\delta_{jk} = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases}$$

Thus, the inner product of two basis vectors is equal to unity when they are the same vector and vanishes when they are different vectors.

An orthonormal basis possesses another important property known as the completeness relation, which is given by

$$\sum_j | \phi_j \rangle \langle \phi_j | = I,$$

where I denotes the identity operator. The identity operator leaves every vector in the linear vector space unchanged when it acts on it. The summation extends over all vectors in the basis set.

The completeness relation implies that every vector in the linear vector space can be represented uniquely as a linear combination of the basis vectors. Thus, an orthonormal basis provides a complete description of the vector space. It is important to note that orthonormality and completeness are distinct concepts. The orthonormality relation,

$$\langle \phi_j \mid \phi_k \rangle = \delta_{jk},$$

is a scalar relation involving inner products, whereas the completeness relation,

$$\sum_j | \phi_j \rangle \langle \phi_j | = I,$$

is an operator relation. Orthogonal and orthonormal bases considerably simplify calculations involving vector expansions, inner products, and operators. Moreover, any linearly independent basis can be converted into an orthonormal basis through the Gram–Schmidt orthonormalization procedure, which systematically constructs a set of mutually orthogonal and normalized basis vectors from a given set of linearly independent vectors.

2.6 INNER PRODUCT

In a linear vector space, the quantity $\langle\phi \mid \psi\rangle$, formed from a bra $\langle\phi \mid$and a ket $\mid \psi\rangle$, is called the inner product of the vectors $\mid \phi\rangle$ and $\mid \psi\rangle$. The inner product is, in general, a complex number and is therefore also known as a scalar product.

Properties of the Inner Product

1. **Linearity in the second argument (Ket)**

$$\langle\phi \mid (\mid \psi\rangle + \mid \chi\rangle) = \langle\phi \mid \psi\rangle + \langle\phi \mid \chi\rangle$$

2. **Linearity in the first argument (bra)**

$$(\langle\psi \mid + \langle\chi \mid) \mid \phi\rangle = \langle\psi \mid \phi\rangle + \langle\chi \mid \phi\rangle$$

3. **Scalar multiplication in the ket**

$$\langle\phi \mid a \mid \psi\rangle = a\,\langle\phi \mid \psi\rangle$$

4. **Scalar multiplication in the bra**

$$a \mid \phi\rangle)^{\dagger} \mid \psi\rangle = a^{*}\,\langle\phi \mid \psi\rangle$$

or equivalently,

$$a\langle\phi \mid \psi\rangle = a^{*}\,\langle\phi \mid \psi\rangle$$

where a^{*}denotes the complex conjugate of the scalar a.

5. **Conjugate symmetry**

$$\langle\phi \mid \psi\rangle = \langle\psi \mid \phi\rangle^{*}$$

These properties define the mathematical structure of the inner product in a complex linear vector space. In particular, the appearance of the complex conjugate in properties (4) and (5) distinguishes complex vector spaces from real vector spaces. If a ket vector $\mid \psi\rangle$is represented by an $(n \times 1)$column matrix with complex components $x_1, x_2, \ldots, x_n$, then the corresponding bra vector $\langle\psi \mid$is obtained by taking the complex conjugate transpose (Hermitian conjugate) of the ket. Thus,

$$\mid \psi\rangle = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and

$$\langle\psi \mid = (x_1^{*} \quad x_2^{*} \quad \cdots \quad x_n^{*}),$$

where $x_1, x_2, \ldots, x_n$are generally complex numbers and the superscript $(^{*})$ denotes complex conjugation. The vectors $\mid \phi\rangle$and $\mid \psi\rangle$ are orthogonal to each other if and only if

$$\langle\phi \mid \psi\rangle = 0.$$

Thus, the vanishing of the inner product serves as the condition for orthogonality in a linear vector space.

The inner product of a vector with itself is used to define the norm of the vector. The norm is a generalization of the familiar concept of the length or magnitude of a vector in Euclidean space. The norm of the vector $|\psi\rangle$is defined as

$$\| \psi \| = (\langle \psi | \psi \rangle)^{1/2}$$

Equivalently, $$\| \psi \| = \sqrt{\langle \psi | \psi \rangle}$$

The norm is always a non-negative real number. It vanishes only when the vector itself is the null vector. Thus,

$$\| \psi \| = 0 \text{if and only if } |\psi\rangle = |\omega\rangle.$$

Therefore, the norm of a non-null vector is always positive
In more general settings, one may encounter vector spaces in which an inner product between two arbitrary vectors is not defined, while a norm is specified independently through some other prescription. Such spaces are beyond the scope of the present discussion. Here, attention is restricted to linear vector spaces in which the norm is derived from the inner product as defined above.

Basis Independence of the Inner Product

An important property of the inner product is that its value does not depend on the orthonormal basis chosen to represent the vectors. This property is known as the basis independence of the inner product.

Let $|\psi\rangle$ and $|\eta\rangle$be two vectors in a linear vector space. Suppose these vectors are expanded in two different orthonormal basis sets, $|\phi_k\rangle$and $|\chi_j\rangle$, as

$$|\psi\rangle = \sum_k c_k \ |\phi_k\rangle = \sum_j d_j \ |\chi_j\rangle$$

and

$$|\eta\rangle = \sum_k C_k \ |\phi_k\rangle = \sum_j D_j \ |\chi_j\rangle$$

Using the orthonormality of each basis, the inner product of the vectors $|\psi\rangle$and $|\eta\rangle$ can be expressed as

$$\langle \psi | \eta \rangle = \sum_k c_k^* \, C_k = \sum_j d_j^* \, D_j$$

Although the expansion coefficients depend on the basis chosen, the value of the inner product remains unchanged. This shows that the inner product is an intrinsic property of the vectors and is independent of the basis used to represent the

2.7 HILBERT SPACE

Hilbert spaces form an important class of vector spaces in modern mathematics. They provide a general framework for studying vectors, inner products, norms, orthogonality,

and convergence in both finite- and infinite-dimensional settings. Many mathematical concepts involving functions, sequences, and linear operators can be described naturally within the structure of a Hilbert space.

A Hilbert space may be viewed as a generalization of ordinary Euclidean space to higher, and even infinite, dimensions. While retaining familiar geometric notions such as length, angle, distance, and orthogonality using an inner product, it also possesses the property of completeness, which ensures that the limit of every convergent sequence of vectors remains within the space. Because of these features, Hilbert spaces provide a powerful and rigorous framework for the study of vector spaces and their applications. In this section, we shall examine the definition and fundamental properties of Hilbert spaces.

Definition:A Hilbert space, denoted by H, is a complex vector space (i.e., a vector space over the field $\mathbb{C}$ of complex numbers) endowed with an inner product. Furthermore, it is complete with respect to the norm generated by this inner product.
Since H is a vector space, it consists of elements called vectors and supports an operation of vector addition. For any vectors α, β, and $\gamma \in H$, the addition operation satisfies the following properties:
Commutativity $\alpha + \beta = \beta + \alpha$
Associativity $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

Existence of an Additive Identity
There exists a null vector $o \in H$such that

$$\alpha + o = \alpha, \forall\, \alpha \in H.$$

Since H is a vector space over the field of complex numbers $\mathbb{C}$, vectors can be multiplied by scalars $a, b, c, \ldots \in \mathbb{C}$. Scalar multiplication satisfies the following distributive properties:
Distributivity over Vector Addition

$$c(\alpha + \beta) = c\alpha + c\beta$$

Distributivity over Scalar Addition

$$(a + b)\alpha = a\alpha + b\alpha.$$

Inner Product in a Hilbert Space
In addition to the vector space structure, a Hilbert space is equipped with an **inner product**. The inner product is a mapping

$$(\alpha, \beta): H \times H \to \mathbb{C},$$

which associates a complex number with every ordered pair of vectors α and β belonging to the Hilbert space H.
Here, $H \times H = \{(\alpha, \beta) \mid \alpha, \beta \in H\}$,
denotes the set of all ordered pairs of vectors in H, and $\mathbb{C}$,denotes the field of complex numbers. Thus, the inner product assigns to each pair of vectors α and β a unique complex number (α, β).The inner product satisfies the following properties:
Conjugate Symmetry

$$(\alpha, \beta) = (\beta, \alpha)^*$$

Linearity in the Second Argument

$$(\alpha, a\beta) = a(\alpha, \beta)$$

where $a \in \mathbb{C}$.
Additivity$(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$
Positive Definiteness$(\alpha, \alpha) \geq 0, \forall\, \alpha \in H,$
with equality only when$\alpha = o$.
The first two properties imply that

$$(a\alpha, \beta) = a^*(\alpha, \beta),$$

where a^*denotes the complex conjugate of a. This shows that the inner product is antilinear in its first argument and linear in its second argument.

Furthermore, from the conjugate symmetry property,

$(\alpha, \alpha) = (\alpha, \alpha)^*$, which implies that $(\alpha, \alpha)$is always a real number.
Thus, the inner product provides a means of defining geometric concepts such as length, angle, and orthogonality in a Hilbert space.

Norm of a Vector
The existence of an inner product allows us to define the **norm** (or magnitude) of a vector. If $\alpha \in H$, then its norm is defined by

$$\| \alpha \| = \sqrt{(\alpha, \alpha)}.$$

The norm provides a measure of the length of a vector and generalizes the familiar concept of magnitude in Euclidean space. Since the inner product is positive definite, the norm is always a non-negative real number and vanishes only for the null vector.

Cauchy–Schwarz Inequality
An important consequence of the properties of the inner product is the Cauchy–Schwarz inequality**,** which states that for any vectors $\alpha, \beta \in H$,

$$| (\alpha, \beta) |^2 \leq (\alpha, \alpha)(\beta, \beta).$$

Equivalently, $| (\alpha, \beta) | \leq \| \alpha \| \| \beta \|.$

The Cauchy–Schwarz inequality is one of the fundamental results associated with inner-product spaces. It plays an important role in establishing several properties of Hilbert spaces and is widely used in the study of vector spaces, functional analysis, and related mathematical structures.

2.8 GRAM–SCHMIDT ORTHONORMALISATION

In elementary coordinate geometry, points in space may be described using either Cartesian coordinates or oblique coordinates. Since Cartesian coordinates are based on mutually perpendicular axes, they often simplify calculations and geometric interpretations. Similarly, in a linear vector space (LVS), it is advantageous to work with an orthonormal basis rather than an arbitrary basis.

Let, $\{ \ | \phi_k \rangle \}$ denotes an orthonormal basis of a linear vector space. The basis vectors satisfy

$$\langle \phi_j \mid \phi_k \rangle = 0, \text{for } j \neq k,$$

and $\langle \phi_k \mid \phi_k \rangle = 1, \forall\, k$

These two conditions may be combined into the compact orthonormality relation

$$\langle \phi_j \mid \phi_k \rangle = \delta_{jk},$$

Where $$\delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

An orthonormal basis also satisfies the completeness relation

$$\sum_j \mid \phi_j \rangle\langle \phi_j \mid = I,$$

where I denotes the identity operator. The action of the identity operator on any vector leaves the vector unchanged, i.e.,

$$I \mid \psi \rangle = \mid \psi \rangle.$$

The summation extends over all basis vectors in the orthonormal set. In an infinite-dimensional vector space, the summation contains infinitely many terms and therefore represents an infinite series.

It is important to distinguish between orthonormality and completeness. The orthonormality condition

$$\langle \phi_j \mid \phi_k \rangle = \delta_{jk},$$

is a scalar relation, since the inner product of two vectors yields a scalar quantity.

In contrast, the completeness relation

$$\sum_j \mid \phi_j \rangle\langle \phi_j \mid = I,$$

is an operator equation, since the quantity $\mid \phi_j \rangle\langle \phi_j \mid$represents an operator and the summation of these operators yields the identity operator I.

A quantity of the form$\langle \phi \mid \psi \rangle$is a **scalar** (generally a complex number). It may be viewed as analogous to the product of a $(1 \times n)$row matrix and an $(n \times 1)$column matrix.

In contrast,$\mid \phi \rangle\langle \psi \mid$ represents an **operator**. It is analogous to the product of an $(n \times 1)$column matrix and a $(1 \times n)$row matrix, which yields an $(n \times n)$matrix. Such an operator can act on ket vectors to produce other vectors in the linear vector space. We can write in matrix form too as:

$$\langle \phi \mid \psi \rangle = (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_n^*) \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}$$

which results in a scalar, whereas

$$\mid \phi \rangle\langle \psi \mid = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix} (\psi_1^* \quad \psi_2^* \quad \cdots \quad \psi_n^*),$$

which results in an $(n \times n)$matrix (operator).

An arbitrary basis. $\{|\psi_1\rangle, |\psi_2\rangle, \ldots\}$can always be transformed into an orthonormal basis

$$\{|\phi_1\rangle, |\phi_2\rangle, \ldots\}$$

through the Gram–Schmidt orthonormalization procedure.
Step 1: Normalization of the First Vector

$$|\phi_1\rangle = \frac{|\psi_1\rangle}{\sqrt{\langle\psi_1|\psi_1\rangle}}$$

Step 2: Orthogonalization and Normalization of the Second Vector
The second vector is made orthogonal to $|\phi_1\rangle$ by subtracting from it its projection along $|\phi_1\rangle$. The resulting vector is then normalized to obtain

$$|\phi_2\rangle = \frac{|\psi_2\rangle - \langle\phi_1|\psi_2\rangle\ |\phi_1\rangle}{[\langle\psi_2|\psi_2\rangle - |\langle\phi_1|\psi_2\rangle|^2]^{1/2}}$$

Here, $\langle\phi_1|\psi_2\rangle\ |\phi_1\rangle$ represents the projection of $|\psi_2\rangle$ along the direction of $|\phi_1\rangle$.

Step 3: Construction of the Remaining Basis Vectors
Next, the projections of $|\psi_3\rangle$ along both $|\phi_1\rangle$and $|\phi_2\rangle$ are removed. The resulting vector is then normalized to obtain $|\phi_3\rangle$. Repeating this procedure successively generates a complete orthonormal basis.

For $k \geq 2$, the general Gram–Schmidt orthonormalization formula is

$$|\phi_k\rangle = \frac{|\psi_k\rangle - \sum_{i=1}^{k-1}\langle\phi_i|\psi_k\rangle\ |\phi_i\rangle}{\left[\langle\psi_k|\psi_k\rangle - \sum_{i=1}^{k-1}|\langle\phi_i|\psi_k\rangle|^2\right]^{1/2}}, k \geq 2$$

Thus, the Gram–Schmidt procedure provides a systematic method for transforming any linearly independent set of vectors into an orthonormal basis spanning the same linear vector space.

2.9 OUTER PRODUCT

The **outer product** is an important operation in linear vector spaces that allows the construction of operators from vectors. Unlike the inner product, which produces a scalar, The outer product produces a **linear operator** acting on vectors of the space.. Let ψ and ϕ be two vectors in a vector space. The outer product of these vectors is denoted by$|\psi\rangle\langle\phi|$and is defined as an operator acting on any vector $|\chi\rangle$ in the space as:

$$(|\psi\rangle\langle\phi|)\ |\chi\rangle = |\psi\rangle\langle\phi|\chi\rangle$$

Here, $\langle\phi|\chi\rangle$represents the inner product, which is a scalar. Thus, the result of the operation is a vector proportional to $|\psi\rangle$.

An important application of outer products arises in the description of orthonormal basis sets. If $\{ \mid \phi_j \rangle \}$ is an orthonormal basis of a linear vector space, then the sum of the outer products of all basis vectors satisfies the completeness relation

$$\sum_j \mid \phi_j \rangle\langle \phi_j \mid = I$$

where $I$is the identity operator. This relation shows that the identity operator can be expressed as a sum of outer products of the basis vectors. Consequently, outer products play a fundamental role in the representation of vectors and operators in a linear vector space.

Matrix Representation: If vectors are represented in component form as:

$$\mid \psi \rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}, \qquad \langle \phi \mid = (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_n^*).$$

Then the outer product is given by:

$$\mid \psi \rangle\langle \phi \mid = \begin{bmatrix} \psi_1\phi_1^* & \psi_1\phi_2^* & \cdots & \psi_1\phi_n^* \\ \psi_2\phi_1^* & \psi_2\phi_2^* & \cdots & \psi_2\phi_n^* \\ \vdots & \vdots & \cdots & \vdots \\ \psi_n\phi_1^* & \psi_n\phi_2^* & \cdots & \psi_n\phi_n^* \end{bmatrix}$$

Thus, the outer product results in an $n \times n$ matrix.

Properties of Outer Product

- The outer product produces a linear operator, meaning it defines a mapping that acts linearly on vectors in the vector space.
- The outer product is not commutative, that is,
  $$\mid \phi \rangle\langle \psi \mid \neq \mid \psi \rangle\langle \phi \mid,$$
  because interchanging the vectors changes the resulting operator.
- The outer product depends on the order of the vectors, since swapping their positions leads to a different matrix or operator.
- The outer product acts as a linear operator that transforms one vector into another vector within the same vector space.

2.10 APPLICATIONS

Linear vector spaces constitute one of the fundamental mathematical structures used in physics, mathematics, and engineering. They provide a unified framework for representing and analyzing a wide variety of physical and mathematical systems. Some important applications of linear vector spaces are given below.

1. **Representation of Physical States**
   Many physical systems can be represented by vectors in an appropriate vector space. The state of a system is described by a vector, while measurable quantities are represented by operators acting on these vectors. This approach provides a systematic framework for analyzing physical phenomena.

3. **Solution of Differential Equations**
   The set of solutions of a homogeneous linear differential equation forms a linear vector space. Any solution can be expressed as a linear combination of a set of independent basis solutions. This property is extensively used in mathematical physics and engineering.
3. **Fourier Series and Function Spaces**
   Functions can be regarded as elements of infinite-dimensional vector spaces. Using orthogonal or orthonormal basis functions, complicated functions can be expanded in terms of simpler functions through Fourier series and Fourier transforms. Such expansions are widely used in signal analysis, wave phenomena, and boundary-value problems.
4. **Matrix Representation of Linear Transformations**
   Linear transformations acting on vector spaces can be represented by matrices once a basis is chosen. This correspondence enables the study of transformations through matrix algebra and provides powerful computational techniques for solving physical and mathematical problems.
5. **Hilbert Spaces and Quantum Theory**
   Hilbert spaces are special types of linear vector spaces equipped with an inner product. They provide the mathematical setting for the description of states, observables, and operators in modern theoretical physics. Concepts such as orthonormal bases, completeness relations, and projection operators arise naturally in this framework.
6. **Eigenvalue Problems**
   Many important physical problems reduce to eigenvalue equations. Examples include the determination of normal modes of vibration, energy levels of systems, and wave propagation. Linear vector spaces provide the natural framework for formulating and solving such problems.
7. **Orthonormalization Techniques**
   The Gram–Schmidt orthonormalization procedure enables the construction of orthonormal basis sets from linearly independent vectors. Orthonormal bases simplify calculations involving vector expansions, inner products, and matrix representations, making them indispensable in both theoretical and computational applications.

CHECK YOUR PROGRESS

Q 1**.** Determine whether each of the following sets forms a linear vector space (LVS) over the appropriate field (real or complex). If it does, find its dimension and justify your answer.

(a) The set of all $(n \times n)$ matrices with complex entries.
(b) The set of all rank-2 tensors with real components in three-dimensional space.
(c) The set of all antisymmetric rank-2 tensors with complex components in three-dimensional space.
(d) The set of all ((2 \times 2)) matrices whose trace is zero.
(e) The set of all real solutions of the differential equation

$$y'' - 3y' + 2y = 0,$$

where (y) is a function of the real variable (x), and the prime denotes differentiation with respect to (x).

Q 2. Does the set $\{(1,2), (3,4)\}$constitute a vector space under the usual operations of addition and scalar multiplication? Justify your answer by checking the closure properties.

2.11 LET US SUM UP

- A linear vector space (LVS) is a set of vectors that is closed under vector addition and scalar multiplication and satisfies specific algebraic axioms.
- The concept of an LVS extends ordinary vectors to more general objects such as matrices, tensors, functions, polynomials, and solutions of differential equations.
- A set of vectors is linearly independent if no vector can be expressed as a linear combination of the others.
- A basis is a linearly independent set that spans the entire vector space, and the number of basis vectors gives the dimension of the space
- Vector spaces may be finite-dimensional or infinite-dimensional, depending on whether a finite basis exists.
- Associated with every vector space $V$is a dual space V^*, whose elements are represented by bra vectors.
- In Dirac notation, vectors are represented by kets$|\ \psi\rangle$ and dual vectors by bras $\langle\psi\ |$.
- The inner product $\langle\ \phi\ |\ \psi\ \rangle$ is a scalar quantity used to define orthogonality and the norm of vectors.
- Two vectors are said to be **orthogonal** if $\langle\phi\ |\ \psi\rangle = 0$.
- An orthonormal basis satisfies the orthonormality condition $\langle\phi_j\ |\ \phi_k\rangle = \delta_{jk}$, where $\delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$
- An orthonormal basis also satisfies the completeness relation $\sum_j |\ \phi_j\rangle\langle\phi_j\ | = I$, where $I$is the identity operator.
- The value of the inner product and the norm of a vector are independent of the orthonormal basis chosen.
- Any linearly independent basis can be transformed into an orthonormal basis using the Gram–Schmidt orthonormalization procedure.
- The outer product$|\ \phi\rangle\langle\psi\ |$ produces a linear operator, whereas the inner product $\langle\ \phi\ |\ \psi\ \rangle$ produces a scalar quantity.
- A Hilbert space is a complete inner-product space and forms the mathematical framework for many advanced applications.
- Linear vector spaces are widely used in differential equations, Fourier analysis, matrix methods, eigenvalue problems, and the study of physical systems.

✓2.12 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1(a): Yes, it forms a linear vector space over the field of complex numbers, $\mathbb{C}$. A basis may be chosen as the set of matrices in which exactly one entry is equal to 1 and all other entries are equal to 0.
Since there are n^2 independent entries in an $(n \times n)$ matrix, the dimension of the vector space is

$$\text{Dimension} = n^2$$

For example, for 2×2matrices, a basis is

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\},$$

and therefore Dimension $= 4$.

Answer to Q1(b): Yes, it forms a linear vector space over the field of real numbers, $\mathbb{R}$.
A rank-2 tensor in three-dimensional space has two indices, each of which can take three independent values. Therefore, the total number of independent components is

$$3 \times 3 = 9.$$

Hence, the dimension of the vector space is 9.
A possible basis is

$$\{\, e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23}, e_{31}, e_{32}, e_{33} \,\}$$

where e_{ij} denotes the tensor whose (i, j)-th component is 1 and all other components are 0.

Answer to Q1(c): Yes, it forms a linear vector space over the field of complex numbers, $\mathbb{C}$.
For an antisymmetric tensor, $T_{ij} = -T_{ji}$
and consequently, $T_{ii} = 0$.
In three dimensions, the independent components are

$$T_{12}, T_{23}, T_{31}.$$

Therefore, there are three independent components, and hence dimension $= 3$

Answer to Q1(d): The set of all (2×2)matrices whose trace is zero
A general (2×2)matrix is given by. $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$
Since the trace of $A$is zero, $\mathrm{Tr}(A) = a + d = 0$

which implies that $d = -a$.
Therefore, any matrix in the set can be written as

$$A = \begin{pmatrix}a & b\\ c & -a\end{pmatrix}.$$

The matrix is completely determined by the three independent parameters a, b, and c. Hence, the dimension is 3. Since the set is closed under matrix addition

and scalar multiplication and satisfies all the axioms of a linear vector space, it forms a linear vector space.

Answer to Q1(e): The set of all real solutions of

$$y'' - 3y' + 2y = 0$$

forms a linear vector space. The auxiliary equation is

$$m^2 - 3m + 2 = 0,$$

which factors as

$$(m-1)(m-2) = 0.$$

Hence, the roots are $m = 1, m = 2$.

The general solution is $y(x) = C_1 e^x + C_2 e^{2x}$.

Since every solution is a linear combination of the two independent solutions e^x and e^{2x}, the set$\{ e^x, e^{2x} \}$forms a basis. Therefore, dimension$= 2$.
Thus, the set forms a linear vector space over $\mathbb{R}$.

Answer to Q2: Consider the set $\{(1,2), (3,4)\}$

This set contains only two vectors. Under vector addition,

$$(1,2) + (3,4) = (4,6),$$

But $(4,6) \notin \{(1,2), (3,4)\}$. Hence, the set is not closed under addition.
Similarly, $2(1,2) = (2,4)$
But,$(2,4) \notin \{(1,2), (3,4)\}$
Thus, the set is not closed under scalar multiplication. Therefore, the set does not form a vector space.

2.13 FURTHER READING

- Balakrishnan.V.,(2020), *Mathematical Physics (Applications and Problems*), An eBook Pvt.Ltd. Springer
- Joglekar S.D, (2024), *Mathematical Physics: The Basic*, Universities Press
- Dass H.K, Verma Rama (2022), *Mathematical Physics (1st Edition),* S Chand Publication
- Arfken, G. B., Weber, H. J., & Harris, F. E. (2013). *Mathematical Methods for Physicists* (7th ed.). Academic Press, USA.
- Skinner, D. (2021). *Principles of Quantum Mechanics* [Lecture notes], University of Cambridge.
- Griffiths, D. J. (2018). *Introduction to Quantum Mechanics* (3rd ed.). Cambridge University Press.

⍰2.14 MODEL QUESTIONS

1. Explain the properties that define a linear vector space.
2. Discuss the concept of basis and dimension with suitable examples.
3. Explain the properties of inner product and their significance.
4. Describe orthogonal and orthonormal basis in a vector space.
5. Explain the concept of Hilbert space and its properties.
6. Describe the Gram–Schmidt orthonormalization process.
7. Test whether the vectors $(1,2,3), (2,4,6), (1,0,1)$ are linearly independent.
8. Find the inner product of (1, 2, 3) and (4, 5, 6).
9. Determine whether the vectors (1, 0, 1) and (2, 0, 2) are orthogonal.
10. Find the magnitude of the vector (3, 4, 0) using inner product.
11. Find the outer product of the vectors

$$\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}$$

*** *** ***

UNIT 3: LINEAR OPERATORS

UNIT STRUCTURE

3.1 Learning Objectives
3.2 Introduction
3.3 Overview of Linear Operators
3.4 Algebra of Operators
3.5 Matrix Representation of Linear Operators
3.6 Eigenvalue Problems (Introductory)
3.7 Let Us Sum Up
3.8 Answers to Check Your Progress
3.9 Further Reading
3.10 Model Questions

3.1 LEARNING OBJECTIVES

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

- explain the concept and properties of linear operators
- test the linearity of operators
- perform algebraic operations on operators
- represent linear operators in matrix form
- relate operator algebra to matrix representation
- analyse eigenvalues and eigenfunctions and solve eigenvalue problems.

3.2 INTRODUCTION

In the previous unit, we studied scalar and vector fields and the differential operators—gradient, divergence, and curl which are used to describe their spatial variation and physical significance. Building on this, the present unit introduces linear operators, which provide a more general mathematical framework for describing transformations of functions and vectors. The concepts of vector representation and algebra developed earlier are extended here to operator form. In this unit, we study the definition and properties of linear operators, their algebra, and their matrix representation. The concept of eigenvalues and eigenfunctions is also introduced, which is fundamental in many areas of physics, especially quantum mechanics.

3.3 OVERVIEW OF LINEAR OPERATORS

In the study of linear vector spaces, it is often necessary to describe transformations that map one vector into another while preserving the vector space structure. Such transformations are called linear operators. A linear operator acts on vectors belonging to a vector space and produces another vector according to a specified rule. Linear operators play a fundamental role in mathematics and physics, particularly in the representation of physical observables and transformations.

Consider a linear vector space V. An operator A on this vector space acts on the vectors belonging to V to yield (possibly) another vector in V, and it is completely defined by giving the rule for its action:

$$A \mid x\rangle = \mid y\rangle$$

for all $\mid x\rangle \in S \subset V$. Here $\mid y\rangle \in V$. The set of vectors S belonging to V on which the operation of A is defined is called the domain of A, and the set of vectors $\mid y\rangle$ generated from all $\mid x\rangle \in S$ is called the range of A.
Consider the three-dimensional vector space V_3consisting of all real polynomials of degree two in x. A typical vector is given by

$$\mid \alpha\rangle = a_0 + a_1 x + a_2 x^2.$$

Consider the operator

$$A = \frac{d}{dx}$$

The operation of A on every element of $V_3$is defined by

$$A(a_0 + a_1 x + a_2 x^2) = a_1 + 2a_2 x \in V_3.$$

The range of A is, however, not V_3, but the two-dimensional subspace V_2 consisting of linear polynomials in x. Thus, the domain of A is $V_3$and the range of A is V_2.

Unless otherwise specified, we shall assume that the domain of an operator A is $V$and that its range also belongs to V. This assumption is particularly important when representing operators by matrices, since a matrix representation of A can be constructed only when the operator acts within the same vector space.

Definition of a Linear Operator
An operator A is said to be **linear** if, for every $\mid x\rangle, \mid y\rangle \in S$and every $\alpha, \beta \in \mathbb{C}$,

$$A[\alpha \mid x\rangle + \beta \mid y\rangle] = \alpha A \mid x\rangle + \beta A \mid y\rangle.$$

This property is known as the principle of superposition and expresses the fact that the operator preserves both vector addition and scalar multiplication.

For example, the operator

$$A = \frac{d}{dx}$$

introduced above satisfies the linearity condition. For any two polynomials $f(x)$ and $g(x)$, and any scalars α and β,

$$\frac{d}{dx}[\alpha f(x) + \beta g(x)] = \alpha \frac{df(x)}{dx} + \beta \frac{dg(x)}{dx}.$$

Hence, $A = \frac{d}{dx}$is a linear operator.

Therefore, a **linear operator** is a mapping from a vector space to itself (or to a subspace of it) that preserves vector addition and scalar multiplication. Linear operators provide the mathematical framework for describing transformations in linear vector spaces and form the basis for matrix representations and many applications in mathematical physics.

Special Linear Operators

In the study of linear vector spaces and linear transformations, certain classes of operators occur frequently because of their important mathematical properties and physical applications. These operators play a central role in areas such as Quantum Mechanics, Mathematical Physics, and Linear Algebra. Some of the most important special linear operators are self-adjoint (Hermitian) operators, orthogonal (unitary) operators, skew-Hermitian operators, and antilinear operators**.** Their definitions and properties are discussed below:

1. Self-Adjoint (Hermitian) Operators

A linear operator M acting on a complex vector space is called a **Hermitian operator** if it is equal to its adjoint operator. Mathematically, this condition can be written as

$$M^{\dagger} = M$$

The symbol † denotes the adjoint (or Hermitian conjugate) of the operator.A Hermitian operator possesses the important property that its eigenvalues are always real. Because measurable physical quantities such as position, momentum, energy, and angular momentum must have real values, these quantities are represented by Hermitian operators in Quantum Mechanics. For example, if E represents the energy operator (Hamiltonian), then

$$E^{\dagger} = E$$

Thus, the operator is Hermitian, and its eigenvalues correspond to observable energy values of the system.

2. Orthogonal (Unitary) Operators

A linear operator U is called a **unitary operator** if its inverse is equal to its adjoint. In other words,

$$U^{-1} = U^{\dagger}.$$

Equivalently, the above condition may be expressed as

$$U^{\dagger}U = UU^{\dagger} = I$$

where I is the identity operator.

One of the most important properties of a unitary operator is that it preserves the length (or norm) of a vector. By this we mean that $|\chi\rangle \mathrm{and} U|\chi\rangle \equiv |y\rangle$ have the same for any $|\chi\rangle \in V$. Suppose a vector $|\phi\rangle$ is transformed by U into another vector

$$|\chi\rangle = U|\phi\rangle$$

The norm of the transformed vector is

$$\langle\chi|\chi\rangle = \langle\phi|U^{\dagger}U|\phi\rangle.$$

Using the unitary condition $U^\dagger U = I$

$$\langle \chi \mid \chi \rangle = \langle \phi \mid I \mid \phi \rangle = \langle \phi \mid \phi \rangle.$$

Thus, the original and transformed vectors have the same norm. This property is extremely important in Quantum Mechanics because it ensures the conservation of total probability during the evolution of a quantum system. Rotations and symmetry transformations are often represented by unitary operators.

Skew-Hermitian (Anti-Hermitian) Operators

A linear operator K is called **anti-Hermitian** if its adjoint is equal to the negative of the operator itself:

$$K^\dagger = -K$$

This condition is like that of a Hermitian operator, except for the negative sign. There exists a close relationship between Hermitian and anti-Hermitian operators. If M is Hermitian, then multiplying it by the imaginary unit i produces an anti-Hermitian operator:

$$(iM)^\dagger = -iM.$$

Conversely, if K is anti-Hermitian, then multiplying it by i yields a Hermitian operator:

$$(iK)^\dagger = iK$$

An important result is that every linear operator can be expressed uniquely as the sum of a Hermitian part and an anti-Hermitian part**.**

Let T be any linear operator. Then

$$T = \frac{T + T^\dagger}{2} + \frac{T - T^\dagger}{2}.$$

The first term, $T_H = \frac{T+T^\dagger}{2}$,is Hermitian because

$$T_H^\dagger = T_H$$

The second term,$T_A = \frac{T-T^\dagger}{2}$,is anti-Hermitian because

$$T_A^\dagger = -T_A$$

Therefore, $$T = T_H + T_A$$

This decomposition is widely used in mathematical physics and operator theory.

Conjugate-Linear (Antilinear) Operators

So far, we have considered linear operators that satisfy the usual linearity condition

$$L(a\mathbf{x} + b\mathbf{y}) = aL(\mathbf{x}) + bL(\mathbf{y})$$

where a and b are complex scalars. However, certain physical transformations require a different type of operator known as an **antilinear operator**.

An operator A is said to be antilinear if, for any vectors $\mathbf{x}$and $\mathbf{y}$, and any complex numbers a and b,

$$A(a\mathbf{x} + b\mathbf{y}) = a^* A(\mathbf{x}) + b^* A(\mathbf{y})$$

where $a^*$and b^*denote the complex conjugates of a and b, respectively. The important difference between a linear and an antilinear operator is that the scalars are replaced by their complex conjugates during the transformation.

For example, if $a = p + iq$,then its complex conjugate is $a^* = p - iq$.

Therefore, an antilinear operator acts on scalar coefficients through complex conjugation. A well-known example from Quantum Mechanics is the **time-reversal operator**, which reverses the direction of time evolution of a quantum system. The time-reversal transformation is represented by an antilinear operator because it necessarily involves complex conjugation.

Special linear operators possess properties that make them indispensable in mathematical and physical applications. Hermitian operators represent measurable physical quantities and have real eigenvalues. Unitary operators preserve vector norms and ensure probability conservation. Anti-Hermitian operators are closely related to Hermitian operators and arise naturally in symmetry transformations. Antilinear operators extend the concept of linearity by introducing complex conjugation and are important in describing time-reversal symmetry in Quantum Mechanics. Together, these operators form the mathematical foundation for many advanced topics in modern physics.

3.4 ALGEBRA OF OPERATORS

Having defined a **linear operator (LO)**, we now study the algebraic operations that can be performed on linear operators. Unless otherwise specified, we shall assume that all operators under consideration are linear operators acting on a vector space V.

(i) Equality of Operators

Two operators $A$and $B$are said to be equal if their action on every vector in the vector space Vproduces identical results. Thus, equality of operators is determined entirely by their action on vectors.

$$A = B \text{ if } A \mid x\rangle = B \mid x\rangle \text{ for every } \mid x\rangle \in V$$

In other words, if two operators transform every vector of V in the same way, they represent the same operator.

(ii) Addition of Operators

Given two linear operators $A$and B, their sum $A + B$is defined as the operator whose action on any vector $\mid x\rangle \in V$is obtained by adding the results of the individual actions of $A$and B:

$$(A + B) \mid x\rangle = A \mid x\rangle + B \mid x\rangle \quad \text{for every } \mid x\rangle \in V$$

Since vector addition is commutative,

$$A \mid x\rangle + B \mid x\rangle = B \mid x\rangle + A \mid x\rangle,$$

therefore, $A + B = B + A$.
Thus, the addition of linear operators is commutative.

(iii) Multiplication (Composition) of Operators
The multiplication of two operators is defined through successive action. The product $A \cdot B$ means that operator B acts first, followed by operator A.
The action of the product operator on a vector $| x\rangle$ is

$$A \cdot B \mid x\rangle = A(B \mid x\rangle)$$

In general, $A(B \mid x\rangle) \neq B(A \mid x\rangle)$,

and therefore, $A \cdot B \neq B \cdot A$
This shows that operator multiplication is generally **non-commutative**, and the operators $A$and B are said not to commute.

(iv) Powers of an Operator
For a positive integer m, the power $A^m$of a linear operator $A$is defined as repeated application of Aon a vector:

$$A^m \mid x\rangle = A[A[\dots A \mid x\rangle] \dots]$$

for every, $| x\rangle \in V$
where, the operator Aacts m times successively.

Example 1: Rotation Operators in Three-Dimensional Space
Consider the usual three-dimensional vector space. Let R_1 represent a rotation through $\pi/2$about the z-axis. Its action on the unit vectors is

$$R_1\hat{\imath} = \hat{\jmath}, R_1\hat{\jmath} = -\hat{\imath}, R_1(\hat{k}) = \hat{k}.$$

For a general vector $A = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}$,

the action of $R_1$is $R_1A = (-a_2\hat{\imath} + a_1\hat{\jmath} + a_3\hat{k}) \equiv B$.

The operator $R_1$is clearly linear, and both its domain and range are the entire vector space.
Now introduce another rotation operator R_2, corresponding to a rotation through $-\pi/2$about the x-axis. Its action on the unit vectors is

$$R_2\hat{\imath} = \hat{\imath}, R_2\hat{\jmath} = -\hat{k}, R_2\hat{k} = \hat{\jmath}$$

Thus, $R_2A = a_1\hat{\imath} + a_3\hat{\jmath} - a_2\hat{k}$

Product $\boldsymbol{R_1 \cdot R_2}$: The operator $R_1 \cdot R_2$ is obtained by applying R_2first and then R_1.

$$R_1 \cdot R_2\hat{\imath} = R_1(\hat{\imath}) = \hat{\jmath},$$
$$R_1 \cdot R_2\hat{\jmath} = -\hat{k},$$
$$R_1 \cdot R_2\hat{k} = -\hat{\imath}.$$

Therefore, $R_1 \cdot R_2A = -a_3\hat{\imath} + a_1\hat{\jmath} - a_2\hat{k}$
Product $\boldsymbol{R_2 \cdot R_1}$: Similarly,

$$R_2 \cdot R_1\hat{\imath} = R_2\hat{\jmath} = -\hat{k},$$
$$R_2 \cdot R_1\hat{\jmath} = -\hat{\imath},$$
$$R_2 \cdot R_1\hat{k} = \hat{\jmath}.$$

Hence, $R_2 \cdot R_1 A = -a_2\hat{\imath} + a_3\hat{\jmath} - a_1\hat{k}.$

Since $R_1 \cdot R_2 \neq R_2 \cdot R_1$,the two operators do not commute.

Square of R_1

The square of the rotation operator is

$$R_1^2 = R_1 \cdot R_1.$$

Its action on the basis vectors is

$$R_1^2\hat{\imath} = R_1(R_1\hat{\imath}) = R_1\hat{\jmath} = -\hat{\imath},$$
$$R_1^2\hat{\jmath} = -\hat{\jmath},$$
$$R_1^2\hat{k} = \hat{k}.$$

Thus, R_1^2represents a rotation about the z-axis through an angle π.

Example 2: Differential Operators on V_3

Consider the vector space $V_3$of all quadratic polynomials.

Let $A = \frac{d}{dx}$ and $B = x\frac{d^2}{dx^2}$

Sum of Operators

Using the definition of operator addition,

$$A + B = \frac{d}{dx} + x\frac{d^2}{dx^2}.$$

Product of Operators AB

Applying B first and then A,

$$AB = \frac{d}{dx}\left(x\frac{d^2}{dx^2}\right) = \frac{d^2}{dx^2} + x\frac{d^3}{dx^3}.$$

Product of Operators BA

Applying A first and then B,

$$BA = x\frac{d^2}{dx^2}\frac{d}{dx} = x\frac{d^3}{dx^3}.$$

Since $AB \neq BA$,the operators A and B do not commute.

Further, $A^3 = \frac{d^3}{dx^3}$

and for every $|\alpha\rangle \in V_3, A^3|\alpha\rangle = 0$

Thus, on the vector space V_3, the third derivative operator acts as the null operator.

Exercise

Show that $A + B$ and $AB$are linear operators.

3.5 MATRIX REPRESENTATION OF LINEAR OPERATORS

In the study of linear vector spaces, linear operators are often defined abstractly through their action on vectors. While this abstract formulation provides a general and

elegant framework, practical computations are greatly facilitated by representing linear operators in terms of matrices. The process of associating a matrix with a linear operator is known as the matrix representation of a linear operator. This representation is based on a one-to-one correspondence between linear operators acting on an n-dimensional vector space V and $n \times n$matrices, established with respect to a chosen basis of the vector space. Through this correspondence, the action of an abstract operator can be expressed in a concrete numerical form, enabling calculations to be carried out using the familiar methods of matrix algebra. Since the representation preserves the algebraic properties of linear operators, operations such as addition, scalar multiplication, composition, and adjoint formation can be performed directly through their corresponding matrix operations. Consequently, matrix representation provides a powerful and practical framework for studying linear operators in mathematics and physics.

Construction of a Matrix Representation

Consider an n-dimensional linear vector space $V$and a linear operator A acting on V.
To construct the matrix representation, choose a basis where the basis need not to be orthonormal

$$B = \{| e_1\rangle, | e_2\rangle, \dots, | e_n\rangle\}$$

Representation of a Vector

Any vector $| x\rangle \in V$ can be expanded in terms of the chosen basis as

$$| x\rangle = \sum_{i=1}^{n} x^i | e_i\rangle$$

The vector $| x\rangle$ is represented by the column vector

$$| x\rangle \leftrightarrow X = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}$$

Thus, the coordinates of $| x\rangle$ in the chosen basis form its matrix representation.

Matrix Representation of a Linear Operator

To obtain the matrix representation of the operator A, determine the action of A on each basis vector. For every basis vector $| e_i\rangle$,

$$A | e_i\rangle = \sum_{j=1}^{n} a_{ji} | e_j\rangle$$

The coefficients a_{ji}completely characterize the operator A in the chosen basis. Since there are n^2 such coefficients, they form the matrix

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

The matrix $\bar{A}$ is called the **matrix representation (MR)** of the operator A in the basis B.

Matrix Representation of the Product of Operators

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two linear operators whose matrix representations are $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{B}}$, respectively. Using

$$B \mid e_i\rangle = \sum_j b_{ji} \mid e_j\rangle,$$

we obtain. $$AB \mid e_i\rangle = A\left(\sum\nolimits_j b_{ji} \mid e_j\right) = \sum\nolimits_j b_{ji}\, A \mid e_j\rangle.$$

Substituting the expansion of $A \mid e_j\rangle$,

$$AB \mid e_i\rangle = \sum_j b_{ji} \sum_k a_{kj} \mid e_k\rangle = \sum_{k=1}^{n} \left[\sum_{j=1}^{n} a_{kj}\, b_{ji}\right] \mid e_k\rangle.$$

The quantity in brackets is the *(k, i)* th element of the matrix product AB. Hence,

$$AB \mid e_i\rangle = \sum_k (\bar{A}\bar{B})_{ki} \mid e_k\rangle.$$

Therefore, the matrix representation of the product operator AB is simply

$$AB \leftrightarrow \bar{A}\bar{B}.$$

This result explains the usual rule of matrix multiplication.

Action of an Operator on a Vector

Let $$\mid x\rangle = \sum\nolimits_i x^i \mid e_i\rangle$$

Then $$A \mid x\rangle = A\big(\sum\nolimits_i x^i \mid e_i\big) = \sum\nolimits_i x^i\, A \mid e_i\rangle$$

Using the matrix representation of A,

$$A \mid x\rangle = \sum_i x^i \sum_j a_{ji} \mid e_j\rangle = \sum_j \left(\sum_i a_{ji}\, x^i\right) \mid e_j\rangle.$$

Recognizing the expression in brackets as the components of the matrix product $\bar{A}X$,

$$A \mid x\rangle = \sum_j (\bar{A}X)_j \mid e_j\rangle.$$

Hence, if X is the column-vector representation of $\mid x\rangle$, then the matrix representation of $A \mid x\rangle$is

$$\boxed{\bar{A}X}$$

Thus, the action of a linear operator on a vector is represented by ordinary matrix multiplication.

Dependence of Matrix Representation on Basis

The matrix representation of a linear operator depends on the basis chosen for the vector space. The operator itself remains unchanged, but its matrix form changes when a different basis is selected.

Suppose a linear operator $\hat{A}$has matrix representation Ain one basis and A'in another basis. Then the two matrices represent the same operator in different coordinate systems. Thus, matrix representations are basis-dependent, whereas the operator itself is basis-independent.

Similarity Transformation and Change of Basis

As discussed earlier, the matrix representation of a linear operator depends on the basis chosen for the vector space. Although the operator itself remains unchanged, its matrix representation changes when a different basis is selected.

Let A be the matrix representation of a linear operator in a basis B, and let A' be its matrix representation in another basis B'. Suppose the two bases are related through a non-singular matrix S. Then the matrices A and $A'$are related by

$$A' = S^{-1}AS$$

This relation is known as a **similarity transformation**.

The matrices A and A' represent the same linear operator in two different bases. Therefore, they describe the same transformation even though their matrix elements may differ. Similar matrices have the same eigenvalues and hence the same characteristic equation. Similarity transformations are particularly useful because they allow an operator to be represented in a simpler form. If a complete set of eigenvectors exists, the matrix representation can be transformed into a diagonal matrix by a suitable similarity transformation.

Example: Differential Operator on Polynomial Space

Consider the vector space $V$of quadratic real polynomials

$$f(x) = b_1 + b_2x + b_3x^2$$

Choosing the basis$\{1, x, x^2\}$

there is a one-to-one correspondence $f(x) \leftrightarrow \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

Now consider the linear operator $A = \frac{d}{dx}$

Its action on the basis vectors is

$$A \mid e_1\rangle = \frac{d}{dx}(1) = 0,$$

$$A \mid e_2\rangle = \frac{d}{dx}(x) = 1 = \mid e_1\rangle,$$

$$A \mid e_3\rangle = \frac{d}{dx}(x^2) = 2x = 2 \mid e_2\rangle.$$

Therefore, the matrix representation of A is

$$\bar{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

For the polynomial $1 + 5x + 3x^2$,the corresponding column vector is. $\begin{bmatrix}1\\5\\3\end{bmatrix}$

Applying the operator,

$$\bar{A}\begin{bmatrix}1\\5\\3\end{bmatrix} = \begin{bmatrix}5\\6\\0\end{bmatrix}$$

which represents the polynomial $5 + 6x$,the derivative of $1 + 5x + 3x^2$

Adjoint, Unitary and Hermitian Operators

Hermitian and Unitary Operators

In the study of linear operators, certain classes of operators possess special mathematical properties and play an important role in physics. Two such important classes are **Hermitian operators** and **Unitary operators**. Recall that for an operator A, the adjoint operator $A^\dagger$is defined through the relation

$$\langle \phi \mid A\psi \rangle = \langle A^\dagger \phi \mid \psi \rangle$$

for any vectors $\mid \phi\rangle$ and $\mid \psi\rangle$ in the vector space.

Hermitian Operators: An operator A is said to be **Hermitian** if it is equal to its own adjoint, that is,

$$\mathrm{A}^\dagger = \mathrm{A}.$$

In matrix language, a matrix is Hermitian if it is equal to its conjugate transpose. Hermitian operators are important because they possess several useful properties.

1. **All eigenvalues of a Hermitian operator are real.**
   This means that if. $A \mid \psi\rangle = \lambda \mid \psi\rangle$, then the eigenvalue λ is always a real number.

2. **Eigenvectors corresponding to different eigenvalues are orthogonal.**

   If $\mid \psi_1\rangle$ and $\mid \psi_2\rangle$ are eigenvectors associated with distinct eigenvalues, then
   $$\langle \psi_1 \mid \psi_2 \rangle = 0.$$

3. **Hermitian operators can be diagonalised.**
   This means that there exists a suitable basis consisting entirely of eigenvectors of the operator. In this basis, the matrix representation of the operator becomes diagonal, making calculations much simpler.

Because physical measurements always yield real values, Hermitian operators are used in quantum mechanics to represent observable quantities such as energy, momentum, angular momentum, and position.

Example of a Hermitian Matrix

Consider a Pauli Matrix, $\sigma_x = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$

Since $\sigma_x^\dagger = \sigma_x$, the matrix is Hermitian.

The eigenvalue equation gives $\lambda = \pm 1$,which are real eigenvalues. The corresponding normalized eigenvectors are

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix},$$

which are orthogonal.

Unitary Operators

An operator U is called **Unitary** if

$$U^{\dagger}U = UU^{\dagger} = I,$$

where I is the identity operator.

For a unitary operator,

$$U^{-1} = U^{\dagger}.$$

Thus, the adjoint of a unitary operator is also its inverse.

Unitary operators are important because they preserve the geometric properties of vectors.

Their important properties are:

1. **Preservation of vector lengths (norms)**

   If $|\psi\rangle$ is any vector, then

   $$\| U\psi \| = \| \psi \|$$

   Thus, a unitary transformation does not change the length of a vector.

2. **Preservation of inner products**

   For any two vectors $|\phi\rangle$ and $|\psi\rangle$,

   $$\langle U\phi \mid U\psi \rangle = \langle \phi \mid \psi \rangle$$

   Therefore, angles and orthogonality between vectors remain unchanged.

3. **Eigenvalues of a unitary operator have unit magnitude**

   If λ is an eigenvalue of a unitary operator, then

   $$|\lambda| = 1$$

Unitary operators describe transformations that preserve the overall structure of a vector space. Examples include rotations of coordinate systems and changes of basis between orthonormal sets of vectors.

Example of a Unitary Matrix

Consider the matrix

$$U = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

This matrix represents a rotation through an angle θ in a two-dimensional plane.

Since, $U^{T}U = I$,the matrix is orthogonal. For real matrices, orthogonal matrices are unitary matrices. Thus, U is a unitary operator.

Significance of Hermitian and Unitary Operators

Hermitian and unitary operators are among the most important classes of linear operators. Hermitian operators are associated with measurable physical quantities because they possess real eigenvalues. Unitary operators describe transformations that preserve lengths and angles and therefore maintain the geometric structure of the vector space. Together, these operators form the foundation of many applications in mathematical physics, particularly in matrix theory, quantum mechanics, and symmetry transformations.

Diagonalization of Linear Operators

The matrix representation of a linear operator depends on the basis chosen for the vector space. A particularly important choice is the basis formed by the eigenvectors of the operator. In such a basis, the matrix representation often takes a very simple form known as a **diagonal matrix**.

Consider a linear operator A acting on an n-dimensional vector space V. Suppose A possess a complete set of linearly independent eigenvectors

$$| e_1' \rangle,\ | e_2' \rangle,\ \ldots,\ | e_n' \rangle$$

with corresponding eigenvalues

$$\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n.$$

Then,

$$A | e_i' \rangle = \lambda_i | e_i' \rangle.$$

Since the eigenvectors form a basis of the vector space, they can be chosen as the basis for the matrix representation of A. In this basis, the matrix representation of the operator becomes

$$A' = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Thus, the matrix representation contains the eigenvalues on the diagonal, and all off-diagonal elements vanish. If A is represented in an arbitrary basis and P is the matrix whose columns are the eigenvectors of A, then the matrix representation in the eigenvector basis is obtained through the similarity transformation

$$D = P^{-1}AP$$

where D is the diagonal matrix

$$D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Therefore, a matrix can be diagonalized if it possesses a complete set of linearly independent eigenvectors. The columns of the matrix P are the eigenvectors of A, while the diagonal elements of D are the corresponding eigenvalues. Diagonalization is important because it simplifies many calculations involving powers of matrices, matrix functions, and operator transformations. In Quantum Mechanics, Hermitian operators can be diagonalized, and the diagonal elements correspond to the possible measured values of the associated physical observable.

3.6 EIGENVALUE PROBLEMS (INTRODUCTORY)

Let A be a linear operator acting on a linear vector space V. Consider the equation

$$A \mid x\rangle = \lambda \mid x\rangle$$

If there exists a non-zero vector $\mid x\rangle \in V$ satisfying this relation for some scalar λ (which may be real or complex), then λ is called an eigenvalue of the operator A, and the corresponding vector $\mid x\rangle$ is called an eigenvector associated with that eigenvalue.

The above equation indicates that the action of A on $\mid x\rangle$ does not alter its direction in the vector space; it merely scales the vector by the factor λ. Therefore, eigenvectors identify those special vectors whose orientation remains unchanged under the transformation represented by A.

Example (1): Eigenvectors of a Matrix Operator

Let V be the linear vector space of two-component column vectors and consider the linear operator

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

To determine the eigenvectors of A, we look for non-zero vectors that satisfy the eigenvalue equation

$$A \mid x\rangle = \lambda \mid x\rangle$$

Consider first the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

When the operator A acts on this vector, the two components are interchanged. Since both components are equal, the vector remains unchanged:

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The result is the original vector multiplied by the scalar 1. Comparing this with the eigenvalue equation, we conclude that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is an eigenvector of A, and the corresponding eigenvalue is

$$\lambda = 1.$$

Let us now examine another vector,

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Applying the operator Aonce again interchanges the two components, producing

$$A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Notice that the resulting vector is simply -1times the original vector. Hence,

$$A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

This shows that. $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ satisfies the eigenvalue equation with eigenvalue -1. Therefore, it is also an eigenvector of the operator A.

From the above calculations, we find that the operator A possesses two eigenvalue–eigenvector pairs:

$$\lambda_1 = 1, |\, x_1\rangle = \begin{bmatrix}1\\1\end{bmatrix},$$

and

$$\lambda_2 = -1, |\, x_2\rangle = \begin{bmatrix}1\\-1\end{bmatrix}$$

These results illustrate the basic idea of eigenvectors. Under the action of the operator A, the vectors

$$\begin{bmatrix}1\\1\end{bmatrix} \text{ and } \begin{bmatrix}1\\-1\end{bmatrix}$$

retain their original directions; only their magnitudes are scaled by the factors 1 and -1, respectively. Such vectors are precisely the eigenvectors of the operator, while the corresponding scaling factors are the eigenvalues.

Q1: Verify that $\psi(x) = e^{2x}$is an eigenfunction of $\frac{d}{dx}$.

Solution: $\frac{d}{dx}(e^{2x}) = 2e^{2x} = 2\,\psi(x)$

Eigenvalue $\lambda = 2$

Q2: Show that $\psi(x) = \sin 3x$ is an eigenfunction of $\frac{d^2}{dx^2}$.

Solution: $\frac{d^2}{dx^2}(\sin 3x) = -9\sin 3x = -9\,\psi(x)$

Comparing with the eigenvalue equation$\hat{A}\Psi = \lambda\Psi$

We have $\lambda = -9$

Therefore, $\psi(x)$ is an eigenfunction of $\dfrac{d^2}{dx^2}$

Q3: Find if $\binom{1}{0}$is an eigenvector of $A = \begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}$

Solution: $A\binom{1}{0} = \binom{2}{0} = 2\binom{1}{0}$

Thus, Eigenvalue$\lambda = 2$

Hence, $\binom{1}{0}$ is an eigonvector with $\lambda = 2$

Q4: Check if $\psi(x) = e^{-x}$ is an eigenfunction. Use operator $\hat{A} = \frac{\mathrm{d}}{dx}$+1

Solution: $\hat{A}\psi = \frac{d}{dx}(e^{-x}) + e^{-x} = -e^{-x} + e^{-x} = 0$

Since,$0 = 0 \cdot \psi(x)$

Eigen Value is $\lambda = 0$

Q 5: Find the eigenvalues of the operator $\hat{A} = \frac{d^2}{dx^2}$ subject to boundary conditions $\psi(0) = 0, \psi(L) = 0$

Solution:

Eigenvalue equation: $\frac{d^2\psi}{dx^2} = \lambda\psi$

The general solution is given by $\psi(x) = A\sin kx + B\cos kx, \lambda = -k^2$

Applying boundary conditions at $x = 0$: $\psi = 0 \Rightarrow B = 0$, we have

At $x = L$ $\psi(L) = A\sin kL = 0$

$$\Rightarrow \quad kL = n\pi$$

$$=> \ k = \frac{n\pi}{L}$$

Thus, $\lambda_n = -\frac{n^2\pi^2}{L^2}$

Now, eigenfunctions: $\psi_n(x) = \sin\frac{n\pi x}{L}$

Q 6: Find the eigenvalues of the matrix

$$[A] = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}.$$

Solution:

$$det([A] - \lambda I) = (4 - \lambda)(9 - \lambda) = 0$$

$$\lambda = 4, \ 9$$

Q7: Verify that the function $\psi(x) = e^{ikx}$ is an eigenfunction of the operator

$$\hat{A} = \frac{d}{dx}$$

Solution:

$$\hat{A}\psi = \frac{d}{dx}(e^{ikx}) = ike^{ikx}$$

$$\Rightarrow \hat{A}\psi = ik\psi$$

Thus, eigenvalue $\lambda = ik$.

Commuting Operators and Common Eigenvectors

Let A and B be two linear operators acting on a vector space V. If

$$AB = BA$$

or equivalently,

$$[A, B] = AB - BA = 0$$

the operators are said to commute. Suppose $|\ \psi\rangle$ is an eigenvector of A with eigenvalue λ, so that

$$A\ |\ \psi\rangle = \lambda\ |\ \psi\rangle$$

Operating with B and using the commutation relation,

$$A(B \mid \psi\rangle) = B(A \mid \psi\rangle) = B(\lambda \mid \psi\rangle) = \lambda(B \mid \psi\rangle)$$

Therefore, $B \mid \psi\rangle$ is also an eigenvector of A corresponding to the same eigenvalue λ. Under suitable conditions, commuting operators possess a common set of eigenvectors and can be represented simultaneously in a basis formed by these eigenvectors. If a complete set of common eigenvectors exists, both operators can be represented by diagonal matrices in the same basis. Such operators are said to be simultaneously diagonalizable**.** The existence of a common eigen basis greatly simplifies calculations involving several operators and plays an important role in the matrix representation of linear operators.

In Quantum Mechanics, observables represented by commuting operators possess a common set of eigenfunctions and can have simultaneous definite values. Conversely, non-commuting operators generally do not possess a complete common eigenbasis. This fact forms the mathematical basis of the Heisenberg uncertainty principle, which limits the simultaneous determination of certain pairs of physical observables.

CHECK YOUR PROGRESS

Q1. Define a linear operator. State and explain the conditions of linearity with suitable examples.

Q2. Distinguish between a linear operator and a non-linear operator with examples.

Q3. Verify whether the operator $\hat{A} = \frac{d}{dx} + x$ is linear.

Q4. Show that the operators x and $\frac{d}{dx}$do not commute and find their commutator.

Q5. Find the matrix representation of the operator

$$\hat{A}(x, y) = (3x + 2y,\ x + 4y)$$

with respect to the standard basis.

Q6. Verify that $\psi(x) = e^{ikx}$is an eigenfunction of $\frac{d}{dx}$and find the eigenvalue.

3.7 LET US SUM UP

- Linear operators are mappings that transform vectors or functions while preserving addition and scalar multiplication.
- They generalize the concept of transformations studied in vector algebra and are fundamental in mathematical physics.
- The defining property of a linear operator is linearity, expressed through additivity and homogeneity conditions.
- Linear operators can act on finite-dimensional vector spaces (matrix form) or infinite-dimensional spaces (differential operators).
- The algebra of operators includes addition, scalar multiplication, and multiplication (composition).

- Operator addition is commutative, whereas operator multiplication is generally non-commutative.
- The non-commutativity of operators is described using the commutator $[A, B] = AB - BA$, which has important physical implications.
- In quantum mechanics, non-commuting operators lead to fundamental results such as the uncertainty principle.
- A linear operator can be represented as a matrix once a basis is chosen, making calculations more convenient.
- Each column of the matrix represents the transformation of a corresponding basis vector under the operator.
- Eigenvalue problems arise naturally in the study of operators and are expressed as $A\psi = \lambda\psi$.
- Eigenvalues represent scaling factors, while eigenvectors or eigenfunctions represent invariant directions.
- In matrix form, eigenvalues are obtained from the characteristic equation $\det(A - \lambda I) = 0$.
- Eigenvalue analysis is essential in understanding physical systems, especially in quantum mechanics and wave phenomena.
- Overall, linear operators provide a unified framework connecting vector spaces, matrices, and physical observables.

✓3.8 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: A linear operator is a mapping $\hat{A}$ that transforms elements of a vector space such that it satisfies:

$$\hat{A}(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \hat{A} x_1 + \alpha_2 \hat{A} x_2$$

Conditions: 1. Additivity: $\hat{A}(x_1 + x_2) = \hat{A}x_1 + \hat{A}x_2$
2. Homogeneity: $\hat{A}(\alpha x) = \alpha \hat{A} x$

For example: $\hat{A} = \frac{d}{dx}$, $\hat{A}(x^2 + x) = 2x + 1$

Answer to Q2:

Linear Operator	**Non-Linear Operator**
Satisfy linearity Condition	Does not satisfy linearity condition
$A(x + y) = A(x) + A(y)$	$A(x + y) \neq A(x) + A(y)$
A(cx)=cA(x)	A(cx) ≠ cA(x)
Preserves addition and scalar multiplication	Does not preserve addition and scalar multiplication
Example: $A = \frac{d}{dx}$	Example: $A(x) = x^2$

Answer to Q3: Given operator: $\hat{A} = \frac{d}{dx} + x$

To check linearity, we verify: $\hat{A}(c_1 f + c_2 g) = c_1 \hat{A}(f) + c_2 \hat{A}(g)$ (A)

Applying operator to a linear combination

$$\hat{A}(c_1 f + c_2 g) = \frac{d}{dx}(c_1 f + c_2 g) + x(c_1 f + c_2 g)$$

$$= c_1 \frac{df}{dx} + c_2 \frac{dg}{dx} + c_1 x f + c_2 x g$$

$$= c_1 \left(\frac{df}{dx} + xf\right) + c_2 \left(\frac{dg}{dx} + xg\right)$$

Comparing with RHS of eq (A), we have

$$c_1 \hat{A}(f) + c_2 \hat{A}(g) = c_1 \left(\frac{df}{dx} + xf\right) + c_2 \left(\frac{dg}{dx} + xg\right)$$

Thus, $\hat{A}(c_1 f + c_2 g) = c_1 \hat{A}(f) + c_2 \hat{A}(g)$

Hence, the operator $\boxed{\hat{A} = \frac{d}{dx} + x}$ is linear**.**

Answer to Q4: Let the operators be: $\hat{A} = x, \hat{B} = \frac{d}{dx}$

Applying operators successively on a function $f(x)$

First, $\hat{A}\hat{B}f = x\frac{d}{dx}f = xf'$

Next. $\hat{B}\hat{A}f = \frac{d}{dx}(xf) = xf' + f$

The commutator is defined as: $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$

So, $[\hat{A}, \hat{B}]f = xf' - (xf' + f) = -f$

Answer to Q5: Given operator: $\hat{A}(x, y) = (3x + 2y,\ x + 4y)$

We need to find its matrix representation with respect to the standard basis:

$$\mathrm{e}_1 = (1,0), \mathrm{e}_2 = (0,1)$$

Applying operator to basis vectors

For $\mathrm{e}_1 = (\mathbf{1}, \mathbf{0})$**:** $\hat{A}(1,0) = (3 \cdot 1 + 2 \cdot 0,\ 1 + 4 \cdot 0) = (3,1)$

For $\mathrm{e}_2 = (\mathbf{0}, \mathbf{1})$**:** $\hat{A}(0,1) = (3 \cdot 0 + 2 \cdot 1,\ 0 + 4 \cdot 1) = (2,4)$

Thus, the matrix is formed by taking these as columns

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

Answer to Q6: Applying the operator $\frac{d}{dx}\left(e^{ikx}\right) = ik\, e^{ikx}$

Comparing with the eigenvalue equation, we have

An eigenfunction satisfies: $\hat{A}\psi = \lambda\psi$

Here, $\frac{d}{dx}(e^{ikx}) = ik\, e^{ikx} = \lambda\psi(x)$

Thus, $\lambda = ik$

Thus, $\psi(x) = e^{ikx}$ is an eigenfunction of $\frac{d}{dx}$, with eigenvalue ik

3.9 FURTHER READING

- Riley, K. F., Hobson, M. P., and Bence, S. J. (2006), *Mathematical Methods for Physics and Engineering*, 3rd Edition, Cambridge University Press.
- Dass, H. K., and Verma, R. (2010), *Mathematical Physics*, S. Chand & Company Ltd.
- Gupta, B. D. (2005), *Mathematical Physics*, Vikas Publishing House.
- Boas, M. L. (2006), *Mathematical Methods in the Physical Sciences*, Wiley.
- Arfken, G. (2013), Weber, H., and Harris, F., *Mathematical Methods for Physicists*, Academic Press.
- Kreyszig, E. (2011), *Advanced Engineering Mathematics*, Wiley.
- Chattopadhyay, P. K. (2012), *Mathematical Physics*, New Age International.

3.10 MODEL QUESTIONS

1. Show that the operators x and $\frac{d}{dx}$ do not commute and find their commutator.
2. Find the matrix representation of the operator $\hat{A}(x, y) = (x + y,\, 2x - y)$ with respect to the standard basis.
3. Explain why linearity of operators is important in physical systems.
4. Check if
$$\psi(x) = x^2$$
is an eigenfunction of the operator $\hat{A} = x\frac{d}{dx}$
5. If $\hat{A} = \frac{d}{dx}, \hat{B} = x$find $\hat{A}\hat{B}(f)$ and $\hat{B}\hat{A}(f)$
6. Verify whether the operator $\hat{A}(f) = x\frac{d}{dx}f$ is linear.
7. Define a linear operator. State and explain the conditions for linearity with examples.
8. Explain the algebra of linear operators, including addition, multiplication, and commutation, with suitable examples.
9. What is an eigenvalue problem for a linear operator? Explain the concept of eigenvalues and eigenfunctions with a simple example.

*** *** ***

UNIT 4: MATRICES

UNIT STRUCTURE

4.1 Learning Objectives
4.2 Introduction
4.3 Special Matrices
    4.3.1 Symmetric Matrix
    4.3.2 Anti-Symmetric Matrix
    4.3.3 Orthogonal Matrix
    4.3.4 Hermitian Matrix
    4.3.5 Anti-Hermitian Matrix
    4.3.6 Unitary
4.4 Eigenvalues and Eigenvectors
4.5 Cayley-Hamilton Theorem
4.6 Diagonalization of Matrix
4.7 Let Us Sum Up
4.8 Answers to Check Your Progress
4.9 Further Reading
4.10 Model Questions

4.1 LEARNING OBJECTIVES

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

- understand the mathematical structure and properties of matrices
- identify and classify special matrices used in physics
- determine eigenvalues and eigenvectors of matrices
- apply the Cayley–Hamilton theorem to simplify matrix calculations
- diagonalize matrices using similarity transformations.

4.2 INTRODUCTION

In the previous unit, the concept of linear operators was introduced as mappings that transform vectors or functions while preserving linearity. It was shown that these operators can be represented in matrix form once a suitable basis is chosen, thereby establishing a close connection between operator theory and linear algebra. The study of operator algebra and eigenvalue problems further highlighted the importance of matrices in understanding transformations and physical systems.

Building upon this foundation, the present unit focuses on the detailed study of matrices and their properties. Special classes of matrices such as symmetric, anti-symmetric, orthogonal, Hermitian, anti-Hermitian, and unitary matrices are introduced, which play an important role in mathematical physics. The unit also develops the theory of eigenvalues and eigenvectors in greater depth and introduces the Cayley–Hamilton theorem as a powerful tool in matrix analysis. Further, the process of diagonalization is discussed, which simplifies matrices and makes it easier to analyze linear transformations and physical systems. These concepts are essential in many areas of

physics, especially quantum mechanics, where matrices and operators are used to represent physical observables and transformations.

4.3 SPECIAL MATRICES

Matrices provide a systematic and compact way of representing linear transformations and sets of linear equations. In mathematical physics, they serve as concrete representations of linear operators, allowing abstract transformations to be expressed in a form suitable for computation and analysis. By arranging elements in rows and columns, matrices enable efficient manipulation of vectors and facilitate the study of transformations in vector spaces.

A matrix is defined as a rectangular array of numbers or functions arranged in rows and columns. In general, a matrix having m rows and n columns are called an $m \times n$matrix and is denoted by

$$A = [a_{ij}]$$

where $i = 1,2, \dots, m$ and $j = 1,2, \dots, n$.

Here, a_{ij}represents the element located in the ith row and jth column of the matrix. Matrices provide a convenient way of representing systems of linear equations.
For instance, consider the system of equations

$$x + 3y + 2z + 5t = 0,$$

$$4x + 2y + 5z + 3t = 0,$$

$$3x + 4y + 2z + 6t = 0.$$

By taking the coefficients of the variables x, y, z and t from these equations and arranging them in rows, we obtain the matrix

$$A = \begin{bmatrix} 1 & 3 & 2 & 5 \\ 4 & 2 & 5 & 3 \\ 3 & 4 & 2 & 6 \end{bmatrix}$$

This matrix has 3 rows and 4 columns, and hence it is called a 3×4matrix, containing a total of $3 \times 4 = 12$ elements. Each element of a matrix is uniquely identified by two subscripts: the first indicates the row number, and the second indicates the column number. Thus, the element a_{ij} denotes the entry in the ith row and jth column of the matrix.

However, not all matrices possess the same properties. Depending on the nature of their elements and the relationships between them, matrices can be classified into different types, each with distinct mathematical and physical significance. These special classes of matrices play an important role in simplifying problems and understanding physical systems.

4.3.1 Symmetric Matrix

A square matrix is said to be symmetric if its transpose is equal to the matrix itself. In other words, a matrix $A = [a_{ij}]$ is symmetric if

$$a_{ij} = a_{ji} \quad \text{for all } i, j,$$

or equivalently,

$$A^T = A$$

This means that the elements of the matrix are symmetric about the main diagonal (principal diagonal).

For example,

$$A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$$

is a symmetric matrix since each element above the main diagonal is equal to the corresponding element below it.

Important Points to remember:

- A symmetric matrix must always be a square matrix**.**
- The main diagonal elements (i.e., $a_{11}, a_{22}, a_{33}, \ldots$) can be any real numbers.
- If all elements above (or below) the diagonal are known, the entire matrix is determined.
- The sum of two symmetric matrices is also symmetric**.**
- The product of two symmetric matrices is not necessarily symmetric**.**
- If a matrix is symmetric, its eigenvalues are real.

4.3.2 Antisymmetric Matrix

A square matrix is said to be anti-symmetric (or skew-symmetric) if its transpose is equal to the negative of the matrix, that is,

$$A^T = -A.$$

In terms of elements, a matrix $A = [a_{ij}]$is anti-symmetric if $a_{ij} = -a_{ji}$for all values of $i$and j. This implies that the elements are antisymmetric about the main (principal) diagonal. An immediate consequence of this definition is that all the diagonal elements must be zero, since $a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$. For example, the matrix

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

is an anti-symmetric matrix.

Important points to remember:

- All diagonal elements are zero, since $a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$.
- The sum of two anti-symmetric matrices is also anti-symmetric.
- The product of two anti-symmetric matrices is not necessarily anti-symmetric.

- Anti-symmetric matrices are important in physics, especially in describing rotational quantities and angular momentum**.**
- They are associated with direction-dependent properties in physical systems.

4.3.3 Orthogonal Matrix

A non-singular square matrix is called an orthogonal matrix if its transpose is equal to its inverse, that is,

$$A^T = A^{-1}$$

From this relation, it follows that the inverse of an orthogonal matrix is also orthogonal, since taking the transpose of the inverse gives

$$(A^{-1})^T = (A^T)^{-1} = (A^{-1})^{-1} = A$$

Hence, the transpose of the inverse of an orthogonal matrix is the matrix itself.

Further, for an orthogonal matrix, we have

$$A^T A = I$$

where I is the identity matrix. Taking determinants on both sides, we obtain

$$|\ A^T A\ | = |\ A^T\ ||\ A\ | = |\ A\ |^2 = |\ I\ | = 1$$

Hence, the determinant of an orthogonal matrix must be

$$|\ A\ | = \pm 1$$

An important property of orthogonal matrices is that they preserve the length (or norm) of vectors. If a vector x is transformed into another vector y by the relation $y = Ax$, then

$$y^T y = x^T A^T A x = x^T x$$

Thus, the magnitude of the vector remains unchanged under the transformation. This shows that orthogonal matrices represent transformations such as rotations and reflections that preserve distances and angles in real vector spaces.

Example: Verification of an Orthogonal Matrix
Consider the rotation matrix

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Its transpose is $A^T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

Now, $A^T A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Multiplying row by column,

$$= \begin{pmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\sin\theta\cos\theta + \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{pmatrix}$$

Using $\sin^2\theta + \cos^2\theta = 1$

we obtain $A^T A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$

Hence, $A^T A = I$, therefore A is an orthogonal matrix.

Next, $\det(A) = \cos^2\theta + \sin^2\theta = 1$
Thus, $|\det(A)| = 1$ verifies the determinant property of an orthogonal matrix.

4.3.4 Hermitian Matrix

A square matrix $A = (a_{ij})$is said to be a Hermitian matrix if each element of the matrix is equal to the complex conjugate of the corresponding element obtained by interchanging the row and column indices. That is,

$$a_{ij} = \overline{a_{ji}}$$

Thus, a matrix is Hermitian if it is equal to its conjugate transpose, which is written as

$$A = A^{\dagger} = \overline{A^T}$$

This condition is both necessary and sufficient for a matrix to be Hermitian.

In a Hermitian matrix, the diagonal elements are always real, since each diagonal element must be equal to its own complex conjugate. The off-diagonal elements occur in complex conjugate pairs, symmetrically placed about the main diagonal.

For Example:
Consider the matrix

$$A = \begin{bmatrix} 1 & 2+3i & 3+i \\ 2-3i & 2 & 1-2i \\ 3-i & 1+2i & 5 \end{bmatrix}.$$

The transpose of A is

$$A^T = \begin{bmatrix} 1 & 2-3i & 3-i \\ 2+3i & 2 & 1+2i \\ 3+i & 1-2i & 5 \end{bmatrix},$$

and taking the complex conjugate gives

$$\overline{A^T} = \begin{bmatrix} 1 & 2+3i & 3+i \\ 2-3i & 2 & 1-2i \\ 3-i & 1+2i & 5 \end{bmatrix}.$$

Since $A = \overline{A^T}$the matrix A is Hermitian.
Hermitian matrices are of great importance in physics, particularly in quantum mechanics, because they have real eigenvalues and represent observable physical quantities such as energy and momentum.

4.3.5 Anti-Hermitian Matrix

A square matrix $A = (a_{ij})$ is said to be anti-Hermitian (or skew-Hermitian) if its conjugate transpose is equal to the negative of the matrix itself, that is,

$$A^{\dagger} = -A$$

or equivalently, in terms of elements,

$$a_{ij} = \overline{-a_{ji}}$$

This condition is both necessary and sufficient for a matrix to be anti-Hermitian. An important consequence of this definition is that the diagonal elements of an anti-Hermitian matrix are either zero or purely imaginary. In fact, if a diagonal element is written as $a_{ii} = a + ib$, then using the condition $a_{ii} = -\overline{a_{ii}}$, we obtain $a = 0$, showing that a_{ii}must be purely imaginary. Anti-Hermitian matrices can be viewed as a generalization of real skew-symmetric matrices to the complex case. Moreover, if a matrix is anti-Hermitian, its inverse (if it exists) is also anti-Hermitian.

Any square matrix can be expressed as the sum of a Hermitian and an anti-Hermitian matrix, where the anti-Hermitian part is given by $\frac{1}{2}(A - A^{\dagger})$. For example, the matrix

$$\begin{pmatrix} i & 2-3i & 4+5i \\ -(2+3i) & 0 & 2i \\ -(4-5i) & 2i & -3i \end{pmatrix}$$

is an anti-Hermitian matrix, since it satisfies $A^{\dagger} = -A$.

4.3.6 Unitary Matrix

A square matrix A is said to be a unitary matrix if its conjugate transpose is equal to its inverse, that is,

$$A^{\dagger} = A^{-1}$$

or equivalently,

$$A^{\dagger}A = I$$

where I is the identity matrix. This condition implies that the rows and columns of a unitary matrix form an orthonormal set in a complex vector space. If the elements of $A$are real, then $A^{\dagger} = A^{T}$, and hence a real unitary matrix reduces to an orthogonal matrix. The inverse of a unitary matrix is also unitary, since

$$(A^{-1})^{\dagger} = (A^{\dagger})^{-1} = (A^{-1})^{-1}$$

Further, taking determinants of the relation $A^{\dagger}A = I$, we obtain

$$|\, A^{\dagger} \,||\, A \,| = |\, A \,|^{*} |\, A \,| = 1$$

Hence, $$|\, \det(A) \,|^{2} = 1,$$

or equivalently, $$\boxed{|\, \det(A) \,| = 1.}$$

Therefore**,** the determinant of a unitary matrix has unit modulus. The set of all unitary matrices forms the group $U(n)$, while the subgroup satisfying det U=1 is the special unitary group $SU(n)$.

An important property of unitary matrices is that they preserve the norm of complex vectors. If a vector x is transformed as $y = Ax$, then

$$y^\dagger y = x^\dagger A^\dagger A x = x^\dagger x$$

which shows that the length of the vector remains unchanged under the transformation. Thus, unitary matrices represent transformations that preserve inner products and norms in complex vector spaces, playing a fundamental role in quantum mechanics and related areas of physics. For example, the matrix

$$A = \begin{pmatrix} \frac{1+i}{2} & \frac{-1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{pmatrix}$$

is unitary since it satisfies $A^\dagger A = I$.

Proof: For the given matrix

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix},$$

its conjugate transpose is

$$A^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.$$

Now, $A^\dagger A = \frac{1}{2}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$

Computing each element, $(1,1) = 1 + (-i)(i) = 1 + 1 = 2,$

$$(1,2) = i - i = 0,$$

$$(2,1) = -i + i = 0,$$

$$(2,2) = (-i)(i) + 1 = 1 + 1 = 2.$$

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

Therefore, $A^\dagger A = I$which confirms that A is a unitary matrix.

4.4 EIGENVALUES AND EIGENVECTORS

Definition of Eigenvalues and Eigenvectors

Let A be a square matrix of order$n$. If there exists a **non-zero vector**X such that

$$AX = \lambda X$$

where λ is a scalar, then λ is called an eigenvalueof the matrix A, and the corresponding non-zero vector X is called the eigenvector associated with that eigenvalue.

In other words, when a matrix operates on a vector and the result is a scalar multiple of the same vector, the scalar is called the eigenvalue, and the vector is called the eigenvector**.**

Let, $A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix}$

$\Rightarrow \quad A - \lambda I = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-\lambda & 3-0 & 2-0 \\ 3-0 & 1-\lambda & 1-0 \\ 2-0 & 1-0 & 2-\lambda \end{bmatrix}$

This is known as the characteristic matrix.

Characteristic Equation: The equation $|A - \lambda I|$=0 is known as characteristic equation of the matrix A . The determinant $|A - \lambda I|$=0 can be expanded into a polynomial known as the characteristic polynomial of matrix A. For eg:

$$A = \begin{vmatrix} 1-\lambda & 3 & 2 \\ 3 & 1-\lambda & 1 \\ 2 & 1 & 2-\lambda \end{vmatrix}$$

$A = \lambda^3 - 7\lambda^2 + 11\lambda - 5$ is known as the characteristic polynomial

and $A = \lambda^3 - 7\lambda^2 + 11\lambda - 5 = 0$is known as the characteristic equation**.**

Characteristic Roots or Eigenvalues: The roots of characteristic equation $|A - \lambda I| = 0$ is known as the characteristic or eigen values of matrix A.

For example: $\lambda^3 - 7\lambda^2 + 11\lambda - 5 = 0$

$\Rightarrow \quad (\lambda - 1)(\lambda - 1)(\lambda - 5) = 0$

Thus, $\lambda = (1, 1, 5)$ are the characteristic roots or eigenvalues of matrix A.

Trace of a Matrix: The **trace** of a square matrix $A = [a_{ij}]$ of order n, denoted by $\mathrm{Tr}(A)$, is defined as the sum of its principal diagonal elements. Thus,

$$\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

Properties of Eigenvalues

Important properties of Eigenvalues:

1. Any square A and its transpose A' have the same eigenvalues.
2. The sum of the eigenvalues of a matrix is equal to its trace, i.e.,

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{Tr}(A)$$

3. The product of the eigenvalues of a matrix A is equal to the determinant of A.
4. If eigenvalues of matrix A are $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the eigenvalues of

$i)$ kA are $k\lambda_1, k\lambda_2, \ldots, k\lambda_n$ ii) A^m are $\lambda_1^m, \lambda_2^m, \ldots, \lambda_n^m$ iii) $A^{-1}$are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$

Derivation of the Trace–Eigenvalue and Determinant–Eigenvalue Relations

Suppose A is an $n \times n$ square matrix whose eigenvalues are. $\lambda_1, \lambda_2, \ldots, \lambda_n$.
The characteristic polynomial of A is

$$\det(A - \lambda I) = 0$$

Expanding the determinant gives the characteristic polynomial in the form

$$\det(A - \lambda I) = (-\lambda)^n + \mathrm{Tr}(A)(-\lambda)^{n-1} + \cdots + \det(A)$$

where,

- $\mathrm{Tr}(A)$ denotes the trace of the matrix,
- $\det(A)$ denotes the determinant of the matrix.

On the other hand, since the roots of the characteristic polynomial are the eigenvalues,

$$\det(A - \lambda I) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

According to **Vieta's formulae**, comparison of coefficients gives

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{Tr}(A)$$

and $$\lambda_1\lambda_2\cdots\lambda_n = \det(A)$$

Hence,

- The sum of the eigenvalues is equal to the trace of the matrix.
- The product of the eigenvalues is equal to the determinant of the matrix.

Let us understand with an example:

Example 1: Find the eigenvalues of the matrix $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

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

The characteristic equation of A is $|A - \lambda I| = 0$

$$\Rightarrow \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0$$

Expanding the determinant with the help of third row, we have

$\Rightarrow \quad (1-\lambda)[(2-\lambda)^2 - 1] = 0$

$\Rightarrow \quad (1\text{-}\lambda)(\lambda^2 - 4\lambda + 4 - 1) = 0$

$\Rightarrow \quad (1\text{-}\lambda)(\lambda^2 - 4\lambda + 3) = 0$

$\Rightarrow \quad (1\text{-}\lambda)(\lambda - 3)(\lambda - 1) = 0$

$\Rightarrow \quad \lambda = 1,1,3$

Thus, the eigenvalues of matrix $A = 1, 1, 3$. Now, let us verify its trace property:
The trace of a matrix is the sum of its principal diagonal elements.

$$\mathrm{Tr}(A) = 2 + 2 + 1 = 5.$$

The sum of the eigenvalues is $\lambda_1 + \lambda_2 + \lambda_3 = 1 + 1 + 3 = 5.$

Hence, $\mathrm{Tr}(A) = \lambda_1 + \lambda_2 + \lambda_3 = 5$

Thus, the sum of the eigenvalues is equal to the trace of the matrix.

Let us verify its determinant property:

The determinant of the matrix is

$$\det(A) = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= 1 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 1(4-1) = 3$$

The product of the eigenvalues is $\lambda_1\lambda_2\lambda_3 = 1 \times 1 \times 3 = 3$

Therefore, $\det(A) = \lambda_1\lambda_2\lambda_3 = 3$.Thus, the product of the eigenvalues is equal to the determinant of the matrix.

Example 2: Prove that the matrix A and its transpose A' has the same characteristic roots.

Proof: The characteristic equation of the matrix A is $\det(A - \lambda I) = 0$.

The characteristic equation of the transpose matrix $A^T$is

$$\det(A^T - \lambda I) = 0.$$

Now, $A^T - \lambda I = A^T - \lambda I^T = (A - \lambda I)^T$

Since $I^T = I$

Using the property that the determinant of a matrix is equal to the determinant of its transpose,

$$\det(A^T - \lambda I) = \det[(A - \lambda I)^T] = \det(A - \lambda I).$$

Hence, $\det(A^T - \lambda I) = \det(A - \lambda I)$

Therefore, the characteristic equations of A and $A^T$are identical and have the same roots.

Hence**,** a matrix and its transpose have the same eigenvalues**.**

Properties of Eigenvectors

1. The eigen vector X of a matrix A is not unique.
2. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct eigen values of $n \times n$ matrix then corresponding eigen vectors $X_1, X_2, \ldots, X_n$ form a linearly independent set.
3. If two or more eigen values are equal it may or may not be possible to get linearly independent eigen vectors corresponding to the equal roots.
4. Two eigen vectors X_1 and X_2 are called orthogonal vectors if $X_1'X_2$= 0
5. Eigen vectors of a symmetric matrix corresponding to different eigen values are orthogonal.

Normalized form of vectors: To find normalized form of a matrix $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, each element is to be divided by $\sqrt{a^2 + b^2 + c^2}$.

For example, Normalized form of matrix $\begin{bmatrix}2\\3\\4\end{bmatrix}$ is $\begin{bmatrix}2/5\\3/5\\4/5\end{bmatrix}$, since $\sqrt{2^2+3^2+4^2}=5$

Example 3: Find the eigen vectors of the matrix $A=\begin{bmatrix}1&-3&3\\3&-5&3\\6&-6&4\end{bmatrix}$ with eigen values (-2), (-2) and 4.

Solution: The characteristic equation of A is $|A-\lambda I|=0$. Let the eigen vectors for the eigen values (-2), (-2) and 4 are $X_1, X_2$and X_3 respectively.

For $\lambda=-2$, $|A-\lambda I|X_1=0$

$|A+2I|X_1=0$

$$\begin{bmatrix}1&-3&3\\3&-5&3\\6&-6&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

⇨ $3x_1-3x_2+3x_3=0$

⇨ $x_1-x_2+x_3=0$

If $x_1=x_2=1, \quad x_3=0$

And $x_1=1, x_2=-1, x_3=-2$

∴ $X_1=\begin{bmatrix}1\\1\\0\end{bmatrix}$ and $X_2=\begin{bmatrix}1\\-1\\-2\end{bmatrix}$

For $\lambda=4$, $|A-\lambda I|X_3=0$

⇨ $|A-4I|X_3=0$

⇨ $$\begin{bmatrix}-3&-3&3\\3&-9&3\\6&-6&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

⇨ $-x_1-x_2+x_3=0$

⇨ $x_1-3x_2+x_3=0$

⇨

⇨ $\frac{x_1}{-1+3}=\frac{x_2}{1+1}=\frac{x_3}{3+1}$

⇨

⇨ $\frac{x_1}{2}=\frac{x_2}{2}=\frac{x_3}{4}$

⇨

⇨ $\frac{x_1}{1}=\frac{x_2}{1}=\frac{x_3}{2}$ $\Rightarrow X_3=\begin{bmatrix}1\\1\\2\end{bmatrix}$

4.5 CAYLEY-HAMILTON THEOREM

This theorem states that '*every square matrix satisfies its own characteristic equation'.*

If $|A-\lambda I| = (-1)^n(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_n)$ be the characteristic polynomial of $n \times n$ matrix $A = (a_{ij})$, then the equation,

$$X^n + a_1X^{n-1} + a_2X^{n-2} + \cdots + a_nI = 0 \text{ is satisfied by } X = A$$

That is $A^n + a_1A^{n-1} + a_2A^{n-2} + \cdots + a_nI = 0$

Proof of Cayley-Hamilton Theorem

Proof: Since the elements of the matrix $A-\lambda I$ are linear functions of λ (i.e., at most first degree in λ), the elements of the adjoint matrix adj$(A-\lambda I)$will be polynomials in λ of degree not exceeding $n-1$. Therefore, the adjoint matrix adj$(A-\lambda I)$can be expressed as a matrix polynomial in λ.

$$Adj.(A-\lambda I) = B_0\lambda^{n-1} + B_1\lambda^{n-2} + \cdots . + B_{n-1}$$

where, $B_0, B_1, \ldots, B_{n-1}$ are $n \times n$ matrices, their elements being polynomial in λ.

We know that, $(A-\lambda I)\,Adj(A-\lambda I) = |A-\lambda I|I$

$(A-\lambda I)(B_0\lambda^{n-1} + B_1\lambda^{n-2} + \cdots .. + B_{n-1}) = (-1)^n(\lambda^n + a_1\lambda^{n-1} + \cdots . + a_n)I$

Equating coefficient of like power of λ on both sides, we get

$$-IB_0 = (-1)^nI$$
$$AB_0 - IB_1 = (-1)^na_1I$$
$$AB_1 - IB_2 = (-1)^na_2I$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$AB_{n-1} = (-1)^na_nI$$

On multiplying the equation by $A^n, A^{n-1}, \ldots, I$ respectively and adding, we get

$$0 = (-1)^n[A^n + a_1A^{n-1} + \cdots . + a_nI]$$

Thus, $$[A^n + a_1A^{n-1} + \cdots . + a_nI] = 0$$

Let us understand with an example:

Let, A be a square matrix and its characteristic equation be

$$\lambda^3 - 2\lambda^2 + 3\lambda - 4 = 0 \qquad (1)$$

From the characteristic equation, it seems that the Cayley Hamilton Equation is satisfied by matrix A given by

$$A^3 - 2A^2 + 3A - 4 = 0 \qquad (2)$$

Let us find A^{-1}. Multiplying equation (2) by A^{-1}, we have

$$A^2 - 2A + 3I - 4A^{-1} = 0$$

⇨ $$A^{-1} = \frac{1}{4}[A^2 - 2A + 3I]$$

Example 4: Verify Cayley-Hamilton theorem for the matrix A and hence find A^{-1}

Solution: Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Thus, its characteristic equation is $|A - \lambda I| = 0$

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0$$

$$(2-\lambda)(1-\lambda) - 1 = 0$$
$$2 - 3\lambda + \lambda^2 - 1 = 0$$
$$\lambda^2 - 3\lambda + 1 = 0 \qquad (1)$$

According to Cayley-Hamilton theorem, matrix A satisfy its own characteristic equation:

$$A^2 - 3A + I = 0 \qquad (2)$$

Thus, $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

$$3A = 3\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 3 & 3 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore, $\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 6 & 3 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1+1 & 3-3 \\ 3-3 & 3-3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Thus, Cayley Hamilton theorem is satisfied.
Now, multiplying equation (2) by A^{-1}, we have:

$$A^2A^{-1} - 3AA^{-1} + IA^{-1} = 0$$

⇨ $A - 3I + A^{-1} = 0$

⇨ $A^{-1} = 3I - A$

⇨ $A^{-1} = 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

⇨ $A^{-1} = \begin{bmatrix} 3-2 & 0-1 \\ 0-3 & 3-1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

Ans: $A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

Example 5: Verify Cayley Hamilton Theorem for the matrix A and find A^{-1}

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Proof: Consider, the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

Now, the characteristic equation is given by

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{bmatrix}$$

Since the matrix is triangular, the determinant is the product of diagonal elements,

$$(1-\lambda)(1-\lambda)(2-\lambda) = 0$$
$$(1-\lambda)^2(2-\lambda) = 0$$

Expanding, $\lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0$

Thus, the characteristic equation is

$$\lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0 \qquad (1)$$

According to the theorem, the matrix satisfies its own characteristic equation

$$A^3 - 4A^2 + 5A - 2I = 0 \qquad (2)$$

$$\text{Now, } A^3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 7 \\ 0 & 0 & 8 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}[c] = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Now, } A^3 - 4A^2 + 5A - 2I = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 7 \\ 0 & 0 & 8 \end{bmatrix} - 4\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix} + 5\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, Cayley–Hamilton theorem is verified.

From the characteristic equation (2), we have

$$A^3 - 4A^2 + 5A - 2I = 0$$

Multiplying by A^{-1}, we have

$$A^2 - 4A + 5I - 2A^{-1} = 0$$

$$A^{-1} = \frac{1}{2}(A^2 - 4A + 5I)$$

Substituting the values, we have. $A^2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$

$$4A = \begin{bmatrix} 4 & 4 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 8 \end{bmatrix}$$

$$5I = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\therefore \quad A^2 - 4A + 5I = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{or,} \quad A^{-1} = \frac{1}{2}\begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

4.6 DIAGONALIZATION OF MATRIX

Diagonalization of a matrix Arefers to the process of transforming the matrix A into a diagonal matrix D. If the matrix A is related to the diagonal matrix D through a modal matrix P by the relation

$$D = P^{-1}AP$$

then the matrix $D$is known as the spectral matrix of A. The columns of the modal matrix P are formed by the eigenvectors of the matrix A.

Theorem: *If a square matrix A of order n has n linearly independent eigen vectors, then a matrix P can be found such that $P^{-1}AP$ is a diagonal matrix*

NOTE:

1) The square matrix P that diagonalizes the matrix $A$is obtained by arranging the eigenvectors of A as the columns of the matrix P. When these eigenvectors form a square matrix, the resulting diagonal matrix contains the eigenvalues of Aalong its principal diagonal.
2) The transformation of a matrix A into $P^{-1}AP$ is referred to as a similarity transformation**.**
3) The process of reducing a matrix A to a diagonal matrix is therefore a special case of similarity transformation.
4) The matrix P, which diagonalizes A, is called the modal matrix of A, and the resulting diagonal matrix $D$is known as the spectral matrix of A.

Example 6: Diagonalize the matrix

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}.$$

Solution:

The characteristic equation is

$$|A - \lambda I| = 0,$$

$$\begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)(3-\lambda)-2=0$$

$$\lambda^2-7\lambda+10=0$$
$$(\lambda-5)(\lambda-2)=0$$

Hence, the eigenvalues are

$$\boxed{\lambda_1=5,\lambda_2=2.}$$

For $\lambda=5$,

$$(A-5I)X=0\Rightarrow\begin{bmatrix}-1 & 1\\ 2 & -2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0,$$

which gives $x=y$. Hence, an eigenvector is. $X_1=\begin{bmatrix}1\\1\end{bmatrix}$.

For $\lambda=2$, $(A-2I)X=0\Rightarrow\begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0,$

which gives $2x+y=0$. Taking $x=1$, we obtain

$$X_2=\begin{bmatrix}1\\-2\end{bmatrix}$$

Thus, the modal matrix is

$$P=\begin{bmatrix}1 & 1\\ 1 & -2\end{bmatrix}$$

and $P^{-1}=\frac{1}{3}\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}$

Finally, $D=P^{-1}AP=\begin{bmatrix}5 & 0\\ 0 & 2\end{bmatrix}$

Hence, the matrix is diagonalized as

$$\boxed{P^{-1}AP=D=\begin{bmatrix}5 & 0\\ 0 & 2\end{bmatrix}}$$

Example 7: Find the eigen values, eigen vectors and the modal matrix and diagonalized the matrix below:

$$A=\begin{bmatrix}4 & 1 & 1\\ 1 & 4 & 1\\ 1 & 1 & 4\end{bmatrix}$$

Solutions: Characteristic equation is given by: $|A-\lambda I|=0$

$$|A-\lambda I|=\begin{bmatrix}4-\lambda & 1 & 1\\ 1 & 4-\lambda & 1\\ 1 & 1 & 4-\lambda\end{bmatrix}=0$$

Expanding, we have

$$(4-\lambda)^3-3(4-\lambda)+2=0$$

$$\lambda^3-12\lambda^2+45\lambda-54=0$$

Factorizing, we have,

$$(\lambda-6)(\lambda-3)^2=0$$

Thus, the eigenvalues are

$$\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 3$$

Now, for $\lambda_1 = 6$ $\quad (A - 6\,I)X = 0$

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

Solving, $\quad x_1 = x_2 = x_3$

$$\therefore X_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 3$ $(A - 3I)X = 0$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} X = 0$$

Two independent eigenvectors are

$$X_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, X_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

The modal matrix P is formed by placing eigenvectors as columns

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

P^{-1} is determined by the formula $\quad P^{-1} = \frac{1}{|P|} adjP$

where $|\,P\,|$= determinant of P, adj(P)= adjoint of P

Diagonalizing the given matrix, we have

$$D = P^{-1}AP = P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Thus, the matrix A is diagonalized.

CHECK YOUR PROGRESS

Q1. Find all the Eigen values the matrix $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$

Q2. Find the eigen values and eigen vectors of $\begin{bmatrix} 5 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 5 \end{bmatrix}$.

Q3. Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

i. Find the characteristic equation of the matrix A.
ii. Verify that the matrix A satisfies its own characteristic equation (Cayley–Hamilton theorem).
iii. Using the Cayley–Hamilton theorem, find A^{-1}

Q4. Consider, the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$

i. Determine thecharacteristic equation of the matrix A.
ii. Verify the Cayley–Hamilton theorem for this matrix.
iii. Use the theorem to evaluate A^{-1}.
iv. Hence find A^4 using the characteristic equation.

Q5: Diagonalize the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find eigenvalues, eigenvectors, modal matrix P, Diagonal matrix D.

Q6: Diagonalize the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}$

Find eigenvalues, eigen vectors, modal matrix P, Diagonal matrix D.

4.7 LET US SUM UP

- Matrices are rectangular arrays used to represent linear transformations and systems of equations.
- Each element a_{ij} of a matrix is identified by its row and column position.
- A symmetric matrix satisfies $A^T = A$ and has elements symmetric about the main diagonal.
- An anti-symmetric matrix satisfies $A^T = -A$, and all diagonal elements are zero.
- An orthogonal matrix satisfies $A^T A = I$, and it preserves lengths and angles of vectors.
- A Hermitian matrix satisfies $A^\dagger = A$, and its eigenvalues are always real.

- An anti-Hermitian matrix satisfies $A^{\dagger} = -A$, and its diagonal elements are purely imaginary or zero.
- A unitary matrix satisfies $A^{\dagger}A = I$ and preserves the norm of complex vectors.
- Eigenvalues and eigenvectors are defined by $AX = \lambda X$.
- Eigenvalues are obtained from the characteristic equation $| A - \lambda I | = 0$.
- The sum of eigenvalues equals the trace, and the product equals the determinant.
- Eigenvectors corresponding to distinct eigenvalues are linearly independent.
- Every square matrix satisfies its own characteristic equation.
- It is useful for finding matrix powers and inverses.
- A matrix is diagonalizable if it has sufficient independent eigenvectors.
- Diagonalization is given by $D = P^{-1}AP$.
- The matrix Pcontains eigenvectors, and D contains eigenvalues.
- Diagonalization simplifies matrix calculations and physical analysis.

✓4.8 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: 5, -3, -3

Answer to Q2: Given, $A = \begin{bmatrix} 5 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 5 \end{bmatrix}$

The eigenvalues are $\lambda = -2, 4, 6$

1. For $\lambda = -2$ $(A + 2I)X = 0.$

   Gives $x = 0, y =$ arbitrary, $z = 0.$

   Taking $y = 1, \quad X_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

2. For $\lambda = 4,$ $(A - 4I)X = 0$

   Gives $x = -z.$

   Taking$z = 1, \quad X_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

3. For $\lambda = 6,$ $(A - 6I)X = 0$

   gives. $x = z,$

   Taking $z = 1, \; X_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Verification, For $X_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$AX_3 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 6X_3.$$

Hence, the eigenvector corresponding to $\lambda = 6$ is verified.

Answer to Q3: (i) $\lambda^2 - 4\lambda + 3 = 0$ (ii) Yes. (iii) $A^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

Answer to Q4: (i) $\lambda^3 - 5\lambda^2 + 7\lambda - 3 = 0$

(ii) Given, $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$

The characteristic equation obtained in part (i) is

$$\lambda^3 - 5\lambda^2 + 7\lambda - 3 = 0.$$

According to the Cayley–Hamilton theorem, every square matrix satisfies its own characteristic equation. Therefore,

$$A^3 - 5A^2 + 7A - 3I = 0.$$

Calculating A^2

$$A^2 = \begin{bmatrix} 1 & 4 & 8 \\ 0 & 1 & 8 \\ 0 & 0 & 9 \end{bmatrix}$$

Calculate A^3

$$A^3 = A^2A = \begin{bmatrix} 1 & 6 & 33 \\ 0 & 1 & 26 \\ 0 & 0 & 27 \end{bmatrix}$$

Computing, we have $A^3 - 5A^2 + 7A - 3I$

Substituting the matrices,

$$= \begin{bmatrix} 1 & 6 & 33 \\ 0 & 1 & 26 \\ 0 & 0 & 27 \end{bmatrix} - 5\begin{bmatrix} 1 & 4 & 8 \\ 0 & 1 & 8 \\ 0 & 0 & 9 \end{bmatrix} + 7\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 6 & 33 \\ 0 & 1 & 26 \\ 0 & 0 & 27 \end{bmatrix} + \begin{bmatrix} -5 & -20 & -40 \\ 0 & -5 & -40 \\ 0 & 0 & -45 \end{bmatrix} + \begin{bmatrix} 7 & 14 & 7 \\ 0 & 7 & 14 \\ 0 & 0 & 21 \end{bmatrix} + \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Adding the corresponding elements,

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, $\boxed{A^3 - 5A^2 + 7A - 3I = 0.}$

Therefore, the matrix satisfies its own characteristic equation, verifying the Cayley–Hamilton theorem.

(iii) $A^{-1} = \frac{1}{3}\begin{bmatrix} 3 & -6 & 3 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ (iv) $A^4 = \begin{bmatrix} 1 & 8 & 80 \\ 0 & 1 & 80 \\ 0 & 0 & 81 \end{bmatrix}$

Answer to Q5: The eigenvalues are $\lambda_1 = 3, \lambda_2 = 1$, $X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$,

$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

Answer to Q6: The eigenvalues.$\lambda_1 = 1, \lambda_2 = 4, \lambda_3 = 2$, $X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

4.9 FURTHER READING

- Arfken, G. B., Weber, H. J., and Harris, F. E. (2013), *Mathematical Methods for Physicists*, Academic Press, Amsterdam.
- Boas, M. L. (2006), *Mathematical Methods in the Physical Sciences*, Wiley India, New Delhi.
- Riley, K. F., Hobson, M. P., and Bence, S. J. (2006), *Mathematical Methods for Physics and Engineering*, Cambridge University Press, Cambridge.
- Spiegel, M. R., Lipschutz, S., and Liu, J. (2009), *Schaum's Outline of Linear Algebra*, McGraw-Hill, New York.
- Dass, H. K. and Verma, Rama (2011), *Mathematical Physics*, S. Chand & Company Ltd., New Delhi.
- Jordan, T. F. (1969), *Linear Operators for Quantum Mechanics*, John Wiley & Sons, New York.

4.10 MODEL QUESTIONS

1. Find the characteristic roots of $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$
2. If $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 \\ -1 & 0 \\ 2 & -1 \end{bmatrix}$. Obtain the product AB and explain why BA is not defined.

3. Verify $A = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$ is orthogonal.

4. Show that the eigenvalues of a Hermitian matrix are always real.

5. If A is a unitary matrix, show that A^{-1} is also unitary.

6. State the Cayley–Hamilton theorem**.**

7. Prove that eigenvectors of a symmetric matrix corresponding to distinct eigenvalues are orthogonal.

8. Find e^A for A = $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ using the Cayley-Hamilton theorem.

9. Prove that the eigenvalues of a unitary matrix have unit modulus.

10. Find the modal matrix and diagonalize $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

11. Using the Cayley-Hamilton theorem, express A^4 in terms of lower powers of A for A = $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

*** *** ***

UNIT 5: TENSOR ANALYSIS

UNIT STRUCTURE

5.1 Learning Objectives
5.2 Introduction
5.3 Basics of Tensors
5.4 Covariant and Contravariant Tensors
5.5 Symmetric and Anti-symmetric Tensor
5.6 Four Vectors
5.7 Line Elements
5.8 Metric Tensor
5.9 Associated Tensors
5.10 Let Us Sum Up
5.11 Answers to Check Your Progress
5.12 Further Reading
5.13 Model Questions

5.1 LEARNING OBJECTIVES

After going through this unit, you will be able to

- explain the concepts and significance of scalars, vectors, and tensors in physics
- distinguish between covariant and contravariant tensors and their transformations
- explain the line element and distance in general coordinate systems
- describe the metric tensor and its geometrical significance
- apply the metric tensor in tensor operations and physical representations.

5.2 INTRODUCTION

In the previous unit, we studied the mathematical foundations of vector spaces, including basis, dimension, orthogonality, orthonormality, and the Gram–Schmidt orthonormalization process. These concepts provided the framework for understanding mathematical structures used in physical systems. Building on that foundation, this unit introduces the concepts of scalars, vectors, and tensors, which are essential tools for describing physical quantities in mathematics and physics.

Some physical quantities, such as mass, temperature and pressure, are completely described by a single numerical value and are called **scalars**, whereas quantities such as displacement, velocity and force require both magnitude and direction and are represented by **vectors**. Tensors provide a more general mathematical framework for describing physical quantities that depend on several directions or coordinate systems; in fact, scalars and vectors are tensors of rank zero and rank one, respectively. In this unit, we shall study the basic concepts of tensors, covariant and contravariant tensors, their transformation laws, the line element, the metric tensor and associated tensors, which form the foundation of tensor analysis and have important applications in mechanics, electromagnetism and general relativity.

5.3 BASICS OF TENSORS

The study of tensors begins with the description of space using coordinates. In ordinary three-dimensional space, any point is uniquely specified by three real numbers, commonly written as (x, y, z). For convenience and generalization, these coordinates are often denoted as (x_1, x_2, x_3), where $x_1 = x$, $x_2 = y$, and $x_3 = z$. In compact notation, we represent them as x_i, where the index $i = 1,2,3$ indicates the coordinate. This indexed form is particularly useful in advanced mathematical formulations, as it allows expressions to be written in a concise and generalized manner. The idea of representing quantities using indices serves as the foundation for tensor analysis.

Extension to N-Dimensional Space

The concept of coordinates can be extended beyond three dimensions to an n-dimensional space. In such a space, a point is described by n independent numbers written as

$$(x_1, x_2, x_3, \dots, x_n)$$

Each coordinate specifies the position of the point along a particular axis in the chosen coordinate system. This generalization is essential in tensor analysis, as many physical and mathematical problems involve spaces with more than three dimensions. The use of indexed notation x_i, where $i = 1,2, \dots, n$, provides a unified way to describe points in any number of dimensions and simplifies the formulation of transformation laws and tensor equations.

Coordinate Transformation

In tensor analysis, the components of a physical quantity depend on the chosen coordinate system. Therefore, it is important to understand how the coordinates of a point and the corresponding tensor components transform from one coordinate system to another. Let $x_i$and $\hat{x}_i$denote the coordinates of the same point in two different coordinate systems related by continuous and one-to-one transformation equations.
Thus, the relation between the two coordinate systems is given by the transformation equations:

$$\hat{x}_i = \hat{x}_i(x_1, x_2, \dots, x_n), i = 1,2, \dots, n \qquad (5.1)$$

The equation (5.1) is said to be a transformation of the coordinates from one coordinate $(x_1, x_2, x_3, \dots, x_n)$to $(\hat{x}_1, \hat{x}_2, \hat{x}_3, \dots \hat{x}_i)$. The transformation equations ensure that the physical description of a system remains consistent, even though the coordinate representation may vary.

Inverse Transformation

The transformation between two coordinate systems can also be reversed. That is, the original coordinates can be expressed in terms of the new coordinates as

$$x_i = x_i(\hat{x}_1, \hat{x}_2, \hat{x}_3, .., \hat{x}_n), i = 1,2, \dots, n \qquad (5.2)$$

These relations are known as the inverse transformation equations and ensure that there exists a one-to-one correspondence between the two coordinate systems. Together with the direct transformation equations, they provide a complete description of how coordinates are related under a change of reference frame.
It is important to note that, in general,

$$(x_1, x_2, \dots, x_n) \neq (\hat{x}_1, \hat{x}_2, \hat{x}_3, .., \hat{x}_n) \qquad (5.3)$$

since the numerical values of the coordinates depend on the chosen coordinate system. However, both sets of coordinates represent the same physical point in space. This distinction highlights a fundamental concept in tensor analysis: while the components of a quantity may change under a coordinate transformation, the underlying physical entity remains invariant. Thus, coordinate transformations alter only the description of a point, not the point itself, and this idea forms the basis for defining tensor quantities in a coordinate-independent manner.

Index Notation and Einstein Summation Convention

In tensor analysis, many mathematical expressions involve adding many similar terms. Writing all these terms explicitly can make expressions long and difficult to read. For example, a sum like

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \qquad (5.4)$$

Equation (5.4) can be written in a shorter form using sigma notation as

$$\sum_{i=1}^{n} a_i\, x_i \qquad (5.5)$$

To make equation (5.5) even simpler, tensor calculus uses a compact rule called the Einstein summation convention**.**

According to convention**,** whenever an index appears twice in a single term, summation over that index is automatically understood. Therefore, the equation (5.5) can be written simply as. $a_ix_i$without explicitly writing the summation symbol. In this notation, the index that appears twice is called a dummy index (or summation index), since it is summed over all its values. The choice of this index is not important.

For example, $a_ix_i = a_jx_j = a_px_p$, all represent the same sum.

On the other hand, an index that appears only once in a term is called a free index. A free index is not summed over and indicates a specific component of a quantity. For example, consider the relation

$$A_k = g_{kj}B^j \qquad (5.6)$$

Here, the index j is repeated and hence summed over, so it is a dummy index, while k appears only once and is therefore a free index. This equation represents n separate equations, one for each value of k.Thus, the Einstein summation convention provides a simple and efficient way to write mathematical expressions, making tensor equations more compact and easier to handle in physics and engineering.

5.4 COVARIANT AND CONTRAVARIANT TENSORS

The components of a tensor generally depend on the chosen coordinate system and therefore change when the coordinates are transformed. However, different physical quantities follow different transformation laws under a change of coordinates. For example, displacement and velocity behave as contravariant quantities, whereas the gradient of a scalar function behaves as a covariant quantity. Understanding these two transformation laws is essential in tensor analysis, particularly in curvilinear coordinate systems, where the metric tensor provides the relationship between them.

Let us consider a point in space described using two different coordinate systems. In one system, say the X-coordinate system, the coordinates of the point are given by

$$(x_1, x_2, \dots, x_n)$$

while in another system, say the Y-coordinate system, the same point is represented by

$$(\hat{x}_1, \hat{x}_2, \hat{x}_3, .., \hat{x}_n)$$

Although the numerical values of the coordinates are different in the two systems, the point itself remains unchanged. This emphasizes an important idea: coordinates depend on the reference system, but the physical point in space is independent of it.

Now, suppose there is a set of quantities $A_1, A_2, \dots, A_n$associated with the point, which depend on its coordinates. When the coordinate system is changed, these quantities also change their values. However, this change is not arbitrary; it follows a definite mathematical rule determined by the transformation between the coordinate systems. The way these quantities transform under a change of coordinates is of fundamental importance in tensor analysis.

In fact, the classification of a physical quantity as a scalar, vector, or higher-rank tensor depends entirely on how its components transform when the coordinate system is changed. If a quantity remains unchanged, it is a scalar; if its components transform in a linear manner, it is a vector; and if it follows more general transformation rules involving multiple indices, it is identified as a tensor of higher rank. Thus, the transformation behavior of quantities provides the key to understanding the concept of tensors in mathematical physics.

Contravariant Tensor

A tensor whose components transform according to the same law as the coordinate differentials under a change of coordinates is called a **contravariant tensor**.

If the components transform according to

$$\hat{A}^i = \frac{\partial \hat{x}^i}{\partial x^j} A^j \quad \text{or} \quad A^j = \frac{\partial x^j}{\partial \hat{x}^i} \hat{A}^i \tag{5.7}$$

then the set A^i with $i = 1,2,3,4, \dots$, is called a contravariant vector or contravariant tensor of the first rank. In simple terms, the components change in the same way as the coordinates themselves. Contravariant components are written with superscripts.

Again, suppose, A^{ij} $(i, j = 1, 2, \dots, n)$ be n^2 functions of coordinates $x^1, x^2, x^3, \dots, x^n$in X-coordinate system. If the quantities A^{ij} are transformed to $\hat{A}^{ij}$ in Y-coordinate system having coordinates, $\hat{x}^1, \hat{x}^2, \hat{x}^3 .. \hat{x}^n$, then according to the law of transformation

$$\hat{A}^{ij} = \frac{\partial \hat{x}^i}{\partial x^k} \frac{\partial \hat{x}^j}{\partial x^l} A^{kl} \tag{5.8}$$

Then A^{ij} are called components of contravariant tensor of rank two.

For example: If x_i are coordinates, then under transformation:

$$\hat{x_i} = \frac{\partial \hat{x}_i}{\partial x_j} x_j,$$

so, coordinates themselves behave like contravariant quantities.

Covariant Tensor

If the components transform according to

$$\hat{A}_i = \frac{\partial x^j}{\partial \bar{x}^i} A_j \text{ or } A_j = \frac{\partial \hat{x}^i}{\partial x^j} \hat{A}_i \tag{5.9}$$

then the set A_i with $i = 1,2,3,4, \ldots$ is called a covariant vector or covariant tensor of the first rank. Contravariant components are written with superscripts, whereas covariant components are written with subscripts. The coordinates x themselves are written with superscripts because they transform countravariantly.

Similarly let, $A_{ij}\ (i,j = 1,2,\ldots,n)$ be n^2 functions of coordinates $x_1, x_2, \ldots, x_n$in X-coordinate system. If the quantities A_{ij} are transformed to $\hat{A}_{ij}$ in Y-coordinate system having coordinates, $\hat{x}^1, \hat{x}^2, \hat{x}^3 .. \hat{x}^n$, then according to the law of transformation

$$\bar{\hat{A}}_{ij} = \frac{\partial x^k}{\partial \hat{x}^i} \frac{\partial x^l}{\partial \hat{x}^j} A_{kl} \tag{5.10}$$

In this case, the quantities A_{ij} are called the components of a covariant tensor of rank two.

For example: A standard example of a covariant tensor of rank one is the gradient of a scalar function. Let $\phi(x_1, x_2, \ldots, x_n)$be a scalar function. The components of its gradient are $\frac{\partial \phi}{\partial x_i}$. Under a change of coordinates $x_i \to \hat{x}_i$, these components transform according to

$$\frac{\partial \phi}{\partial \hat{x}_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial \phi}{\partial x_j}$$

which is precisely the transformation law of a covariant tensor. Hence, the gradient of a scalar function is a covariant vector (or tensor of rank one).

Mixed Tensor

A mixed tensor is a tensor that contains both contravariant and covariant indices at the same time. Consider a coordinate system X with coordinates $x_1, x_2, \ldots, x_n$. Suppose there are n^2 quantities denoted by A^i_j where $i,j = 1,2,\ldots,n$. These quantities depend on the coordinates of the system. Now, when we change to another coordinate system Y with coordinates $\hat{x}^1, \hat{x}^2, \hat{x}^3 .. \hat{x}^n$the components A^i_j also change. If the new components $\hat{A}^i_j$ are related to the old components by the transformation rule

$$\hat{A}^i_j = \frac{\partial x_i}{\partial x_k} \frac{\partial x_l}{\partial \hat{x}_j} A^k_l$$

then the set of quantities A^i_j is called a mixed tensor. Since each component has one upper index and one lower index, the tensor is said to be of rank two. Here i and $j$are the free indices, k and l are the dummy (summation) indices**.**

Note:

- *The rank of the tensor is defined as the total number of indices per component.*
- *Instead of saying that "A^{ij} are the components of a tensor of rank two" we shall often say " A^{ij} is a tensor of rank two."*

The Kronecker Delta

The Kronecker deltaδ_j^i φ s defined as

$$\delta_j^i = \quad 0, \text{ if } j \quad \neq i\# \qquad (5.11)$$
$$= \quad 1, \qquad \text{if } j = i$$

From its notation, we see that it has one upper index and one lower index. Therefore, it behaves like a mixed tensor of rank two and acts as the identity tensor under contraction.

The Kronecker delta acts as the identity tensor in tensor algebra and is useful in index manipulation and tensor contraction.

Basic Tensor Operations

1. **Addition**: Two tensors can be added only if they have the same rank and the same type (both covariant, both contravariant or both mixed). The sum is obtained by adding their corresponding components. Thus, if $A_{ij}$and $B_{ij}$are two second-rank covariant tensors, then
$$C_{ij} = A_{ij} + B_{ij}$$
2. **Subtraction:** Two tensors can be subtracted only if they have the same rank and the same type (both covariant, both contravariant, or both mixed). The subtraction is performed by subtracting the corresponding components of the tensors. If $A_{ij}$and B_{ij} are two second-rank covariant tensors, then
$$C_{ij} = A_{ij} - B_{ij},$$
where C_{ij} is also a second-rank covariant tensor.
3. **Outer (Direct) Product:** The outer (or direct) product of two tensors is obtained by multiplying their components. The resulting tensor has a rank equal to the sum of the ranks of the two tensors. For example, if $A_i$and $B_j$are two first-rank covariant tensors, then their outer product is
$$C_{ij} = A_i B_j$$
which is a second-rank covariant tensor. Similarly, if A^i and $B^j$are two first-rank contravariant tensors, then $C^{ij} = A^i B^j$, which is a second-rank contravariant tensor.
4. **Contraction:** It is the process of equating one covariant index and one contravariant index and summing over the repeated index. This operation reduces the rank of a tensor by two**.** For example, if $A^i{}_j$ is a mixed tensor of rank two, then its contraction
$$A^i{}_i = \sum_i A^i{}_i,$$
which is a scalar (tensor of rank zero)**.** Similarly, for a third-rank mixed tensor, mixed tensor $T^{ij}{}_k$,$T^{ij}{}_j$ is a first-rank contravariant tensor, since one upper and one lower index have been contracted.
5. **Inner Product:** The inner product of two tensors is obtained by taking their outer product and then contracting one pair of corresponding covariant and contravariant indices. For example, if $A^i$and B_i are two first-rank tensors of contravariant and covariant types, respectively, then
$$A^i B_i$$
is the inner product, which is a scalar (tensor of rank zero)**.**

5.5 SYMMETRIC AND ANTI-SYMMETRIC TENSOR

Symmetric Tensor

A tensor is said to be symmetric with respect to two contravariant (or two covariant) indices if its components remain unchanged on an interchange of the two indices.

For example:

1. A tensor A_{ij} is symmetric if
$$A_{ij} = A_{ji}.$$
2. A tensor A^{ijk}_{lm} is symmetric in the indices i and j if
$$A^{ijk}_{lm} = A^{jik}_{lm}.$$

Theorem: A symmetric tensor of rank two in an n-dimensional space has
$$\frac{n(n+1)}{2}$$
independent components.

Proof: Let A_{ij} be a symmetric tensor of rank two. Then
$$A_{ij} = A_{ji}$$
A rank-two tensor in n-dimensional space has n^2 total components.

Now, the diagonal components are given by
$$A_{11}, A_{22}, A_{33}, \dots, A_{nn}$$
are all independent which give n components. The remaining components occur in equal pairs because of symmetry given by
$$A_{12} = A_{21}, A_{13} = A_{31}, A_{23} = A_{32}, \ etc.$$
Hence only half of the remaining $(n^2 - n)$ components are independent.

Therefore, the number of independent components is
$$= n + \frac{n^2 - n}{2}$$
$$= \frac{2n + n^2 - n}{2}$$
$$= \frac{n(n+1)}{2}$$
Hence proved.

Anti-Symmetric Tensors

A tensor is called skew-symmetric (or antisymmetric) with respect to two contravariant indices or two covariant indices if its components change sign when the indices are interchanged.

For examples

1. A tensor A_{ij} is anti-symmetric if
$$A_{ij} = -A_{ji}.$$
2. A tensor A^{ijk}_{lm} is anti-symmetric in the indices i and j if
$$A^{ijk}_{lm} = -A^{jik}_{lm}.$$

Theorem

A skew-symmetric tensor of rank two in an n-dimensional space has
$$\frac{n(n-1)}{2}$$
independent components.

Proof

Let A_{ij} be a skew-symmetric tensor of rank two. Then

$$A_{ij} = -A_{ji}$$

Now putting $i = j$,

$$A_{ii} = -A_{ii}$$

which gives

$$A_{ii} = 0$$

Therefore, all diagonal components are zero. The remaining components satisfy relations such as

$$A_{12} = -A_{21}, A_{13} = -A_{31}, A_{23} = -A_{32}$$

Thus, only half of the off-diagonal components are independent. The number of off diagonal components is

$$n^2 - n$$

Hence the number of independent components is

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

Hence proved.

5.6 FOUR-VECTORS

A 4-tuple $A = (A^0, A^1, A^2, A^3)$ is defined as a Four-vector if its components transform under Lorentz transformations in the same manner as the spacetime coordinates (cdt, dx, dy, dz).

A 4-vector is denoted as

$$A^\mu = (A^0, A^1, A^2, A^3).$$

A quantity is defined to be a 4-vector if its components obey the same transformation rules as the spacetime coordinates (ct, x, y, z).

For a Lorentz boost along the x −direction with relative velocity v, the components transform as

$$A^{0'} = \gamma\left(A^0 + \frac{v}{c}A^1\right),$$

$$A^{1'} = \gamma\left(A^1 + \frac{v}{c}A^0\right),$$

$$A^{2'} = A^2,$$

$$A^{3'} = A^3,$$

where

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}.$$

The same form of transformation applies for boosts along the $y-$ and z −directions, with the corresponding components mixing in an analogous manner.

Thus, the defining feature of a 4-vector is that its time component A^0 and spatial components (A^1, A^2, A^3) transform into each other under Lorentz transformations in the same way as spacetime coordinates, while preserving the structure of the transformation laws. This property ensures that physical laws written using 4-vectors retain the same form in all inertial reference frames.

Example:

1. **4-Velocity:** The 4-velocity is defined as the rate of change of 4-position with respect to proper time τ

$$u^{\mu} = \frac{dx^{\mu}}{d\tau}$$

So, its components are

$$u^{\mu} = (\gamma c,\ \gamma v_x,\ \gamma v_y,\ \gamma v_z)\ ,$$

where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. The quantity u^{μ}transforms under Lorentz transformation exactly like spacetime coordinates, so it is a four-vector.

5.7 LINE ELEMENT

In rectangular (Cartesian) coordinates (x, y, z) the small distance between two nearby points is called the line elementds. It is given by

$$ds^2 = dx^2 + dy^2 + dz^2 \quad (5.12)$$

This expression shows that the square of the infinitesimal distance is equal to the sum of the squares of the small changes along the x, y, and z directions. It is a direct consequence of the Euclidean geometry of space, where the coordinate axes are mutually perpendicular.

The line element plays an important role in tensor analysis and differential geometry, as it provides a way to measure distances that remains invariant under coordinate transformations. It also forms the basis for defining more general expressions of distance in curvilinear coordinate systems, where the form of ds^2may involve additional terms and coefficients.

When we move from rectangular (Cartesian) coordinates to a more general curvilinear coordinate system $(u_1, u_{2,} u_3)$, the expression for the line element becomes more general. In such a coordinate system, the infinitesimal distance between two nearby points is given by

$$ds^2 = \sum_{p=1}^{3} \sum_{q=1}^{3} g_{pq}\ du_p\, du_q. \qquad (5.13)$$

Here, the quantities $g_{pq}$are known as the metric coefficients. These coefficients determine how distances are measured in the new coordinate system and account for the scale factors and possible non-orthogonality of the coordinate axes. Unlike Cartesian coordinates, where the metric is simple and constant, in curvilinear coordinates the metric coefficients may vary from point to point.

Thus, the line element expressed in this form provides a generalized way of measuring distances, which remains valid under coordinate transformations. Such a formulation is fundamental in tensor analysis and describes the geometry of three-dimensional Euclidean space in arbitrary coordinate systems.

5.8 METRIC TENSOR

For an N-dimensional space with coordinates $(x_1, x_2, \dots, x_N)$, the infinitesimal distance between two nearby points is given by the generalized line element

$$ds^2 = \sum_{p=1}^{N} \sum_{q=1}^{N} g_{pq}\ dx_p\, dx_q. \qquad (5.14)$$

This quadratic expression is known as the metric form or simply the metric of the space. Using Einstein's summation convention, it can be written in a more compact form as

$$ds^2 = g_{pq}\, dx_p dx_q. \qquad (5.15)$$

The quantities g_{pq} are called the metric coefficients, and they represent the components of a covariant tensor of rank two, known as the metric tensor or fundamental tensor. The metric tensor plays a central role in describing the geometry of a space, as it determines how distances and angles are measured. In general, the coefficients g_{pq} may depend on the coordinates and reflect the nature of the coordinate system being used.

Since the products of coordinate differentials satisfy

$$dx_p\, dx_q = dx_q\, dx_p,$$

the metric tensor is symmetric, that is $g_{pq} = g_{qp}$.Therefore, in an N-dimensional space, the metric tensor has only

$$\frac{N(N+1)}{2}$$

independent components.

Example: - In a three-dimensional Cartesian coordinate system, the line element is

$$ds^2 = dx^2 + dy^2 + dz^2$$

In tensor form, this can be written as

$$ds^2 = g_{pq} dx_p dx_q$$

where the metric tensor is

$$[g_{pq}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here, $g_{11} = g_{22} = g_{33} = 1$ and all other components are zero.

This shows that in Cartesian coordinates, the metric tensor is simply the identity matrix, reflecting that the coordinate axes are mutually perpendicular and equally scaled.

Example: In cylindrical coordinates (r, ϕ, z), the line element is

$$ds^2 = dr^2 + r^2 d\phi^2 + dz^2.$$

Comparing this expression with Eq. (5.15),

$$ds^2 = g_{pq}\, dx_p dx_q,$$

we obtain

$$g_{rr} = 1,\ g_{\phi\phi} = r^2,\ g_{zz} = 1,$$

while all the remaining components are zero. Thus, the metric tensor in cylindrical coordinates is diagonal.

5.9 ASSOCIATED TENSORS

The **inverse metric tensor**, denoted by g^{pq}, is defined as the inverse of the metric tensor g_{pq}. Its components satisfy the relation

$$g^{pr}g_{rq} = \delta_q^p,$$

where δ_q^p is the Kronecker delta. The inverse metric tensor is also symmetric and is used to convert covariant tensor components into contravariant components.

The process of raising and lowering indices can be understood with the help of the metric tensor and its inverse. All tensors obtained from a given tensor by forming inner products with the metric tensor g_{pq}or its inverse g^{pq} are called the associated tensors of the original tensor.

Raising an Index

A covariant component can be converted into a contravariant component by multiplying it with the inverse metric tensor. Thus,

$$A^p = g^{pq}A_q. \qquad (5.16)$$

Here, the lower index of $A_q$is raised to the upper index p.

Lowering an Index

Similarly, a contravariant component can be converted into a covariant component by multiplying it with the metric tensor. Thus,

$$A_p = g_{pq}A^q. \qquad (5.17)$$

Here, the upper index of $A^q$is lowered to the lower index p. Hence, the tensors $A^p$and A_p are called **associated tensors**. Equation (5.16) gives the contravariant components, whereas Eq. (5.17) gives the corresponding covariant components.

CHECK YOUR PROGRESS

Q1. Define the line element in a three-dimensional Cartesian coordinate system.

Q2. What is a metric tensor? Why is it called a covariant tensor of rank two?

Q3. Write each of the following using the summation convention.

a) $d\Phi = \frac{\partial\Phi}{\partial x^1}\,dx^1 + \frac{\partial\Phi}{\partial x^2}\,dx^2 + \cdots + \frac{\partial\Phi}{\partial x^n}\,dx^n$

b) $(x^1)^2 + (x^2)^2 + (x^3)^2 + \ldots + (x^N)^2$

Q 4. Write the law of transformation for the tensors

(a) A_k^k (b) B_k (c) C^m

Q 5. Evaluate

(a) $\delta_q^p A_b^{qa}$ (b) $\delta_p^q \delta_q^r$

Q 6. In 3-D Cartesian coordinates (x, y, z) the line element is

$$ds^2 = dx^2 + dy^2 + dz^2.$$

1) Write this expression using index notation and the Einstein summation convention.
2) Identify the metric tensor components g_{ij}

Q7. Show that $ds^2 = dx^2 + dy^2 + dz^2$ can be written in the form $ds^2 = g_{pq}dx_p dx_q$.

Q8. Verify that the metric tensor for Cartesian coordinates satisfies

$$g_{pq} = \delta_{pq},$$

where δ_{pq} is the Kronecker delta.

Q9. If $dx_1 = 2, dx_2 = 3, dx_3 = 6$,find the value of ds using the Cartesian metric tensor.

5.10 LET US SUM UP

- A point in space can be described using coordinates, which may vary depending on the chosen coordinate system.
- The same physical point can be represented by different coordinate values in different coordinate systems, but the point itself remains unchanged.
- Coordinate transformations relate one system of coordinates to another through functional relations.
- The inverse transformation allows the original coordinates to be expressed in terms of the new coordinates, ensuring a one-to-one correspondence.
- The transformation behaviour of quantities determines whether they are scalars, vectors, or higher-rank tensors.
- Scalars remain invariant under coordinate transformation, whereas vectors and tensors change according to specific transformation laws.
- The Einstein summation convention provides a compact way of writing expressions by implying summation over repeated indices.
- A repeated index is called a dummy index, while a non-repeated index is called a free index**.**
- The concept of tensors generalizes physical quantities to higher dimensions and provides a coordinate-independent description of physical laws.
- The line element ds^2 represents the infinitesimal distance between two nearby points and is fundamental in describing geometry.
- In Cartesian coordinates, the line element is given by $ds^2 = dx^2 + dy^2 + dz^2$, which follows from Euclidean geometry.
- In general, curvilinear coordinates, the line element is expressed as

  $$ds^2 = g_{pq}dx_p dx_q,$$

  where g_{pq} are the metric coefficients.
- The quantities g_{pq}form the metric tensor, which is a covariant tensor of rank two.
- Basic tensor operations such as addition, subtraction, outer product, contraction and inner product are useful in manipulating tensors and obtaining new tensors of different ranks.
- The metric tensor is symmetric $(g_{pq} = g_{qp})$ and, in an N-dimensional space, it has $\frac{N(N+1)}{2}$ independent components.
- The metric tensor and its inverse are used to raise and lower tensor indices, giving rise to associated tensors.

✓5.11 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q 1. The line element is the infinitesimal distance between two nearby points.

In cartesian coordinates (x, y, z), it is given by

$$ds^2 = dx^2 + dy^2 + dz^2$$

Answer to Q2: The metric tensor g_{pq} is a set of quantities that define how distances are measured in a coordinate system through the relation

$$ds^2 = g_{pq} dx_p dx_q$$

It is called covariant because its components transform with derivatives of old coordinates with respect to new coordinates. It is of rank two because it has two indices p and q

Answer to Q3: (a) $d\Phi = \frac{\partial \Phi}{\partial x^i} dx^i$.

(b) $x^i x^i$.

Answer to Q 4. (a) A_k^k is a scalar (invariant under transformation).

(b) $\hat{B}_i = \frac{\partial x^j}{\partial \hat{x}^i} B_j$

(c) $\hat{C}^i = \frac{d\hat{x}^i}{dx^j} C^j$

Answer to Q 5. (a) $\delta_q^p A_b^{qa} = A_b^{pa}$

(b) $\delta_p^q \delta_q^r = \delta_p^r$

Answer to Q 6. (1) $ds^2 = g_{ij} dx^i dx^j$

(2) $g_{ij} = \delta_{ij}$

Answer to Q7: In index notation: $dx^2 + dy^2 + dz^2 = dx_1^2 + dx_2^2 + dx_3^2$.

This can be written as $ds^2 = g_{pq} dx_p dx_q$

where $g_{pq} = \begin{cases} 1, & p = q \\ 0, & p \neq q \end{cases}$

Thus, $ds^2 = dx_1^2 + dx_2^2 + dx_3^2$

Answer to Q8: The Kronecker delta is defined as

$$\delta_{pq} = \begin{cases} 1, & p = q \\ 0, & p \neq q \end{cases}$$

In Cartesian coordinates: $g_{11} = g_{22} = g_{33} = 1, g_{pq} = 0$ for $p \neq q$.

Hence, $g_{pq} = \delta_{pq}$

Answer to Q9: Using. $ds^2 = dx_1^2 + dx_2^2 + dx_3^2$,

$$ds^2 = 2^2 + 3^2 + 6^2 = 4 + 9 + 36 = 49$$

$$ds = \sqrt{49} = 7$$

🔍5.12 FURTHER READING

- Dass, H. K. and Verma, Rama (2022), *Mathematical Physics*, 8th edition. S. Chand Publishing, New Delhi.
- Arfken, G. B., Weber, H. J., and Harris, F. E. (2013), *Mathematical Methods of Physicists*: *A comprehensive guide* 7th edition, Academic Press, Elsevier.
- Boas, Marg L. (2006), *Mathematical Methods in the Physical Sciences*, *(3rd ed.)*, Wiley.

- Lawden D. F. (1982). *Introduction to Tensor Calculus, Relativity and Cosmology*, Dover Publications.

⍰5.13 MODEL QUESTIONS

1. Define Tensor. Give Example.
2. If x^i be the coordinate of a point in n-dimensional space show that dx^i are component of a contravariant vector.
3. Show that $\frac{\partial \phi}{\partial x^i}$ is a covariant vector where ϕ is a scalar function.
4. Show that the Kronecker delta (δ_j^i) is a mixed tensor of rank two.
5. Show that δ_j^i is an invariant i.e., it has same components in every coordinate system.
6. If $A_i$is a covariant tensor, then prove that $\frac{\partial A_i}{\partial x^j}$ do not form a tensor.
7. If A_j^i is mixed tensor of rank 2 and B_m^{kl} is mixed tensor of rank 3. Prove that $A_j^i B_m^{jl}$ is a mixedtensor of rank 3.
8. Show that the product of two tensors A_j^i and B_m^{kl} is a tensor of rank five.
9. If $ds^2 = g_{jk}\, dx^j\, dx^k$ is an invariant, show that g_{jk} is a symmetric covariant tensor of rank two.
10. If $A_i = g_{ij}A^j$, then show that $A^j = g^{ij}A_i$.

*** *** ***

UNIT 6: TENSOR CALCULUS AND GEOMETRY

UNIT STRUCTURE

6.1 Learning Objectives
6.2 Introduction
6.3 Christoffel Symbols
6.4 Covariant Derivative
6.5 Geodesics
6.6 Riemann-Christoffel Curvature Tensor
6.7 Bianchi Identities
6.8 Let Us Sum Up
6.9 Answers to Check Your Progress
6.10 Further Reading
6.11. Model Questions

6.1 LEARNING OBJECTIVES

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

- explain curvilinear coordinate systems and the role of the metric tensor
- define Christoffel symbols and describe their role in curved coordinate systems
- apply covariant differentiation to scalars, vectors, and higher-rank tensors
- derive the equations of geodesics and interpret the Riemann–Christoffel curvature tensor
- state and apply the Bianchi identities and explain their significance in tensor analysis.

6.2 INTRODUCTION

In the previous unit, we studied the fundamental concepts of scalars, vectors, and tensors, along with covariant and contravariant tensors, transformation properties, line elements, and the metric tensor. These concepts provided the mathematical foundation for describing physical and geometrical quantities in different coordinate systems. Building on this foundation, the present unit extends tensor analysis to more general coordinate systems and curved spaces.

In many physical and mathematical problems, Cartesian coordinates are not always sufficient, especially when dealing with curved surfaces or non-uniform coordinate systems. In such cases, it becomes necessary to use curvilinear coordinates, where the basis vectors vary from point to point and ordinary differentiation is no longer adequate. This requires more general mathematical tools for describing geometry and physical laws.

In this unit, we shall study the concept of curvilinear coordinate systems and the role of the metric tensor in measuring distances and describing geometry. We will introduce Christoffel symbols, which account for the variation of basis vectors, and use them to define the covariant derivative, a generalized form of differentiation that preserves

tensor character. Further, we shall study geodesics, the Riemann–Christoffel curvature tensor, and the Bianchi identities, which form the mathematical basis for understanding curvature and geometrical relations in curved spaces. These concepts are essential in advanced areas of physics such as differential geometry and general Relativity.

6.3 CHRISTOFFEL SYMBOLS

In the study of curvilinear coordinates and tensor analysis, the derivatives of the metric tensor play an important role in describing the geometry of space. In 1869, Elwin Bruno Christoffel introduced certain special combinations of the partial derivatives of the fundamental (metric) tensor. These quantities are now known as Christoffel symbols. They help us describe how the coordinate basis changes from point to point in a curved space.

Christoffel symbols are important because, in general curvilinear coordinates, the basis vectors are not constant but vary with position. As a result, ordinary partial derivatives of vectors and tensors do not transform properly as tensors. Christoffel symbols provide the necessary correction terms that allow us to define the covariant derivative, which preserves the tensorial nature of physical and geometrical quantities.

They are therefore essential for

1. Defining covariant differentiation of vectors and tensors.
2. Describing parallel transport of vectors.
3. Determining geodesics (shortest paths or free-particle motion).
4. Constructing curvature tensors such as the Riemann and Ricci tensors.
5. Studying the geometry of curved spaces and spacetime, particularly in general relativity.

Thus, Christoffel symbols act as the mathematical tools that connect the metric of a space with its geometric and physical properties, making them fundamental to differential geometry and tensor analysis. There are two kinds of Christoffel symbols.

1. Christoffel symbols of first kind.
2. Christoffel symbols of second kind.

These two forms arise from the derivatives of the metric tensor and are used to describe how the coordinate basis vectors vary from point to point in a curved space.

Christoffel Symbols of the first kind

The Christoffel symbols of the first kind are defined by

$$[ij,k] = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right). \tag{6.1}$$

These quantities depend solely on the first–order partial derivatives of the metric tensor g_{ij}.Since all three indices are in the lower position, they are often referred to as the covariant form of the Christoffel symbols. They are symmetric in their first two indices, that is,

$$[ij,k] = [ji,k].$$

The Christoffel symbols of the first kind arise naturally in theoretical derivations and serve as intermediate expressions in the development of covariant differentiation and curvature tensors. However, they are not frequently used directly in practical calculations or physical applications. Instead, they are primarily employed to obtain the Christoffel symbols of the second kind, which play a more direct role in geometrical and physical formulations.

Christoffel Symbols of the second kind

Another important set of quantities is obtained from the Christoffel symbols of the first kind by raising one index with the contravariant metric tensor $g^{k\alpha}$. These are called the Christoffel symbols of the second kind, originally introduced by Elwin Bruno Christoffel, and are defined by

$$\Gamma_{ij}^{k} = g^{k\alpha}\,[ij,\alpha]. \tag{6.2}$$

Equivalently, they may be written explicitly in terms of the metric tensor as

$$\Gamma_{ij}^{k} = \frac{1}{2} g^{k\alpha}\left(\frac{\partial g_{i\alpha}}{\partial x^{j}} + \frac{\partial g_{j\alpha}}{\partial x^{i}} - \frac{\partial g_{ij}}{\partial x^{\alpha}}\right). \tag{6.3}$$

These symbols possess one contravariant (upper) index and two covariant (lower) indices and are symmetric in the lower pair of indices, that is,

$$\Gamma_{ij}^{k} = \Gamma_{ji}^{k}. \tag{6.4}$$

The Christoffel symbols of the second kind are of primary importance in applications. They arise naturally in the definition of the covariant derivative, where they provide the correction terms required to account for the variation of the coordinate basis. Consequently, they play a fundamental role in describing parallel transport, geodesics, and the curvature of space. For this reason, they are widely used in differential geometry, tensor calculus, and gravitational theories such as general relativity.

Consider the polar coordinate metric $ds^2 = dr^2 + r^2 d\theta^2$. To determine the Christoffel symbol of the first kind $[\theta\theta, r]$, we use

$$[ij,k] = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^{j}} + \frac{\partial g_{jk}}{\partial x^{i}} - \frac{\partial g_{ij}}{\partial x^{k}}\right).$$

Now the metric components are $g_{rr} = 1, g_{\theta\theta} = r^2$ and

$$[\theta\theta, r] = \frac{1}{2}\left(\frac{\partial g_{\theta r}}{\partial x^{\theta}} + \frac{\partial g_{\theta r}}{\partial x^{\theta}} - \frac{\partial g_{\theta\theta}}{\partial x^{r}}\right),$$

$$= \frac{1}{2}\left(0 + 0 - \frac{\partial (r^2)}{\partial r}\right),$$

$$[\theta\theta, r] = -r.$$

Now to determine the Christoffel symbol of the second kind $\Gamma_{\theta\theta}^{r}$ we use the relation (6.2) and $g^{rr} = 1$. Now

$$\Gamma_{\theta\theta}^{r} = g^{r\alpha}\,[\theta\theta,\alpha],$$

Replacing α with r we get

$$\Gamma_{\theta\theta}^{r} = g^{rr}\,[\theta\theta, r],$$
$$= 1(-r),$$
$$\Gamma_{\theta\theta}^{r} = -r,$$

Transformation Laws of Christoffel's symbols

In the theory of curvilinear coordinates and tensor analysis, it is important to understand how the Christoffel symbols behave under a change of coordinates. Unlike tensors, Christoffel symbols do not follow the standard tensor transformation law. This fact highlights that they are not tensors but rather connection coefficients that describe how the coordinate basis vectors vary from point to point.

Let $[ij,k]$ and Γ_{ij}^{k} denote the Christoffel symbols of the first and second kinds, respectively, in a coordinate system x^k. Suppose we introduce a new coordinate system $\bar{x}^k$. Quantities expressed in the new system will be denoted by a bar.

Transformation of the first kind

Under the coordinate transformation $x^k \rightarrow \bar{x}^k$, the Christoffel symbols of the first kind transform as

$$[\overline{ij},k] = \frac{\partial x^p}{\partial \overline{x}^{\,i}} \frac{\partial x^q}{\partial \overline{x}^{\,j}} \frac{\partial x^r}{\partial \overline{x}^{\,k}} [pq,r] + g_{pr} \frac{\partial^2 x^p}{\partial \bar{x}^i \, \partial \bar{x}^j} \frac{\partial x^r}{\partial \overline{x}^{\,k}} \quad (6.5)$$

Transformation of the second kind

Similarly, the Christoffel symbols of the second kind transform as

$$\overline{\Gamma}_{ij}^{k} = \frac{\partial \overline{x}^{\,k}}{\partial x^r} \frac{\partial x^p}{\partial \bar{x}^i} \frac{\partial x^q}{\partial \overline{x}^{\,j}} \Gamma_{pq}^{r} + \frac{\partial \overline{x}^{\,k}}{\partial x^r} \frac{\partial^2 x^r}{\partial \bar{x}^i \, \partial \overline{x}^{\,j}} \quad (6.6)$$

We see that each transformation law (6.5 and 6.6) contains two terms. Because of this additional term, Christoffel symbols do not transform as tensors. They behave as non-tensorial geometric objects that depend on the coordinate system.

6.4 COVARIANT DERIVATIVE

In curvilinear coordinates, the components of vectors and tensors depend not only on position but also on the variation of the coordinate basis. Consequently, ordinary partial derivatives do not, in general, transform as tensors. To preserve tensorial character under coordinate transformations, a modified derivative, called the covariant derivative, is introduced. This derivative incorporates correction terms involving the Christoffel symbols introduced by Elwin Bruno Christoffel.

Covariant derivative of a contravariant vector

Let A^p be a contravariant vector. The covariant derivative of A^p with respect to the coordinate x^q denoted by A_q^p is defined as

$$A_q^p = \frac{\partial A^p}{\partial x^q} + \Gamma_{sq}^{p} A^s \quad (6.7)$$

This quantity transforms as a mixed tensor of rank two (one contravariant and one covariant index).

Covariant derivative of a covariant vector

If A_p is a covariant vector, its covariant derivative is defined by

$$A_{p,q} = \frac{\partial A_p}{\partial x^q} - \Gamma^s_{pq} A_s \,. \tag{6.8}$$

Here the negative sign appears because the index is covariant. This derivative is a covariant tensor of rank two.

Covariant Derivative of Higher–Rank Tensors

The definition extends naturally to tensors of higher rank. For each contravariant index, a positive Christoffel term is added, and for each covariant index, a negative Christoffel term is subtracted. For example, for a mixed tensor $A^{p_1 \ldots p_m}_{r_1 \ldots r_n}$,

$$A^{p_1 \ldots p_m}_{r_1 \ldots r_n\ ,q} = \frac{\partial}{\partial x^q}\left(A^{p_1 \ldots p_m}_{r_1 \ldots r_n}\right) + \sum_{i=1}^{m} \Gamma^{\mathrm{p_i}}_{sq} A^{p_1 \ldots s \ldots p_m}_{r_1 \ldots r_n} - \sum_{j=1}^{n} \Gamma^{s}_{r_j q} A^{p_1 \ldots p_m}_{r_1 \ldots s \ldots r_n} \,. \tag{6.9}$$

Thus, each index contributes a corresponding correction term.

6.5 GEODESICS

In a Riemannian space, the distance between two points lying on a smooth curve $x^r = x^r(t)$, joining the parameter values t_1 and t_2, is determined by the metric tensor g_{pq}.The infinitesimal arc length is given by

$$ds^2 = g_{pq} dx^p dx^q, \tag{6.10}$$

and hence the total length of the curve is

$$s = \int_{t_1}^{t_2} \sqrt{g_{pq} \frac{dx^p}{dt} \frac{dx^q}{dt}}\, dt \,. \tag{6.11}$$

Among all possible curves connecting the two points, the one that makes this distance stationary (usually a minimum) is called a geodesic of the space. Thus, a geodesic represents the shortest or extremal path between two points and generalizes the concept of a straight line to curved spaces. Using the methods of the calculus of variations, the condition for extremizing the arc length leads to the differential equation of a geodesic. This equation can be written in terms of the Christoffel symbols as

$$\frac{d^2 x^r}{ds^2} + \Gamma^r_{pq} \frac{dx^p}{ds} \frac{dx^q}{ds} = 0, \tag{6.12}$$

where s is the arc-length parameter. This equation shows that the curvature of the path is governed by the Christoffel symbols, which account for the variation of the coordinate basis. In the absence of these terms (as in Cartesian coordinates of flat space), the equation reduces to

$$\frac{d^2 x^r}{ds^2} = 0, \tag{6.13}$$

whose solutions are straight lines. Hence, geodesics may be interpreted as the natural generalization of straight lines in curved geometry. For example, the geodesics on a plane are straight lines, while the geodesics on a sphere are arcs of great circles. In physical applications, particularly in differential geometry and general relativity, geodesics describe the motion of freely moving particles under no external forces.

Derivation of equation of Geodesics

The distance between any two points A and B lying on surface is given by

$$ds^2 = g_{pq}dx^p dx^q \quad (6.14)$$

Taking the **variation** on both sides,

$$\delta(ds^2) = \left(\frac{\partial g_{pq}}{\partial x^k}\delta x^k\right)dx^p dx^q + g_{pq}\,\delta(dx^p)\,dx^q + g_{pq}\,dx^p\,\delta(dx^q)$$

Using the symmetry $g_{pq} = g_{qp}$,

$$2ds\,\delta(ds) = \left(\frac{\partial g_{pq}}{\partial x^k}\delta x^k\right)dx^p dx^q + 2g_{pq}\,\delta(dx^p)\,dx^q$$

Hence, $\delta(ds) = \left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\delta x^k\frac{dx^p}{ds}\frac{dx^q}{ds} + g_{pq}\frac{d(\delta x^p)}{ds}\frac{dx^q}{ds}\right]ds$

Therefore,

$$\int_A^B \delta\,(ds) = \int_A^B \left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\delta x^k\frac{dx^p}{ds}\frac{dx^q}{ds} + g_{pq}\frac{d(\delta x^p)}{ds}\frac{dx^q}{ds}\right]ds \quad (6.15)$$

Now,

$$\int_A^B g_{pq}\frac{d(\delta x^p)}{ds}\frac{dx^q}{ds}\,ds = \left[g_{pq}\frac{dx^q}{ds}\delta x^p\right]_A^B - \int_A^B \frac{d}{ds}\left(g_{pq}\frac{dx^q}{ds}\right)\delta x^p\,ds.$$

Since, $\delta x^p = 0$ at A and B,
the boundary term vanishes, giving

$$= -\int_A^B \frac{d}{ds}\left(g_{pq}\frac{dx^q}{ds}\right)\delta x^p\,ds.$$

Substituting into Eq. (6.15),

$$\delta\int_A^B d\,s = \int_A^B \left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\frac{dx^p}{ds}\frac{dx^q}{ds}\delta x^k - \frac{d}{ds}\left(g_{pq}\frac{dx^q}{ds}\right)\delta x^p\right]ds.$$

Renaming the dummy index $p \to k$,

$$0 = \int_A^B \left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\frac{dx^p}{ds}\frac{dx^q}{ds} - \frac{d}{ds}\left(g_{kq}\frac{dx^q}{ds}\right)\right]\delta x^k\,ds.$$

Expanding,

$$\int_A^B \left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\frac{dx^p}{ds}\frac{dx^q}{ds} - \frac{\partial g_{qk}}{\partial x^p}\frac{dx^p}{ds}\frac{dx^q}{ds} - g_{qk}\frac{d^2x^q}{ds^2}\right]\delta x^k\, ds = 0.$$

Since

$$\frac{d}{ds}(g_{qk}) = \frac{\partial g_{qk}}{\partial x^p}\frac{dx^p}{ds},$$

and because the variations δx^k are arbitrary,

$$\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k}\frac{dx^p}{ds}\frac{dx^q}{ds} - \frac{\partial g_{qk}}{\partial x^p}\frac{dx^p}{ds}\frac{dx^q}{ds} - g_{qk}\frac{d^2x^q}{ds^2} = 0.$$

Rearranging,

$$\left[\frac{1}{2}\frac{\partial g_{pq}}{\partial x^k} - \frac{1}{2}\frac{\partial g_{qk}}{\partial x^p} - \frac{1}{2}\frac{\partial g_{pk}}{\partial x^q}\right]\frac{dx^p}{ds}\frac{dx^q}{ds} - g_{qk}\frac{d^2x^q}{ds^2} = 0.$$

Or, $$\frac{1}{2}\left[\frac{\partial g_{pk}}{\partial x^q} + \frac{\partial g_{qk}}{\partial x^p} - \frac{\partial g_{pq}}{\partial x^k}\right]\frac{dx^p}{ds}\frac{dx^q}{ds} + g_{qk}\frac{d^2x^q}{ds^2} = 0.$$

Multiplying by $g^{\alpha k}$,$g^{\alpha k}g_{qk}\frac{d^2x^q}{ds^2} + g^{\alpha k}\Gamma_{pq,k}\frac{dx^p}{ds}\frac{dx^q}{ds} = 0,$

which gives $\delta_q^\alpha\frac{d^2x^q}{ds^2} + \Gamma_{pq}^\alpha\frac{dx^p}{ds}\frac{dx^q}{ds} = 0.$

Hence, $$\boxed{\frac{d^2x^\alpha}{ds^2} + \Gamma_{pq}^\alpha\frac{dx^p}{ds}\frac{dx^q}{ds} = 0} \qquad (6.16)$$

This equation is called the **geodesic equation**.

Example: Geodesics in Flat Cartesian Space

As an illustration, let us show that the geodesics in flat Cartesian space are straight lines. From Eq. (6.16), the geodesic equation is

$$\frac{d^2x^\alpha}{ds^2} + \Gamma_{pq}^\alpha\frac{dx^p}{ds}\frac{dx^q}{ds} = 0.$$

For two-dimensional Cartesian coordinates,

$$ds^2 = dx^2 + dy^2$$

the metric tensor is

$$g_{ij} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

whose components are constants. Hence, all first derivatives of the metric vanish, and therefore,

$$\Gamma_{pq}^{\alpha} = 0$$

Consequently, the geodesic equation reduces to

$$\frac{d^2x^{\alpha}}{ds^2} = 0$$

In two dimensions, this gives $\frac{d^2x}{ds^2} = 0,\ \frac{d^2y}{ds^2} = 0$

Integrating each equation twice with respect to the arc-length parameter s,

$$x = A_1 s + B_1$$

$$y = A_2 s + B_2$$

where A_1, A_2, B_1, and B_2 are constants of integration.

Eliminating the parameter s,$S = \frac{x-B_1}{A_1}$
and substituting into the equation for y,

$$y = \frac{A_2}{A_1}(x - B_1) + B_2$$

Hence, $y = mx + c$

where, $m = \frac{A_2}{A_1},\ c = B_2 - \frac{A_2}{A_1}B_1$

Thus, the geodesics in flat Cartesian space are straight lines, confirming that Eq. (6.16) correctly generalizes the equation of a straight line to curved spaces.

6.6 RIEMANN–CHRISTOFFEL CURVATURE TENSOR

The concept of curvature is fundamental in differential geometry and tensor analysis. In a curved space, the result of transporting a vector around a closed loop depends on the path taken. This path dependence is a manifestation of curvature. The mathematical quantity that measures this intrinsic curvature is the Riemann–Christoffel curvature tensor, developed through the work of Bernhard Riemann and Elwin Bruno Christoffel.

The curvature tensor is defined in terms of the Christoffel symbols Γ_{pq}^{α} and their derivatives. Its components are given by

$$R_{pqk}^{\alpha} = \frac{\partial \Gamma_{pk}^{\alpha}}{\partial x^q} - \frac{\partial \Gamma_{pq}^{\alpha}}{\partial x^k} + \Gamma_{\beta q}^{\alpha}\Gamma_{pk}^{\beta} - \Gamma_{\beta k}^{\alpha}\Gamma_{pq}^{\beta}. \qquad (6.17)$$

This tensor measures the failure of second covariant derivatives to commute. For any vector A^{α}

$$A^{\alpha}_{,pq} - A^{\alpha}_{,qp} = R^{\alpha}_{kpq}A^{k}. \qquad (6.18)$$

Thus, in a flat space, where curvature is absent, the above difference vanishes and the curvature tensor is zero. The Riemann tensor describes how much the geometry deviates from flat (Euclidean) space. If a vector is parallel transported around an infinitesimal closed loop, the change in the vector is proportional to the curvature tensor. Therefore, it provides a precise measure of intrinsic curvature that does not depend on the choice of coordinates.

Consider a covariant vector A_p. The covariant derivative with respect to x^q is given by

$$A_{p,q} = \frac{\partial A_p}{\partial x^q} - A_{\alpha}\Gamma^{\alpha}_{pq}\,. \qquad (6.19)$$

On differentiating equation (6.19) the equation takes the form

$$\left(A_{p,q}\right)_{,k} = \frac{\partial A_{pq}}{\partial x^k} - A_{\alpha q}\Gamma^{\alpha}_{pk} - A_{p\alpha}\Gamma^{\alpha}_{qk},$$

Dropping the comma in $A_{p,q}$ in $A_{p,q} = A_{pq}$,

$$\left(A_{p,q}\right)_{,k} = \frac{\partial A_{p,q}}{\partial x^k} - A_{\alpha,q}\Gamma^{\alpha}_{pk} - A_{p,\alpha}\Gamma^{\alpha}_{qk},$$

$$= \frac{\partial}{\partial x^k}\left(\frac{\partial A_p}{\partial x^q} - A_{\alpha}\Gamma^{\alpha}_{pq}\right) - \left(\frac{\partial A_{\alpha}}{\partial x^q} - A_{\beta}\Gamma^{\beta}_{\alpha q}\right)\Gamma^{\alpha}_{pk} - \left(\frac{\partial A_p}{\partial x^{\alpha}} - A_{\beta}\Gamma^{\beta}_{p\alpha}\right)\Gamma^{\alpha}_{jk},$$

$$= \frac{\partial^2 A_p}{\partial x^k \partial x^q} - \frac{\partial A_{\alpha}}{\partial x^k}\Gamma^{\alpha}_{pq} - A_{\alpha}\frac{\partial}{\partial x^k}\left(\Gamma^{\alpha}_{pq}\right) - \frac{\partial A_{\alpha}}{\partial x^q}\left(\Gamma^{\alpha}_{pk}\right) + A_{\beta}\Gamma^{\beta}_{\alpha p}\Gamma^{\alpha}_{pk} - \frac{\partial A_p}{\partial x^{\alpha}}\left(\Gamma^{\alpha}_{qk}\right) + A_{\beta}\Gamma^{\beta}_{p\alpha}\Gamma^{\alpha}_{pk},$$

$$A_{pq,k} = \left[\frac{\partial^2 A_p}{\partial x^k \partial x^q} - \frac{\partial A_{\alpha}}{\partial x^k}\Gamma^{\alpha}_{pq} - \frac{\partial A_{\alpha}}{\partial x^q}\left(\Gamma^{\alpha}_{pk}\right) - \frac{\partial A_p}{\partial x^{\alpha}}\left(\Gamma^{\alpha}_{qk}\right) + A_{\beta}\Gamma^{\beta}_{p\alpha}\Gamma^{\alpha}_{qk}\right] - A_{\alpha}\frac{\partial}{\partial x^k}\left(\Gamma^{\alpha}_{pq}\right) + A_{\beta}\Gamma^{\beta}_{\alpha q}\Gamma^{\alpha}_{pk} \qquad (6.20)$$

Now interchanging $(q \leftrightarrow k)$ in equation (6.20) we get

$$A_{pk,q} = \left[\frac{\partial^2 A_p}{\partial x^k \partial x^p} - \frac{\partial A_{\alpha}}{\partial x^q}\Gamma^{\alpha}_{pk} - \frac{\partial A_{\alpha}}{\partial x^k}\left(\Gamma^{\alpha}_{pq}\right) - \frac{\partial A_p}{\partial x^{\alpha}}\left(\Gamma^{\alpha}_{kq}\right) + A_{\beta}\Gamma^{\beta}_{p\alpha}\Gamma^{\alpha}_{kq}\right] - A_{\alpha}\frac{\partial}{\partial x^q}\left(\Gamma^{\alpha}_{pk}\right) + A_{\beta}\Gamma^{\beta}_{\alpha k}\Gamma^{\alpha}_{pq} \qquad (6.21)$$

Now subtracting equation (6.21 from (6.20)

$$A_{pq,k} - A_{pk,q} = A_{\alpha}\left[\frac{\partial \Gamma^{\alpha}_{pk}}{\partial x^q} - \frac{\partial \Gamma^{\alpha}_{pq}}{\partial x^k} + \Gamma^{\alpha}_{\beta q}\Gamma^{\beta}_{pk} - \Gamma^{\alpha}_{\beta k}\Gamma^{\beta}_{pq}\right]$$

$$A_{pq,k} - A_{pk,q} = A_\alpha R^\alpha_{pqk}.$$

Here

$$R^\alpha_{pqk} = \frac{\partial \Gamma^\alpha_{pk}}{\partial x^q} - \frac{\partial \Gamma^\alpha_{pq}}{\partial x^k} + \Gamma^\alpha_{\beta q}\Gamma^\beta_{pk} - \Gamma^\alpha_{\beta k}\Gamma^\beta_{pq}. \qquad (6.22)$$

Equation (6.22) is known as Riemann Christoffel curvature tensor.

Consider a two-dimensional sphere (2-sphere) of radius a, described by the metric

$$ds^2 = a^2(d\theta^2 + sin^2\theta d\phi^2)$$

Using the non-vanishing Christoffel symbols for the 2-sphere,

$$\Gamma^\theta_{\phi\phi} = -sin\,\theta cos\theta$$

And $$\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = cot\theta$$

we evaluate the component $R^\theta_{\phi\theta\phi}$.

Substituting $\alpha = \theta, p = \phi, q = \theta$, and $k = \phi$ into equation (6.22), we obtain

$$R^\theta_{\phi\theta\phi} = \frac{\partial \Gamma^\theta_{\phi\phi}}{\partial \theta} - \frac{\partial \Gamma^\theta_{\phi\theta}}{\partial \phi} + \Gamma^\theta_{\beta\phi}\,\Gamma^\beta_{\phi\phi} - \Gamma^\theta_{\beta\phi}\Gamma^\beta_{\phi\theta}.$$

Since $\Gamma^\theta_{\phi\theta} = 0$ and $\Gamma^\theta_{\theta\theta} = 0$ (β is a dummy index, $\beta = \theta, \phi$)
the expression reduces to

$$R^\theta_{\phi\theta\phi} = \frac{\partial}{\partial\theta}(-sin\theta\, cos\theta) - \Gamma^\theta_{\phi\phi}\Gamma^\phi_{\phi\theta},$$
$$R^\theta_{\phi\theta\phi} = \sin^2\theta - \cos^2\theta + \cos^2\theta,$$
$$R^\theta_{\phi\theta\phi} = \sin^2\theta.$$

Now Lowering the index with the metric tensor,

$$R_{\theta\phi\theta\phi} = g_{\theta\theta}R^\theta_{\phi\theta\phi}.$$

we obtain

$$R_{\theta\phi\theta\phi} = a^2\sin^2\theta.$$

Hence, the non-zero Riemann tensor confirms that the 2-sphere is an intrinsically curved surface.

Symmetry Properties

The Riemann tensor possesses several important symmetry relations

1. Antisymmetric in the last two indices
$$R^\alpha_{kpq} = -R^\alpha_{kqp}$$
2. Antisymmetric in the first two covariant indices (after lowering)
$$R_{\alpha kpq} = -R_{k\alpha pq}$$
3. Pair symmetry
$$R_{\alpha kpq} = R_{pqk\alpha.}$$

These symmetries greatly reduce the number of independent components.

Contractions

By contracting indices of the Riemann tensor, we obtain other important tensors:

1. Ricci Tensor

$$R_{kq} = R^{p}_{kpq}.$$

The Ricci tensor measures how the volume of a small ball of particles changes as it moves through curved spacetime.

2. Ricci Scalar (Scalar Curvature)

$$R = g^{kq}R_{kq}.$$

It gives a single number representing the average curvature of spacetime at a point.

For the 2-sphere, the non-zero component of the Riemann curvature tensor is

$$R^{\theta}_{\phi\theta\phi} = \sin^2\theta.$$

Hence,

$$R_{\phi\phi} = R^{p}_{\phi p\phi} = R^{\theta}_{\phi\theta\phi} = \sin^2\theta.$$

Similarly,

$$R_{\theta\theta} = R^{p}_{\theta p\theta} = R^{\phi}_{\theta\phi\theta} = 1.$$

Therefore, the non-zero Ricci tensor components for the 2-sphere are

$$R_{\theta\theta} = 1, \qquad R_{\phi\phi} = \sin^2\theta.$$

The non-zero Ricci tensor components indicate that the 2-sphere is an intrinsically curved space with positive curvature everywhere on its surface.

The Ricci scalar is obtained by contacting the Ricci tensor with the inverse metric tensor

$$R = g^{kq}R_{kq}.$$

Using

$$g^{\theta\theta} = \frac{1}{a^2}, \qquad g^{\phi\phi} = \frac{1}{a^2\sin^2\theta},$$

we obtain

$$R = g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi},$$

which gives

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

Hence,

$$R = \frac{2}{a^2}$$

Thus, the Ricci scalar for a 2-sphere of radius a is positive and constant everywhere on the surface, indicating that the sphere has uniform positive curvature at every point.

6.7 BIANCHI IDENTITIES

The curvature tensor satisfies certain algebraic and differential relations known collectively as the Bianchi identities. The curvature tensor also satisfies certain differential identities known as the Bianchi identities.

First Bianchi Identity (Algebraic)

The first Bianchi identity expresses a cyclic symmetry of the curvature tensor and is given by

$$R^{r}_{spq} + R^{r}_{qsp} + R^{r}_{psq} = 0. \qquad (6.23)$$

This identity follows directly from the symmetry and antisymmetric properties of the Riemann curvature tensor. It is therefore purely algebraic in nature and does not involve derivatives.

The second Bianchi identity involves the covariant derivatives of the curvature tensor and is written as

$$R^r_{spq,t} + R^r_{sqt,p} + R^r_{stp,q} = 0. \qquad (6.24)$$

This relation expresses a fundamental differential property of curvature and may be interpreted as a conservation law for the geometric structure of the space. By contracting indices in this identity, one obtains important results such as the conservation of the Ricci tensor and scalar curvature.

CHECK YOUR PROGRESS

Q1. Define the Christoffel symbols of the first and second kinds. Show their mathematical expressions and state their symmetry properties.

Q 2. Prove (a) $[ab,r] = [ba,r]$ (b) (c) $[pq,r] = g_{rs}\Gamma^s_{pq}$.

Q 3. Write the covariant derivative with respect to x^q of each of the following tensors

(a) A_{jk} (b)A^k_j

Q 4. Define Geodesics.

Q 5. Prove that $R^\alpha_{pqk} = -R^\alpha_{pkq}$

Q6. Show that Christoffel symbols of the second kind are symmetric in the lower indices, i.e., $\Gamma^k_{ij} = \Gamma^k_{ji}$

Q7. Show that if the Riemann curvature tensor vanishes, the space is flat.

6.8 LET US SUM UP

- In general, curvilinear coordinates, the basis vectors vary with position, so the metric tensor is introduced to define distances, angles, and the geometric structure of space.
- Special combinations of derivatives of the metric, first introduced by Elwin Bruno Christoffel, describe how coordinate bases change from point to point. These Christoffel symbols are not tensors but act as connection coefficients and are essential for defining covariant differentiation, parallel transport, and geodesics.
- Ordinary partial derivatives do not preserve tensor character in curved or curvilinear coordinates.
- The covariant derivative, which includes correction terms involving Christoffel symbols, ensures that derivatives of vectors and tensors transform properly as tensors, extending naturally to higher-rank tensors.
- Geodesics represent the shortest or extremal paths between points and generalize straight lines to curved spaces. Their equations involve Christoffel

symbols, showing how curvature influences motion; in physics, they describe the trajectory of freely moving particles.

- The intrinsic curvature of space is measured by the curvature tensor developed from the ideas of Bernhard Riemann. This tensor, along with its contractions (Ricci tensor and scalar curvature) and the Bianchi identities, characterizes how a space deviates from flatness and forms the mathematical foundation of differential geometry and gravitational theory.

✓ 6.9 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q 1. Christoffel symbol of the first kind:

$$[ab,r]=\frac{1}{2}\left(\frac{\partial g_{ar}}{\partial x^b}+\frac{\partial g_{br}}{\partial x^a}-\frac{\partial g_{ab}}{\partial x^r}\right)$$

Its symmetry property is $[ab,r]$=$[ba,r]$ i.e. it is symmetric in the first two indices.

Christoffel symbol of the second kind:

$$\Gamma^i_{jk}=\frac{1}{2}g^{il}\left(\frac{\partial g_{lj}}{\partial x^k}+\frac{\partial g_{lk}}{\partial x^j}-\frac{\partial g_{jk}}{\partial x^l}\right)$$

Its symmetry property is $\Gamma^i_{jk}=\Gamma^i_{kj}$, i.e. it is symmetric in the two lower indices.

Answer to Q2: (a) The Christoffel symbols of the first kind are defined by

$$[a,br]=\frac{1}{2}\left(\frac{\partial g_{ar}}{\partial x^b}+\frac{\partial g_{br}}{\partial x^a}-\frac{\partial g_{ab}}{\partial x^r}\right)$$

Since the metric tensor is symmetric,

$$g_{ab}=g_{ba}$$

we also have

$$\frac{\partial g_{ar}}{\partial x^b}=\frac{\partial g_{ra}}{\partial x^b},\ \frac{\partial g_{br}}{\partial x^a}=\frac{\partial g_{rb}}{\partial x^a}$$

Interchanging the indices $a$and b,

$$[b,ar]=\frac{1}{2}\left(\frac{\partial g_{br}}{\partial x^a}+\frac{\partial g_{ar}}{\partial x^b}-\frac{\partial g_{ba}}{\partial x^r}\right)$$

Since $g_{ab}=g_{ba}$,

$$[b,ar]=\frac{1}{2}\left(\frac{\partial g_{ar}}{\partial x^b}+\frac{\partial g_{br}}{\partial x^a}-\frac{\partial g_{ab}}{\partial x^r}\right)=[a,br]$$

Hence, $$[a,br]=[b,ar]$$

(b) The Christoffel symbols of the second kind are related to those of the first kind by

$$\Gamma_{pq}^{s} = g^{sr}[p, qr]$$

Multiplying both sides by g_{rs},

$$g_{rs}\Gamma_{pq}^{s} = g_{rs}g^{s\alpha}[p, q\alpha].$$

Using the identity$g_{rs}g^{s\alpha} = \delta_r^{\alpha}$

we obtain$g_{rs}\Gamma_{pq}^{s} = \delta_r^{\alpha}[p, q\alpha] = [p, qr]$.

Therefore, $\boxed{[p, qr] = g_{rs}\Gamma_{pq}^{s}.}$

Answer to Q 3. (a)$A_{jk,q} = \frac{\partial A_{jk}}{\partial x^q} - \Gamma_{jk}^{s}A_{sk} - \Gamma_{kq}^{s}A_{js}$

(b)$A_{j,q}^{k} = \frac{\partial A_j^k}{\partial x^q} - \Gamma_{jq}^{s}A_s^k + \Gamma_{qs}^{k}A_j^s$

Answer to Q 4. A curve on a surface joining two points in an n-dimensional space is called the geodesics if its length is extremum**.**

Answer to Q 5. We have $R_{pqk}^{\alpha} = \frac{\partial \Gamma_{pk}^{\alpha}}{\partial x^q} - \frac{\partial \Gamma_{pq}^{\alpha}}{\partial x^k} + \Gamma_{\beta q}^{\alpha}\Gamma_{pk}^{\beta} - \Gamma_{\beta k}^{\alpha}\Gamma_{pq,}^{\beta}$

Now $(q \leftrightarrow k)$

$$R_{pkq}^{\alpha} = \frac{\partial \Gamma_{pq}^{\alpha}}{\partial x^k} - \frac{\partial \Gamma_{pk}^{\alpha}}{\partial x^k} + \Gamma_{\beta k}^{\alpha}\Gamma_{pq}^{\beta} - \Gamma_{\beta q}^{\alpha}\Gamma_{pk,}^{\beta}$$

$$R_{pkq}^{\alpha} = -\left[\frac{\partial \Gamma_{pk}^{\alpha}}{\partial x^q} - \frac{\partial \Gamma_{pq}^{\alpha}}{\partial x^k} + \Gamma_{\beta q}^{\alpha}\Gamma_{pk}^{\beta} - \Gamma_{\beta k}^{\alpha}\Gamma_{pq}^{\beta}\right] = -R_{pqk}^{\alpha}$$

Answer to Q6: We start with the definition of the Christoffel symbols of the second kind:

$$\Gamma_{ij}^{k} = \frac{1}{2}g^{k\alpha}\left(\frac{\partial g_{i\alpha}}{\partial x^j} + \frac{\partial g_{j\alpha}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^{\alpha}}\right)$$

$$\Gamma_{ji}^{k} = \frac{1}{2}g^{k\alpha}\left(\frac{\partial g_{j\alpha}}{\partial x^i} + \frac{\partial g_{i\alpha}}{\partial x^j} - \frac{\partial g_{ji}}{\partial x^{\alpha}}\right)$$

Since the metric tensor is symmetric:

$$g_{ij} = g_{ji}$$

So, $\frac{\partial g_{ji}}{\partial x^{\alpha}} = \frac{\partial g_{ij}}{\partial x^{\alpha}}$

Now, $\Gamma_{ji}^{k} = \frac{1}{2}g^{k\alpha}\left(\frac{\partial g_{i\alpha}}{\partial x^j} + \frac{\partial g_{j\alpha}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^{\alpha}}\right) = \Gamma_{ij}^{k}$

Thus, $\Gamma_{ij}^{k} = \Gamma_{ji}^{k}$

Answer to Q7: The Riemann curvature tensor is given by:

$$R_{pqk}^{\alpha} = \frac{\partial \Gamma_{pk}^{\alpha}}{\partial x^q} - \frac{\partial \Gamma_{pq}^{\alpha}}{\partial x^k} + \Gamma_{\beta q}^{\alpha}\Gamma_{pk}^{\beta} - \Gamma_{\beta k}^{\alpha}\Gamma_{pq}^{\beta}$$

Given: $R^{\alpha}_{pqk} = 0$

The Riemann curvature tensor measures the failure of covariant derivatives to commute:

$$A^{\alpha}_{,pq} - A^{\alpha}_{,qp} = R^{\alpha}_{kpq}A^{k}$$

If $R^{\alpha}_{pqk} = 0.$ then $A^{\alpha}_{,pq} = A^{\alpha}_{,qp}$

This implies that covariant derivatives commute. Furthermore, when $R^{\alpha}_{pqk} = 0$, then there exists a coordinate system in which $\Gamma^{\alpha}_{pq} = 0$. In such a system, metric becomes constant and geometry reduces to Euclidean (flat) space.

Thus, if $R^{\alpha}_{pqk} = 0$, the space is flat.

6.10 FURTHER READING

- Dass, H. K., and Verma, R., *Mathematical Physics*, (2022) 8th Edition, S. Chand & Company.
- Arfken, G. B., Weber, H. J., and Harris, F. E. (2013), *Mathematical Methods for Physicists: A Comprehensive Guide*, 7th Edition, Elsevier.
- Boas, M. L. (2006), *Mathematical Methods in the Physical Sciences*, 3rd Edition, Wiley.
- Kreyszig, E. (2011), *Advanced Engineering Mathematics*, 10th Edition, Wiley.

6.11 MODEL QUESTIONS

1. Evaluate the Christoffel symbols of (a) the first kind, (b) the second kind, for spaces where $g_{pq} = 0$ if $p \neq q$.
2. Find the differential equation of the geodesics in the space with the metric
$$ds^2 = e^{kt}(dx^2 + dy^2 + dz^2) + dt^2.$$
3. Derive the expression for differential equation of geodesics.
4. Show that the geodesics in a Riemannian space are given by $\frac{d^2x^r}{ds^2} + \Gamma^r_{pq}\frac{dx^p}{ds}\frac{ds^q}{ds} = 0.$
5. Define Riemann–Christoffel curvature tensor and discuss its properties.
6. If A_p and A^P are tensors show that (a) $A_{p,q} = \frac{\partial A_p}{\partial x^q} - \Gamma^s_{pq}A_s$, and
(b) $A^p_{,q} = \frac{\partial A^p}{\partial x^q} + \Gamma^p_{qs}A^s$ are tensors.
7. Establish Bianchi identities $R^r_{spq} + R^r_{qsp} + R^r_{psq} = 0.$

*** *** ***

UNIT 7: DIFFERENTIAL EQUATIONS

UNIT STRUCTURE

7.1 Learning Objectives
7.2 Introduction
7.3 Order of a Differential Equation
7.4 Degree of a Differential Equation
7.5 First Order Linear Differential Equation
7.6 Second order Linear Differential Equation
7.7 Let Us Sum Up
7.8 Answers to Check Your Progress
7.9 Further Reading
7.10 Model Questions

7.1 LEARNING OBJECTIVES

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

- define differential equations and their types
- identify order and degree
- classify linear/nonlinear and homogeneous/non-homogeneous equations
- solve first and second order differential equations
- apply differential equations to simple physical problems.

7.2 INTRODUCTION

In the previous unit, we studied Tensor Calculus and Geometry, which provided important mathematical tools for describing physical and geometrical systems. In this unit, we introduce differential Equations, which describe the rate of change of a quantity and play an important role in understanding many physical phenomena. A differential equation establishes a relation between an unknown function and its derivatives and is widely used in physics to study motion, oscillations, heat transfer, and electrical circuits. The solution of a differential equation helps us determine the behaviour of a physical system under given conditions.

In this unit, we shall study the basic concepts and types of differential equations, their order and degree, and methods of solving first-order and second-order differential equations, including the variable separation method, linear method, and the methods of finding the complementary function (C.F.) and particular integral (P.I.).

7.3 ORDER OF A DIFFERENTIAL EQUATION

Differential Equation

A differential equation is an equation that involves an unknown function and one or more of its derivatives with respect to one or more independent variables. It expresses a relationship between the function and its rate of change, and is used to describe many physical, biological, and engineering processes.

A differential equation tells us how a physical quantity changes and gives the basic law that governs the system. When we solve the differential equation, we get the actual function that describes the system's behaviour. In most cases, the solution contains constants, so it represents a family of curves rather than a single curve. Each curve corresponds to a different possible situation depending on the initial or boundary conditions. A classic example of a differential equation in physics is the equation of linear harmonic motion, also known as simple harmonic motion, such as the oscillation of a mass attached to a spring. The differential equation describing this motion is given by

$$\frac{d^2y}{dt^2} + \omega^2 y = 0$$

The general solution of the differential equation is $y(t) = A\cos(\omega t + \phi)$. This solution represents the oscillatory motion of the particle with time. Here, A is called the amplitude of the motion. Physically, it gives the maximum displacement of the particle from its equilibrium position, meaning it tells how far the particle moves on either side during oscillation. The quantity ϕ is known as the phase constant (or initial phase). It determines the initial position of the particle at time t=0 and depends on the starting conditions of the motion. Thus, different values of A and ϕ correspond to different possible oscillations, while the angular frequency ω decides how rapidly the particle oscillates. Based on their characteristics, differential equations can be broadly classified as follows:

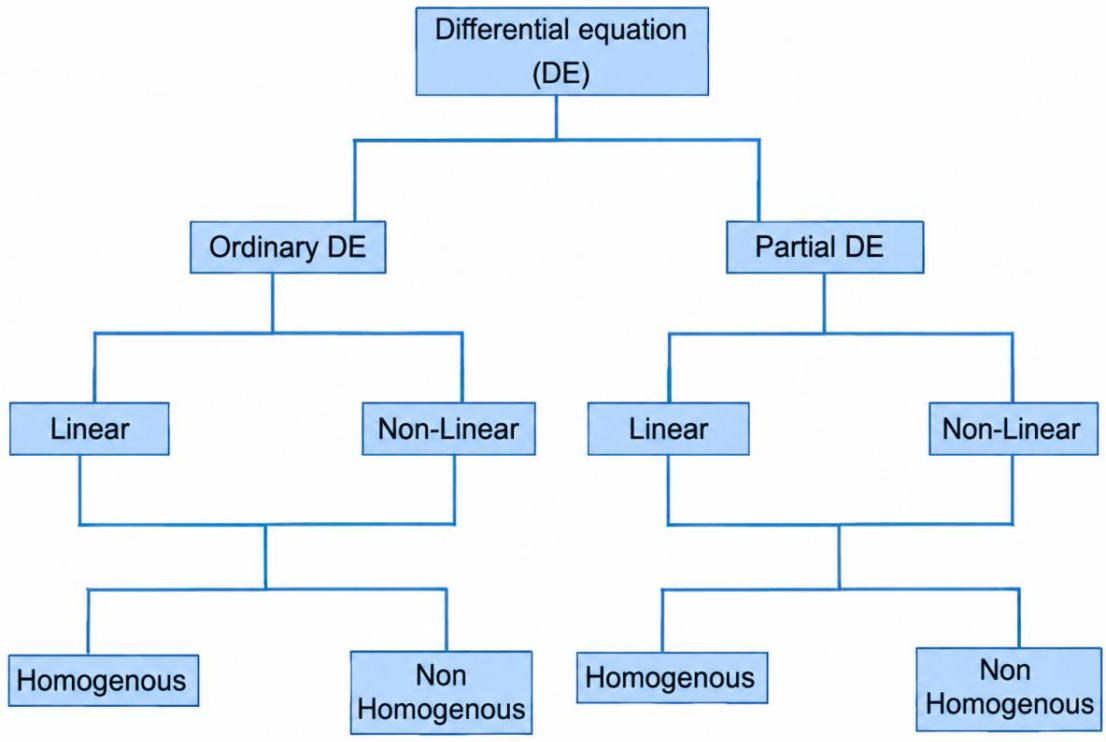


Figure 7.1: Classification of Differential Equations

Ordinary Differential Equation (ODE)

An ordinary differential equation contains derivatives of the dependent variable with respect to only one independent variable. The general form of an ODE can be written as

$$F\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \dots\dots\dots\dots\dots\right) = 0$$

Where y is the unknown function of x, and $x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}$ etc., represent its derivatives.

For example, the equation

$$\frac{d^2y}{dx^2} + 2\,\frac{dy}{dx} + y = f(x)$$

is an ordinary differential equation because it includes the derivatives of y with respect to x. Ordinary differential equations are widely used in science and engineering to model physical phenomena such as motion, growth, heat transfer, and electrical circuits.

Linear And Non Linear Differential Equation

A linear differential equation is a differential equation in which:

- The dependent variable y and its derivative $\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}$ should have power 1.
- There should not be trigonometric, logarithmic, exponential function of dependent variable and its derivatives.
- There should not be cross multiplication terms of dependent variable and its derivatives i.e. $y\frac{dy}{dx}, y\frac{d^2y}{dx^2}$ and $\left(\frac{dy}{dx}\right)\left(\frac{d^2y}{dx^2}\right)$
- *y* is not multiplied with itself or with its derivatives.
- Terms like $y^2, \left(\frac{dy}{dx}\right)^2, \left(\frac{d^2y}{dx^2}\right)^2, \left(\frac{dy}{dx}\right)\left(\frac{d^2y}{dx^2}\right)$ should not occur.
- Example of Linear Equation

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

This is linear because y and its derivatives are in power 1.

A non-linear differential equation is a differential equation that does not follow the criteria of a linear differential equation.

Example of Non-Linear Equation

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

$$\frac{d^2y}{dx^2} + x^2\frac{dy}{dx} + ySinx = \ln(1+x)$$

$$\frac{d^2y}{dx^2} + Sin(x+y) = \ln(1+x)$$

Homogeneous and Non-Homogeneous Differential Equation

If, on the left-hand side of a differential equation, all the terms contain only the dependent variable and its derivatives, and the right-hand side is zero, then it is called a homogeneous differential equation. If the right-hand side is non-zero, then it is called a non-homogeneous differential equation.

Example of Linear Equation

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

Example of Non-Linear Equation

$$\frac{d^2y}{dx^2} + 2x\,\frac{dy}{dx} + y = \sin(x+y)$$

This equation is non-homogeneous because the right-hand side is not zero. It is also non-linear because the dependent variable y appears inside the trigonometric function $\sin(x+y)$.

The order of a differential equation is defined as the order of the highest derivative present in the equation. It indicates the maximum number of times the dependent variable has been differentiated with respect to the independent variable.

Consider the nonlinear homogeneous differential equation:

$$\frac{d^3y}{dx^3} + \left(\frac{d^2y}{dx^2}\right)^4 + 2x\frac{dy}{dx} + \sin y = 0$$

highest order derivative is: $\frac{d^3y}{dx^3}$. Hence, the order of the differential equation is 3.

In this equation, derivatives of different orders appear:

- First derivative: $\frac{dy}{dx}$
- Second derivative: $\frac{d^2y}{dx^2}$
- Third derivative: $\frac{d^3y}{dx^3}$

Example 1: In $\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + y = 0$

Highest derivative: $\frac{d^2y}{dx^2}$. Therefore, Order = 2

Example 2. In $\frac{dy}{dx} + x^2y = e^x$

Highest derivative: $\frac{dy}{dx}$. Therefore, Order = 1

Example 3: In $\left(\frac{d^2y}{dx^2}\right)^2 + \frac{dy}{dx} + y = 0$

Although the second derivative is squared, the highest derivative present is still $\frac{d^2y}{dx^2}$.

Hence, Order = 2

Example 4: (Partial Differential Equation)

$$\frac{\partial^2u}{\partial x^2} + \frac{\partial u}{\partial t} = 0$$

Highest derivative: $\frac{\partial^2u}{\partial x^2}$. Therefore, Order = 2

7.4 DEGREE OF A DIFFERENTIAL EQUATION

The degree of a differential equation is defined only when the equation can be expressed as a polynomial in its derivatives. It is the highest power of the highest-order derivative after removing radicals and fractional powers involving the derivatives.

If the highest-order derivative appears inside a transcendental function (such as $\sin$, $\cos$, $\log$, or e^x) or under a fractional power, the degree of the differential equation is not defined.
For example,

i) $\left(\frac{d^2y}{dx^2}\right)^3 + \frac{dy}{dx} + y = 0$, the highest-order derivative is $\frac{d^2y}{dx^2}$

Order = 2, degree =3

ii) $\frac{dy}{dx} = \tan(x + y)$

Order =1, Degree=1

iii) $tan\left(\frac{dy}{dx}\right) = x + y$

Order =1, Degree = not defined**,** because the equation cannot be expressed as a polynomial in the derivatives.

7.5 FIRST ORDER LINEAR DIFFERENTIAL EQUATION

If a differential equation can be written in the form $\frac{dy}{dx} = f(x, y)$ then we can use various methods to solve. Here we will discuss only two methods: the method of separation of variables and the method for solving first order linear equations.

Variable Separation Method
This method works when the equation can be written like:

$$\frac{dy}{dx} = f(x)g(y)$$

Then we can separate

$$\frac{1}{g(y)}dy = f(x)dx$$

And integrate both sides

$$\int \frac{1}{g(y)}dy = \int f(x)dx$$

Partial Differential Equation:
A partial differential equation (PDE) is a differential equation involving partial derivatives of an unknown function with respect to two or more independent variables. Unlike an ordinary differential equation (ODE), which contains derivatives with respect to only one independent variable, a PDE describes functions of several variables. The concept of order is defined in the same way for both ODEs and PDEs. Although this unit primarily focuses on ODEs, the following example illustrates how the order of a PDE is determined.

Example 1: Solve $\frac{dy}{dx} = xy$

Solution:

$$\frac{dy}{dx} = xy$$
$$\Rightarrow \int \frac{dy}{dx} = \int xy$$
$$\Rightarrow \int \frac{dy}{y} = \int xdx$$
$$\Rightarrow lny = \frac{x^2}{2} + ln\, C$$
$$\Rightarrow ln\frac{y}{C} = \frac{x^2}{2}$$
$$\Rightarrow \frac{y}{c} = e^{\frac{x^2}{2}}$$
$$\Rightarrow y = Ce^{\frac{x^2}{2}}$$

Example 2: Solve $\frac{dy}{dx} - ytan2x = 0; y(0) = 2$

Solution:

$$\frac{dy}{dx} - ytan2x = 0$$
$$\Rightarrow \frac{dy}{dx} = ytan2x$$
$$\Rightarrow \frac{dy}{y} = tan2xdx$$
$$\Rightarrow \int \frac{dy}{y} = \int tan2xdx$$
$$\Rightarrow lny = \frac{1}{2}\ln(\sec 2x) + lnC$$
$$\Rightarrow lny = \ln(\sec 2x)^2 + lnC$$
$$\Rightarrow lny - lnC = \ln\sqrt{Sec\ 2x}$$
$$\Rightarrow \ln\frac{y}{C} = ln\sqrt{Sec\ 2x}$$
$$\Rightarrow y = C\sqrt{Sec\ 2x}$$

Putting $x = 0, y = 2$

$$2 = C\ ln\sqrt{Sec\ 0}$$
$$\Rightarrow 2 = Cln\sqrt{1}$$
$$\Rightarrow C = 2$$

Hence, $y = 2\sqrt{Sec\ 2x}$

Example 3: Solve $\frac{dy}{dx} = (4x + y + 1)^2$

Solution:

$$\frac{dy}{dx} = (4x + y + 1)^2$$
$$Let; 4x + y + 1 = v$$
$$\frac{dv}{dx} = 4 + \frac{dy}{dx}$$
$$\Rightarrow \frac{dv}{dx} - 4 = \frac{dy}{dx}$$

$$\Longrightarrow \frac{dv}{dx} - 4 = v^2$$
$$\Longrightarrow \frac{dv}{dx} = v^2 + 4$$
$$\Longrightarrow \frac{dv}{v^2 + 4} = dx$$
$$\Longrightarrow \int \frac{dv}{v^2 + 4} = \int dx$$

$$\Longrightarrow \frac{1}{2} tan^{-1}\left(\frac{v}{2}\right) = x + c$$
$$\Longrightarrow tan^{-1}\left(\frac{v}{2}\right) = 2x + c$$
$$\Longrightarrow v = 2\tan\,(2\,x + c)$$

Since, $v = 4x + y + 1$

$$\boxed{y = 2\tan(2x + C) - 4x - 1}$$

Example 4: Solve $\frac{dy}{dx} = 1 + \tan(y - x)$

Solution:

Let $y - x = z$

Differentiating it

$$\frac{dy}{dx} - 1 = \frac{dz}{dx}$$
$$\Rightarrow \frac{dy}{dx} = 1 + \frac{dz}{dx}$$
$$\Rightarrow 1 + \tan(y - x) = 1 + \frac{dz}{dx}$$
$$\Rightarrow 1 + \tan z = 1 + \frac{dz}{dx}$$
$$\Rightarrow \tan z = \frac{dz}{dx}$$
$$\Rightarrow \int \frac{dz}{tanz} = \int dx$$

$$\Rightarrow \int \cot z\, dz = \int dx$$
$$\Rightarrow \ln(\sin z) = x + c$$
$$\Rightarrow \sin \mathrm{z} = Ae^{x+c}$$
$$\Longrightarrow \sin(y - x) = Ae^{x}$$

Note: Other Types of First-Order Differential Equations

Besides equations that can be solved by separation of variables, there are other important classes of first-order differential equations. Two commonly encountered types are exact differential equations and Bernoulli differential equations**.**

An equation of the form

$$M(\mathrm{x}, \mathrm{y})\, dx + N(\mathrm{x}, \mathrm{y})\, dy = 0$$

is called an exact differential equation if it satisfies the condition

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

When this condition is satisfied, the solution is obtained by finding a function $F(\mathrm{x},\mathrm{y})$ such that

$$dF = M\,dx + N\,dy,$$

giving the general solution

$$F(\mathrm{x},\mathrm{y}) = C.$$

For example, the equation

$$(2x + y)\,dx + x\,dy = 0$$

is exact because

$$\frac{\partial(2x+y)}{\partial y} = \frac{\partial(x)}{\partial x} = 1,$$

and its solution is

$$x^2 + xy = C.$$

Another important type is the Bernoulli differential equation, whose general form is

$$\frac{dy}{dx} + P(x)y = Q(x)y^n,\ n \neq 0,1.$$

Although this equation is nonlinear, it can be transformed into a linear differential equation by introducing the substitution

$$z = y^{\,1-n}$$

For example, the equation

$$\frac{dy}{dx} + y = y^2$$

can be solved by taking

$$z = \frac{1}{y}$$

which converts it into a linear differential equation. The final solution is

$$y = \frac{1}{1 + Ce^x}$$

Linear Differential Equations

A differential equation of the form $\frac{dy}{dx} + Py = Q$ is called a linear differential equation, where P and Q are functions of x (but not of y) or constants

In such case, we have to multiply both sides of the equation by $e^{\int P dx}$

Hence,

$$e^{\int P dx}\left(\frac{dy}{dx} + Py\right) = Qe^{\int P dx}$$

Here the term $e^{\int P dx}$ is called integrating factor (I.F). The solution of this equation is

$$y \times [I.F] = \int Q[I.F]dx + C$$

Example 5: Solve $(x+1)\frac{dy}{dx} - y = e^x(x+1)^2$

Solution:

$$\frac{dy}{dx} - \frac{y}{1+x} = e^x(x+1)$$

Here the IF

$$e^{-\int \frac{dx}{1+x}} = e^{-\log(1+x)} = e^{\log(1+x)^{-1}} = \frac{1}{1+x}$$

The solution is

$$y.\frac{1}{x+1} = \int e^x.(x+1).\frac{1}{x+1}dx$$
$$\Rightarrow \frac{y}{1+x} = \int e^x dx$$
$$\Rightarrow \frac{y}{1+x} = e^x + C$$

Example 6: Solve $x(x-1)\frac{dy}{dx} - (x-2)y = x^2(2x-1)$

Solution:

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

Hence the integrating factor

$$e^{-\int \frac{x-2}{x(x-1)}dx} = e^{\int \left(\frac{1}{x-1} - \frac{2}{x}\right)dx} = e^{\log(x-1) - 2logx} = e^{log\frac{x-1}{x^2}} = \frac{x-1}{x^2}$$

The solution is

$$y.\frac{x-1}{x^2} = \int \frac{x^2(2x-1)}{x(x-1)}.\frac{x-1}{x^2}dx + C$$
$$= \int \frac{2x-1}{x}dx + C$$
$$= \int \left(2 - \frac{1}{x}\right)dx + C$$
$$= 2x - logx + C$$

Example 7: Solve $\frac{dy}{dx} + \frac{1}{x}y = x^3 - 3$

Solution:

$$IF = e^{\int \frac{1}{x}dx} = e^{lnx} = x$$

The solution is

$$yx = \int (x^3 - 3)xdx + C$$
$$\Rightarrow yx = \int (x^4 - 3x)dx + C$$
$$\Rightarrow yx = \frac{x^5}{5} - \frac{3x^2}{2} + C$$
$$\Rightarrow y = \frac{1}{x}\left(\frac{x^5}{5} - \frac{3x^2}{2} + C\right)$$
$$\Rightarrow y = \left(\frac{x^4}{5} - \frac{3x}{2} + \frac{C}{x}\right)$$

Example 8: Solve $\frac{dy}{dx} + ycot\, x = cos\, x$

Solution: $IF = e^{\int cotxdx} = e^{\ln(sinx)} = \sin x$

The solution is

$$ysinx = \int (cosx)(sinx)dx + C$$
$$\Rightarrow ysinx = \int sinxcosxdx + C$$

$$\Rightarrow ysinx = \frac{1}{2}\int 2sinxcosdx + C$$
$$\Rightarrow ysinx = \frac{1}{2}\int sin2xdx + C$$
$$\Rightarrow ysinx = -\frac{1}{4}cos2x + C$$
$$y = -\frac{1}{4}\frac{cos2x}{sinx} + \frac{C}{Sinx}$$

Example 9: A capacitor of capacitance C is initially charged to a charge Q_0. At time $t = 0$, it is connected across a resistor of resistance R and allowed to discharge. Determine the charge remaining on the capacitor at any time t.

Solution: Let Q be the charge on the capacitor at any instant t. As the capacitor discharges, an electric current flows through the resistor. Since the charge on the capacitor decreases with time, the current is given by

$$I = -\frac{dQ}{dt}$$

where the negative sign indicates that the charge is decreasing. According to Kirchhoff's Voltage Law, the algebraic sum of all potential differences around a closed circuit is zero. During the discharge process, the potential drop across the resistor is IR, while the potential difference across the capacitor is Q/C.

Therefore, $IR + \frac{Q}{C} = 0$

Substituting $I = -\frac{dQ}{dt}$ into the above equation, we get

$$-R\frac{dQ}{dt} + \frac{Q}{C} = 0.$$

Rearranging the equation, $R\frac{dQ}{dt} = -\frac{Q}{C}$,

Or $\boxed{\frac{dQ}{dt} + \frac{Q}{RC} = 0.}$

This is a first-order differential equation that can be solved by the method of separation of variables.

Separating the variables, $\frac{dQ}{Q} = -\frac{dt}{RC}$

Integrating both sides, we have, $\int \frac{dQ}{Q} = -\frac{1}{RC}\int d\,t$

$$\Rightarrow \ln|Q| = -\frac{t}{RC} + C_1$$

where C_1 is the constant of integration.

Taking the exponential of both sides,

$$Q = e^{C_1}e^{-t/RC}.$$

Let $e^{C_1} = A$, where A is a constant. Then,

$$Q = Ae^{-t/RC}.$$

To determine the value of A, we use the initial condition. At the instant the discharge begins $(t = 0)$, the charge on the capacitor is Q_0. Hence,

$$Q_0 = Ae^0 = A.$$

Therefore, $A = Q_0$,

and the required solution becomes

$$\boxed{Q(t) = Q_0 e^{-t/RC}}$$

Thus, the charge remaining on the capacitor decreases exponentially with time. This equation shows that the charge on the capacitor decreases exponentially with time. Initially, the charge decreases rapidly, but the rate of discharge gradually becomes slower.

In this equation, $Q(t)$ is the charge remaining on the capacitor at time t, Q_0 is the initial charge, R is the resistance, and C is the capacitance of the capacitor. The product

$$\tau = RC$$

is called the time constant of the RC circuit. It determines how quickly the capacitor discharges. A larger value of RC results in a slower discharge, while a smaller value leads to a faster discharge.

After one time constant $(t = RC)$,

$$Q = Q_0 e^{-1} \approx 0.368 Q_0,$$

which means that only about **36.8% of the initial charge** remains on the capacitor.

7.6 SECOND ORDER LINEAR DIFFERENTIAL EQUATION

The differential equation of the following form is called second order and first degree linear ordinary differential equation

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = R$$

Or

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = f(x)$$

Where P, Q and R are constants or functions of x.

Here, the differentiation can be written as

$$D^2y + PDy + Qy = R \Rightarrow (D^2 + PD + Q)y = R$$

Where $D = \frac{d}{dx}$ and $D^2 = \frac{d^2}{dx^2}$

The solution of this equation is written as

$$Complete\ solution = Complemenatary\ function + Particular\ intergral$$

$$\Rightarrow y = C.F. + P.I.$$

Method of Finding the Complementary Function (C.F)

To find the C.F. the right-hand side of the given equation is replaced by zero i.e. 0 i.e. $D^2 + PD + Q = 0$

Let mbe the roots of the equation i.e D=m

Case I: If we obtain one two or more roots which are real and different. Say roots are m_1, m_2, m_3 ……. etc then C.F. is

$$y = C_1 e^{m_1 x} + C_2 e^{m_2 x} + C_2 e^{m_2 x} + \ \ldots\ldots$$

Case II: If we obtain one two or more roots which are real and equal. Say roots are $m_1 = m_2 = m_3$ …….= m (Say) etc then C.F. is

$$y = (C_1 + C_2 x + C_3 x^2 + C_4 x^3 + \ \ldots\ldots\ldots)e^{mx}$$

Case III: If the roots are imaginary i.e. the roots are $\alpha \pm i\beta$ then the solution will be

$$y = e^{\alpha x}(C_1 e^{i\beta x} + C_2 e^{-i\beta x})$$

Or

$$y = e^{\alpha x}(A\cos\beta x + B \sin\beta x)$$

Example 9: Solve $\frac{d^2y}{dx^2} - 8\frac{dy}{dx} + 15\, y = 0$

Solution:

The given equation can be written as $(D^2 - 8D + 15)y = 0$

Here auxiliary equation is

$$D^2 - 8D + 15 = 0$$
$$\Rightarrow D^2 - 3D - 5D + 15 = 0$$
$$\Rightarrow (D - 3)(D - \ 5) = 0$$

And, D=-3, 5

Hence the required solution is $y = C_1 e^{3x} + C_2 e^{5x}$

Example 10: Solve $4\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + y = 0$

Solution:

The auxiliary equation is

$$4D^2 - 4D + 1 = 0$$
$$\Rightarrow 4D^2 - 2D - 2D + 1 = 0$$
$$\Rightarrow (2D - 1)(2D - 1) = 0$$

So, D= $\frac{1}{2}, \frac{1}{2}$

Hence the required solution $y = (C_1 + C_2 x)e^{\frac{x}{2}}$

Example 11: Solve $\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 5y = 0$

Solution:

The auxiliary equation is

$$D^2 + 4D + 5 = 0$$

So,

$$D = \frac{-4 \pm \sqrt{4^2 - 4 \times 5}}{2} = \frac{-4 \pm \sqrt{-4}}{2} = \frac{-4 \pm \sqrt{4(-1)}}{2} = \frac{-4 \pm 2i}{2} - 2 \pm i$$

Hence the required solution $y = e^{-2x}\ (A\cos x + B\sin x)$

Method of Finding the Particular Integral (P.T.)

$$(D^2 + PD + Q)y = R$$
$$\Rightarrow f(D).y = R$$

Where $f(D) = D^2 + PD + Q$

And $P.I. = \frac{R}{f(D)}$

Case I: When $R = e^{ax}$ then

$$\frac{1}{f(D)} e^{ax} = \frac{1}{f(D)} e^{ax}$$

If $f(a) = 0$ then $\frac{1}{f(D)}.e^{ax} = x.\frac{1}{f'(a)}.e^{ax}$. If $f'(a) = 0$ then $\frac{1}{f(D)}.e^{ax} = x^2\frac{1}{f''(a)}.e^{ax}$

Example 12: Solve $\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 5e^{3x}$

Solution:

$$(D^2 + 6D + 9)y = 5e^{3x}$$

Auxiliary equation is

$$D^2 + 6D + 9 = 0$$
$$\Rightarrow D^2 + 3D + 3D + 9 = 0$$
$$\Rightarrow (D + 3)(D + 3) = 0$$

So D= -3, -3

C.F= $(C_1 + C_2x)e^{-3x}$

$$P.I. = \frac{1}{D^2 + 6D + 9} 5e^{3x} = 5.\frac{1}{(3)^2 + 6(3) + 9} e^{3x} = \frac{5e^{3x}}{36}$$

Example 13: Solve $\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 9y = 6e^{3x} + 7e^{-2x} - log2$

Solution:

$$(D^2 - 6D + 9)y = 6e^{3x} + 7e^{-2x} - log2$$

Auxiliary equation is

$$(D^2 - 6D + 9) = 0$$
$$\Rightarrow (D - 3)(D - 3) = 0$$
D= 3, 3

Hence C.F. $(C_1 + C_2x)e^{3x}$

$$P.I. = \frac{1}{(D^2 - 6D + 9)}(6e^{3x} + 7e^{-2x} - log2)$$

$$= \frac{1}{(D^2 - 6D + 9)} 6e^{3x} + \frac{1}{(D^2 - 6D + 9)} 7e^{-2x} - \frac{1}{(D^2 - 6D + 9)} log2$$

$$= x\frac{1}{2D - 6} 6e^{3x} + \frac{1}{4 + 12 + 9} 7e^{-2x} - \log 2 \frac{1}{(D^2 - 6D + 9} e^{0x}$$

$$= x^2\frac{1}{2}.6.e^{3x} + \frac{7}{25}e^{-2x} - \log 2\left(\frac{1}{9}\right)e^{0x}$$

$$= 3x^2e^{3x} + \frac{7}{25}e^{-2x} - \frac{1}{9}\log 2$$

Case-II: When $R = \mathrm{x}^n$ then

$$\frac{1}{f(D)} x^n = [f(D)]^{-1}x^n$$

Expand $[f(D)]^{-1}$ and then operate

Example 14: Solve

$$\frac{d^2y}{dx^2} + a^2y = \frac{a^2R}{p}(l - x)$$

Solution:
The equation can be written as

$$D^2y + a^2y = \frac{a^2R}{p}(l - x)$$

The auxiliary equation is

$$D^2y + a^2y = 0$$
$$\Rightarrow D^2y + a^2y = 0$$
$$\Rightarrow D^2 + a^2 = 0$$

So, $D = \pm ia$

$$C.F. = C_1 \cos ax + C_2 sinax$$
$$P.I. = \frac{1}{D^2 + a^2} \cdot \frac{a^2}{p} R\,(l - x)$$
$$= \frac{1}{a^2\left(1 + \frac{D^2}{a^2}\right)} \cdot \frac{a^2}{p} R\,(l - x)$$
$$= \frac{R}{p}\left[1 + \frac{D^2}{a^2}\right]^{-1}(l - x)$$
$$= \frac{R}{p}\left[1 - \frac{D^2}{a^2}\right](l - x)$$
$$= \frac{R}{p}(l - x) - \frac{R}{p}\frac{D^2}{a^2}(l - x)$$
$$= \frac{R}{p}(l - x)$$

Hence $y = C_1 \cos ax + C_2 sinax + \frac{R}{p}(l - x)$

Example 15: Solve

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

Solution: The auxiliary equation is

$$D^2y + 2Dy + y = 0$$
$$\Rightarrow D^2 + 2D + 1 = 0$$
$$\Rightarrow (D + 1)(D + 1) = 0$$

So D= -1, -1

$$C.F = (C_1 + C_2x)e^{-x}$$
$$P.I. = \frac{1}{(D + 1)^2}x$$
$$= (1 + D)^{-2}x$$
$$= (1 - 2D + 3D^2 - \ \ldots\ldots)x$$
$$= x - 2$$

Hence, $y = (C_1 + C_2x)e^{-x} + (x - 2)$

Case-III: When $R = sinax$ and $R = cosax$ then

$$\frac{1}{f(D^2)} sin\, ax = \frac{1}{f(-a^2)} sinax$$

And

$$\frac{1}{f(D^2)} cos\, ax = \frac{1}{f(-a^2)} cos\, ax$$

If $f(-a^2) = 0$ then

$$\frac{1}{f(D^2)} sin\ ax = x.\frac{1}{f'(-a^2)}.sin\ ax$$

Example 16: Solve $(D^2 + 4)y = \sin 3x$

Solution:

Auxiliary equation $D^2 + 4 = 0$

So $D = \pm 2i$

$$C.F. = Acos2x + Bsin2x$$

$$P.I. = \frac{1}{D^2 + 4}.\sin 3x$$

$$\Rightarrow \frac{sin3x}{(-3^2) + 4} = \frac{1}{13}\sin 3x$$

Complete solution $y = Acos2x + Bsin2x + \frac{1}{13}\sin 3x$

Example 17: Solve

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

Solution:

The auxiliary equation is $D^2 + D + 1 = 0$

$$D = \frac{-1 \pm \sqrt{-3}}{2}$$

$$C.F = e^{-\frac{x}{2}}[Acos\frac{\sqrt{3}}{2}x + Bsin\ \frac{\sqrt{3}}{2}x]$$

$$P.I. = \frac{1}{D^2 + D + 1}cos2x$$

$$= \frac{1}{(-2^2) + D + 1}cos2x$$

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

$$\therefore P.I = \frac{2\sin 2x - 3\cos 2x}{13}$$

$$= \frac{D + 3}{D^2 - 9}.\cos 2x = \frac{D + 3}{(-2^2) - 9}cos2x = -\frac{1}{13}(D + 3)cos2x = -\frac{1}{13}(-2sin2x + 3cos2x)$$

Case-IV: When R= e^{ax}.Φ(x) then

$$\frac{1}{f(D)}e^{ax}.\phi(x) = e^{ax}.\frac{1}{f\ (D + a)}\phi(x)$$

Example 18: Solve $(D^2 - 4D + 4)y = x^3e^{2x}$

Solution:

Auxiliary equation:

$$D^2 - 4D + 4 = 0$$

$$\Rightarrow (D - 2)(D - 2) = 0$$

$$D = 2,2$$

$$C.F. = (C_1 + C_2x)e^{2x}$$

$$P.I. = \frac{1}{D^2 - 4D + 4}x^3.e^{2x}$$

$$= e^{2x}\frac{1}{(D+2)^2-4(D+2)+4}x^3$$
$$= e^{2x}\frac{1}{D^2}x^3$$
$$= e^{2x}\frac{1}{D}\left(\frac{x^4}{4}\right)$$
$$= \frac{e^{2x}x^5}{20}$$
$$y = (C_1 + C_2x)e^{2x} + \frac{e^{2x}x^5}{20}$$

Example 19: Solve $(D^2 - 5D + 6)y = e^x cos2x$

Solution:

Auxiliary equation

$$D^2 - 5D + 6 = 0$$
$$\Rightarrow (D-2)(D-3) = 0$$
$$D = 2, 3$$

Therefore, CF is

$$y = C_1e^{2x} + C_2e^{3x}$$
$$P.I. = \frac{1}{D^2-5D+6}e^x cos2x$$
$$= e^x\frac{1}{(D+1)^2-5(D+1)+6}cos2x$$
$$= e^x\frac{1}{-2^2-3D+2}cos2x$$
$$= -e^x\frac{1}{3D+2}cos2x$$
$$= -e^x\frac{3D-2}{3D+2}cos2x$$

$$= -e^x\frac{3D-2}{9D^2-4}cos2x$$
$$= -e^x\frac{3D-2}{9(-4)-4}cos2x$$
$$= \frac{e^x}{40}(3D-2)cos2x$$
$$= \frac{e^x}{40}(-6sin2x - 2cos2x)$$
$$= -\frac{e^x}{20}(3sin2x + cos2x)$$
$$y = C_1e^{2x} + C_2e^{3x} - \frac{e^x}{20}(3sin2x + cos2x)$$

Application of Second-Order Differential Equations

The second-order differential equations discussed in this section are used to describe many physical systems involving vibrations and oscillations. One of the most common examples is the damped harmonic oscillator. In practical situations, oscillating bodies experience friction or air resistance, which gradually reduces the amplitude of oscillation. Such motion is called damped motion.

The equation describing a damped harmonic oscillator is

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \omega_0^2 x = 0,$$

where x is the displacement of the particle from its equilibrium position, γ is the damping constant, and ω_0 is the natural angular frequency of the oscillator.

The nature of the motion depends on the value of the damping constant.

- If $\gamma < \omega_0$, the system continues to oscillate, but the amplitude gradually decreases with time. This is known as **underdamped motion**.
- If $\gamma = \omega_0$, the system returns to its equilibrium position as quickly as possible without oscillating. This condition is called **critical damping**.
- If $\gamma > \omega_0$, the system returns to equilibrium without oscillating, but more slowly than in the critically damped case. This is known as **overdamped motion**.

Thus, the solution of the second-order differential equation not only gives the mathematical expression for the displacement but also helps us understand the physical behaviour of the system. Damped harmonic motion is encountered in many practical situations, such as the suspension system of vehicles, door closers, measuring instruments, and mechanical vibration control. This example illustrates how second-order differential equations are used to model real physical systems.

CHECK YOUR PROGRESS

Q1: Check whether the following differential equations are homogeneous or non-homogeneous.

(i) $\frac{d^2y}{dx^2} + e^x = cosy$

(ii) $\frac{d^2y}{dx^2} + e^x = siny$

(iii) $\frac{d^2y}{dx^2} + e^x = cosy$

(iv) $\frac{d^2y}{dx^2} + y^2 = 0$

Q2: $(2y - 3x)dx + xdy = 0$

Q3: $\frac{dy}{dx} + 2xy = 2e^{-x^2}$

Q4: $\frac{dy}{dx} + 3y = e^{2x}$

Q5: $\frac{d^3y}{dx^3} - \frac{2d^2y}{dx^2} + 4\frac{dy}{dx} - 8y = 0$

Q6: $(D^2 + 2D + 2)y = 0$

Q7: $(D - 1)^2 y = 16e^{3x}$

Q8: $(D^2 - 4D + 3)y = x^3$

7.7 LET US SUM UP

- A differential equation relates a function with its derivatives and describes rate of change.
- Differential equations are classified as ordinary, linear/nonlinear, homogeneous/non-homogeneous.
- The order is the highest derivative present; the degree is the power of the highest derivative (if polynomial form).
- First order and first-degree equations can be solved by variable separation and integrating factor methods.
- Second order linear differential equations are solved using auxiliary equations.
- The complete solution consists of Complementary Function (C.F.) and Particular Function (P.F.).
- Differential equations are widely used to model physical problems such as oscillatory motion.

7.8 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q 1:

(i) Non-homogeneous, because the term e^x is a function of x only and does not involve y or its derivatives.

(ii) Non-homogeneous, because the term e^x appears independently and is not related to y or its derivatives.

(iii) Non-homogeneous, because it contains the independent term e^x, which depends only on x.

(iv) Homogeneous, because every term involves y or its derivatives and there is no independent function of x.

Answer to Q 2:

$$(2y - 3x) + x\frac{dy}{dx} = 0$$

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

$$\Rightarrow \frac{dy}{dx} + \frac{2}{x}y = 3$$

$$I.F. = e^{\int \frac{2}{x}dx} = e^{2lnx} = x^2$$

Using the formula

$$yx^2 = \int 3x^2 dx + C = x^3 + C$$

$$\Rightarrow y = x + \frac{C}{x^2}$$

Answer to Q3:

$$I.F. = e^{\int 2xdx} = e^{x^2}$$

Using the formula

$$ye^{x^2} = \int 2e^{-x^2}e^{x^2}dx + C = \int 2dx + C = 2x + C$$

$$y = (2x + C)e^{-x^2}$$

Answer to Q4:

$$I.F. = e^{\int 3dx} = e^{3x}$$

Using the formula

$$ye^{3x} = \int e^{2x}e^{3x}dx + C$$
$$= \int e^{5x}dx + C$$
$$= \frac{e^{5x}}{5} + C$$
$$y = \frac{e^{2x}}{5} + Ce^{-3x}$$

Answer to Q5:

The A.E

$$D^3 - 2D^2 + 4D - 8 = 0$$
$$\Rightarrow (D-2)(D^2+4) = 0$$
$$D - 2 = 0 \Rightarrow D = 2$$
$$D^2 + 4 = 0 \Rightarrow D = \pm 2i$$
$$C.F. = C_1e^{2x} + (C_2\cos 2x + C_3\sin 2x)$$

Answer to Q6:

The A.E

$$D^2 + 2D + 2 = 0$$
$$D = \frac{-2 \pm \sqrt{4-8}}{2} = -1 \pm i$$
$$C.F. = e^{-x}(C_1\cos x + C_2\sin x)$$

Answer to Q7:

The A.E

$$(D-1)^2 = 0$$
$$D = 1, 1$$
$$C.F. = (C_1 + C_2x)e^x$$
$$P.I. = \frac{1}{(D-1)^2}16e^{3x}$$
$$= \frac{1}{(3-1)^2}16e^{3x}$$

The complete solution is $y = (C_1 + C_2x)e^x + 4e^{3x}$

Answer to Q8:

The AE

$$D^2 - 4D + 3 = 0$$
$$\Rightarrow (D-1)(D-3) = 0$$
$$D = 1, 3$$
$$C.F. = C_1e^x + C_2e^{3x}$$
$$P.I. = \frac{1}{D^2 - 4D + 3}x^3$$
$$= \frac{1}{3}.\frac{1}{1 - \frac{4}{3}D + \frac{1}{3}D^2}x^3$$

Using the series

$$\frac{1}{1-A} = 1 + A + A^2 + A^3 + \quad \ldots\ldots$$

Consider

$$A = \left(\frac{4}{3}D - \frac{1}{3}D^2\right)$$

Since the derivatives of x^3 vanishes after D^4, expand to D^3
Hence,

$$P.I. = \frac{1}{3}.\frac{1}{1-\left(\frac{4}{3}D-\frac{1}{3}D^2\right)}x^3$$

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

Therefore, we will get

$$P.I. = \frac{x^3}{3}+\frac{4x^2}{3}+\frac{26x}{9}+\frac{80}{27}$$

Complete solution will be

$$y = C_1e^x + C_2e^{3x} + \frac{1}{3}x^3 + \frac{4}{3}x^2 + \frac{26}{9}x + \frac{80}{27}$$

7.9 FURTHER READING

- Dass, H. K., and Gupta, R. (2014), *Differential Equations and Their Applications*, 2nd Edition, Pearson India.
- Kreyszig, E. (2011), *Advanced Engineering Mathematics*, 10th Edition, Wiley India.
- Boyce, W. E., and DiPrima, R. C. (2017), *Elementary Differential Equations and Boundary Value Problems*, 11th Edition, Wiley.
- Rainville, E. D. (1964), *Elementary Differential Equations*, 2nd Edition, Macmillan.
- Simmons, G. F. (1991), *Differential Equations with Applications and Historical Notes*, 2nd Edition, McGraw-Hill.

7.10 MODEL QUESTIONS

1. Define a differential equation. Classify differential equations into linear/nonlinear and homogeneous/non-homogeneous with one example each.
2. Explain the order and degree of a differential equation. Find the order and degree of:
   i) $\frac{d^2y}{dx^2}+3\frac{dy}{dx}+y=0$ ii) $\frac{dy}{dx}=xy$
3. Solve the initial value problem $\frac{dy}{dx}+y\tan x=\sec x,\ y(0)=0.$
4. Solve the linear differential equation using the integrating factor method:
$$\frac{dy}{dx}+y=e^x$$
5. Find the complete solution of $(D^2-4D+4)y=e^{2x}$.
6. Using the Wronskian, verify that. $y_1=e^{2x}$ and. $y_2=e^{-3x}$ are linearly independent solutions. Hence, write their general linear combination.
7. Find the complete solution of the differential equation:
$$\frac{d^2x}{dt^2}+2\frac{dx}{dt}+3x=\sin t$$

*** *** ***

UNIT 8: SPECIAL FUNCTIONS

UNIT STRUCTURE

8.1 Learning Objectives
8.2 Introduction
8.3 Types of Polynomials
    8.3.1 Legendre Polynomials
    8.3.2 Bessel Functions
    8.3.3 Hermite Polynomials
    8.3.4 Laguerre Polynomials
8.4 Rodrigues Formula
8.5 Generating Functions of Differential Polynomials
8.6 Orthogonality of Polynomials
8.7 Applications
8.8 Let Us Sum Up
8.9 Further Reading
8.10 Answers to Check Your Progress
8.11 Model Questions

8.1 LEARNING OBJECTIVES

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

- Understand the role and significance of special functions in mathematical physics
- analyse the properties of Legendre, Bessel, Hermite, and Laguerre polynomials
- apply Rodrigues' formula and generating functions to construct special functions
- examine the orthogonality properties of special functions and their mathematical implications
- evaluate Bessel functions and their behaviour in physical problems
- apply and analyse special functions in solving boundary value problems in physics.

8.2 INTRODUCTION

In the previous unit on Differential Equations, we studied methods for solving equations that describe various physical phenomena. In many cases, especially under specific boundary conditions, the solutions of these equations are not expressible in terms of elementary functions but lead to special functions. Thus, the study of differential equations naturally leads to the concept of special functions, which arise as their solutions and provide physically meaningful results in many applications. In this unit, we introduce important special functions such as Legendre, Bessel, Hermite, and Laguerre polynomials, along with their key properties, generating functions, and orthogonality relations. These functions play a crucial role in solving a wide range of problems in mathematical physics.

8.3 TYPES OF POLYNOMIALS

In mathematical physics, many differential equations arising in physical problems have solutions in the form of special polynomials. These polynomials possess important mathematical properties such as orthogonality, recurrence relations, and generating functions, which make them useful in solving boundary value problems. Depending on the form of the differential equation and the physical situation involved, different types of polynomials arise. Some of the most important among them are Legendre polynomials**,** Hermite polynomials, and Laguerre polynomials, each associated with specific physical applications such as potential theory, quantum mechanics, and wave equations. This section introduces these important classes of polynomials and their basic properties.

8.3.1 Legendre Polynomial

The Legendre equation is given by

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0 \qquad (8.1)$$

This equation can be solved by assuming a series solution in powers of x, either ascending or descending. Since the equation admits a solution in descending powers of x, we assume a solution of the form

$y = x^m(a_0 + a_1x^{-1} + a_2x^{-2} + a_3x^{-3} + \cdots)$(8.2) where $a_0 \neq 0$

The above equation can be written as

$$y = \sum_{r=0}^{\infty} a_r\, x^{m-r}. \qquad (8.3)$$

Differentiating term by term, we have

$$\frac{dy}{dx} = \sum_{r=0}^{\infty} a_r\,(m-r)x^{m-r-1}. \qquad (8.4)$$

Differentiating once again,

$$\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} a_r\,(m-r)(m-r-1)x^{m-r-2}. \qquad (8.5)$$

Substituting (8.3), (8.4) and (8.5) into (8.1), we have

$$(1-x^2)\sum_{r=0}^{\infty} a_r\,(m-r)(m-r-1)x^{m-r-2} - 2x\sum_{r=0}^{\infty} a_r\,(m-r)x^{m-r-1} + n(n+1)\sum_{r=0}^{\infty} a_r\,x^{m-r} = 0. \qquad (8.6)$$

Expanding the first term of (8.6), we have

$$\sum_{r=0}^{\infty} a_r\,(m-r)(m-r-1)x^{m-r-2} - \sum_{r=0}^{\infty} a_r\,(m-r)(m-r-1)x^{m-r} - 2\sum_{r=0}^{\infty} a_r\,(m-r)x^{m-r} + n(n+1)\sum_{r=0}^{\infty} a_r\,x^{m-r} = 0.$$

Combining the last three summations of above equation, we have

$$\sum_{r=0}^{\infty}(m-r)(m-r-1)a_r x^{m-r-2} + \sum_{r=0}^{\infty}[n(n+1) - 2(m-r) - (m-r)(m-r-1)]\,a_r x^{m-r} = 0 \quad (8.7)$$

Since, $2(m-r) + (m-r)(m-r-1) = (m-r)(m-r+1)$,

Therefore, equation (8.7) becomes

$$\sum_{r=0}^{\infty}(m-r)(m-r-1)a_r x^{m-r-2} + \sum_{r=0}^{\infty}[n(n+1) - (m-r)(m-r+1)]\,a_r x^{m-r} = 0 \quad (8.8)$$

Since, equation (8.8) is an identity, the coefficient of each power of x must vanish.

In equation (8.8), the highest power is x^m, therefore we have,

$$a_0[n(n+1) - m(m+1)] = 0$$

Since, $a_0 \neq 0$, we obtain,

$$[n(n+1) - m(m+1)] = 0$$

$$\Rightarrow\ n^2 + n - m^2 - m = 0$$
$$\Rightarrow\ (n-m)(n+m+1) = 0.$$
$$\Rightarrow\ m = n \text{ or } m = -(n+1) \quad (8.9)$$

This is known as the indicial equation. It yields two distinct values of the exponent m, indicating that Legendre's differential equation possesses two linearly independent series solutions corresponding to these values.

Now, equating the coefficient of $x^{m-1}(second\ summation\ with\ r=1)$, we have

$$a_1[n(n+1) - (m-1)m] = 0$$

Using, $n(n+1)-(m-1)m=(m+n)(n-m+1)$

we get $a_1(m+n)(n-m+1)=0$

For both values of m, $(m+n)(n-m+1)\neq 0$

Therefore, $a_1=0$

Thus, all odd coefficients will vanish,

$$a_1=a_3=a_5=\cdots=0$$

Now, equating the coefficient of x^{m-r-2}, we have from first summation,

$$(m-r)(m-r-1)a_r$$

Now, to obtain the same power for the second summation, we replace $r\ by\ r+2$,

Hence,

$$[n(n+1)-(m-r-2)(m-r-1)]a_{r+2}.$$

Therefore,

$$(m-r)(m-r-1)a_r+[n(n+1)-(m-r-2)(m-r-1)]a_{r+2}=0. \qquad (8.10)$$

Now simplifying, we have,

$$n(n+1)-(m-r-2)(m-r-1).$$

It factors as

$$-(m-r+n-1)(m-r-n-2).$$

Substituting the above equation in (8.10), we have

$$(m-r)(m-r-1)a_r-(m-r+n-1)(m-r-n-2)a_{r+2}=0.$$

Hence the recurrence relation is

$$a_{r+2}=\frac{(m-r)(m-r-1)}{(m-r+n-1)(m-r-n-2)}a_r$$

Considering two cases we have,

Case 1 : Substituting (m=n) into (8.10), we have

$$a_{r+2}=-\frac{(n-r)(n-r-1)}{(2n-r-1)(r+2)}a_r$$

Successive coefficients are

a. For $r=0$,

$$a_2 = -\frac{n(n-1)}{2(2n-1)}a_0.$$

b. For $r = 2$,

$$a_4 = \frac{n(n-1)(n-2)(n-3)}{2 \cdot 4(2n-1)(2n-3)}a_0.$$

Similarly,

$$a_6 = -\frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{2 \cdot 4 \cdot 6(2n-1)(2n-3)(2n-5)}a_0,$$

and so on.

Therefore,

$$y = a_0[x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2\cdot4(2n-1)(2n-3)}x^{n-4} - \cdots\cdots] \quad (8.11)$$

When n is a positive integer, the series terminates after a finite number of terms, giving the Legendre polynomial $P_n(x)$.

Case 2: Substituting $m = -(n+1)$ into (8.10), we have

$$a_{r+2} = \frac{(n+r+1)(n+r+2)}{(r+2)(2n+r+3)}a_r$$

Thus, $$a_2 = \frac{(n+1)(n+2)}{2(2n+3)}a_0$$

$a_4 = \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot4(2n+3)(2n+5)}a_0$ and so on.

Hence,

$$y = a_0[x^{-(n+1)} + \frac{(n+1)(n+2)}{2(2n+3)}x^{-(n+3)} + \frac{(n+1)(n+2)(n+3)(n+4)}{2 \cdot 4(2n+3)(2n+5)}x^{-(n+5)} + \cdots]. \quad (8.12)$$

This is the second independent solution of Legendre's differential equation.

Hence, the Legendre differential equation possesses two linearly independent series solutions:

1. For$m = n$**:** the recurrence relation generates a series that terminates when n is a non-negative integer. This finite series represents the Legendre polynomial (Legendre function of the first kind), denoted by $P_n(x)$.

2. For $m = -(n+1)$**:** the recurrence relation generates a non-terminating series. This series represents the second independent solution of Legendre's differential equation. By applying a suitable normalization, this solution leads to the Legendre function of the second kind**,** denoted by $Q_n(x)$.

Thus, the complete solution of Legendre's differential equation consists of two linearly independent functions, $P_n(x)$ and $Q_n(x)$, whose linear combination gives the general solution.

A) Legendre's Polynomial of First Kind $P_n(x)$

Legendre polynomials are defined as solutions of Legendre's differential equation:

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

A solution obtained in Case 1 (8.11) can be obtained in the form of a series in descending powers of x:

$$y = a_0\left[x^n - \frac{n(n-1)}{(2n-1)2}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{(2n-1)(2n-3)\,2\cdot 4}x^{n-4} + \cdots\right]$$

where a_0 is an arbitrary constant.

Formation of $P_n(x)$: If n is a positive integer and we choose

$$a_0 = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{n!},$$

then the above series represents the Legendre polynomial $P_n(x)$. Hence,

$$P_n(x) = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{n!}\left[x^n - \frac{n(n-1)}{(2n-1)2}x^{n-2} + \cdots\right] \qquad (8.13)$$

Note 1: This series terminates after a finite number of terms.

- When n is even, the number of terms is $\frac{n}{2}+1$, and the last term is

$$(-1)^{\frac{n}{2}}\frac{n(n-1)(n-2)\cdots 1}{(2n-1)(2n-3)\cdots(n+1)\,2\cdot 4\cdot 6\cdots(n-1)}.$$

- When n is odd, the series contains $\frac{n+1}{2}$ terms, and the final term is

$$(-1)^{\frac{n-1}{2}}\frac{n(n-1)(n-2)\cdots 3\cdot 2}{(2n-1)(2n-3)\cdots(n+2)\,2\cdot 4\cdot 6\cdots(n-1)}\,x.$$

The function $P_n(x)$ is called the Legendre function of the first kind. $P_n(x)$ is that solution of Legendre's equation (8.1) which satisfies $P_n(1) = 1$.

B) Legendre's Function of the Second Kind $Q_n(x)$

The series solution obtained in Case II [equation (8.12)] represents the second linearly independent solution of Legendre's differential equation. Since Legendre's differential

equation is a second-order linear differential equation, its complete solution consists of two linearly independent solutions. The first solution is the Legendre function of the first kind, $P_n(x)$, while the second independent solution is called the Legendre function of the second kind, denoted by $Q_n(x)$. For a positive integer n, the second series solution obtained in Case II is

$$y = a_0\left[x^{-(n+1)} + \frac{(n+1)(n+2)}{2(2n+3)}x^{-(n+3)} + \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot 4(2n+3)(2n+5)}x^{-(n+5)} + \cdots\right].$$

If we choose

$$a_0 = \frac{n!}{1\cdot 3\cdot 5\cdots(2n+1)}$$

the above solution is denoted by $Q_n(x)$. Thus,

$$Q_n(x) = \frac{n!}{1\cdot 3\cdot 5\cdots(2n+1)}\left[x^{-(n+1)} + \frac{(n+1)(n+2)}{2(2n+3)}x^{-(n+3)} + \frac{(n+1)(n+2)(n+3)(n+4)}{2\cdot 4(2n+3)(2n+5)}x^{-(n+5)} + \cdots\right]$$

The series for $Q_n(x)$ is a non-terminating series. Unlike the Legendre function of the first kind, $P_n(x)$, which reduces to a polynomial for positive integral values of n, the function $Q_n(x)$ does not terminate after a finite number of terms. Consequently, $Q_n(x)$ generally becomes singular at $x = \pm 1$. Therefore, in many physical boundary value problems defined over the interval [-1,1], the bounded solution $P_n(x)$ is usually selected, whereas $Q_n(x)$ is used only when the boundary conditions permit its singular behaviour.

Hence, $P_n(x)$ and $Q_n(x)$ are two linearly independent solutions of Legendre's differential equation. Therefore, the most general solution of Legendre's differential equation is

$$\boxed{y = A\,P_n(x) + B\,Q_n(x),}$$

where A and B are arbitrary constants determined by the boundary conditions.

Having obtained the complete solution of Legendre's differential equation, it is useful to understand the importance of Legendre polynomials in solving physical problems.

Physical Significance of Legendre Polynomials

Legendre polynomials arise naturally in many branches of mathematical physics. They occur whenever Laplace's equation or Poisson's equation is solved in spherical coordinates for problems possessing azimuthal symmetry. Their important applications include:

- **Electrostatics:** Determination of the electric potential due to spherically symmetric charge distributions.
- **Gravitational Potential:** Evaluation of the gravitational potential of planets, stars and other celestial bodies.

- **Multipole Expansion:** Representation of the potential due to localized charge or mass distributions.
- **Quantum Mechanics:** Solution of the angular part of the Schrödinger equation for central-force problems.
- **Orbital Angular Momentum:** Description of eigenfunctions associated with orbital angular momentum.
- **Spherical Harmonics:** Construction of spherical harmonics used in quantum mechanics, electromagnetism and geophysics.

Other important properties of Legendre polynomials, including Rodrigues' formula, generating function, orthogonality and recurrence relations, are discussed collectively in Sections 8.4–8.6.

8.3.2 Bessel Functions

a. Bessel Differential Equation

The **Bessel differential equation** of order n is

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0 \qquad (8.14)$$

is known as Bessel's differential equation, and its solutions are called Bessel functions of order$n$.This equation naturally appears while solving Laplace's equation in polar coordinates using the method of separation of variables. It has wide applications in engineering and physics. They are also referred to as cylindrical and spherical functions

Bessel functions arise in problems such as:

- Oscillations of a hanging chain
- Vibrations of a circular membrane
- Planetary motion
- Wave propagation
- Elasticity problems
- Fluid dynamics
- Potential theory
- Cylindrical and spherical wave motion
- Plane wave theory

Since $x = 0$ is a regular singular point, the equation is conveniently solved by the Frobenius (power-series) method.

b. Solution by the Frobenius Method

Assuming the solution in the form

$$y = \sum_{r=0}^{\infty} a_r\, x^{m+r},\ a_0 \neq 0, \qquad (8.15)$$

Where m is to be determined.

Differentiating equation (8.15) term by term, we have

$$\frac{dy}{dx} = \sum_{r=0}^{\infty} a_r\,(m+r)x^{m+r-1} \qquad (8.16)$$

and $\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} a_r\,(m+r)(m+r-1)x^{m+r-2}$ (8.17)

Substituting these expressions into Bessel's differential equation given by (8.14), we have

$$x^2\sum_{r=0}^{\infty} a_r\,(m+r)(m+r-1)x^{m+r-2} + x\sum_{r=0}^{\infty} a_r\,(m+r)x^{m+r-1} + (x^2-n^2)\sum_{r=0}^{\infty} a_r\,x^{m+r} = 0.$$

On simplifying

$$\sum_{r=0}^{\infty} a_r\,[(m+r)^2-n^2]x^{m+r} + \sum_{r=0}^{\infty} a_r\,x^{m+r+2} = 0. \qquad (8.18)$$

Let us, combine the like powers of x, thus replacing r by r-2 in the second summation, we have

$a_r[(m+r)^2-n^2] + a_{r-2} = 0$ (8.19)

Therefore, the recurrence relation is

$$a_r = -\frac{a_{r-2}}{(m+r)^2-n^2},$$

or equivalently,

$$a_{r+2} = -\frac{a_r}{(m+r+2)^2-n^2} \qquad (8.20)$$

Indicial Equation:

From equation (8.19), equating the lowest power of x, we have

$$a_0(m^2-n^2) = 0$$

Since, $a_0 \neq 0$, we have

$(m^2-n^2) = 0$ (8.21)

Hence, $m = \pm n$ (8.22)

Equation (8.22) is known as the indicial equation.
Also,

$$a_1[(m+1)^2-n^2] = 0.$$

Since $(m+1)^2 \neq n^2$, $a_1 = 0$

Thus,

$$a_1 = a_3 = a_5 = \cdots = 0$$

Let us consider two cases:

Case 1: $m = n$

Substituting $m = n$ into the recurrence relation, we have

$$a_{r+2} = -\frac{a_r}{(r+2)(2n+r+2)}$$

Successive coefficients are

$$a_2 = -\frac{a_0}{4(n+1)}$$

$a_4 = \frac{a_0}{4^2\, 2!(n+1)(n+2)}$ and so on.

Therefore,

$$y = a_0 x^n \left[1 - \frac{x^2}{4(n+1)} + \frac{x^4}{4^2 2!\,(n+1)(n+2)} - \cdots\right]. \qquad (8.23)$$

Case II: $m = -n$
For $m = -n$,

$$y = a_0 x^{-n} \left[1 - \frac{x^2}{4(1-n)} + \frac{x^4}{4^2 2!\,(1-n)(2-n)} - \cdots\right]. \qquad (8.24)$$

Thus, Bessel's differential equation possesses two linearly independent series solutions (8.23) and (8.24) corresponding to $m = n$ and $m = -n$ respectively.The solution corresponding to $m = n$ remains finite at the origin for non-negative integral values of n and is therefore of greater physical significance. By choosing a suitable normalization constant, this solution assumes the standard form known as the *Bessel function of the first kind*, denoted by $J_n(x)$.

Gamma Function
Before defining the Bessel function of the first kind, it is useful to introduce the **Gamma function**, which frequently appears in the study of special functions. The Gamma function is defined by

$$\Gamma(z) = \int_0^\infty t^{\,z-1}\, e^{-t}\, dt, \text{ for } \mathrm{Re}\,(z) > 0.$$

The Gamma function extends the concept of the factorial to real and complex numbers (except negative integers and zero). For every positive integer n,

$$\boxed{\Gamma(n+1) = n!}$$

Thus, $\Gamma(1) = 1$, $\Gamma(2) = 1!$, $\Gamma(3) = 2!$, $\Gamma(4) = 3!$, and so on.
The Gamma function is widely used in the theory of special functions because it provides a compact representation of coefficients appearing in infinite series. It is used to express the normalization constant in the standard series representation of the Bessel function.

Bessel Function of the First Kind $J_n(x)$
The Frobenius method yields two linearly independent series solutions corresponding to $m = n$ and $m = -n$. Among these, the solution corresponding to $m = n$remains finite at the origin for non-negative integral values of n and is therefore of greater physical

significance. By choosing a suitable normalization constant, this solution assumes the standard form known as the **Bessel function of the first kind**, denoted by $J_n(x)$. Choosing

$$a_0 = \frac{1}{2^n \Gamma(n+1)}, \ [\Gamma(n+1) = n!] \qquad (8.25)$$

the series solution obtained in equation (8.23) becomes

$$J_n(x) = \frac{1}{2^n \Gamma(n+1)} \sum_{r=0}^{\infty} \frac{(-1)^r x^{n+2r}}{2^{2r} r!\,(n+1)(n+2)\cdots(n+r)} \qquad (8.26)$$

Using the identity

$$(n+r)! = n!\,(n+1)(n+2)\cdots(n+r)$$

equation (8.26) may be written in the more compact form

$$J_n(x) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\,(n+r)!} \left(\frac{x}{2}\right)^{n+2r} \qquad (8.27)$$

Equation (8.27) is called the standard series representation of the Bessel function of the first kind of order n.

Special Cases

(i) For$n = 0$: Substituting $n = 0$ in equation (8.27**)**, we obtain

$$J_0(x) = \sum_{r=0}^{\infty} \frac{(-1)^r}{(r!)^2} \left(\frac{x}{2}\right)^{2r} = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \frac{x^6}{2^2 \cdot 4^2 \cdot 6^2} + \cdots .$$

(ii) For $\mathrm{n} = 1$: Similarly, substituting $\mathrm{n} = 1$ in equation (8.27), we obtain

$$J_1(x) = \frac{x}{2} - \frac{x^3}{2^2 \cdot 4} + \frac{x^5}{2^2 \cdot 4^2 \cdot 6} - \cdots .$$

The functions $J_0(x)$ and $J_1(x)$ are the lowest-order Bessel functions and are frequently encountered in practical problems involving cylindrical symmetry.

Bessel Function of Negative Order

Replacing n by $-n$ in equation (8.11), we obtain

$$J_{-n}(x) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\,(-n+r)!} \left(\frac{x}{2}\right)^{-n+2r} \qquad (8.28)$$

This series represents the Bessel function of the first kind of negative order.

General Solution of Bessel's Differential Equation

The general solution of Bessel's differential equation depends upon the nature of the order n.

Case I: If n is neither an integer nor zero,

$$y = A\,J_n(x) + B\,J_{-n}(x) \qquad (8.29)$$

Case II: If $n = 0$,

$$y = A\,J_0(x) + B\,Y_0(x). \qquad (8.30)$$

where $Y_0(x)$ is the Bessel function of the second kind of order zero.

Case III: If $n$is a positive integer, the functions $J_n(x)$ and $J_{-n}(x)$ are not linearly independent since

$$J_{-n}(x) = (-1)^n J_n(x). \qquad (8.31)$$

Therefore, a second linearly independent solution is required. This leads to the introduction of the **Bessel function of the second kind**, which will be discussed in the next section.

Thus, Bessel functions provide the complete set of solutions of Bessel's differential equation and play a fundamental role in describing physical systems possessing cylindrical symmetry, particularly in problems related to wave propagation, vibrations, heat conduction, acoustics, and electromagnetic theory.

Bessel Function of the Second Kind

For non-integral values of n, the functions $J_n(x)$ and $J_{-n}(x)$ are linearly independent and together form the general solution of Bessel's differential equation. However, when n is a positive integer,

$$J_{-n}(x) = (-1)^n J_n(x),$$

so that $J_n(x)$ and $J_{-n}(x)$ are no longer linearly independent. Hence, a second independent solution is required to obtain the complete solution of Bessel's differential equation. This solution is called the Bessel function of the second kind, denoted by $Y_n(x)$, and is also known as the Neumann function.

The Bessel function of the second kind is defined as

$$Y_n(x) = \frac{1}{\sin n\pi}[J_n(x)\cos n\pi - J_{-n}(x)],\ n \notin \mathbb{Z}. \qquad (8.32)$$

When $n$is an integer, the above expression becomes indeterminate. In this case, $Y_n(x)$ is defined by the limiting form

$$Y_n(x) = \lim_{m\to n} \frac{J_m(x)\cos m\pi - J_{-m}(x)}{\sin m\pi}. \qquad (8.33)$$

Unlike the Bessel function of the first kind, $J_n(x)$, the function $Y_n(x)$ is singular at the origin $(x = 0)$. Nevertheless, together they form the complete set of linearly independent solutions of Bessel's differential equation.

Hence, for integer values of n, the general solution of Bessel's differential equation is

$$\boxed{y = A\,J_n(x) + B\,Y_n(x),} \qquad (8.34)$$

where $A$and B are arbitrary constants determined by the boundary conditions.

The Bessel functions $J_n(x)$ and $Y_n(x)$ are widely used in solving boundary value problems involving cylindrical symmetry, such as heat conduction in cylindrical bodies, vibration of circular membranes, electromagnetic wave propagation in cylindrical waveguides, and acoustic wave propagation.

Other important properties of Bessel functions, including generating functions, orthogonality and recurrence relations, are discussed collectively in Sections 8.5 and 8.6.

8.3.3 Hermite Polynomials

The Hermite equation is given by

$$\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2ny = 0 \qquad (8.35)$$

The solution of this equation is called the Hermite polynomial.

Let us assume a series solution of the form

$$y = a_0x^m + a_1x^{m+1} + a_2x^{m+2} + \cdots + a_kx^{m+k} = \sum_{k=0}^{\infty} a_k\, x^{m+k} \qquad (8.36)$$

Then, $\frac{dy}{dx} = \sum_{k=0}^{\infty} a_k\,(m+k)x^{m+k-1}$

$$\frac{d^2y}{dx^2} = \sum_{k=0}^{\infty} a_k\,(m+k)(m+k-1)x^{m+k-2}$$

Substituting into (8.35), we obtain

$$\sum a_k(m+k)(m+k-1)x^{m+k-2} - 2x\sum a_k(m+k)x^{m+k-1} + 2n\sum a_kx^{m+k} = 0$$

Simplifying,

$$\sum a_k(m+k)(m+k-1)x^{m+k-2} - 2\sum a_k(m+k)x^{m+k} + 2n\sum a_kx^{m+k} = 0$$

$$\sum a_k(m+k)(m+k-1)x^{m+k-2} - \sum 2a_k[(m+k)-n]x^{m+k} = 0 \qquad (8.37)$$

To obtain the lowest power term x^{m-2}, we take $k = 0$ in the first summation:

$$a_0m(m-1) = 0$$

$$\Rightarrow m = 0,\ 1 \qquad (8.38)$$

Equating coefficients of the general term x^{m+k}, we get

$$a_{k+2} = \frac{2(m+k-n)}{(m+k+2)(m+k+1)}a_k \qquad (8.39)$$

Case I: When$m = 0$

$$a_{k+2} = \frac{2(k-n)}{(k+2)(k+1)}a_k \qquad (8.40)$$

Now, $a_2 = \frac{2(0-n)}{2\cdot 1}a_0 = -na_0$

$$a_4 = \frac{2(2-n)}{4\cdot 3}a_2 = \frac{2(2-n)}{12}(-na_0) = \frac{2^2n(n-2)}{4!}a_0$$

Proceeding similarly,

$$a_{2r} = (-1)^r \frac{2^r n(n-2)(n-4)\cdots(n-2r+2)}{(2r)!} a_0$$

And $a_{2r+1} = 0$
Thus, the solution becomes

$$y = a_0 \left[1 - \frac{2n}{2!}x^2 + \frac{2^2 n(n-2)}{4!}x^4 - \cdots + (-1)^r \frac{2^r n(n-2)\cdots(n-2r+2)}{(2r)!}x^{2r} + \cdots \right] \quad (8.41)$$

Case II: When$m = 1$

$$a_{k+2} = \frac{2(k+1-n)}{(k+3)(k+2)} a_k \quad (8.42)$$

Now,$a_2 = \frac{2(1-n)}{3!} a_0$

$$a_4 = \frac{2^2(n-1)(n-3)}{5!} a_0$$

and in general,

$$a_{2r+1} = (-1)^r \frac{2^r (n-1)(n-3)\cdots(n-2r+1)}{(2r+1)!} a_0$$

Thus, the solution becomes

$$y = a_0 x \left[1 - \frac{2(n-1)}{3!}x^2 + \frac{2^2(n-1)(n-3)}{5!}x^4 - \cdots + (-1)^r \frac{2^r (n-1)(n-3)\cdots(n-2r+1)}{(2r+1)!}x^{2r} + \cdots \right] \quad (8.43)$$

Combining both cases, the general solution of equation (8.35) is

$$y = A\left[1 - \frac{2n}{2!}x^2 + \frac{2^2 n(n-2)}{4!}x^4 - \cdots \right] + B\left[x - \frac{2(n-1)}{3!}x^3 + \frac{2^2(n-1)(n-3)}{5!}x^5 - \cdots \right]$$

where A and B are arbitrary constants. (8.44)

When $n$is a non-negative integer, the series terminates and yields a polynomial. These finite solutions are called Hermite polynomials**.**

Physical Significance of Hermite Polynomials

Hermite polynomials have important applications in mathematical physics, including:

- **Quantum Harmonic Oscillator:** Solutions of the one-dimensional Schrödinger equation.
- **Gaussian Beam Optics:** Description of Hermite–Gaussian laser modes.
- **Probability Theory:** Study of the Gaussian (normal) distribution and stochastic processes.
- **Mathematical Physics:** Solution of differential equations involving Gaussian weight functions.

Like the Legendre and Bessel functions, Hermite polynomials possess Rodrigues' formula, generating functions, orthogonality and recurrence relations. These properties are presented collectively in Sections 8.4–8.6.

8.3.4 Laguerre Polynomial

The Laguerre differential equation is given by

$$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + ny = 0 \qquad (8.45)$$

This can also be written as

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

Since $x = 0$ is a regular singular point, we solve it using the series method.

Let $$y = \sum_{k=0}^{\infty} a_k\, x^{m+k}$$

$$y = a_0x^m + a_1x^{m+1} + a_2x^{m+2} + \cdots\cdots \qquad (8.46)$$

Then, $$\frac{dy}{dx} = \sum_{k=0}^{\infty}(m+k)\, a_k x^{m+k-1}$$

$$\frac{d^2y}{dx^2} = \sum_{k=0}^{\infty}(m+k)(m+k-1)\, a_k x^{m+k-2}$$

Substituting into equation (8.45), we obtain

$$x\sum(m+k)(m+k-1)a_kx^{m+k-2} + (1-x)\sum(m+k)a_kx^{m+k-1} + n\sum a_kx^{m+k} = 0$$

Simplifying,

$$\sum a_k[(m+k)(m+k-1) + (m+k)]x^{m+k-1} - \sum a_k(m+k-n)x^{m+k} = 0$$

Equating the lowest power term,

$$a_0m^2 = 0 \Rightarrow m = 0$$

From the coefficients of x^{m+k},

$$(m+k+1)^2a_{k+1} - (m+k-n)a_k = 0$$

$$\Rightarrow a_{k+1} = \frac{m+k-n}{(m+k+1)^2}a_k$$

For $m = 0$, $$a_{k+1} = \frac{k-n}{(k+1)^2}a_k$$

Thus, $$a_1 = -na_0, a_2 = \frac{n(n-1)}{(2!)^2}a_0, a_3 = -\frac{n(n-1)(n-2)}{(3!)^2}a_0$$

Hence, $$a_k = (-1)^k\frac{n(n-1)(n-2)\cdots(n-k+1)}{(k!)^2}a_0$$

Laguerre Polynomial

Substituting back,

$$y = a_0 \sum_{k=0}^{\infty} (-1)^k \frac{n(n-1)\cdots(n-k+1)}{(k!)^2} x^k$$

Taking $a_0 = 1$

$$L_n(x) = n! \sum_{k=0}^{n} \frac{(-1)^k}{(k!)^2} (n)_k x^k$$

or equivalently,

$$L_n(x) = \sum_{k=0}^{n} \frac{(-1)^k \mathrm{n!}}{(k!)^2 (n-k)!} x^k$$

Values for small n

$$L_0 = 1$$
$$L_1 = 1 - x$$
$$L_2 = 1 - 2x + \frac{x^2}{2}$$
$$L_3 = 1 - 3x + \frac{3x^2}{2} - \frac{x^3}{6}$$

Physical Significance of Laguerre Polynomials

Laguerre polynomials have important applications in problems involving radial symmetry, including:

- **Hydrogen Atom:** Solution of the radial Schrödinger equation.
- **Central Force Problems:** Determination of radial wavefunctions.
- **Laser Beam Modes:** Description of Laguerre–Gaussian beams carrying orbital angular momentum.
- **Mathematical Physics:** Solution of differential equations in cylindrical and spherical coordinates.

Other important properties of Laguerre polynomials, including Rodrigues' formula, generating function, orthogonality and recurrence relations, are discussed collectively in Sections 8.4–8.6.

8.4 RODRIGUE'S FORMULA

Rodrigues' formula gives a compact and direct representation for generating important classes of orthogonal polynomials. Instead of deriving them through series expansions, it expresses these polynomials in terms of derivatives of a suitable weight function. This formulation is particularly useful in simplifying calculations and establishing their properties. In what follows, we shall derive Rodrigues' formula for the Legendre, Laguerre, and Hermite polynomials, thereby illustrating their unified representation.

A) Rodrigues' Formula for Legendre Polynomial

$$P_n(x) = \frac{1}{2^n\, n!}\frac{d^n}{dx^n}(x^2-1)^n$$

Proof: Let $v = (x^2-1)^n$ (A)

Differentiating with respect to x, we get

$$\frac{dv}{dx} = n(x^2-1)^{n-1}(2x)$$

Multiplying both sides by (x^2-1), we obtain

$$(x^2-1)\frac{dv}{dx} = 2nx(x^2-1)^n$$

Using (A), this becomes

$$(x^2-1)\frac{dv}{dx} = 2nxv \qquad (8.47)$$

Now differentiating equation (8.47), $(n+1)$ times and applying Leibniz's theorem, we get an expression which simplifies to

$$(x^2-1)\frac{d^{n+2}v}{dx^{n+2}} + 2x\frac{d^{n+1}v}{dx^{n+1}} - n(n+1)\frac{d^nv}{dx^n} = 0 \qquad (8.48)$$

If we put

$$y = \frac{d^nv}{dx^n}$$

Then equation (8.47) reduces to

$$\boxed{(1-x^2)\tfrac{d^2y}{dx^2} - 2x\tfrac{dy}{dx} + n(n+1)y = 0,}$$

which is the Legendre differential equation.

Hence, it can be shown that

$$y = \frac{d^n}{dx^n}(x^2-1)^n$$

is a solution of Legendre's equation.

Therefore, introducing a constant C, we write

$$C\frac{d^n}{dx^n}(x^2-1)^n = P_n(x),$$

where $P_n(x)$ denotes the Legendre polynomial of degree n.

Choosing the constant appropriately, we obtain Rodrigues' formula as

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n \qquad (8.49)$$

This formula provides a convenient method for generating Legendre polynomials of any order by repeatedly differentiating $(x^2-1)^n$.

Evaluation of Initial Polynomials

Using Rodrigues' formula given by equation (8.49), we obtain the first few Legendre polynomials.

For $n = 0$: $$P_0(x) = \frac{1}{2^0 \cdot 0!}(x^2-1)^0 = 1 \qquad (8.50)$$

For $n = 1$: $$P_1(x) = \frac{1}{2^1 \cdot 1!}\frac{d}{dx}(x^2-1) = \frac{1}{2}(2x) = x \qquad (8.51)$$

For $n = 2$:
$$P_2(x) = \frac{1}{2^2 \cdot 2!}\frac{d^2}{dx^2}(x^2-1)^2$$
$$= \frac{1}{8}\frac{d}{dx}[2(x^2-1)(2x)]$$
$$= \frac{1}{2}(3x^2-1) \qquad (8.52)$$

Higher-order polynomials

$$P_3(x) = \frac{1}{2}(5x^3-3x) \qquad (8.53)$$

$$P_4(x) = \frac{1}{8}(35x^4-30x^2+3) \qquad (8.54)$$

$$P_5(x) = \frac{1}{8}(63x^5-70x^3+15x) \qquad (8.55)$$

$$P_6(x) = \frac{1}{16}(231x^6-315x^4+105x^2-5) \qquad (8.56)$$

Expanding $(x^2-1)^n$ using the binomial theorem,

$$\left(x^2-1\right)^n = \sum_{r=0}^{n}\binom{n}{r}(x^2)^{n-r}(-1)^r$$
$$= \sum_{r=0}^{n}\frac{n!}{r!\,(n-r)!}(-1)^r x^{2n-2r} \qquad (8.57)$$

Differentiating n times,

$$\frac{d^n}{dx^n}(x^2-1)^n = \sum_{r=0}^{n}\frac{n!}{r!\,(n-r)!}(-1)^r\frac{d^n}{dx^n}x^{2n-2r} \qquad (8.58)$$
$$= \sum_{r=0}^{n}\frac{n!}{r!\,(n-r)!}(-1)^r\frac{(2n-2r)!}{(n-2r)!}x^{n-2r}$$

Substituting into Rodrigues' formula,

$$P_n(x) = \frac{1}{2^n n!}\sum_{r=0}^{n}(-1)^r \frac{n!}{r!\,(n-r)!}\frac{(2n-2r)!}{(n-2r)!}x^{n-2r}$$

$$P_n(x) = \sum_{r=0}^{\lfloor n/2 \rfloor}\frac{(-1)^r(2n-2r)!}{2^n r!\,(n-r)!\,(n-2r)!}x^{n-2r} \qquad (8.59)$$

The last term depends on whether $n$is even or odd:

- If the last term is x^0:

$$n - 2r = 0 \Rightarrow r = \frac{n}{2}, (n \text{ even}) \qquad (8.60)$$

- If the last term is x^1:

$$n - 2r = 1 \Rightarrow r = \frac{n-1}{2}, (n \text{ odd}) \qquad (8.61)$$

Thus, the series terminates after a finite number of terms.

C) Rodrigues Formula of Hermite Polynomials

We start with the relation

$$e^{x^2}\frac{\partial^n}{\partial t^n}\left[e^{-(t-x)^2}\right] = H_n(x) + H_{n+1}(x)t + H_{n+2}(x)\frac{t^2}{2!} + \cdots \qquad (8.62)$$

Differentiating $e^{-(t-x)^2}$with respect to t, we obtain

$$\frac{\partial}{\partial t}e^{-(t-x)^2} = -2(t-x)e^{-(t-x)^2}$$

Taking the limit as $t \to 0$,

$$\lim_{t\to 0}\frac{\partial}{\partial t}e^{-(t-x)^2} = 2xe^{-x^2} \qquad (8.63)$$

Now differentiating with respect to x,

$$\frac{\partial}{\partial x}e^{-(t-x)^2} = 2(t-x)e^{-(t-x)^2}$$

Taking the limit as $t \to 0$,

$$\lim_{t\to 0}\frac{\partial}{\partial x}e^{-(t-x)^2} = -2xe^{-x^2} \qquad (8.64)$$

From (8.63) and (8.64), we obtain

$$\lim_{t\to 0}\frac{\partial}{\partial t}e^{-(t-x)^2} = (-1)\lim_{t\to 0}\frac{\partial}{\partial x}e^{-(t-x)^2}$$

Proceeding similarly,

$$\lim_{t\to 0}\frac{\partial^2}{\partial t^2}e^{-(t-x)^2} = (-1)^2\lim_{t\to 0}\frac{\partial^2}{\partial x^2}e^{-(t-x)^2}$$

and in general,

$$\lim_{t\to 0}\frac{\partial^n}{\partial t^n}e^{-(t-x)^2} = (-1)^n\lim_{t\to 0}\frac{\partial^n}{\partial x^n}e^{-(t-x)^2}$$

$$= (-1)^n\frac{d^n}{dx^n}e^{-x^2} \qquad (8.65)$$

Putting $t = 0$in equation (8.62), we get

$$\lim_{t\to 0}e^{x^2}\frac{\partial^n}{\partial t^n}e^{-(t-x)^2} = H_n(x) \qquad (8.66)$$

Substituting from (8.65) into (8.66),

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} \qquad (8.67)$$

Thus, Rodrigue's formula for Hermite Polynomial

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}(8.68)$$

Verification for First Few Polynomials

1) For$n = 0$. $H_0(x) = 1$ (8.69)

2) $For\ n = 1$

$$H_1(x) = (-1)e^{x^2}\frac{d}{dx}e^{-x^2} = -e^{x^2}(-2x)e^{-x^2} = 2x \qquad (8.70)$$

3) $For\ n = 2$

$$H_2(x) = e^{x^2}\frac{d^2}{dx^2}e^{-x^2} = e^{x^2}\frac{d}{dx}(-2xe^{-x^2})$$
$$= e^{x^2}\left[-2e^{-x^2} + 4x^2e^{-x^2}\right] = 4x^2 - 2 \qquad (8.71)$$

4) For $n = 3$

$$H_3(x) = (-1)^3 e^{x^2}\frac{d^3}{dx^3}e^{-x^2} = 8x^3 - 12x \qquad (8.72)$$

The other higher-order polynomials

$$H_4(x) = 16x^4 - 48x^2 + 12 \qquad (8.73)$$
$$H_5(x) = 32x^5 - 160x^3 + 120x \qquad (8.74)$$
$$H_6(x) = 64x^6 - 480x^4 + 720x^2 - 120 \qquad (8.75)$$
$$H_7(x) = 128x^7 - 1344x^5 + 3360x^3 - 1680x \qquad (8.76)$$

C) Rodrigues' formula for Laguerre Polynomial

The Laguerre differential equation is given by

$$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + ny = 0 \qquad (8.77)$$

Derivation of Rodrigues' Formula

Let $y = e^x u$ (8.78)

Differentiating, we have

$$\frac{dy}{dx} = e^x\left(u+\frac{du}{dx}\right) \qquad (8.79)$$

$$\frac{d^2y}{dx^2} = e^x\left(u+2\frac{du}{dx}+\frac{d^2u}{dx^2}\right) \qquad (8.80)$$

Substituting these values into equation (8.77), we get

$$x\left(u+2\frac{du}{dx}+\frac{d^2u}{dx^2}\right)+(1-x)\left(u+\frac{du}{dx}\right)+nu=0 \qquad (8.81)$$

Expanding and simplifying,

$$x\frac{d^2u}{dx^2}+(x+1)\frac{du}{dx}+(n+1)u=0$$

Assuming a solution of the form

$$u=\frac{d^n}{dx^n}(x^n v)$$

On substituting and simplifying (using Leibniz's theorem), this equation is satisfied if

$$v=e^{-x}$$

Thus,

$$u=\frac{d^n}{dx^n}(x^n e^{-x})$$

Substituting into $y=e^x u$, we obtain

$$y=e^x\frac{d^n}{dx^n}(x^n e^{-x})$$

Thus, Rodrigue's Formula for Laguerre Polynomial is obtained byintroducing the factor $\frac{1}{n!}$, we get

$$\boxed{L_n(x)=\frac{e^x}{n!}\frac{d^n}{dx^n}(e^{-x}x^n)} \qquad (8.82)$$

8.5 GENERATING FUNCTIONS OF DIFFERENT POLYNOMIAL

A) General functions of Legendre Polynomial

The generating function of Legendre polynomials $P_n(x)$is given by

$$\frac{1}{\sqrt{1-2xt+t^2}}=P_0(x)+P_1(x)t+P_2(x)t^2+\cdots=\sum_{n=0}^{\infty}P_n(x)t^n \qquad (8.83)$$

Derivation:

Let $$G(x,t)=\frac{1}{\sqrt{1-2xt+t^2}} \qquad (8.84)$$

We expand this expression in powers of t using the binomial expansion.

Rewriting the expression (8.84), we have

$$G(x,t) = (1 - 2xt + t^2)^{-1/2}$$

Using binomial expansion for $(1+z)^m$,

$$(1+z)^m = 1 + mz + \frac{m(m-1)}{2!}z^2 + \cdots$$

Let $$z = -2xt + t^2$$

Then, $$G(x,t) = 1 + \left(-\frac{1}{2}\right)(-2xt + t^2) + \frac{(-1/2)(-3/2)}{2!}(-2xt + t^2)^2 + \cdots$$

Expanding and collecting terms of same powers of t, we obtain

$$G(x,t) = 1 + xt + \frac{1}{2}(3x^2 - 1)t^2 + \frac{1}{2}(5x^3 - 3x)t^3 + \cdots$$

Comparing with the series

$$G(x,t) = \sum_{n=0}^{\infty} P_n(x)t^n$$

We identify

$$P_0(x) = 1$$
$$P_1(x) = x$$
$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

Thus, the generating function is given by

$$\boxed{\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n} \qquad (8.85)$$

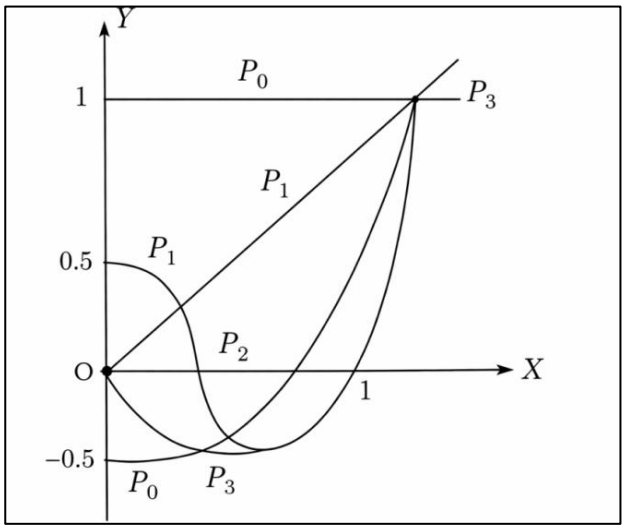


Figure 8.1: Graph illustrating the Legendre polynomial $P_0(x), P_1(x), P_2(x)\ and\ P_3(x)$

Recurrence Formula:

The following recurrence relations are obtained from the generating function of Legendre polynomials:

(1) $(n+1)\,P_{n+1}(x) = (2n+1)\,x\,P_n(x) - n\,P_{n-1}(x)$

(2) $n\,P_n(x) = x\,\frac{d}{dx}P_n(x) - \frac{d}{dx}P_{n-1}(x)$

(3) $(2n+1)\,P_n(x) = \frac{d}{dx}P_{n+1}(x) - \frac{d}{dx}P_{n-1}(x)$

(4) $\frac{d}{dx}P_n(x) = x\,\frac{d}{dx}P_{n-1}(x) + n\,P_{n-1}(x)$

(5) $(x^2-1)\,\frac{d}{dx}P_n(x) = n[x\,P_n(x) - P_{n-1}(x)]$

(6) $(x^2-1)\,\frac{d}{dx}P_n(x) = (n+1)[P_{n+1}(x) - x\,P_n(x)]$

B) Generating Function of Bessel's functions $J_n(x)$

To show that $J_n(x)$ is the coefficient of z^n in the expansion of

$$e^{\frac{x}{2}\left(z-\frac{1}{z}\right)} \qquad (8.86)$$

Proof: We begin by writing the exponential function as a product:

$$e^{\frac{x}{2}\left(z-\frac{1}{z}\right)} = e^{\frac{x}{2}z}\cdot e^{-\frac{x}{2z}}$$

Let us expand each exponential of the above equations

$$e^{\frac{x}{2}z} = 1 + \frac{x}{2}z + \frac{1}{2!}\left(\frac{x}{2}z\right)^2 + \frac{1}{3!}\left(\frac{x}{2}z\right)^3 + \cdots \qquad (8.87)$$

$$e^{-\frac{x}{2z}} = 1 - \frac{x}{2z} + \frac{1}{2!}\left(\frac{x}{2z}\right)^2 - \frac{1}{3!}\left(\frac{x}{2z}\right)^3 + \cdots \qquad (8.88)$$

Multiplying (8.87) and (8.88), we get

$$e^{\frac{x}{2}\left(z-\frac{1}{z}\right)} = \left(\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{x}{2}z\right)^r\right)\left(\sum_{s=0}^{\infty}\frac{(-1)^s}{s!}\left(\frac{x}{2z}\right)^s\right)$$

$$= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{(-1)^s}{r!\,s!}\left(\frac{x}{2}\right)^{r+s} z^{\,r-s} \qquad (8.89)$$

For the coefficient of z^n, we require

$$r - s = n \Rightarrow r = n + s$$

Substituting the condition $r - s = n$ into the expanded series, the coefficient of z^n in equation (8.89) is obtained as

$$\sum_{s=0}^{\infty}\frac{(-1)^s}{s!\,(n+s)!}\left(\frac{x}{2}\right)^{n+2s},$$

which is precisely the series representation of the Bessel function of the first kind of order n, namely

$$J_n(x) = \sum_{s=0}^{\infty} \frac{(-1)^s}{s!\,(n+s)!}\left(\frac{x}{2}\right)^{n+2s}.$$

Hence, comparing with the expansion, we obtain the generating function for Bessel functions as

$$e^{\frac{x}{2}\left(z-\frac{1}{z}\right)} = \sum_{n=-\infty}^{\infty} J_n(x)\, z^n. \qquad (8.90)$$

Thus, $J_n(x)$ appears as the coefficient of z^n in the above expansion. In particular, the coefficient of z^0 gives

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 \cdot 4^2} - \cdots,$$

which agrees with the known series expansion of the Bessel function of order zero.

Recurrence Formulae (Bessel Functions)

The following recurrence relations are extremely useful in solving problems involving Bessel functions. Due to their importance, they should be remembered.

Recurrence Relations

1. $xJ_n^{'}(x) = nJ_n(x) - xJ_{n+1}(x)$
2. $xJ_n'(x) = -nJ_n(x) + xJ_{n-1}(x)$
3. $2J_n'(x) = J_{n-1}(x) - J_{n+1}(x)$
4. $2nJ_n(x) = x[J_{n-1}(x) + J_{n+1}(x)]$
5. $\frac{d}{dx}(x^{-n}J_n(x)) = -x^{-n}J_{n+1}(x)$
6. $\frac{d}{dx}(x^nJ_n(x)) = x^nJ_{n-1}(x)$

These recurrence relations establish connections between Bessel functions of different orders and their derivatives. They are particularly helpful for simplifying expressions and solving differential equations involving Bessel functions.

D) Generating Function of Laguerre Polynomial

We start with the generating relation

$$(1-t)^{-1}\sum_{n=0}^{\infty}\frac{L_n(x)}{n!}t^n = e^{-\frac{xt}{1-t}} \qquad (8.91)$$

Derivation:

From equation (8.91), we have

$$\sum_{n=0}^{\infty} \frac{L_n(x)}{n!} t^n = (1-t)e^{-\frac{xt}{1-t}}$$

Now, consider the exponential term:

$$e^{-\frac{xt}{1-t}} = \exp\left(-x\frac{t}{1-t}\right)$$

Using the expansion

$$e^{-y} = \sum_{k=0}^{\infty} \frac{(-1)^k y^k}{k!}$$

we obtain

$$e^{-\frac{xt}{1-t}} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\frac{xt}{1-t}\right)^k$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k x^k t^k}{k!\,(1-t)^k} \qquad (8.92)$$

Multiplying equation (8.92) by $(1-t)^{-1}$, we get

$$(1-t)^{-1}e^{-\frac{xt}{1-t}} = \sum_{k=0}^{\infty} \frac{(-1)^k x^k t^k}{k!\,(1-t)^{k+1}} \qquad (8.93)$$

Now expand

$$\frac{1}{(1-t)^{k+1}} = \sum_{r=0}^{\infty} \frac{(k+r)!}{k!\;r!} t^r \qquad (8.94)$$

Substituting (8.94) into (8.93),

$$(1-t)^{-1}e^{-\frac{xt}{1-t}} = \sum_{k=0}^{\infty} \frac{(-1)^k x^k}{k!} \sum_{r=0}^{\infty} \frac{(k+r)!}{k!\;r!} t^{k+r} \qquad (8.95)$$

Rearranging the summation by putting $n = k + r$,

$$= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} \frac{(-1)^k (k+n-k)!}{(k!)^2 (n-k)!} x^k\right] t^n$$

$$= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} \frac{(-1)^k n!}{(k!)^2 (n-k)!} x^k\right] t^n$$

Equating coefficients of t^n from both sides of equation (8.91), we obtain

$$\frac{L_n(x)}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{(k!)^2} \frac{n!}{(n-k)!} x^k \qquad (8.96)$$

Thus, the Laguerre polynomial is

$$L_n(x) = \sum_{k=0}^{n} \frac{(-1)^k n!}{(k!)^2 (n-k)!} x^k$$

Hence, the generating function

$$\frac{1}{(1-t)} \exp\left(\frac{-xt}{1-t}\right) = \sum_{n=0}^{\infty} \mathrm{L_n(x)t^n} \qquad (8.97)$$

produces the Laguerre polynomials $L_n(x)$, confirming that they are polynomials of degree n in x with leading coefficient proportional to $(-1)^n$.

8.6 ORTHOGONALITY OF POLYNOMIALS

1) Orthogonality of Legendre Polynomials

Orthogonality of Legendre Polynomials is given by

$$\int_{-1}^{1} P_n(x)\, P_m(x)\, dx = 0 \text{for } n \neq m$$

Proof:

Let $P_n(x)$be a solution of Legendre's differential equation:

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0 \qquad (8.98)$$

Similarly, $P_m(x)$satisfies:

$$(1-x^2)\frac{d^2z}{dx^2} - 2x\frac{dz}{dx} + m(m+1)z = 0 \qquad (8.99)$$

Combining the Two Equations

Multiplying equation (8.94) by z and equation (8.95) by y, and then subtracting, we obtain:

$$(1-x^2)\left[z\frac{d^2y}{dx^2} - y\frac{d^2z}{dx^2}\right] - 2x\left[z\frac{dy}{dx} - y\frac{dz}{dx}\right] + [n(n+1) - m(m+1)]yz = 0$$

Rewriting the above expression,

$$(1-x^2)\left[\frac{d}{dx}\left(z\frac{dy}{dx}\right) - \frac{d}{dx}\left(y\frac{dz}{dx}\right)\right] - 2x\left[z\frac{dy}{dx} - y\frac{dz}{dx}\right] + (n-m)(n+m+1)yz = 0$$

This simplifies to

$$\frac{d}{dx}\left[(1-x^2)\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)\right]+(n-m)(n+m+1)yz=0$$

Integrating over the interval $[-1,1]$, we get:

$$\left[(1-x^2)\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)\right]_{-1}^{1}+(n-m)(n+m+1)\int_{-1}^{1}yz\,dx=0$$

Since $(1-x^2)=0$ at $x=\pm1$, the boundary term vanishes. Therefore,

$$0+(n-m)(n+m+1)\int_{-1}^{1}yz\,dx=0$$

Substituting $y=P_n(x)$ and $z=P_m(x)$, we obtain:

$$(n-m)(n+m+1)\int_{-1}^{1}P_n(x)\,P_m(x)\,dx=0$$

For $n\neq m$,

$$\int_{-1}^{1}P_n(x)\,P_m(x)\,dx=0$$

Thus, Legendre polynomials of different degrees are orthogonal over the interval $[-1,1]$

B) Orthogonality of Bessel Function

$$\int_{0}^{1}x\,J_n(\alpha x)\,J_n(\beta x)\,dx=0$$

where α and β are distinct zeros of $J_n(x)=0$.

Proof

We begin with Bessel's differential equation:

$$x^2\frac{d^2y}{dx^2}+x\frac{dy}{dx}+(\alpha^2x^2-n^2)y=0 \qquad (8.100)$$

Similarly,

$$x^2\frac{d^2z}{dx^2}+x\frac{dz}{dx}+(\beta^2x^2-n^2)z=0 \qquad (8.101)$$

The solutions of (8.100) and (8.101) are

$$y=J_n(\alpha x),z=J_n(\beta x)$$

Multiplying equation (8.100) by $\frac{z}{x}$ and equation (8.101) by $\frac{y}{x}$, and subtracting, we obtain

$$x\left(z\frac{d^2y}{dx^2}-y\frac{d^2z}{dx^2}\right)+\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)+(\alpha^2-\beta^2)xyz=0$$

This can be rewritten as

$$\frac{d}{dx}\left[x\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)\right]+(\alpha^2-\beta^2)xyz=0 \qquad (8.102)$$

Integrating equation (8.102) with respect to x from 0 to 1, we get

$$\left[x\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)\right]_0^1+(\alpha^2-\beta^2)\int_0^1 xyz\,dx=0$$

Rewriting, $$(\beta^2-\alpha^2)\int_0^1 xyz\,dx=\left[x\left(z\frac{dy}{dx}-y\frac{dz}{dx}\right)\right]_0^1 \qquad (8.103)$$

Substituting

$$y=J_n(\alpha x),\frac{dy}{dx}=\alpha J_n'(\alpha x)$$

$$z=J_n(\beta x),\frac{dz}{dx}=\beta J_n'(\beta x)$$

into (8.103), we obtain

$$(\beta^2-\alpha^2)\int_0^1 x\,J_n(\alpha x)J_n(\beta x)\,dx=[\alpha J_n'(\alpha x)J_n(\beta x)-\beta J_n'(\beta x)\,J_n(\alpha x)]_0^1 \qquad (8.104)$$

Since α and β are roots of $J_n(x)=0$, we have

$$J_n(\alpha)=0,J_n(\beta)=0$$

Thus, the right-hand side of (8.99) vanishes, giving

$$(\beta^2-\alpha^2)\int_0^1 x\,J_n(\alpha x)J_n(\beta x)\,dx=0$$

For $\alpha\neq\beta$,

$$\int_0^1 x\,J_n(\alpha x)J_n(\beta x)\,dx=0$$

Hence, Bessel functions corresponding to different zeros are orthogonal over the interval $[0,1]$ with weight function x.

C) Orthogonality of Hermite Polynomials

The orthogonality relation for Hermite polynomials is given by

$$\int_{-\infty}^{\infty}e^{-x^2}H_m(x)H_n(x)\,dx=\begin{cases}0, & m\neq n\\ 2^n n!\sqrt{\pi}, & m=n\end{cases}$$

Solution : We begin with the generating function

$$e^{x^2-(t_1-x)^2}=\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}t_1^n \qquad (8.105)$$

and similarly, $e^{x^2-(t_2-x)^2}=\sum_{m=0}^{\infty}\frac{H_m(x)}{m!}t_2^m$ (8.106)

Multiplying (8.105) and (8.106), we obtain

$$e^{x^2-(t_1-x)^2}\cdot e^{x^2-(t_2-x)^2}=\left[\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}t_1^n\right]\left[\sum_{m=0}^{\infty}\frac{H_m(x)}{m!}t_2^m\right]$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{H_n(x) H_m(x)}{n!\, m!} t_1^n t_2^m$$

Multiplying both sides by e^{-x^2} and integrating from $-\infty$ to ∞, we get

$$\sum_{n,m} \left[\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx \right] \frac{t_1^n t_2^m}{n!\, m!}$$

$$= \int_{-\infty}^{\infty} e^{-x^2} e^{x^2-(t_1-x)^2} e^{x^2-(t_2-x)^2} dx$$

$$= \int_{-\infty}^{\infty} e^{x^2-(t_1-x)^2-(t_2-x)^2}\, dx$$

Simplifying the exponent,

$$= e^{-(t_1^2+t_2^2)} \int_{-\infty}^{\infty} e^{-x^2+2x(t_1+t_2)}\, dx \cdots \qquad (8.107)$$

We use the known result

$$\int_{-\infty}^{\infty} e^{-ax^2+2bx}\, dx = \sqrt{\frac{\pi}{a}}\, e^{\frac{b^2}{a}} \text{(standard formula)}$$

Putting $a = 1$ and $2b = 2(t_1 + t_2)$, we obtain

$$\int_{-\infty}^{\infty} e^{-x^2+2x(t_1+t_2)}\, dx = \sqrt{\pi}\, e^{(t_1+t_2)^2} \cdots \qquad (8.108)$$

Substituting (8.108) into (8.107), we get

$$\sum_{n,m} \left[\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx \right] \frac{t_1^n t_2^m}{n!\, m!} = \sqrt{\pi}\, e^{-(t_1^2+t_2^2)} e^{(t_1+t_2)^2}$$

$$= \sqrt{\pi}\, e^{2t_1 t_2}$$

Expanding,

$$= \sqrt{\pi} \left[1 + 2t_1 t_2 + \frac{(2t_1 t_2)^2}{2!} + \frac{(2t_1 t_2)^3}{3!} + \cdots \right]$$

$$= \sqrt{\pi} \sum_{r=0}^{\infty} \frac{2^r}{r!} (t_1 t_2)^r$$

$$= \sqrt{\pi} \sum_{r=0}^{\infty} \frac{2^r}{r!} t_1^r t_2^r$$

From equation (8.107), we have

$$\sum_{n,m} \left[\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx \right] \frac{t_1^n t_2^m}{n!\, m!} = \sqrt{\pi} \sum_{r=0}^{\infty} \frac{2^r}{r!} t_1^r t_2^r$$

Equating coefficients of $t_1^n t_2^m$, we obtain

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx = \sqrt{\pi}\, 2^n n!\ \delta_{n,m}$$

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx = \begin{cases} 0, & m \neq n \\ 2^n n! \sqrt{\pi}, & m = n \end{cases}$$

Where $\delta_{n,m} = \begin{cases} 0, & n \neq m \\ 1, & n = m \end{cases}$

Thus, Hermite polynomials are orthogonal over $(-\infty, \infty)$ with respect to the weight function e^{-x^2}.

Recurrence Formulae for $H_n(x)$ of Hermite Equation

There are four important recurrence relations for Hermite polynomials:

1. $2n\, H_{n-1}(x) = H_n'(x)$
2. $2x\, H_n(x) = 2n\, H_{n-1}(x) + H_{n+1}(x)$
3. $H_n'(x) = 2x\, H_n(x) - H_{n+1}(x)$
4. $H_n''(x) - 2x\, H_n'(x) + 2n\, H_n(x) = 0$

D) Orthogonal Property of Laguerre Polynomial

Consider $\int_0^{\infty} e^{-x} x^m L_n(x)\, dx$

Using integration by parts,

$$\int_0^{\infty} e^{-x} x^m L_n(x)\, dx = (-1)^n m! \int_0^{\infty} \frac{d^{n-m}}{dx^{n-m}} (x^n e^{-x}) dx$$

$$= 0 \text{ if } n > m$$

Thus, $\int_0^{\infty} e^{-x} L_m(x) L_n(x)\, dx = 0$ for $m \neq n$

Case 1: $m = n$ $\quad L_n(x) = (-1)^n x^n$

$$\int_0^{\infty} e^{-x} [L_n(x)]^2 dx = (n!)^2$$

Normalized Form

$$\int_0^{\infty} e^{-x/2} \frac{L_m(x)}{m!} e^{-x/2} \frac{L_n(x)}{n!} dx = \delta_{mn} (8.109)$$

The final orthogonality relation is given by

$$\int_0^{\infty} e^{-x} L_m(x) L_n(x)\, dx = \delta_{mn} \begin{cases} 0, & m \neq n \\ (n!)^2, & m = n \end{cases}$$

Thus, Laguerre polynomials arise from the solution of Laguerre's differential equation, possess a well-defined generating function, satisfy recurrence relations and form an orthogonal set over $[0, \infty)$ with weight e^{-x}

Recurrence Formula

1) Relation 1: Starting from the generating function

$$e^{-\frac{xt}{1-t}} = (1-t)\sum_{n=0}^{\infty}\frac{L_n(x)}{n!}t^n \qquad (8.110)$$

Differentiating with respect to t, we obtain the important recurrence relation

$$\boxed{L_{n+1}(x) + (x-2n-1)\,L_n(x) + n^2\,L_{n-1}(x) = 0}$$

2) Relation 2: Differentiating equation (1) with respect to x, we get

$$\boxed{\frac{d}{dx}L_n(x) = -L_{n-1}(x)}$$

8.7 APPLICATIONS

We can execute different problems based on the polynomials to know its application

Q1: Using the generating function of Legendre polynomials $P_n(x)$, show that

$$\int_0^1 P_{2n+1}(x)\,dx = (-1)^n\frac{(2n)!}{2^{2n+1}n!\,(n+1)!}$$

and that, except for $n = 0$, $\int_0^1 P_{2n}(x)\,dx = 0$.

Solution: The generating function for Legendre polynomials is

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty}P_n(x)\,t^n \qquad (1)$$

Integrating both sides of equation (1) from $x = 0$ to $x = 1$, we obtain

$$\int_0^1\frac{dx}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty}t^n\int_0^1 P_n(x)\,dx \qquad (2)$$

Consider $I = \int_0^1\frac{dx}{\sqrt{1-2xt+t^2}}$

This integral evaluates to $I = \frac{1}{t}\left[\sqrt{1+t^2} - (1-t)\right]$ (3)

Expanding $\sqrt{1+t^2}$ using binomial expansion:

$$\sqrt{1+t^2} = 1 + \frac{t^2}{2} - \frac{t^4}{8} +$$

Substituting into (3), we get

$$I = \frac{1}{t}\left[1 + \frac{t^2}{2} - \frac{t^4}{8} + \cdots - (1-t)\right]$$
$$= \frac{1}{t}\left[t + \frac{t^2}{2} - \frac{t^4}{8} + \cdots\right]$$
$$= 1 + \frac{t}{2} - \frac{t^3}{8} + \cdots\cdots \qquad (4)$$

From equation (2),

$$\sum_{n=0}^{\infty} t^n \int_0^1 P_n(x)\, dx$$

Comparing with expansion (4), we equate coefficients of powers of t. Since the expansion (4) contains no even powers of t (except the constant term),

$$\int_0^1 P_{2n}(x)\, dx = 0 (n \neq 0)$$

Comparing coefficients of t^{2n+1}, we obtain

$$\int_0^1 P_{2n+1}(x)\, dx = (-1)^n \frac{(2n)!}{2^{2n+1} n!\,(n+1)!}$$

Thus,

$$\int_0^1 P_{2n}(x)\, dx = 0 (n \neq 0)$$

Thus, using the generating function of Legendre polynomials:

1. Odd-degree polynomials yield a closed-form integral.
2. Even-degree polynomials (except P_0) integrate to zero over $[0, 1]$.

Q2: Show that $\int_{-1}^{1} P_n(x)\, dx = 0, n \neq 0$

and $\int_{-1}^{1} P_n(x)\, dx = 2, n = 0$

Solution: (i) We know that $P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$

Integrating both sides over the interval $[-1,1]$, we obtain

$$\int_{-1}^{1} P_n(x)\, dx = \frac{1}{2^n n!}\int_{-1}^{1} \frac{d^n}{dx^n}(x^2-1)^n\, dx$$
$$= \frac{1}{2^n n!}\left[\frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n\right]_{-1}^{1}$$

$$= \frac{1}{2^n n!}[0 - 0] = 0$$

(ii) When $n = 0$, $\int_{-1}^{1} P_0(x)\,dx = \int_{-1}^{1} 1\,dx = [x]_{-1}^{1} = 2$

Sturm–Liouville Equation

The general form of the Sturm–Liouville differential equation is

$$\frac{d}{dx}\left[P(x)\frac{dy}{dx}\right] + [\lambda q(x) + r(x)]y = 0$$

We know that Bessel's equation is

$$X^2\frac{d^2y}{dX^2} + X\frac{dy}{dX} + (X^2 - n^2)y = 0 (1)$$

Substituting $X = kx$ in (1), we get

$$\frac{dy}{dX} = \frac{dy}{dx}\frac{dx}{dX} = \frac{1}{k}\frac{dy}{dx}$$
$$k^2x^2\left(\frac{d^2y}{dx^2}\frac{1}{k^2}\right) + (kx)\left(\frac{dy}{dx}\frac{1}{k}\right) + (k^2x^2 - n^2)y = 0$$

which simplifies to $x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (k^2x^2 - n^2)y = 0$

Dividing throughout by x, we obtain

$$x\frac{d^2y}{dx^2} + \frac{dy}{dx} + \left(k^2x - \frac{n^2}{x}\right)y = 0$$

or equivalently, $\left(x\frac{d^2y}{dx^2} + \frac{dy}{dx}\right) + \left(k^2x - \frac{n^2}{x}\right)y = 0$

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) + \left(\lambda x - \frac{n^2}{x}\right)y = 0 \text{ (put } k^2 = \lambda) (2)$$

Equations (1) and (2) are of the standard form

$$\frac{d}{dx}\left[P(x)\frac{dy}{dx}\right] + [\lambda q(x) + r(x)]y = 0 (3)$$

Equation (3) is called the **Sturm–Liouville equation**.

Sturm–Liouville Problem

When equation (3) is considered along with the boundary conditions

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0$$
$$\beta_1 y(b) + \beta_2 y'(b) = 0$$

it is known as the Sturm–Liouville problem**.**

Eigenvalues and Eigenfunctions

The solution of the Sturm–Liouville problem is called an eigenfunction, and the corresponding value of λ is called an eigenvalue**.**

Putting $p = 1$, $q = 1$, $r = 0$ in (3), we get

$$\frac{d^2y}{dx^2} + \lambda y = 0$$

Taking the boundary conditions $\alpha_1 = \beta_1 = 1, \alpha_2 = \beta_2 = 0$

we obtain$y(a) = 0$ and $y(b) = 0$

Thus,$y'' + \lambda y = 0, y(a) = 0,\ y(b) = 0$

which represents the simplest form of the Sturm–Liouville problem.

CHECK YOUR PROGRESS

Q1: Express in terms of Legendre Polynomial

$a)\ x^4 + 3x^3 - x^2 + 5x - 2$ b) $1 + 2x\ - 3x^2 + 4x^3$

Q2: Show that $x^3 = \frac{2}{5}P_3(x) + \frac{3}{5}P_1(x)$

Q3: Evaluate $\int_{-1}^{+1} x^3 P_4(x)\, dx$

Q4: Prove that$(2\mathrm{n} + 1)\mathrm{P_n}(\mathrm{x}) = \mathrm{P'_{n+1}}(\mathrm{x}) - \mathrm{P'_{n-1}}(\mathrm{x})$

Q5: Evaluate, $\int P_n\,(x)dx$, n = 1, 2, 3 where $P_n(x)$ is the Legendre polynomial of degree n [(i) (ii) $\frac{1}{2}$ (iii) 0 (iv) 2

Q6: The value of J_2 (1) is, where J2(x) is the Bessel Function of first kind

(i)0 (ii) 2 (iii) -1 (iv) None of these

Q7: Convert the Hermite polynomial $2H_3\ - 4H_2 + H_1 + H_0$ into ordinary polynomial.

Q8: Find a series solution of the following differential equation-$xy'' + (1 + x)y' + y\ = 0$

8.8 LET US SUM UP

- Special functions arise as solutions of differential equations in mathematical physics under boundary conditions.
- The main functions studied are Legendre, Bessel, Hermite, and Laguerre polynomials/functions.
- Each function satisfies a specific differential equation, namely the Legendre, Bessel, Hermite, and Laguerre equations.
- For integer values of parameters, the solutions reduce to finite polynomials.

- General solutions often involve two linearly independent solutions (e.g., $P_n(x)$ and $Q_n(x)$ for Legendre).
- Rodrigues' formula provides a direct and compact way to generate these polynomials.
- Generating functions represent entire sequences of polynomials and are useful for deriving properties.
- Recurrence relations connect functions of different orders and simplify calculations.
- Series solutions are widely used to derive these functions and understand their structure.
- Orthogonality is a key property in which functions of different orders are orthogonal over specific intervals with respect to an associated weight function.
- Orthogonality allows expansion of functions into series (basis functions in physics problems).
- Bessel functions are oscillatory and appear in problems with cylindrical or spherical symmetry.
- These functions are essential in solving boundary value problems such as heat conduction, wave equations, vibrations, and potential problems. Special functions provide physically meaningful solutions where elementary functions are not sufficient.
- They play a crucial role in mathematical modelling across physics and engineering.

✓ 8.9 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1 (a) $\frac{8}{35}P_4(x)+\frac{6}{5}P_3(x)-\frac{2}{21}P_2(x)+\frac{34}{5}P_1(x)-\frac{32}{15}$

(b) $\frac{8}{5}P_3(x)-2P_2(x)+\frac{22}{5}P_1(x)$

Answer to Q2: We know the Legendre polynomials:

$$P_1(x)=x$$
$$P_3(x)=\frac{1}{2}(5x^3-3x)$$

Now consider the RHS of the equation

$$\frac{2}{5}P_3(x)+\frac{3}{5}P_1(x)$$

Substitute the values

$$\frac{2}{5}P_3(x)+\frac{3}{5}P_1(x)=\frac{2}{5}\left[\frac{1}{2}(5x^3-3x)\right]+\frac{3}{5}(x)$$
$$=\frac{1}{5}(5x^3-3x)+\frac{3}{5}x$$
$$=x^3-\frac{3}{5}x+\frac{3}{5}x$$
$$=x^3$$

Hence, $x^3=\frac{2}{5}P_3(x)+\frac{3}{5}P_1(x)$

Answer to Q3: 0

Answer to Q4: We start with the standard recurrence relation of Legendre polynomials:

$$(2n+1)xP_n(x) = (n+1)P_{n+1}(x) + nP_{n-1}(x) \ldots \quad .(1)$$

Differentiate both sides with respect to x:

$$(2n+1)[P_n(x) + xP_n'(x)] = (n+1)P_{n+1}'(x) + nP_{n-1}'(x) \ldots (2)$$

Now, use the identity:

$$(1-x^2)P_n'(x) = n[P_{n-1}(x) - xP_n(x)] \ldots (3)$$

Rearranging (3): $xP_n'(x) = P_{n-1}(x) - \frac{1-x^2}{n}P_n'(x)$

Substitute appropriately into (2) and simplify. After eliminating terms and arranging, we obtain:

$$(2n+1)P_n(x) = P_{n+1}'(x) - P_{n-1}'(x)$$

Answer to Q5: (iii) 0

Answer to Q6: (iv) None of these

Answer to Q7: 16 x^3-16 x^2-22x+9

Answer to Q8: $y = e^{-x}L_o(x) = e^{-x} = 1 - x + \frac{x^2}{2I_0} - \frac{x^3}{3I_0} + \ldots$

8.10 FURTHER READING

- Dass.H. K. and Verma Rama (2010), *Mathematical Physics*, S. Chand & Company Ltd.
- Grewal. B. S. (2014), *Higher Engineering Mathematics,* Khanna Publishers.
- Boas. M. L. (2006), *Mathematical Methods in the Physical Sciences*, Wiley (3rd Edition).
- Arfken. G. and WeberH. (2013), *Mathematical Methods for Physicists*, Academic Press (7th Edition).
- Kreyszig. E. (2011), *Advanced Engineering Mathematics*, Wiley (10th Edition).

8.11 MODEL QUESTIONS

1. Prove $L_3^1(x) = -18 + 18x - 3x^2$

2. Prove that $2n\,H_{n-1}(x) = H_n'(x)$

3. Prove identities like $x^3 = \frac{2}{5}P_3(x) + \frac{3}{5}P_1(x)$

4. Verify that $H_1(x) = 2x$ satisfies Hermite's equation

$$y'' - 2xy' + 2ny = 0 (n = 1)$$

5. Evaluate $\int_0^\infty e^{-x} L_0(x) L_2(x)\, dx$

6. Verify that $y = J_0(x)$ satisfies Bessel equation of order 0

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

*** *** ***

UNIT 9: LAPLACE'S TRANSFORM

UNIT STRUCTURE

9.1 Learning Objectives
9.2 Introduction
9.3 Laplace's Transform of Elementary Functions
    9.3.1 Properties
    9.3.2 Shifting
    9.3.3 Scaling
9.4 Derivatives and Integrals of Laplace's Transform
9.5 Convolution Theorem
9.6 Inverse Laplace Transformation
9.7 Solutions of Differential Equations
9.8 Let Us Sum Up
9.9 Answers to Check Your Progress
9.10 Further Reading
9.11 Model Questions

9.1 LEARNING OBJECTIVES

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

- define and explain the Laplace transform and its basic properties
- compute Laplace transforms of elementary functions using shifting and scaling rules
- apply derivative, integral, and convolution properties in the s-domain
- determine inverse Laplace transforms using standard techniques
- solve linear differential equations using Laplace transform methods.

9.2 INTRODUCTION

In the previous unit, we studied standard functions and mathematical techniques that provide a useful foundation for evaluating and applying Laplace transforms. In this unit, these concepts are extended to transform functions from the time domain into the complex-frequency domain. Here, we study the Laplace transform, one of the important mathematical tools used in the analysis of linear systems. A transformation converts a mathematical expression from one form into another equivalent but often simpler form. This unit covers the Laplace transform and its derivatives and integrals, along with the convolution theorem and inverse Laplace transformation. Finally, the application of Laplace transforms to the solution of differential equations is introduced.

9.3 LAPLACE'S TRANSFORM OF ELEMENTARY FUNCTIONS

The Laplace transform is particularly useful for converting differential equations into algebraic equations, thereby simplifying their solution. The Laplace transform has wide applications in physics and engineering, particularly in the analysis of electrical circuits, mechanical systems, optical devices and harmonic oscillators. Among the commonly

used integral transforms, the Laplace and Fourier transforms are particularly important. While the Fourier transform represents a function or signal in terms of frequencies, the Laplace transform is especially useful for solving differential and integral equations and analysing systems involving initial conditions.

A transformation converts an operation or function into another equivalent form, often making a complex problem easier to handle. Differentiation and integration are simple examples of such transformations. In general, a linear integral transformation of a function G(t)can be represented as

$$g(s) = \int G(t)k(s,t)dt$$

where g(s) is the transformed function and k(st)is the kernel of the transformation. When the kernel is of the form est, the transformation is known as the Laplace transform. If the kernel is of the form e-tot used in the Fourier transform, the resulting transformation is known as the Fourier transform, which is discussed in the next unit.

The Laplace transform provides significant advantages in analysing continuous-time physical systems, particularly those described by linear differential equations with constant coefficients. By transforming these equations into the complex frequency domain, it simplifies their mathematical treatment and offers a clear representation of the system's behaviour and characteristics. However, the Laplace transform itself does not possess direct physical significance; its primary role is to convert a time-domain signal into its complex frequency-domain representation. It is useful in many areas, especially in the analysis of signals and communication methods.
Let the function f(t) be defined for all values of t, then

$$f(s) = \int_0^\infty e^{-st} F(t)dt \tag{9.1}$$

Provided the integral exists, it is known as the Laplace Transform of f(t). Mathematically, it is denoted by

$$\mathcal{L}[\mathrm{F}(t)] = \int_0^\infty e^{-st}\,\mathrm{F}(t)dt \tag{9.2}$$

Here, $\mathcal{L}$ is called the Laplace transformation operator. The parameter s is generally a real number, but it can also be complex. The Laplace transform is said to exist if the integral is convergent for some value of S. The functions are generally denoted by the capital letter F(t), G(t), etc., and their Laplace transforms are denoted by the corresponding lower case letters f(s), g(s)... or f(p), g(p).. etc.

The Laplace transforms of some standard functions:

a. $\mathcal{L}[1] = \frac{1}{s} \qquad s > 0$ (9.3)

b. $\mathcal{L}[t^n] = \frac{n!}{s^{n+1}} \qquad where, n = 0,1,2 \ldots$ (9.4)

c. $\mathcal{L}[e^{at}] = \frac{1}{s-a} \qquad s > a$ (9.5)

d. $\mathcal{L}[cosh(at)] = \frac{s}{s^2-a^2} \qquad s^2 > a^2$ (9.6)

e. $\mathcal{L}[sinh(at)] = \frac{a^2}{s^2-a^2} \qquad s^2 > a^2$ (9.7)

(9.8)

f. $\mathcal{L}[sinh(at)] = \frac{a}{s^2+a^2}$ $s > 0$

g. $\mathcal{L}[cos(at)] = \frac{s}{s^2+a^2}$ $s > 0$ (9.9)

Some simple operational properties of the Laplace transforms that are important in applications are as follows.

9.3.1 Properties

The Laplace transform of the linear combination of any number of functions satisfies the linearity property given by

$$\mathcal{L}[c_1F_1(t) + c_2F_2(t)] = c_1\mathcal{L}[F_1(t)] + c_2\mathcal{L}[F_2(t)]$$

c_1 and c_2 are constants.

Proof: If Laplace transform for $F_1(t)$ and $F_2(t)$ are $f_1(s)$ and $f_2(s)$ respectively.

$$f_1(s) = \mathcal{L}[F_1(t)] = \int_0^\infty e^{-st}\, F_1(t)dt$$

$$f_2(s) = \mathcal{L}[F_2(t)] = \int_0^\infty e^{-st}\, F_2(t)dt.$$

For constants c_1 and c_2,

$$\mathcal{L}[c_1F_1(t)] = \int_0^\infty c_1e^{-st}\, F_1(t)dt = c_1\int_0^\infty e^{-st}\, F_1(t)dt = c_1f_1(s)$$

$$Similarly, \qquad \mathcal{L}[c_2F_2(t)] = c_2f_2(s)$$

Adding the two equations,

$$\mathcal{L}[c_1F_1(t)] + \mathcal{L}[c_2F_2(t)] = c_1f_1(s) + c_2f_2(s) = c_1\mathcal{L}[F_1(t)] + c_2\mathcal{L}[F_2(t)] \qquad (9.10)$$

Using the linearity property of the Laplace Transform, a complicated function can be expressed as a sum of simpler functions. Each component can be transformed individually, and the results are then combined, which simplifies calculations. This property is especially useful in solving differential equations involving sums of terms and in analysing electrical or control systems with multiple inputs, where the overall response is obtained by adding the individual responses.

9.3.2 Shifting

The **shifting property** helps in dealing with functions that start after a certain delay or are multiplied by exponential terms. It is widely used to model real-life situations where a signal or force does not act immediately, such as delayed inputs in electrical circuits or control systems.

Two important shifting properties of the Laplace transform are

i. **First Shifting Property (First Translation Theorem):** Let $\mathcal{L}[F(t)] = f(s)$. Then, for any real or complex constant k,

$$\mathcal{L}[e^{kt}F(t)] = f(s - k).$$

Proof: Since e^{kt} is a constant using the Linearity property of LT,

$$\mathcal{L}[e^{kt}F(t)] = e^{kt}\mathcal{L}[F(t)] = e^{kt}\int_0^\infty e^{-st}F(t)dt = \int_0^\infty e^{kt}e^{-st}F(t)dt$$
$$= \int_0^\infty e^{-(s-k)t}F(t)dt = f(s-k)$$

This theorem can also be stated as: If f(s) is the Laplace transform of F(t) and k is any constant (real or complex), then f(s+k) is the Laplace transform of $e^{-kt}F(t)$.

$$f(s) = \mathcal{L}[F(t)] => f(s+k) = \mathcal{L}[e^{-kt}F(t)]$$

ii. Second Shifting Property (Second Translation Theorem): This property arises from the multiplication of the transform by an exponential function. Let $\mathcal{L}[F(t)] = f(s)$. Define

$$G(t) = \begin{cases} F(t-k), & t > k, \\ 0, & 0 < t < k, \end{cases}$$

Where $k > 0$ is a constant. Then

$$\mathcal{L}[G(t)] = e^{-ks}f(s).$$

Proof:

$$\mathcal{L}[G(t)] = \int_0^\infty e^{-st}G(t)dt$$
$$= \int_0^k e^{-st}.0dt + \int_0^\infty e^{-st}F(t-k)dt$$
$$= \int_0^\infty e^{-st}F(t-k)dt.$$

Let $m = t - k$. Then $dt = dm$. When $t = k, m = 0$, and when $t \to \infty, m \to \infty$. Thus,

$$\mathcal{L}[G(t)] = \int_0^\infty e^{-s(m+k)}F(m)dm$$
$$= e^{-ks}\int_0^\infty e^{-sm}F(m)dm$$
$$= e^{-ks}f(s).$$

This theorem can also be written as: If $\mathcal{L}[f(t)] = F(s)$, then

$$\mathcal{L}[u(t-a)f(t-a)] = e^{-as}F(s), \qquad a > 0,$$

Where $u(t-a)$ is the unit step (Heaviside) function. This theorem shows that a delay of a units in the time domain results in multiplication by e^{-as} in the Laplace domain. The **Heaviside Function** is a special function used to represent signals that "turn on" at a specific time. It is represented by u(t-a), where u(t-a)= 0 for t<a and follows some function or u(t-a)=1 for t≥a .

9.3.3 Scaling Property (Change of Scale)

The **scaling property** is used when a function is stretched or compressed in time. It allows us to understand how changing the time scale of a signal affects its Laplace transform. This is useful in signal processing and in solving differential equations involving time-dependent changes. This can be stated as follows,

If $\mathcal{L}[F(t)] = f(s)$, then for $k > 0$,

$$\mathcal{L}[F(kt)] = \frac{1}{k} f\left(\frac{s}{k}\right).$$

Proof: Since $\mathcal{L}[F(t)] = f(s)$, we have

$$\mathcal{L}[F(kt)] = \int_0^\infty e^{-st}\, F(kt)dt.$$

Let $x = kt$. Then $dt = \frac{dx}{k}$. When $t = 0, x = 0$, and when $t \to \infty, x \to \infty$. Hence,

$$\mathcal{L}[F(kt)] = \int_0^\infty e^{-s(x/k)}\, F(x)\frac{dx}{k} = \frac{1}{k}\int_0^\infty e^{-(s/k)x}\, F(x)dx.$$

Recognizing the integral as the Laplace transform of $F(x)$ evaluated at s/k, we obtain

$$\mathcal{L}[F(kt)] = \frac{1}{k} f\left(\frac{s}{k}\right).$$

Laplace Transform of Derivatives

Here, we examine how the Laplace Transform works with the fundamental operations of calculus, namely differentiation and integration. The primary focus is on the generalised result, which connects the transform of a function's derivative with the transform of the function itself. This relationship makes it possible to convert many initial value problems into simpler algebraic equations.

If $\mathcal{L}[F(t)] = f(s)$, then

$$\mathcal{L}\left[F^{(n)}(t)\right] = s^n f(s) - s^{n-1}F(0) - s^{n-2}F'(0) - \cdots - F^{(n-1)}(0),$$

Where $F^{(n)}(t)$ stands for $\frac{d^n}{dt^n}F(t)$.

Proof. Let $\mathcal{L}[F(t)] = f(s)$, then

$$\mathcal{L}[F'(\mathrm{t})] = \int_0^\infty e^{-st}\frac{dF}{dt}dt$$

$$= [e^{-st}F(t)]_0^\infty - \int_0^\infty e^{-st}\,(-s)F(t)dt$$

$$= -F(0) + sf(s).$$

$$\mathcal{L}[F'(\mathrm{t})] = sf(s) - F(0).$$

In view of this,

$$\mathcal{L}[G'(t)] = sg(s) - G(0) = s\mathcal{L}[G(t)] - G(0).$$

If $G(t) = F'(t)$, we get

$$\mathcal{L}[F''(t)] = s\mathcal{L}[F'(t)] - F'(0)$$

$$= s[sf(s) - F(0)] - F'(0)$$

$$= s^2 f(s) - sF(0) - F'(0).$$

Similarly, for other orders of derivatives

$$\mathcal{L}[F^{(4)}(t)] = s^4 f(s) - s^3 F(0) - s^2 F'(0) - sF''(0) - F^{(3)}(0).$$

$$\mathcal{L}[F^{(5)}(t)] = s^5 f(s) - s^4 F(0) - s^3 F'(0) - s^2 F''(0) - sF^{(3)}(0) - F^{(4)}(0).$$

and so on..., such that we can generalise and write

$$\mathcal{L}[F^{(n)}(t)] = s^n f(s) - s^{n-1} F(0) - s^{n-2} F'(0) - \cdots - F^{(n-1)}(0),$$

Laplace Theorem of Integral

If the function F(t) is sectionally continuous and of exponential order with $\mathcal{L}[F(t)] = f(s)$, then (9.11)

$$\mathcal{L}\left[\int_0^t F(\tau)d\tau\right] = \frac{1}{s} f(s)$$

Or else if $\mathcal{L}[F(t)] = f(s)$, then

$$\frac{1}{s} f(s) = \mathcal{L}\left[\int_0^t F(\tau)d\tau\right].$$

Proof. Let $\mathcal{L}[F(t)] = f(s)$ and

$$G(t) = \int_0^t F(\tau)d\tau. \quad \text{(A)}$$

We are to prove that $\mathcal{L}[G(t)] = \frac{1}{s} f(s)$.

From equation (A) it is clear that,

$$G(0) = \int_0^0 F(\tau)d\tau = 0$$

$$G'(t) = \frac{d}{dt}\left(\int_0^t F(\tau)d\tau\right)$$

By the theorem of derivatives,

$$\mathcal{L}[G'(t)] = s\mathcal{L}[G(t)] - G(0)$$

i.e.,

$$\mathcal{L}[F(t)] = s\mathcal{L}[G(t)] - 0 = s\mathcal{L}[G(t)]$$

Therefore,

$$\mathcal{L}\left[\int_0^t F(\tau)d\tau\right] = \frac{1}{s} f(s).$$

9.4 DERIVATIVES AND INTEGRALS OF LAPLACE'S TRANSFORM

Derivatives of Laplace Transforms

If the $function\ F(t)$ is sectionally continuous and of exponential order $e^{\alpha t}$, then

$$f^{(n)}(s) = \mathcal{L}[(-t)^n F(t)], \quad n = 1,2,\ldots,\ s > \alpha$$

where

$$f(s) = \mathcal{L}[F(t)].$$

In particular, for $n = 1$,

$$\mathcal{L}[-tF(t)] = f'(s) = \frac{df}{ds}.$$

Example: The Laplace transformation,

$$\mathcal{L}[-t\ sin\ kt] = \frac{d}{ds}\left(\frac{k}{k^2 + s^2}\right) = \frac{-2ks}{(k^2 + s^2)^2}.$$

Hence,

$$\mathcal{L}[t\ sin\ kt] = \frac{2ks}{(k^2 + s^2)^2}.$$

Similarly,

$$\mathcal{L}[-t\ cos\ kt] = \frac{d}{ds}\left(\frac{s}{k^2 + s^2}\right) = \frac{k^2 - s^2}{(k^2 + s^2)^2}.$$

Thus,

$$\mathcal{L}[t\cos kt] = \frac{s^2 - k^2}{(k^2 + s^2)^2}.$$

Integration of Laplace Transforms

If $f(s) = \mathcal{L}[F(t)]$, then

$$\int_s^\infty f(x)\,dx = \mathcal{L}\left[\frac{F(t)}{t}\right].$$

Example: Since

$$\mathcal{L}[\sin t] = \frac{1}{s^2 + 1}, \qquad s > 0,$$

By the integration property of Laplace transforms,

$$\mathcal{L}\left[\frac{\sin t}{t}\right] = \int_s^\infty \frac{dx}{1 + x^2}$$

Evaluating the integral,

$$= [\tan^{-1}]_s^\infty = \frac{\pi}{2} - \tan^{-1} s.$$

Since

$$\frac{\pi}{2} - \tan^{-1} s = \cot^{-1} s,$$

we obtain

$$\mathcal{L}\left[\frac{\sin t}{t}\right] = \cot^{-1} s.$$

If

$$\mathrm{Si}(t) = \int_0^t \frac{\sin\tau}{\tau}\,d\tau,$$

Then using

$$\mathcal{L}\left[\int_0^t F(\tau)d\tau\right] = \frac{1}{s}\mathcal{L}[F(t)],$$

We get

$$\mathcal{L}[\mathrm{Si}(t)] = \frac{1}{s}\cot^{-1} s = \frac{1}{s}\tan^{-1}\frac{1}{s}.$$

Another useful result:

$$\mathcal{L}\left[\frac{e^{-at}-e^{-bt}}{t}\right]=\log\left(\frac{s+b}{s+a}\right),\quad s>-a,\qquad s>-b.$$

Special case: if $a=0, b=1$,

$$\mathcal{L}\left[\frac{1-e^{-t}}{t}\right]=\log\left(1+\frac{1}{s}\right).$$

9.5 CONVOLUTION THEOREM

The theorem is formulated using the *convolution* of two functions $F(t)$ and $G(t)$, defined by

$$F(t)*G(t)=\int_0^t F(\tau)G(t-\tau)d\tau. \tag{9.12}$$

This operation is commutative, that is,

$$F(t)*G(t)=G(t)*F(t). \tag{9.13}$$

Suppose that $f(s)$ and $g(s)$ are the Laplace transforms of $F(t)$ and $G(t)$, respectively. If $F(t)$ and $G(t)$ are piecewise continuous on every finite interval $0\le t\le T$ and are of exponential order $e^{\alpha t}$ as $t\to\infty$, then for $s>\alpha$ the Laplace transform of their convolution exists and satisfies

$$\mathcal{L}[F(t)^*G(t)]=f(s)g(s). \tag{9.14}$$

Consequently, the inverse Laplace transform (refer to section 9.6) of the product $f(s)g(s)$ is

$$\mathcal{L}^{-1}[f(s)g(s)]=F(t)^*G(t). \tag{9.15}$$

Proof: Begin with the definitions of the Laplace transforms:

$$f(s)=\mathcal{L}[F(t)]=\int_0^\infty e^{-st}F(t)dt\,, \tag{9.16}$$

$$g(s)=\mathcal{L}[G(t)]=\int_0^\infty e^{-st}G(t)dt. \tag{9.17}$$

Using the shifting property of the Laplace transform, we write

$$e^{-s\tau}f(s)=\int_0^\infty e^{-st}\,F(t-\tau)U(t-\tau)dt, \tag{9.18}$$

Where $U(t-\tau)$ denotes the unit step function given by

$$U(t-\tau)=\begin{cases}0, \text{if } t>\tau,\\ 1 \text{ if } t\ge\tau\end{cases}$$

Substituting this result into the expression for $f(s)g(s)$ and changing the order of integration, we obtain

$$f(s)g(s)=\int_0^\infty e^{-st}\left[\int_0^t G(\tau)F(t-\tau)d\tau\right]dt \tag{9.19}$$

$$= \mathcal{L}\left[\int_0^t G(\tau)F(t-\tau)d\tau\right]. \tag{9.20}$$

Recognizing the inner integral as the convolution of $G(t)$ and $F(t)$, we have

$$f(s)g(s) = \mathcal{L}[G(t)^*F(t)]. \tag{9.21}$$

Since convolution is commutative,

$$\mathcal{L}[F(t)^*G(t)] = f(s)g(s). \tag{9.22}$$

Taking the inverse Laplace transform of both sides yields

$$\mathcal{L}^{-1}[f(s)g(s)] = F(t)^*G(t), \tag{9.23}$$

which completes the proof.

The convolution theorem simplifies complex problems by converting difficult inverse transforms into manageable integrals, making it a powerful tool in mathematics, physics, and engineering. It finds application in solving linear differential equations, finding inverse Laplace transforms, and analysing systems in engineering (input–output response).

9.6 INVERSE LAPLACE TRANSFORMATION

As mentioned above, the inverse Laplace transform is defined as follows:
If the Laplace transform of a function F(t) is f(s) such that $f(s) = L[F(t)]$, then F(t) can be called as the Inverse Laplace transform of *f*(s)

$$F(t) = \mathcal{L}^{-1}[f(s)]$$

Examples:

i. Find $\mathcal{L}^{-1}\left[\frac{1}{s-a}\right]$.

Solution: $\mathcal{L}[e^{at}] = \frac{1}{s-a}$, $s > a$, therefore, $\mathcal{L}^{-1}\left[\frac{1}{s-a}\right] = e^{at}$.

ii. A few well-known results $\mathcal{L}^{-1}\left[\frac{s}{s^2+a^2}\right] = \cos(at)$

$$\mathcal{L}^{-1}\left[\frac{a}{s^2+a^2}\right] = \sin(at)$$

9.7 SOLUTION OF DIFFERENTIAL EQUATIONS

The Laplace Transform is a powerful mathematical tool that converts a function from the time domain into the s-domain. This transformation significantly simplifies the process of solving problems involving differential equations. It is widely used to obtain solutions of linear differential equations, both ordinary and partial. In addition, it helps in solving systems of simultaneous differential equations, integral equations, and even in evaluating certain definite integrals. Ordinary and partial differential equations are used to describe how physical quantities change with time, for example, electric current in a circuit, vibrations of a membrane, or heat flow through a conductor. These equations are usually accompanied by initial conditions that specify the state of the system at time t=0. The Laplace transform provides an efficient method for solving such problems

by converting differential equations into simpler algebraic equations. Once solved in the s-domain, the solution can be transformed back to obtain the result in the time domain.

Application of Laplace Transform to Differential Equations
Consider the second-order linear differential equation

$$a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + y = G(t), \tag{9.24}$$

where a and b may either be constants or functions of t.

Case I: a and b are constants
First assume that a and b are constants. We apply the Laplace transform to both sides of the equation.
Using the Laplace transform formula for derivatives:

$$\mathcal{L}\left[y^{(n)}(t)\right] = s^nY(s) - s^{n-1}y(0) - s^{n-2}y'(0) - \cdots - y^{(n-1)}(0),$$

where

$$\mathcal{L}[y(t)] = Y(s).$$

Taking the Laplace transform of the differential equation gives:

$$a\mathcal{L}[y''(t)] + b\mathcal{L}[y'(t)] + \mathcal{L}[y(t)] = \mathcal{L}[G(t)].$$

Substituting the derivative formulas:

$$a(s^2Y(s) - sy(0) - y'(0)) + b(sY(s) - y(0)) + Y(s) = G(s).$$

We now collect terms involving $Y(s)$:

$$Y(s)(as^2 + bs + 1) - a(sy(0) + y'(0)) - by(0) = G(s).$$

Solving for $Y(s)$:

$$Y(s) = \frac{G(s) + a(sy(0) + y'(0)) + by(0)}{as^2 + bs + 1}.$$

Finally, the solution $y(t)$ is obtained by taking the inverse Laplace transform:

$$y(t) = \mathcal{L}^{-1}[Y(s)].$$

Case II: a and b are functions of t:
Suppose now that the coefficients depend on t. A typical example is the Cauchy–Euler type equation:

$$t^2\frac{d^2y}{dt^2} + t\frac{dy}{dt} + y = G(t). \tag{9.25}$$

To handle terms involving powers of t, we use the Laplace transform property:

$$\mathcal{L}[t^m f(t)] = (-1)^m \frac{d^m}{ds^m}[F(s)],$$

where $F(s) = \mathcal{L}[f(t)]$.
Applying this rule:

$$\mathcal{L}[t^2 y''(t)] = \frac{d^2}{ds^2}[s^2 Y(s) - sy(0) - y'(0)],$$

$$\mathcal{L}[ty'(t)] = -\frac{d}{ds}[sY(s) - y(0)].$$

Taking the Laplace transform of the entire equation yields:

$$\frac{d^2}{ds^2}[s^2 Y(s) - sy(0) - y'(0)] - \frac{d}{ds}[sY(s) - y(0)] + Y(s) = G(s).$$

This equation is now a differential equation in the variable s. After solving for $Y(s)$, we take the inverse Laplace transform to obtain $y(t)$.

CHECK YOUR PROGRESS

Q1: Find the Laplace transform of $F(t) = 1 + sin(2t)$

Q2: Find the Laplace transform of $f(t)$ defined as

$$f(t) = \begin{cases} \frac{t}{k}, & 0 < t < k, \\ 1 & t > k. \end{cases}$$

Q3: Evaluate $\mathcal{L}[F(t)]$ for

$$F(t) = \begin{cases} cos(t - 2\pi/3), & t > 2\pi/3, \\ 0, & t < 2\pi/3 \end{cases}$$

Q4: Evaluate $\mathcal{L}[cos(kt)]$.

Q5: If

$$f(s) = \mathcal{L}[F(t)] = \int_0^\infty e^{-st} F(t)dt,$$

where $F(t)$ is piecewise continuous and of exponential order. Then, prove that

$$\mathcal{L}[tF(t)] = -\frac{d}{ds} f(s).$$

More generally,

$$\mathcal{L}[t^n F(t)] = (-1)^n \frac{d^n}{ds^n} f(s), \qquad n = 1,2,3,\ldots$$

Q6: If

$$f(s) = \mathcal{L}[F(t)] = \int_0^\infty e^{-st} F(t)dt,$$

Where $F(t)$ is piecewise continuous and of exponential order. Then, prove that

$$\boxed{\int_0^\infty f(x)dx = \mathcal{L}\left[\frac{F(t)}{t}\right]}$$

9.8 LET US SUM UP

- **Definition of Laplace Transform**
  The Laplace transform converts a time-domain function into a function in the complex frequency domain:
  $$\mathcal{L}[f(t)] = F(s) = \int_0^\infty e^{-st} f(t)dt.$$
- **Linearity and Basic Transforms** The Laplace transform is linear:
  $$\mathcal{L}[af(t) + bg(t)] = aF(s) + bG(s).$$
- **Shifting and Scaling Properties** First Shifting Theorem:
  $$\mathcal{L}[e^{at}f(t)] = F(s-a).$$
  Second Shifting Theorem:
  $$\mathcal{L}[f(t-a)u(t-a)] = e^{-as}F(s)$$
  Scaling Property:
  $$\mathcal{L}[f(at)] = \frac{1}{a}F\left(\frac{s}{a}\right)$$
- **Derivative Property**
  Differentiation in the time domain corresponds to multiplication by s in the Laplace domain:
  $$\mathcal{L}[f'(t)] = sF(s) - f(0),$$
  $$\mathcal{L}[f''(t)] = s^2F(s) - sf(0) - f'(0).$$
- **Inverse Transform and Differential Equations**
  The inverse Laplace transform recovers the original function:
  $$f(t) = \mathcal{L}^{-1}[F(s)].$$
  Laplace transforms convert linear differential equations into algebraic equations in s, making them easier to solve.

9.9 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: Here

$$\mathcal{L}[1 + \sin(2t)] = \mathcal{L}[1] + \mathcal{L}[\sin(2t)] = \frac{1}{s} + \frac{2}{s^2+4}$$

Answer to Q2:

$$\mathcal{L}[f(t)] = \int_0^k \frac{t}{k}e^{-st}dt + \int_k^\infty e^{-st}dt.$$

$$= \frac{1}{k}\int_0^k te^{-st}\,dt + \int_k^\infty e^{-st}dt.$$

Using integration by parts,

$$\int te^{-st}dt = \frac{te^{-st}}{-s} - \int \frac{e^{-st}}{-s}dt = \frac{te^{-st}}{-s} - \frac{e^{-st}}{s^2}.$$

Hence,

$$\mathcal{L}[f(t)] = \frac{1}{k}\left[\frac{ke^{-ks}}{-s} - \frac{e^{-ks}}{s^2} + \frac{1}{s^2}\right] + \frac{e^{-ks}}{s}.$$

$$= -\frac{e^{-ks}}{s} - \frac{e^{-ks}}{ks^2} + \frac{1}{ks^2} + \frac{e^{-ks}}{s}.$$

$$= \frac{1}{ks^2}\left(1 - e^{-ks}\right).$$

Answer to Q3: Write the function using the unit step function i.e. $u(t-a)$ and $a = 2\pi/3$:

$$F(t) = u\left(t - \frac{2\pi}{3}\right)\cos\left(t - \frac{2\pi}{3}\right).$$

Using the Second Shifting Property,

$$\mathcal{L}[u(t-a)f(t-a)] = e^{-as}\mathcal{L}[f(t)].$$

Here, $a = \frac{2\pi}{3}$ and $f(t) = \cos t$.

Since

$$\mathcal{L}[\cos t] = \frac{s}{s^2+1},$$

We obtain

$$\mathcal{L}[F(t)] = e^{-\frac{2\pi s}{3}}\frac{s}{s^2+1}.$$

$$\mathcal{L}[F(t)] = \frac{s}{s^2+1}e^{-\frac{2\pi s}{3}}$$

Answer to Q4: Using the change of scale property, we get

$$\mathcal{L}[F(kt] = \frac{1}{k}f(s/k)$$

We have

$$\mathcal{L}[\cos(t)] = \frac{s}{s^2+1}$$

$$\mathcal{L}[\cos(kt)] = \frac{1}{k}\mathcal{L}[\cos(s/k)] = \frac{1}{k}\frac{s/k}{(s/k)^2+1} = \frac{s}{(k^2+s^2)}$$

Answer to Q5: Starting from the definition,

$$f(s) = \int_0^\infty e^{-st}F(t)dt.$$

Differentiate with respect to s:

$$\frac{df}{ds} = \frac{d}{ds}\int_0^\infty e^{-st}F(t)dt.$$

Interchanging differentiation and integration,

$$= \int_0^\infty \frac{\partial}{\partial s}\left(e^{-st}F(t)\right)dt.$$

Since

$$\frac{\partial}{\partial s}(e^{-st}) = -te^{-st},$$

We obtain

$$\frac{df}{ds} = \frac{d}{ds}\int_0^\infty (-te^{-st})F(t)dt.$$
$$= -\int_0^\infty e^{-st}tF(t)dt.$$

Recognizing the Laplace transform,

$$\frac{df}{ds} = -\mathcal{L}[tF(t)].$$

Hence,

$$\boxed{\mathcal{L}[tF(t)] = -\frac{df}{ds}}$$

General Case Differentiating repeatedly, we will get

$$\frac{d^n f}{ds^n} = (-1)^n \int_0^\infty t^n e^{-st}F(t)dt.$$

Therefore,

$$\boxed{\mathcal{L}[t^n F(t)] = (-1)^n \frac{d^n}{ds^n} f(s)}$$

Answer to Q6: From the definition of the Laplace transform,

$$f(s) = \int_0^\infty e^{-st}F(t)dt.$$

Integrate both sides with respect to s from s to ∞:

$$\int_s^\infty f(x)dx = \int_s^\infty \left[\int_0^\infty e^{-xt}F(t)dt\right]dx.$$

Interchanging the order of integration (since the variables are independent),

$$= \int_0^\infty F(t)\left[\int_s^\infty e^{-xt}dx\right]dt.$$

Now evaluate the inner integral:

$$\int_s^\infty e^{-xt}dx = \left[\frac{-1}{t}e^{-xt}\right]_s^\infty = \frac{e^{-xt}}{t}.$$

Substituting,

$$\int_s^\infty f(x)dx = \int_0^\infty e^{-st}\frac{F(t)}{t}dt.$$

Recognizing this as a Laplace transform,

$$\boxed{\int_s^\infty f(x)dx = \mathcal{L}\left[\frac{F(t)}{t}\right]}$$

9.10 FURTHER READING

- Dass, H.K., & Verma, R. (2013). Mathematical physics. S.Chand Publishing.
- Rajput, B.S. (2008). Mathematical physics (20th ed.). Pragati Prakashan.
- Riley, K.F., Hobson, M.P., & Bence, S.J. (2006). Mathematical methods for physics and engineering (3rd ed.). Cambridge University Press.

9.11 MODEL QUESTIONS

1. Find the Laplace transform of (i) t^2e^{-at} and (ii) $\frac{1}{a^2}(1-cos(at))$.
2. Evaluate $\mathcal{L}[\sin(at)-at\cos(at)]$.
3. Evaluate $\mathcal{L}[F(t)]$ where, $\mathcal{L}^{-1}$
4. $F(t)=\begin{cases}0, & 0<t<1\\ 1, & 1<t<2\\ 2, & 2<t\end{cases}$
5. $F(t)=\begin{cases}\sin(t-\pi/3), & t>\pi/3\\ 0, & t<\pi/3\end{cases}$
6. Solve $x''(t)+4x'(t)+4x(t)=4e^{-2t}$ so that $x(0)=-1, x'(0)=4$
7. Find $\mathcal{L}^{-1}[f(s)g(s)]$, where $\mathcal{L}^{-1}[f(s)]=F(t)$ and $\mathcal{L}^{-1}[g(s)]=G(t)$ with
8. $F(t)=\begin{cases}1, & 0<t\le T\\ 0, & t>T\end{cases}$
9. and
10. $G(t)=\begin{cases}\frac{t}{T}, & 0<t\le T\\ 2-\frac{t}{T}, & T<t\le 2T\\ 0, & t>2T\end{cases}$

*** *** ***

UNIT 10: FOURIER INTEGRALS

UNIT STRUCTURE

10.1 Learning Objectives
10.2 Introduction
10.3 Preliminaries of Special Functions
    10.3.1 Gamma
    10.3.2 Error
    10.3.3 Bessel Functions
10.4 Fourier Integral Representations
10.5 Proof of Fourier Integral Theorem
10.6 Let Us Sum Up
10.7 Answers to Check Your Progress
10.8 Further Reading
10.9 Model Questions

10.1 LEARNING OBJECTIVES

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

- explain the concept of Fourier integrals and their necessity for non-periodic functions
- describe the role of special functions such as Gamma, error and Bessel functions
- derive and apply Fourier integral representations of functions
- analyse the Fourier integral theorem and its mathematical structure
- evaluate integrals and solve problems using Fourier integral techniques.

10.2 INTRODUCTION

In the previous unit, we studied special functions and their mathematical properties as solutions of differential equations arising in many physical problems. We discussed important classes of special functions and their properties such as orthogonality, recurrence relations, and generating functions, which provide powerful mathematical tools for solving boundary value problems in physics. These concepts form an important foundation for further analytical methods in mathematical physics.

In this unit, we shall study Fourier integrals and Fourier transforms, which extend the idea of Fourier series to non-periodic functions defined over infinite intervals. Fourier analysis, developed from the work of Joseph Fourier, provides a powerful method for expressing functions in terms of continuous sinusoidal components. The Fourier transform has wide applications in solving differential equations, wave propagation, signal processing, quantum mechanics, and heat conduction problems. In this unit, we shall discuss the Fourier integral representation of functions, some special functions useful in evaluating Fourier integrals, and the Fourier integral theorem.

Definition of Fourier Series with Examples

Let a function $g(x)$ be defined in the interval $(-a, a)$ and suppose it is extended outside this interval in such a way that

$$g(x + 2a) = g(x),$$

that is, $g(x)$ is periodic with period $2a$. The corresponding Fourier representation of $g(x)$ may be written as

$$g(x) = \frac{A_0}{2} + \sum_{m=1}^{\infty} A_m \cos\frac{m\pi x}{a} + B_m \sin\frac{m\pi x}{a}. \tag{10.1}$$

The coefficients A_m and B_m are determined by

$$A_m = \frac{1}{a}\int_{-a}^{a} g(x)\cos\left(\frac{m\pi x}{a}\right)dx, \qquad m = 0,1,2,\ldots \tag{10.2}$$

$$B_m = \frac{1}{a}\int_{-a}^{a} g(x)\sin\left(\frac{m\pi x}{a}\right)dx, \qquad m = 0,1,2,\ldots \tag{10.3}$$

The above expressions give the Fourier coefficients for a function defined over the symmetric interval $(-a, a)$.

If $g(x)$ possesses the period $2a$, the coefficients may also be evaluated over any interval of length $2a$. Thus, for an arbitrary real constant d,

$$A_m = \frac{1}{a}\int_{d}^{d+2a} g(x)\cos\left(\frac{m\pi x}{a}\right)dx, \tag{10.4}$$

$$B_m = \frac{1}{a}\int_{d}^{d+2a} g(x)\sin\left(\frac{m\pi x}{a}\right)dx, \qquad m = 0,1,2,\ldots \tag{10.5}$$

Clearly, choosing $d = -a$ reduces these formulas to the earlier expressions. The constant term in the series is given by

$$\frac{A_0}{2} = \frac{1}{2a}\int_{-a}^{a} g(x)\,dx,$$

which represents the average value of the function over one complete period.

In the particular case when $a = \pi$, the period of the function becomes 2π and the Fourier expansion assumes a simpler form.

It is important to note that the series written above is the Fourier series associated with $g(x)$. However, the existence of this series does not automatically guarantee convergence, nor does it ensure that the series necessarily converges to the original function for all values of x.

Example 1 Fourier Series of $f(x) = x$ in $(-\pi, \pi)$

Consider the function

$$f(x) = x, \quad -\pi < x < \pi$$

Since $f(-x) = -f(x)$, the function is *odd*. Therefore

$$A_n = 0.$$

The coefficients B_n are given by

$$B_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)\,dx.$$

Since the integrand is even,

$$B_n = \frac{2}{\pi}\int_{0}^{\pi} x\sin(nx)\,dx.$$

Evaluating the integral using integration by parts, we obtain

$$B_n = \frac{2(-1)^{n+1}}{n}.$$

Hence the Fourier series becomes

$$x = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(nx).$$

Example 2

Fourier Series of $f(x) = x^2$ in $(-\pi, \pi)$

Let

$$f(x) = x^2, \quad -\pi < x < \pi.$$

Since the function is even,

$$B_n = 0.$$

The coefficient A_0 is

$$A_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\,dx.$$

Since the integrand is even,

$$A_0 = \frac{2}{\pi}\int_{0}^{\pi} x^2\,dx = \frac{2\pi^2}{3}.$$

Thus

$$\frac{A_0}{2} = \frac{\pi^2}{3}.$$

The coefficients A_n are

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos(nx)\,dx$$

which gives

$$A_n = \frac{4(-1)^n}{n^2}.$$

Therefore the Fourier series is

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx).$$

Example 3

Fourier Series of $f(x) = |x|$ in $(-\pi, \pi)$

Consider

$$f(x) = |x|, \quad -\pi < x < \pi.$$

The function is even, hence

$$B_n = 0.$$

The coefficients A_0 is

$$A_0 = \frac{2}{\pi}\int_0^{\pi} x\,dx = \pi.$$

Thus

$$\frac{A_0}{2} = \frac{\pi}{2}.$$

The coefficients A_n are

$$A_n = \frac{2}{\pi}\int_0^{\pi} x\cos(nx)\;dx.$$

Evaluating the integral gives

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

Hence the Fourier series becomes

$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\left(\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right).$$

Example 4

Constant Function

Let

$$f(x) = 1, \qquad -L < x < L.$$

Then

$$A_0 = \frac{1}{L}\int_{-L}^{L} 1\,dx = 2.$$

Thus

$$\frac{A_0}{2} = 1.$$

Also

$$A_n = 0, \quad b_n = 0.$$

Therefore the Fourier series reduces to

$$f(x) = 1.$$

Fourier integrals frequently involve definite integrals that cannot always be evaluated in terms of elementary functions. Such integrals naturally lead to special functions like the Gamma function, Error function and Bessel functions. Before discussing Fourier integral representations, it is therefore useful to review these special functions and their relevance in evaluating Fourier integrals.

10.3 PRELIMINARIES OF SPECIAL FUNCTIONS

In many problems of mathematical physics, phenomena such as heat conduction, electrostatic potential and wave motion are governed by partial differential equations together with appropriate boundary conditions. Fourier series provide an important method for solving many such boundary value problems. During the evaluation of Fourier integrals and Fourier transforms, however, several definite integrals arise whose solutions cannot be expressed in terms of elementary functions. Such integrals naturally lead to special functions, particularly the Gamma function, Error function and Bessel functions. Therefore, before studying Fourier integrals, it is useful to review these special functions and their role in evaluating Fourier integrals.

10.3.1 Gamma

The Gamma function frequently appears in the evaluation of definite integrals and plays an important role in Fourier integrals, probability theory and statistical mechanics.
The Gamma function is defined by the integral

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx, \quad s > 0.$$

For positive integers n, the Gamma function satisfies

$$\Gamma(n) = (n-1)!$$

Thus, the Gamma function can be regarded as a continuous extension of the factorial function.

$$\Gamma(n+1) = n!, \quad n = 1,2,3,\ldots \qquad (10.6)$$

For this reason, $\Gamma(n)$ is sometimes called the factorial function. (10.6)

For Example**:** $\Gamma(3) = 2! = 2, \quad \Gamma(4) = 3! = 6, \quad \frac{\Gamma(7)}{\Gamma(4)} = \frac{6!}{3!} = 120.$

Asymptotic Formula for Γ_n

When n becomes large, the direct evaluation of Γ_n is computationally inconvenient. In such situations, an approximate expression of considerable importance is given (10.7)

$$\Gamma(n+1) = \sqrt{2\pi n}\, n^n e^{-n} e^{\frac{\theta}{12(n+1)}}, \quad 0 < \theta < 1. \qquad (10.7)$$

For sufficiently large n, the exponential factor involving θ is very close to unity and may be neglected for practical calculations. Thus, for integer values of n, we obtain the approximation

$$n! \sim \sqrt{2\pi n}\; n^n e^{-n} \qquad (10.8)$$

The symbol ~ indicates that the two sides are asymptotically equivalent as n→ (10.8) result is commonly known as Stirling's approximation, which provides an approximate value of the factorial function for large n.

Example 5

Use Stirling's approximation to estimate $5!$.

$$5! \approx \sqrt{2\pi \cdot 5}\left(\frac{5}{e}\right)^5.$$

$$5! \approx \sqrt{10\pi}\left(\frac{5}{e}\right)^5 \approx 118.02.$$

(Exact value: $5! = 120$)

Example 6

Approximate 10! using Stirling's formula.

$$10! \approx \sqrt{2\pi \cdot 10}\left(\frac{10}{e}\right)^{10}.$$

$$10! \approx \sqrt{20\pi}\left(\frac{5}{e}\right)^5 \approx 3.60 \times 10^6.$$

(Exact value: 10! = 3628800)

Example 7

Evaluate

$$\frac{\Gamma(7)}{2\Gamma(4)\Gamma(3)}.$$

Solution:

Using $\Gamma(n+1) = n!$,

$$\Gamma(7) = 6!,\ \Gamma(4) = 3!,\ \Gamma(3) = 2!.$$

Thus,

$$\frac{\Gamma(7)}{2\Gamma(4)\Gamma(3)} = \frac{6!}{2.3!.2!} = \frac{720}{2.6.2} = \frac{720}{24} = 30.$$

Example 8

Evaluate

$$\frac{\Gamma(3)\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{9}{2}\right)}.$$

Solution:

We use

$$\Gamma(3) = 2!,\ \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Now,

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}.$$

Also,

$$\Gamma\left(\frac{9}{2}\right) = \frac{7}{2}.\frac{5}{2}.\frac{3}{2}.\frac{1}{2}\sqrt{\pi} = \frac{105}{16}\sqrt{\pi}.$$

Thus,

$$\frac{\Gamma(3)\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{9}{2}\right)} = \frac{2.\frac{\sqrt{\pi}}{2}}{\frac{105}{16}\sqrt{\pi}} = \frac{\sqrt{\pi}}{\frac{105}{16}\sqrt{\pi}} = \frac{16}{105}.$$

Example 9

Evaluate

$$\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{5}{2}\right).$$

Solution:

Using

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \qquad \Gamma = \left(\frac{3}{2}\right) = \frac{1}{2}\pi,$$

$$\Gamma\left(\frac{1}{2}\right) = \frac{3}{2}.\frac{1}{2}\sqrt{\pi} = \frac{3}{4}\sqrt{\pi}.$$

Thus,

$$\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{5}{2}\right) = \sqrt{\pi}.\frac{1}{2}\sqrt{\pi}.\frac{3}{4}\sqrt{\pi} = \frac{3}{8}\pi^{3/2}.$$

Example10

Evaluate

$$\int_0^\infty x^4 e^{-x} dx.$$

Solution:
Using the definition of Gamma function,

$$\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx.$$

Here $n-1=4 \Rightarrow n=5$, so

$$\int_0^\infty x^4 e^{-x}dx = \Gamma(5) = 4! = 24.$$

Example11
Evaluate

$$\int_0^\infty x^2 e^{-2x^2}dx$$

Solution:
Let $t = 2x^2 \Rightarrow dt = 4xdx$.
Then,

$$x^2dx = \frac{t^{1/2}}{4\sqrt{2}}dt.$$

Thus the integral becomes

$$\int_0^\infty x^2 e^{-2x^2}dx = \frac{1}{4\sqrt{2}}\int_0^\infty t^{1/2}e^{-t}\,dt.$$

Using Gamma function,

$$\int_0^\infty t^{1/2}e^{-t}\,dt = \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}.$$

Hence,

$$\int_0^\infty x^2 e^{-2x^2}\,dx = \frac{1}{4\sqrt{\pi}}.\frac{1}{2}\sqrt{\pi} = \frac{\sqrt{\pi}}{8\sqrt{2}}$$

Note: The Gamma function is frequently used to evaluate definite integrals of the form $\int_0^\infty x^n\, e^{-x^2}\,dx$, which commonly arise while evaluating Fourier cosine transforms and Fourier integral representations of exponentially decaying functions.

10.3.2 Error

The normal (Gaussian) distribution plays a fundamental role in mathematical physics and probability theory. It was first introduced by de Moivre and later developed by Laplace in the study of statistical errors. The general form of the normal distribution is given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],\ -\infty < x < \infty, \sigma > 0. \qquad (10.9)$$

For the special case $\mu = 0$ and $\sigma = 1$, we obtain the standard normal distribution

$$p(u) = \frac{1}{\sqrt{2\pi}}exp\left(-\frac{u^2}{2}\right),\ -\infty < u < -\infty. \qquad (10.10)$$

In the study of this distribution, one encounters integrals of form

$$I = \int_{-\alpha}^{\alpha}\frac{1}{\sqrt{2\pi}}exp\left(-\frac{u^2}{2}\right)du \qquad (10.11)$$

.

By a suitable change of variable, this integral can be reduced to

$$I = \frac{1}{\sqrt{\pi}} \int e^{-u^2}\, du. \qquad (10.12)$$

The integral $e^{-u^2}du$ cannot be expressed in terms of elementary functions. This leads to the definition of a special function, called the *error function*, defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du. \qquad (10.13)$$

Using the known result from the Gamma function,

$$\int_{-\infty}^{\infty} e^{-u^2}\, du = \sqrt{\pi}, \qquad (10.14)$$

we obtain

$$\int_0^{\infty} e^{-u^2}\, du = \frac{\sqrt{\pi}}{2}. \qquad (10.15)$$

Thus, the total area under the curve e^{-u^2} over the interval $[0, \infty)$ is finite. This allows us to write

$$\frac{2}{\sqrt{\pi}} \left(\int_0^x e^{-u^2}\, du + \int_x^{\infty} e^{-u^2}\, du \right) = 1 \qquad (10.16)$$

This naturally leads to the definition of the complementary error function, denoted by $\operatorname{erfc}(x)$, as

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-u^2}\, du. \qquad (10.17)$$

Hence, the error function and complementary error function are related by

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x). \qquad (10.18)$$

The functions $\operatorname{erf}(x)$ and $\operatorname{erfc}(x)$ provide a convenient way of expressing integrals involving the Gaussian function and are widely used in problems of heat conduction, diffusion and probability theory.

The function e^{-u^2} is Gaussian in nature. So, the region under the Gaussian curve represents the normalised area from 0 to x. Thus, it measures how much of a quantity has accumulated up to a given point.

The error function physically describes the gradual spreading of a quantity such as heat, particles or probability from an initially localised region. Instead of a sharp boundary, the distribution becomes smooth over time.

1. Applications in Physics

• Signal Processing

It plays an important role in signal processing in problems involving noise, filtering and signal transitions. Because real-world signals and noise processes generally follow Gaussian distributions, the error function naturally arises in their analysis.

For example, consider the study of Gaussian noise that generally arises in communication systems. The probability that a signal exceeds a certain noise threshold is expressed in terms of the error function or its complementary form $\operatorname{erfc}(x)$.

In digital communication systems, the probability of bit error is often given by

$$P_e = \frac{1}{2}\text{erfc}\left(\frac{A}{\sqrt{2\sigma^2}}\right)$$

where A is the signal amplitude and σ^2 is the noise variance.

Even in signal smoothing, the error function arises. When a signal is allowed to pass through a system with a Gaussian response, error function also comes into play.

• Heat Conduction:

The error function also comes into play in the solution of the heat equation. For a semi-infinite medium, the temperature distribution is given by

$$T(x,t) = T_0\,\text{erf}\left(\frac{x}{2\sqrt{\alpha t}}\right)$$

where α is the thermal diffusivity.

• Diffusion Processes:

In particle diffusion problems governed by Fick's law,

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2}$$

The solution often involves the error function, representing the spread of particles in space with time.

• Probability Theory:

The error function is closely related to the cumulative distribution function of the normal (Gaussian) distribution and is used to calculate probabilities of random variables.

• Quantum Mechanics:

The error function appears in the study of Gaussian wave packets and their time evolution.

2. Elementary Properties of the Error Function

1. $\text{erf}(\infty) = 1$
2. $\text{erf}(0) = 0$
3. $\text{erf}(-\infty) = -1$
4. $\text{erfc}(0) = 1$
5. $\text{erfc}(\infty) = 0$
6. $\text{erf}(-x) = -erf(x)$
7. $\text{erfc}(-x) = 1 + erf(x)$
8. $\text{erfc}(-x) + erfc(x) = 2$

Example 12
Evaluate erf(0).
Solution:

$$\text{erf}(0) = \frac{2}{\sqrt{\pi}}\int_0^0 e^{-u^2}\,du = 0.$$

Example 13

Show that $\text{erf}(\infty) = 1$.

Solution:

$$\text{erf}(\infty) = \frac{2}{\sqrt{\pi}}\int_0^{\infty} e^{-u^2}\, du = \frac{2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = 1.$$

Example 14

Evaluate $\text{erf}(-x)$.

Solution:

$$\text{erf}(-x) = \frac{2}{\sqrt{\pi}}\int_0^{-x} e^{-u^2}\, du.$$

Changing limits,

$$= \frac{2}{\sqrt{\pi}}\int_0^{x} e^{-u^2}\, du = -\text{erf}(x).$$

Example 15

Express $\int_0^a e^{-x^2}\, dx$ in terms of $\text{erf}(x)$.

Solution:

From the definition,

$$\text{erf}(a) = \frac{2}{\sqrt{\pi}}\int_0^{a} e^{-x^2}\, dx.$$

Hence,

$$\int_0^{a} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}\text{erf}(a).$$

Example16

$$\text{Evaluate} \int_a^{\infty} e^{-x^2}\, dx.$$

Solution:

Using the complementary form,

$$\text{erf}(a) = 1 - \frac{2}{\sqrt{\pi}}\int_{\alpha}^{\infty} e^{-x^2}\, dx.$$

Thus,

$$\int_{\alpha}^{\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}\ [1 - \text{erf}(a)].$$

3. Connection between Error Function and Gamma Function

The connection between the error function and the Gamma function arises through the evaluation of Gaussian integrals. The Gamma function is defined as

$$\Gamma(s) = \int_0^{\infty} x^{s-1}e^{-x}\, dx.$$

For $s = \frac{1}{2}$,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} x^{-1}e^{-x}\, dx.$$

Using the substitution $x = u^2$, we obtain

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^{\infty} e^{-u^2}\, du.$$

Hence,

$$\int_0^\infty e^{-u^2}\,du = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

The error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}\,du.$$

Thus, the normalization constant in the definition of the error function is obtained using the Gamma function. Therefore, the error function may be regarded as a normalized form of the Gaussian integral whose total value is determined by $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Connection with Fourier Integrals

Integrals of Gaussian functions over finite or semi-infinite intervals lead to the error function. Consequently, the Error function is widely used in solving diffusion and heat conduction problems using Fourier integral methods.

10.3.3 Bessel Functions

Bessel functions arise as solutions of a class of differential equations that frequently occur in problems possessing cylindrical or spherical symmetry. The standard form of Bessel's differential equation is

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0 \qquad (10.19)$$

where n is a constant known as the order of the equation.

This equation naturally appears when solving important physical problems such as the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u, \qquad (10.20)$$

by the method of separation of variables in cylindrical or spherical coordinates. In such coordinate systems, the radial part of the solution leads directly to Bessel's equation.

Because of this, Bessel functions are widely used in the study of vibrations, heat flow, electromagnetic waves, and quantum mechanical systems where the geometry of the problem is circular or spherical.

When the order n is an integer or half-integer, Bessel functions are closely related to cylindrical or spherical harmonics. They also appear in the separable solutions of fundamental equations such as Laplace's equation and the Helmholtz equation.

Since Bessel's equation is a second-order linear differential equation, it has two linearly independent solutions. These are called the Bessel functions of the first kind, $J_n(x)$, and the second kind, $Y_n(x)$. The function $J_n(x)$ remains finite at the origin, whereas $Y_n(x)$ becomes singular there.

Thus, Bessel functions provide the appropriate mathematical tools for describing physical phenomena in systems with radial symmetry.

Relation between Bessel Function and Gamma Function

The connection between Bessel functions and the Gamma function arises primarily through the series representation of Bessel functions and their normalization. The Bessel function of the first kind is defined by the series

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(n+k+1)} \left(\frac{x}{2}\right)^{2k+n}. \tag{10.21}$$

In this expression, the Gamma function generalizes the factorial function. For integer values of n,

$$\Gamma(n\ +k+1) = (n+k)!. \tag{10.22}$$

Thus, the Gamma function allows the definition of Bessel functions for non-integer orders. For integer n, the series reduces to

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,(n+k)!} \left(\frac{x}{2}\right)^{2k+n}. \tag{10.23}$$

For half-integer values of n, the Gamma function leads to simple closed-form expressions. For example,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x, \tag{10.24}$$

which follows from the relation

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{10.25}$$

Thus, the Gamma function plays a crucial role in determining the coefficients and normalization of Bessel functions and provides a natural extension of these functions to non-integer orders.

Basic Properties of Bessel Functions

In addition to their definition as solutions of Bessel's differential equation, Bessel functions possess several important properties that are useful in applications.

Series Representation

The Bessel function of the first kind of order n can be expressed as

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(n+k+1)} \left(\frac{x}{2}\right)^{2k+n}. \tag{10.26}$$

This series is valid for all real values of x and shows explicitly the connection between Bessel functions and the Gamma function.

Recurrence Relations

Bessel functions satisfy the following recurrence relations:

$$J_{n-1}(x) + J_{n-1}(x) = \frac{2n}{x} J_n(x), \tag{10.27}$$

$$J_{n-1}(x) + J_{n-1}(x) = 2\frac{d}{dx} J_n(x). \tag{10.28}$$

These relations are very useful in practical computations.

Orthogonality Property

Bessel functions of the same order corresponding to different zeros satisfy the following orthogonality relation:

$$\int_0^a xJ_n\left(\frac{\alpha_m x}{a}\right)J_n\left(\frac{\alpha_k x}{a}\right)dx = 0, \quad m \neq k, \tag{10.29}$$

Where α_m and α_k are zeros of $J_n(x)$.

Special Cases

For small values of x, the Bessel function behaves as

$$J_n(x) \approx \frac{1}{\Gamma(n+1)}\left(\frac{x}{2}\right)^n. \tag{10.30}$$

For half-integer order, Bessel functions reduce to elementary functions. For example,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x. \tag{10.31}$$

Example 17

Show that $J_0(0) = 1$.

Solution:
Using the series definition,

$$J_0(x) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k!)^2}\left(\frac{x}{2}\right)^{2k}.$$

At $x = 0$, all terms vanish except the first term $(k = 0)$:

$$J_0(0) = 1.$$

Example 18

Obtain the first three terms of $J_1(x)$.

Solution:
The series form is

$$J_1(x) = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,\Gamma(\mathrm{k}+2)}\left(\frac{x}{2}\right)^{2k+1}.$$

Using $\Gamma(k+2) = (k+1)!$, we get

$$J_1(x) = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,(\mathrm{k}+1)!}\left(\frac{x}{2}\right)^{2k+1}.$$

Thus,

$$J_1(x) = \frac{x}{2} - \frac{x^3}{16} + \frac{x^5}{384} - \cdots.$$

Example 19
Verify the recurrence relation

$$J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x)$$

for $n = 1$.

Solution:

Substituting $n = 1$, we get

$$J_0(x) + J_2(x) = \frac{2}{x} J_1(x).$$

Using known series expansions:

$$J_0(x) = 1 - \frac{x^2}{4} + \cdots,$$

$$J_1(x) = \frac{x}{2} - \frac{x^3}{16} + \cdots,$$

$$J_2(x) = \frac{x^2}{8} + \cdots,$$

Adding,

$$J_0(x) + J_2(x) = 1 - \frac{x^2}{4} + \frac{x^2}{8} \ldots = 1 - \frac{x^2}{8} + \cdots.$$

Now,

$$\frac{2}{x} J_1(x) = \frac{2}{x}\left(\frac{x}{2} - \frac{x^3}{16} + \cdots\right) = 1 - \frac{x^2}{8} + \cdots.$$

Thus both sides are equal.

Example 20

Show that for small x,

$$J_n(x) \approx \frac{1}{\Gamma(n+1)}\left(\frac{x}{2}\right)^n.$$

Solution:

From the series definition,

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(\mathrm{n} + \mathrm{k} + 1)}\left(\frac{x}{2}\right)^{2k+1}.$$

For small x, higher powers can be neglected. Taking $k = 0$,

$$J_n(x) \approx \frac{1}{\Gamma(n+1)}\left(\frac{x}{2}\right)^n.$$

Example 21

Evaluate $J_{1/2}(x)$.

Solution:

For half-integer order, Bessel functions reduce to elementary forms:

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x.$$

This result follows from the general series representation and the relation

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Connection with Fourier Integrals

Bessel functions arise in Fourier–Bessel integrals, which are encountered while solving boundary value problems possessing cylindrical symmetry, such as heat conduction, wave propagation and vibration of circular membranes.

10.4 FOURIER INTEGRAL REPRESENTATIONS

The Fourier integral representation is an extension of the Fourier series for functions defined over an infinite interval. While Fourier series express periodic functions as sums of discrete harmonic components, Fourier integrals represent non-periodic functions as a continuous superposition of sinusoidal waves.

A function $f(x)$ is said to satisfy the Dirichlet conditions for Fourier integrals if it satisfies the following conditions:

- $f(x)$is piecewise continuous over every finite interval.
- $f(x)$is absolutely integrable, i.e., $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$
- $f(x)$has only a finite number of maxima, minima and discontinuities in every finite interval.

Hence, the Fourier integral representation of $f(x)$ is given by

$$f(x) = \int_0^{\infty} [A(\omega)\cos(\omega x) + B(\omega)\sin(\omega x)]\, d\omega. \qquad (10.32)$$

The functions $A(\omega)$ and $B(\omega)$ are given by

$$A(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\cos(\omega t)\, dt, \qquad (10.33)$$

$$B(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\sin(\omega t)\, dt. \qquad (10.34)$$

Thus, the function $f(x)$ is expressed as a continuous distribution of frequencies.

Special Cases

If f(x) is an even function, then

$$f(x) = \frac{2}{\pi}\int_0^{\infty} F_c(\omega)\cos(\omega x)\, d\omega, \qquad (10.35)$$

Here, $F_C(\omega) = \int_0^{\infty} f_{(t)}\cos(\omega t)$ for an even function.

and if f(x) is an odd function, then

$$f(x) = \frac{2}{\pi}\int_0^{\infty} F_s(\omega)\sin(\omega x)\, d\omega. \qquad (10.36)$$

Here, $F_S(\omega) = \int_0^{\infty} f_{(t)} sin(\omega)$dtfor an odd function.

The Fourier integral representation provides a powerful method for analysing non-periodic functions and plays an important role in applications such as wave propagation, heat conduction and signal analysis.

Connection of Dirac Delta Function with Fourier Integral Representation

The Dirac delta function arises naturally in the derivation of the Fourier integral representation of a function. In the process of extending the Fourier series to an infinite interval, expressions of the form

$$\int_{-\infty}^{\infty} e^{i\omega(x-t)}\, d\omega \tag{10.37}$$

are encountered. This integral behaves like a function that is zero everywhere except at $x = t$ and whose integral over the entire real line is unity. This leads to the identification

$$\delta(x-t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(x-t)}\, d\omega. \tag{10.38}$$

Using this representation, the Fourier integral form of a function $f(x)$ can be written as

$$f(x) = \int_{-\infty}^{\infty} f(t)\delta(x\ -\ t)\ dt. \tag{10.39}$$

Thus, the delta function acts as an identity kernel that extracts the value of the function at a given point.

This shows that the Fourier integral representation is closely connected to the Dirac delta function, which may be regarded as the Fourier transform of unity.

10.5 PROOF OF FOURIER INTEGRAL THEOREM

If a function $f(x)$ is defined over $(-\infty, \infty)$ and satisfies suitable conditions such as piecewise continuity and absolute integrability, then it can be represented in the form

$$f(x) = \frac{1}{\pi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} f(t)\cos\omega\,(x\ -\ t)dt\right] d\omega. \tag{10.40}$$

At a point of discontinuity, the integral converges to the mean value

$$\frac{f(x^-) + f(x^+)}{2}. \tag{10.41}$$

To prove it, let $f(x)$ be defined in the interval $(-L, L)$ and represented by its Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right), \tag{10.42}$$

where

$$a_n = \tfrac{1}{L}\textstyle\int_{-L}^{L} f(t)\cos\frac{n\pi t}{L}\,dt, b_n = \tfrac{1}{L}\int_{-L}^{L} f(t)\sin\frac{n\pi t}{L}\,dt. \tag{10.43}$$

Substituting these coefficients into the Fourier series, we obtain

$$f(x) = \frac{1}{2L}\int_{-L}^{L} f(t)\,dt + \sum_{n=1}^{\infty}\frac{1}{L}\int_{-L}^{L} f(t)\cos\frac{n\pi(x-t)}{L}\,dt. \tag{10.44}$$

As $L \to \infty$, the constant term $a_0/2$ tends to zero because $f(x)$is integrable. Furthermore, the summation approaches a Riemann sum with interval width $\Delta\omega = \pi/L$, which tends to zero as $L \to \infty$. Hence, the summation may be replaced by the corresponding definite integral.
Let us introduce a new variable

$$\omega = \frac{n\pi}{L}, \Delta\omega = \frac{\pi}{L}. \qquad (10.45)$$

As L→∞, the interval (−L,L) extends to(−∞,∞) and the discrete values of ω become continuous. Hence, the summation can be replaced by an integral:

$$\sum_{n=1}^{\infty} \to \int_0^{\infty} d\omega. \qquad (10.46)$$

Thus, the Fourier series transforms into

$$f(x) = \frac{1}{\pi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} f(t)\cos\omega\,(x\ -\ t)dt\right] d\omega. \qquad (10.47)$$

This is known as the Fourier Integral Theorem.

Complex Form
The theorem can also be written in complex form as

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt\right]e^{-i\omega x}\,d\omega. \qquad (10.48)$$

Example 22

Find the Fourier integral representation of the function

$$f(x) = e^{-a|x|}, \qquad a > 0.$$

Solution:
Since f(x) is an even function, we use the cosine form:

$$f(x) = \frac{2}{\pi}\int_0^{\infty} F_c(\omega)\cos(\omega x)\,d\omega,$$

where

$$F_c(\omega) = \int_0^{\infty} e^{-at}\cos(\omega t)\,dt.$$

Using the standard integral,

$$\int_0^{\infty} e^{-at}\cos(\omega t)\,dt = \frac{a}{a^2+\omega^2},$$

We get

$$F_c(\omega) = \frac{a}{a^2+\omega^2}.$$

Hence,

$$e^{-a|x|} = \frac{2}{\pi}\int_0^{\infty}\frac{a\cos(\omega x)}{a^2+\omega^2}d\omega.$$

Example 23
Obtain the Fourier integral representation of the function

$$f(x) = \begin{cases}1, |x| < a,\\ 0, |x| > a.\end{cases}$$

Solution:
The function is even, so

$$F_c(\omega) = \int_0^a \cos(\omega t)\, dt = \frac{\sin(\omega a)}{\omega}.$$

Thus,

$$f(x) = \frac{2}{\pi}\int_0^\infty \frac{\sin(\omega a)}{\omega}\cos(\omega x)\, d\omega.$$

Example 24

Evaluate

$$\int_0^\infty \frac{\cos(\omega x)}{a^2+\omega^2}\, d\omega.$$

Solution:

From the result of Example22,

$$e^{-a|x|} = \frac{2}{\pi}\int_0^\infty \frac{a\cos(\omega x)}{a^2+\omega^2} d\omega.$$

Thus,

$$\int_0^\infty \frac{\cos(\omega x)}{a^2+\omega^2}\, d\omega = \frac{\pi}{2a} e^{-a|x|}.$$

Example 25

Show that

$$\int_0^\infty \frac{\sin(\omega a)}{\omega} d\omega = \frac{\pi}{2}, \qquad a > 0.$$

Solution:

From Example 23, take $x = 0$:

$$1 = \frac{2}{\pi}\int_0^\infty \frac{\sin(\omega a)}{\omega} d\omega.$$

Thus,

$$\int_0^\infty \frac{\sin(\omega a)}{\omega} d\omega = \frac{\pi}{2}.$$

CHECK YOUR PROGRESS

Q1. State the definition of the Gamma function.

Q2. Write the definition of the error function.

Q3. What is Bessel's differential equation?

Q4. Evaluate $\Gamma\left(\frac{1}{2}\right)$.

Q5. State one application each of:
- (a) Error function
- (b) Bessel function.

Q6. State the Fourier Integral Theorem.

Q7. Write the Fourier integral representation of an even function.

Q8. Write the Fourier integral representation of an odd function.

Q9. State one application of Fourier integrals.

Q10. Under what conditions does a function admit a Fourier integral representation?

10.6 LET US SUM UP

- Fourier integrals extend Fourier series to represent non-periodic functions defined over infinite intervals.
- The Gamma function is a continuous extension of the factorial function and is useful in evaluating definite integrals arising in Fourier analysis.
- The Error function and its complementary function are widely used in heat conduction, diffusion and probability theory.
- Bessel functions arise as solutions of Bessel's differential equation and are important in problems possessing cylindrical or spherical symmetry.
- The Gamma, Error and Bessel functions frequently occur in the evaluation of Fourier integrals and related physical problems.
- A function admits a Fourier integral representation if it satisfies the Dirichlet conditions.
- Fourier cosine and Fourier sine integrals provide the representations of even and odd functions, respectively.
- The Dirac delta function provides a compact representation of the Fourier integral kernel and acts as an identity operator.
- The Fourier Integral Theorem forms the mathematical basis for expressing suitable functions as continuous superpositions of sinusoidal components.
- Fourier integrals have wide applications in heat conduction, wave propagation, signal processing, quantum mechanics and electromagnetic theory**.**

10.7 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: The Gamma function is defined as

$$\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\, dx, \qquad s > 0.$$

Answer to Q2: The error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du.$$

Answer to Q3: Bessel's differential equation is given by

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

Answer to Q4: Using the definition of the Gamma function, it can be shown that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

This is one of the most important properties of the Gamma function and is widely used in mathematical physics.

Answer to Q5: (a) The Error function is used in solving **heat conduction (diffusion) problems**, particularly in determining the temperature distribution in semi-infinite solids.

(b) Bessel functions are used in solving **boundary value problems involving cylindrical symmetry**, such as the vibration of a circular membrane and heat conduction in cylindrical bodies.

Answer to Q6: The Fourier Integral Theorem states that if a function $f(x)$satisfies the Dirichlet conditions, then it can be represented as a Fourier integral over the interval $(-\infty, \infty)$.

Answer to Q7: For an even function,

$$f(x) = \frac{2}{\pi}\int_0^\infty \cos(\omega x)\left[\int_0^\infty f(t)\cos(\omega t)\,dt\right]d\omega.$$

This is called the Fourier cosine integral.

Answer to Q8: For an odd function,

$$f(x) = \frac{2}{\pi}\int_0^\infty \sin(\omega x)\left[\int_0^\infty f(t)\sin(\omega t)\,dt\right]d\omega.$$

This is called the Fourier sine integral.

Answer to Q9: Fourier integrals are widely used in solving heat conduction and wave propagation problems defined over infinite or semi-infinite domains.

Answer to Q10: A function $f(x)$admits a Fourier integral representation if it satisfies the following conditions:

- $f(x)$ is piecewise continuous over every finite interval.
- $f(x)$ is integrable, i.e., $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$.
- $f(x)$ has only a finite number of maxima, minima and discontinuities in every finite interval.

These are commonly referred to as the Dirichlet conditions for Fourier integra

10.8 FURTHER READING

- Sneddon, I. N. (1995) *Fourier Transforms*. New York: Dover Publications**.**
- Arfken, G. B., Weber, H. J. and Harris, F. E.(2013) *Mathematical Methods for Physicists*. 7th ed. Amsterdam: Elsevier.
- Boas, M. L. (2006) *Mathematical Methods in the Physical Sciences*. 3rd ed. New York: John Wiley & Sons,
- Kreyszig, E. (2011**)** *Advanced Engineering Mathematics*. 10th ed. Hoboken, NJ: John Wiley & Sons, 2011.
- Grewal, B. S**.** (2018**)** *Higher Engineering Mathematics*. 44th ed. New Delhi: Khanna Publishers.

- Dass, H. K. and Verma, Rajnish**.**(2013) *Higher Engineering Mathematics*. New Delhi: S. Chand Publishing.

⍰ 10.9 MODEL QUESTIONS

1. Define Fourier integral. Distinguish between Fourier series and integral.
2. Derive the Fourier integral representation of a function defined over $(-\infty, \infty)$.
3. State and prove the Fourier integral theorem.
4. Obtain the Fourier integral representation of $f(x) = xe^{-a|x|},\ a > 0$.
5. Using Fourier integrals, show that
$$\int_0^\infty \frac{\omega \sin(\omega x)}{a^2 + \omega^2}\, d\omega = \frac{\pi}{2} e^{-ax},\ a > 0,\ x > 0.$$
6. Using the recurrence relation of the Gamma function, evaluate $\Gamma\left(\frac{5}{2}\right)$.
7. Show that the Gamma function is a generalization of the factorial function by proving that$\Gamma(n+1) = n!$.
8. Discuss the role of the Gamma, Error and Bessel functions in the evaluation of Fourier integrals.

*** *** ***

UNIT 11: FOURIER TRANSFORM

UNIT STRUCTURE

11.1 Learning Objectives
11.2 Introduction
11.3 Fourier Transformation and Properties
11.4 Fourier Transform of Derivatives
  11.4.1 Shift Theorem
  11.4.2 Change of Scale
  11.4.3 Modulation Theorem
11.5 Convolution Theorem
11.6 Parseval's Identity
11.7 Let Us Sum Up
11.8 Answers to Check Your Progress
11.9 Further Reading
11.10 Model Questions

11.1 LEARNING OBJECTIVES

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

- explain the concept of the Fourier transform and distinguish it from the Fourier series
- define the Fourier sine and cosine transforms and their inverse transforms
- apply the properties of the Fourier transform
- derive the Fourier transform of derivatives and explain Parseval's identity and the convolution theorem
- explain the concept and applications of the Dirac delta function in Fourier analysis
- describe the applications of the Fourier transform in mathematical physics.

11.2 INTRODUCTION

The Fourier transform is a special linear integral transformation used to solve differential equations defined on unbounded domains. The method is associated with the French physicist and mathematician Joseph Fourier whose, whose work opened up a new way to solving linear partial differential equations.

In the previous unit, we studied the Fourier series, which are mainly useful for periodic functions. The Fourier series expresses a periodic function as a sum of sinusoids. In this unit these ideas and techniques extend to a non-periodic function defined on the whole real line by taking the limiting form of Fourier series when the fundamental period is made very large (infinite). The Fourier transform has extensive applications in science and engineering, for example in physical optics, signal processing, linear time invariant systems.

11.3 FOURIER TRANSFORMATION AND PROPERTIES

Fourier transforms have for a long time been a basic tool of applied mathematics, particularly for solving differential equations and in conjunction with integral equations. There are really three Fourier transforms, the Fourier sine and Fourier cosine transforms and a complex form which is usually referred to as the Fourier transform. First, we will define the Fourier transform and then study its various properties.

The Fourier transform of $f(t)$ is defined by

$$\mathcal{F}\{f(t)\} = F(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt \qquad (11.1)$$

as long as the integral exists. The inverse Fourier transform of $F(\omega)$ is defined by

$$\mathcal{F}^{-1}\{F(\omega)\} = f(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,d\omega \qquad (11.2)$$

Thus, the Fourier transform pair may be written as

$$\mathcal{F}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt$$

and, consequently

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,d\omega.$$

Equations (11.1) and (11.2) will be used throughout this unit.

Remark 1. Let $u(t)$ denote unit step function, defined by

$$u(t) = \begin{cases} 0 \text{ if } t < 0 \\ 1 \text{ if} t > 0. \end{cases}$$

It is also referred to as the Heaviside step function and denoted by $H(t)$.

Example 1. Let $f(t) = \begin{cases} 0 & \text{if } t < 0 \\ e^{-t} & \text{if} t > 0. \end{cases}$ then

$$F(\omega) = \frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{-(i\omega+1)t}\,dt = \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-(1+i\omega)t}}{-(1+i\omega)}\right]_{t=0}^{\infty} = \frac{1}{\sqrt{2\pi}(1+i\omega)}.$$

NOTE: The function in Example 1 can also be written as $f(t) = e^{-t}u(t)$.

Hence

$$\mathcal{F}\{e^{-t}u(t)\} = \frac{1}{\sqrt{2\pi}(1+i\omega)}.$$

Example 2. The Fourier transform of the rectangular pulse

$$pa(t) = \begin{cases} 1 & -a \le t \le a \\ 0 & \text{otherwise} \end{cases}$$

Using (11.1) is given as following:

$$\mathcal{F}\{pa(t)\} = \frac{1}{\sqrt{2\pi}}\int_{-a}^{a} e^{-i\omega t}\,dt$$
$$= \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-i\omega t}}{-i\omega}\right]_{t=-a}^{t=a}$$
$$= \frac{2}{\sqrt{2\pi}}\frac{e^{i\omega t} - e^{-i\omega t}}{2i\omega}$$
$$= \sqrt{\frac{2}{\pi}}\frac{\sin\omega a}{\omega}.$$

Hence

$$\mathcal{F}\{p_a(t)\} = \sqrt{\frac{2}{\pi}}\frac{\sin\omega a}{\omega}. \qquad (11.3)$$

The following result gives the sufficient conditions of existence of Fourier transform of a given function:

Theorem 1: (**Dirichlet Conditions)**
Sufficient conditions for the existence of the Fourier transform of a function $f(t)$*is that the function satisfies the following conditions:*

(i) $\int_{-\infty}^{\infty}|f(t)|\mathrm{dt} < \infty$i.e., $f(t)$is absolutely integrable over **(−∞,∞).**
(ii) The function $f(t)$ has a finite number of maxima and minima in every finite interval
(iii) The function has only a finite number of discontinuities in any finite interval.

Example 3. For $f(t) = e^t\int_{-\infty}^{\infty} e^t\,dt$, is not finite. Hence the Fourier transform of e^t does not exist in the ordinary sense.

If the Dirichlet conditions hold then $f(t)$ has a unique Fourier transform. However certain functions, which violate one or more of the Dirichlet conditions still have Fourier transform. To discuss this we will introduce the unit impulse function, also known as Dirac delta function. The Dirac Delta function or unit impulse function is defined by $\delta(t-a) = 0$ if $t \neq a$. Using the fundamental property of Dirac delta function which is given as: If $f(t)$ is any differentiable function then $\int_{-\infty}^{\infty} f(t)\delta(t\ -\ a)dt\,f(a)$, we have

$$\mathcal{F}\{\delta(t\ -a)\}\ = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega t}\,\delta(t\ -\ a)dt = \frac{1}{\sqrt{2\pi}}e^{-i\omega t}.$$

Hence the Fourier transform of Dirac delta or unit impulse function is

$$\mathcal{F}\{\delta(t\ -a)\} = \frac{1}{\sqrt{2\pi}}e^{-i\omega t}, \quad \text{if } a = 0 \text{ then } \mathcal{F}\{\delta(t\ )\} = \frac{1}{\sqrt{2\pi}}.$$

FOURIER SINE TRANSFORM:

A function f is said to be odd if $f(-t) = -f(t)$ and a $\int_{-a}^{a} f(t)dt = 0$. For example: $t, \sin t, t^2 \sin t$ etc. If $f(t)$ is an odd function defined on $(-\infty.\infty)$ then its Fourier transform is given by:

$$\mathcal{F}_s\{f(t)\} = F_s(\omega) = \sqrt{\frac{2}{\pi}}\int_0^{\infty} f(t)\sin(\omega t)\,dt \qquad (11.4)$$

and is known to be Fourier sine transform and the inverse Fourier sine transform is defined by

$$\mathcal{F}_s^{-1}\{f_s(\omega)\} = f(t) = \sqrt{\frac{2}{\pi}}\int_0^{\infty} F_s(\omega)\sin(\omega t)d\omega.$$

FOURIER COSINE TRANSFORM:

A function f is said to be even if $f(-t) = f(t)$ and $\int_{-a}^{a} f(t)dt = 2\int_0^a f(t)dt$. For example: $t^2, t^4, \cos t$ etc. If $f(t)$ is an even function defined on $(-\infty.\infty)$ then its Fourier transform is given by:

$$\mathcal{F}_c\{f(t)\} = F_c(\omega) = \sqrt{\frac{2}{\pi}}\int_0^{\infty} f(t)\cos(\omega t)\,dt \qquad (11.5)$$

and is known to be Fourier cosine transform and the inverse Fourier cosine transform is defined by

$$\mathcal{F}_c^{-1}\{f_c(\omega)\} = f(t) = \sqrt{\frac{2}{\pi}}\int_0^{\infty} F_c(\omega)\cos(\omega t)\,d\omega.$$

11.4 FOURIER TRANSFORM OF DERIVATIVES

In this section, we shall study the Fourier transform of derivatives, which is an important application of the Fourier transform in mathematical physics. The Fourier transform changes the operation of differentiation into a simple algebraic operation in the transform domain. This makes it easier to solve many differential equations arising in science and engineering.

Before discussing the Fourier transform of derivatives, we first study some important properties of the Fourier transform, namely the shift theorem, linearity, change of scale and modulation theorem. These properties simplify the calculation of Fourier transforms and help us understand the Fourier transform of derivatives more easily.

11.4.1 Shift Theorem

Theorem 2.(Shift Theorem)

If $\mathcal{F}\{f(t)\} = F(\omega)$, *then*

(i) $\mathcal{F}\{f(t-\alpha)\} = e^{-i\omega\alpha}F(\omega)$ (Time shifting property)

(ii) $\mathcal{F}\{e^{-i\alpha t}f(t)\} = F(\omega+\alpha)$ (Frequency shifting property).

Proof: Case (i): Using (11.1)

$$\mathcal{F}\{f(t-\alpha)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t-\alpha)e^{-i\omega t}\,dt,$$

by taking $t-\alpha = u$ implies $dt = du$, we have

$$\mathcal{F}\{f(t-\alpha)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(u)e^{-i\omega(u+\alpha)}\,du = e^{-i\omega\alpha}F(\omega).$$

Case (ii): Using (11.1), we find

$$\mathcal{F}\{e^{-i\alpha t}f(t)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\alpha t}\,f(t)e^{-i\omega t\,dt} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i(\omega+\alpha)t}\,dt = F(\omega+\alpha).$$

Example 4. Find the Fourier transform of the function $f(t) = \begin{cases}1 & 3 \le t \le 5\\ 0 & \text{otherwise}\end{cases}$

Solution: Here $f(t)$ can be obtained by shifting the rectangular pulse function $p1^{(t)} = \begin{cases}1 & -1 \le t \le 1\\ 0 & \text{otherwise}\end{cases}$ by 4 units to the right. Hence by taking $\alpha = 4$ in the shifting

property $\mathcal{F}(\omega) = \mathcal{F}\{f(t)\} = e^{-4i\omega}\,\frac{2}{\sqrt{2\pi}}\frac{\sin\omega}{\omega}$.

11.4.2 Change of Scale

Before discussing the change of scale property, we first state the linearity property of the Fourier transform, which is frequently used in deriving other transform properties.

Theorem 3. (Linearity Property)

$$If\ \mathcal{F}\{f(t)\} = \mathcal{F}(\omega)\ and\ \mathcal{F}\{g(t)\} = G(\omega)\ then\ \mathcal{F}\{\alpha f(t) + \beta g(t)\} = \alpha\mathcal{F}(\omega) + \beta G(\omega), where\ \alpha\ and\ \beta\ are\ constants.$$

Proof: Using (11.1) and the properties of improper integral, we have

$$\begin{aligned}\mathcal{F}\{\alpha f(t) + \beta g(t)\} &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\{\alpha f(t) + \beta g(t)\}e^{-i\omega t}dt\\ &= \alpha\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt + \beta\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(t)e^{-i\omega t}dt\\ &= \alpha F(\omega) + \beta G(\omega).\end{aligned}$$

This property shows that the Fourier transform is a linear operator. Using this property, we now discuss the change of scale property.

Example 5. If $f(t) = \begin{cases}6 & -2 \le t \le 2\\ 0 & \text{otherwise}\end{cases}$

then $f(t) = 6p_2(t)$

so by (11.3) $\mathcal{F}(\omega) = 6\sqrt{\frac{2}{\pi}}\frac{\sin 2\omega}{\omega}$.

NOTE: The result can also be obtained directly using Equation (11.1)**.**

Another useful property of the Fourier transform is the change of scale property, which relates the Fourier transform of a scaled function to that of the original function.

Theorem 4.(Change of Scale Property)

$$If\ \mathcal{F}\{f(t)\} = \mathcal{F}(\omega), then\ \mathcal{F}\{f(\alpha t)\} = \frac{1}{\alpha}F\left(\frac{\omega}{\alpha}\right), where\ \alpha\ is\ any\ constant.$$

Proof:

$$\mathcal{F}\{f(\alpha t)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(\alpha t)e^{-i\omega t}\,dt.$$

Putting $\alpha t = u$ implies $dt = \frac{1}{\alpha}du$. Hence

$$\mathcal{F}\{f(\alpha t)\} = \frac{1}{\alpha\sqrt{2\pi}}\int_{-\infty}^{\infty} f(u)e^{-i\frac{\omega}{\alpha}u}\,du = \frac{1}{\alpha}\mathcal{F}\left(\frac{\omega}{\alpha}\right).$$

11.4.3 Modulation Theorem

Theorem 5 (Modulation Theorem)

$$If\ \mathcal{F}(\omega)\ \text{is Fourier transform of}\ f(t), \text{then}$$

$$\mathcal{F}\{f(t)\cos\alpha t)\} = \frac{1}{2}\{\mathcal{F}(\omega+\alpha) + F(\omega-\alpha)\}$$

and

$$\mathcal{F}\{f(t)\sin\alpha t)\} = \frac{1}{2}\{\mathcal{F}(\omega+\alpha) - F(\omega-\alpha)\}.$$

Proof: Using Euler's formula,

$$\cos\alpha\, t = \frac{1}{2}\left(e^{i\alpha t} + e^{-i\alpha t}\right).$$

Therefore, $f(t)\cos\alpha\, t = \frac{1}{2}\left[e^{i\alpha t}f(t) + e^{-i\alpha t}f(t)\right]$.

Taking the Fourier transform of both sides and using the linearity property, we obtain

$$F\{f(t)\cos\alpha\, t\} = \frac{1}{2}\left[F\left\{e^{i\alpha t}f(t)\right\} + F\left\{e^{-i\alpha t}f(t)\right\}\right]$$

Applying the frequency shifting property (Theorem 2),

$$\mathcal{F}\left\{e^{i\alpha t}f(t)\right\} = \mathcal{F}(\omega-\alpha)$$

and

$$F\left\{e^{-i\alpha t}f(t)\right\} = F(\omega+\alpha).$$

Hence, $\mathcal{F}\{f(t)\cos\alpha\, t\} = \frac{1}{2}[\mathcal{F}(\omega+\alpha) + \mathcal{F}(\omega-\alpha)]$.

Thus, the modulation theorem is proved.

Example: Find the Fourier transform of

$$f(t) = e^{-a|t|}\cos\omega_0 t,\ a > 0.$$

Solution: *It is known that*

$$\mathcal{F}\{e^{-a|t|}\} = \sqrt{\frac{2}{\pi}}\frac{a}{a^2+\omega^2}.$$

Using the **Modulation Theorem**,

$$\mathcal{F}\{f(t)\cos\omega_0 t\} = \frac{1}{2}[\mathcal{F}(\omega+\omega_0) + \mathcal{F}(\omega-\omega_0)].$$

Therefore,

$$\mathcal{F}\{e^{-a|t|}\cos\omega_0 t\} = \frac{1}{2}\sqrt{\frac{2}{\pi}}\left[\frac{a}{a^2+(\omega+\omega_0)^2} + \frac{a}{a^2+(\omega-\omega_0)^2}\right].$$

Hence,

$$\boxed{\mathcal{F}\{e^{-a|t|}\cos\omega_0 t\} = \frac{1}{2}\sqrt{\frac{2}{\pi}}\left[\frac{a}{a^2+(\omega+\omega_0)^2} + \frac{a}{a^2+(\omega-\omega_0)^2}\right].}$$

Theorem 6 (Time Differentiation Property)

The Fourier transform of the derivative $f'(t)$ of the function $f(t)$ is obtained by multiplication of its Fourier transform $F(\omega)$ by $i\omega$. i.e.,

$$\mathcal{F}\left\{f'(t)\right\} = \mathcal{F}\left\{\frac{df}{dt}\right\} = i\omega\mathcal{F}(\omega)$$ **(Time differentiation)**.

Proof: Using the definition of the Fourier transform (Equation 11.1),

$$\mathcal{F}\left\{\frac{df}{dt}\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{df}{dt}\,e^{-i\omega t}\,dt.$$

Integrating by parts, let $u = e^{-i\omega t},\ dv = \frac{df}{dt}\,dt.$

Then, $du = -i\omega e^{-i\omega t}\,dt,\ v = f(t).$

Hence, $\mathcal{F}\left\{\frac{df}{dt}\right\} = \frac{1}{\sqrt{2\pi}}\left[f(t)e^{-i\omega t}\right]_{-\infty}^{\infty} + \frac{i\omega}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt.$

Since $f(t) \to 0$ as $t \to \pm\infty$, the boundary term vanishes. Therefore,

$$\mathcal{F}\left\{\frac{df}{dt}\right\} = i\omega\mathcal{F}(\omega).$$

Thus, the Fourier transform of the derivative of a function is obtained by multiplying its Fourier transform by $i\omega$.

Corollary (Higher-Order Derivatives)

Applying the Fourier transform of the first derivative repeatedly, we obtain

$$\mathcal{F}\left\{f^{(n)}(t)\right\} = (i\omega)^n\mathcal{F}(\omega)$$

Theorem 7 (Frequency Differentiation Property)

If $\mathcal{F}\{f(t)\} = \mathcal{F}(\omega)$, then$\mathcal{F}\{tf(t)\} = i\frac{dF}{d\omega}$,or equivalently,

$$\mathcal{F}\{(-it)f(t)\} = \frac{dF}{d\omega}$$

Proof:

Using the definition of the Fourier transform given in Equation (11.1),

$$\mathcal{F}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f\,(t)e^{-i\omega t}\,dt$$

Differentiating both sides with respect to ω, we obtain

$$\frac{dF}{d\omega} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f\,(t)\frac{\partial}{\partial\omega}\left(e^{-i\omega t}\right)dt$$

Since $\frac{\partial}{\partial\omega}\left(e^{-i\omega t}\right) = -it\,e^{-i\omega t}$

we have $\frac{dF}{d\omega} = -\frac{i}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t\,f(t)e^{-i\omega t}\,dt$

Hence, $\mathcal{F}\{tf(t)\} = i\frac{dF}{d\omega}$

Thus, the frequency differentiation property is proved.

11.5 CONVOLUTION THEOREM

Definition 1: The convolution of two functions $f(t)$ and $g(t)$ is defined by

$$(f * g)(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(u)g(t-u)du.$$

Following result gives the Fourier transform of convolution of two functions:

Theorem 8 (Convolution Theorem)
If $\mathcal{F}(\omega)$ and $G(\omega)$are the Fourier transforms of $f(t)$ and $g(t)$, respectively, then the Fourier transform of the convolution of $f(t)$and $g(t)$is given by...

$$\mathcal{F}\{(f\ * g)(t)\} = \mathcal{F}(\omega)G(\omega).$$

Remark 2: Using above result, we have $(f * g)(t) = \mathcal{F}^{-1}\{F(\omega)G(\omega)\}$ so

$$\mathcal{F}^{-1}\{F(\omega)G(\omega)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(u)g(t-u)du.$$

Using the convolution theorem we have the following well known identity. This result is particularly significant in filter theory.

11.6 PARSEVAL'S IDENTITY

If $\mathcal{F}(\omega)$ is the Fourier transform of $f(t)$, then

$$\int_{-\infty}^{\infty}|f(t)|^2dt=\int_{-\infty}^{\infty}|F(\omega)|^2d\omega\,.$$

Example 6: Using Parseval's identity we can compute an integral, $\int_0^{\infty}\frac{1}{(1+x^2)^2}\,dx=\frac{\pi}{4}$.

Consider the function $f(t)=e^{-t}$, then its Fourier cosine transform $\mathcal{F}_c(\omega)=\sqrt{\frac{2}{\pi}}\frac{1}{1+\omega^2}$. So from Parseval's identity we have

$$\int_0^{\infty}(e^{-t})^2\,dt=\int_0^{\infty}\left|\sqrt{\frac{2}{\pi}}\frac{1}{1+\omega^2}\right|^2\,d\omega.$$

The desired result follows as $\int_0^{\infty}(e^{-t})^2\,dt=\int_0^{\infty}e^{-2t}\,dt=\frac{1}{2}$

Application of Fourier Transform in Quantum Mechanics

In quantum mechanics, the wave function may be represented either in position space or momentum space. These two representations are related through the Fourier transform. If $\psi(x)$ represents the wave function in position space, then $\phi(p)$ denotes the corresponding wave function in momentum space. The Fourier transform converts one representation into the other. This enables the study of physical systems in either coordinate space or momentum space according to convenience.

CHECK YOUR PROGRESS

Q1. Find the Fourier transform of $f(t)=e^{-at}u(t)$, here $a>0$ is a constant and $u(t)$ represent the unit step function.

Q2. Find the inverse Fourier transform of $\mathcal{F}(\omega)=\frac{1}{-\omega^2+3i\omega+2}$.

Q3. Find the Fourier sine transform of$f(t)=\frac{1}{t}1$

Q4. Find the Fourier cosine transform of

$$f(t)=\begin{cases}1, & 0<t<a,\\ 0, & t>a,\end{cases}\ (a>0).$$

Q5. Find the Fourier transform of

$$f(t)=2e^{-t}u(t)+3e^{-2t}u(t)$$

where $u(t)$ denotes the unit step (Heaviside) function.

Q6. Find the Fourier transform of $f(t) = \begin{cases} 2 \; if |t| < 1 \\ 0 \; if |t| > 1 \end{cases}$

And hence prove that(i) $\int_0^\infty \frac{\sin t}{t} dt = \frac{\pi}{2} (ii) \int_0^\infty \frac{\sin^2 t}{t^2} dt = \frac{\pi}{2}$.

Q7. Solve the integral equation $\int_0^\infty f(t) \cos \omega \, dt = e^{-\omega}$.

Q8. Find the Fourier transform of the Gaussian function $g(t) = e^{-t^2}$ on $(-\infty, \infty)$.

11.7 LET US SUM UP

- First, we defined the Fourier transform and inverse Fourier transform for non-periodic functions as given in (11.1) and (11.2) respectively.
- We discussed the Fourier sine and Fourier cosine transform formulas as given in (11.4) and (11.5) respectively.
- We stated and proved the various properties such as linearity, change of scale and shifting of Fourier transforms, and illustrated them with examples.
- Using these properties, we solved many problems easily and, in few cases, gave an alternate solution.
- We stated the convolution theorem and solved problems to illustrate its application.
- We discussed Parseval's identity with examples and evaluated many improper integrals.

11.8 ANSWERS TO CHECK YOUR PROGRESS

Answer Q.1: By the definition of unit step function $u(t)$, we have $f(t) = \begin{cases} 0 \quad if \; t < 0 \\ e^{-at} \; if \; t > 0. \end{cases}$

Fourier transform of $f(t)$ is given by

$$\mathcal{F}(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-at} e^{-i\omega t} \, dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(a+i\omega t)} \, dt$$

$$= \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-(a+i\omega t)}}{-(a+i\omega)} \right]_0^\infty$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{a+i\omega},$$

Since $e^{-at} \to 0$ as $t \to \infty$ for $a > 0$. Hence $\mathcal{F}\{e^{-at} u(t)\} = \frac{1}{\sqrt{2\pi}} \frac{1}{a+i\omega}$.

Answer Q.2:

$$\mathcal{F}(\omega) = \frac{1}{-\omega^2 + 3i\omega + 2} = \frac{1}{(1+i\omega)(2+i\omega)} = \frac{1}{(1+i\omega)} - \frac{1}{(2+i\omega)}.$$

So

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \mathcal{F}^{-1}\left\{\frac{1}{1+i\omega}\right\} - \mathcal{F}^{-1}\left\{\frac{1}{2+i\omega}\right\} = e^{-t}u(t) - e^{-2t}u(t).$$

This implies the inverse transform $f(t)$ of $F(\omega)$ is

$$f(t) = \begin{cases} 0 & if\, t < 0 \\ e^{-t} - e^{-2t} & if\, t > 0. \end{cases}$$

Answer Q.3: Using the Fourier sine transform formula (11.4) for $f(t) = \frac{1}{t}$, we get

$$\mathcal{F}s\{f(t)\} = \mathcal{F}_s(\omega) = \sqrt{\frac{2}{\pi}}\int_0^\infty \frac{1}{t}\sin(\omega t)\, dt$$

Taking $\omega t = u$ implies $dt = \frac{du}{\omega}$, hence

$$\mathcal{F}s(\omega) = \sqrt{\frac{2}{\pi}}\int_0^\infty \frac{\sin u}{u}\, du.$$

$$\text{Since} \int_0^\infty \frac{\sin u}{u}\, du = \frac{\pi}{2}, \text{gives} \mathcal{F}s(\omega) = \sqrt{\frac{2}{\pi}}.$$

Answer Q.4: Using the Fourier cosine transform (11.5), we have

$$Fc(\omega) = \sqrt{\frac{2}{\pi}}\int_0^a \cos(\omega t)\, dt = \sqrt{\frac{2}{\pi}}\left[\frac{\sin \omega t}{\omega}\right]_{t=0}^a = \sqrt{\frac{2}{\pi}}\frac{\sin \omega t}{\omega}.$$

Answer Q.5: Using the linearity properties, we have

$$\begin{aligned} \mathcal{F}\{2e^{-t}u(t) + 3e^{-2t}u(t)\} &= \mathcal{F}\{2e^{-t}u(t)\} + \mathcal{F}\{3e^{-2t}u(t)\} \\ &= 2\mathcal{F}\{e^{-t}u(t)\} + 3\mathcal{F}\{e^{-2t}u(t)\} \\ &= \frac{2}{1+i\omega} + \frac{3}{2+i\omega}. \end{aligned}$$

Answer Q.6: Using (11.1),we have

$$\begin{aligned} \mathcal{F}\{f(t)\} &= \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} 2e^{-i\omega t}\, dt \\ &= \frac{2}{\sqrt{2\pi}}\left[\frac{e^{-i\omega t}}{-i\omega}\right]_{t=-1}^{t=1} \\ &= \frac{4}{\sqrt{2\pi}}\frac{[e^{i\omega} - e^{-i\omega}]}{2i\omega} \end{aligned}$$

$$= \frac{4}{\sqrt{2\pi}} \frac{\sin \omega}{\omega}.$$

For (i): By inversion formula (11.2),

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} \frac{4}{\sqrt{2\pi}} \frac{\sin \omega}{\omega} e^{i\omega t} d\omega.$$

Taking $t = 0$, we have

$$f(0) = \frac{2}{\pi} \int_{\infty}^{\infty} \frac{\sin \omega}{\omega} d\omega = \frac{4}{\pi} \int_{0}^{\infty} \frac{\sin \omega}{\omega} d\omega.$$

But $f(0) = 2$ gives $\int_{0}^{\infty} \frac{\sin \omega}{\omega} d\omega = \frac{\pi}{2}$.

(ii) Using Parseval's identity

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\infty}^{\infty} |f(s)|^2 \, ds$$

We have

$$\int_{-1}^{1} 4dt = \int_{-\infty}^{\infty} \left| \frac{4}{\sqrt{2\pi}} \frac{\sin \omega}{\omega} \right|^2 d\omega$$

$$2 \int_{0}^{\infty} \left(\frac{\sin \omega}{\omega} \right)^2 = \pi$$

$$\int_{0}^{\infty} \frac{\sin^2 \omega}{\omega^2} d\omega = \frac{\pi}{2}$$

Hence changing ω to t, we have

$$\int_{0}^{\infty} \frac{\sin^2 t}{t^2} dt = \frac{\pi}{2}.$$

Answer Q.7: Given that $\int_{0}^{\infty} f(t) \cos \omega t \, dt = e^{-\omega}$, so

$$\sqrt{\frac{2}{\pi}} \int_{0}^{\infty} f(t) \cos \omega t \, dt = \sqrt{\frac{2}{\pi}} e^{-\omega}$$

This gives

$$\mathcal{F}c\{f(t)\} = \sqrt{\frac{2}{\pi}} e^{-\omega},$$

hence

$$f(t) = \mathcal{F}_c^{-1} \sqrt{\frac{2}{\pi}} e^{-\omega} = \frac{2}{\pi} \int_{0}^{\infty} e^{-\omega} \cos \omega t \, d\omega.$$

Using the known integral result

$$\int_{0}^{\infty} e^{-at} \cos bt \, dt = \frac{a}{a^2 + b^2}.$$

We obtained $f(t) = \dfrac{2}{\pi(1+t)^2}$.

Answer Q.8: From (11.1),we find

$$G(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2}\, e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(t+i\omega/2)^2\omega^2/4}\, dt$$

Substituting $u = t + \frac{i\omega}{2}$ gives $du = dt$ then

$$G(\omega) = \frac{e^{\omega^2/4}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-u^2}\, du = \frac{2e^{\omega^2/4}}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-u^2}\, du$$

Since $\int_0^\infty e^{-t^2}\, dt = \frac{\sqrt{\pi}}{2}$, we have $G(\omega)\frac{e^{-\omega^2/4}}{\sqrt{2}}$.

Hence $\boxed{\mathcal{F}\{e^{-t^2}\} = \dfrac{e^{-\omega^2/4}}{\sqrt{2}}.}$

11.9 FURTHER READING

- Kreyszig, E. (2006), *Advanced Engineering Mathematics*, Wiley, India.
- Sreenadh S. Ranganatham S. et. all, (2014) *Fourier Series and Integral Transforms*, S. Chand.
- Dass. H. K. and Verma. R. (2013). *Mathematical Physics*. S. Chand.
- V. Serov, (2017). Fourier Series, *Fourier Transform and Their Applications to Mathematical Physics*, Springer.

11.10 MODEL QUESTIONS

1. Find the Fourier transform of
$$f(t) = \begin{cases} 1 \, if\, |t| < 3 \\ 4 \\ 0 \, if |t| > 3. \end{cases}$$
2. Find the inverse Fourier transform of $F(\omega) = 4\frac{\sin 2\omega}{2\omega}$.
3. Find the Fourier cosine transform of $f(t) = \frac{1}{1+t^2}$.
4. Find the Fourier cosine transform of $f(t) = \begin{cases} t & \text{if} 0 < t < 1 \\ 2-t & \text{if} 1 < t < 2 \\ 0 & \text{if } t > 2. \end{cases}$
5. Show that $F\{e^{-at}\} = \sqrt{\frac{2}{\pi}}\frac{a}{a^2+\omega^2}$ if $a > 0$.

6. Use the time shift property to obtain the Fourier transform of $f(t) = \begin{cases} 2 & 1 \le t \le 3 \\ 0 & \text{otherwise} \end{cases}$. Verify your answer using the definition of the Fourier transform.

7. Use the frequency shift property to obtain the Fourier transform of the modulated wave $g(t) = f(t)\cos\omega_0 t$, where $f(t)$ is an arbitrary signal whose Fourier transform is $F(\omega)$.

8. Find $f(t)$ if its Fourier transform $F(\omega) = \frac{2}{1+2(\omega-1)i}$.

9. Verify Parseval's identity for the following function:
$$f(t) = \begin{cases} 0 & if\, t < 0 \\ e^{-2t} & if\, t > 0. \end{cases}$$

10. Find the Fourier transform of e^{-t^2} and hence find Fourier transform of $e^{-t^2}, a > 0$.

11. Find the Fourier sine transform of $f(t) = e^{-t},\ t > 0$. Using this find the value of the following integral:
$$\int_0^\infty \frac{t\sin(at)}{1+t^2}\, dt,$$
For any constant $a > 0$.

12. Using convolution theorem find the inverse Fourier transform of $\frac{1}{12+7i\omega-\omega^2}$.

13. State with reasons whether the Fourier transform of the function $f(t) = \sin\frac{1}{t}$ exist.

14. Use Parseval's identity and Fourier transform of $f(t) = \begin{cases} e^{-\alpha t} & t > 0 \\ -e^{-\alpha t} & t < 0. \end{cases}$ where $\alpha > 0$ to show that
$$\int_{-\infty}^{\infty} \frac{d\omega}{(\alpha^2+\omega^2)^2} = \frac{\pi}{2|\alpha|^3}.$$

15. Prove that $f(t) = e^{-t^2/2}$ is self reciprocal with respect to Fourier transform.

16. If $t > 0$, show that $\int_0^\infty \frac{x\sin xt}{x^2+a^2} = \frac{\pi}{2}e^{-at}, a > 0$.

*** *** ***

UNIT 12: COMPLEX ANALYSIS-I

UNIT STRUCTURE

12.1 Learning Objectives
12.2 Introduction
12.3 Complex Functions
12.4 Cauchy-Riemann Equations
12.5 Analytic Functions
12.6 Harmonic Functions
12.7 Let Us Sum Up
12.8 Answers to Check Your Progress
12.9 Further Reading
12.10 Model Questions

12.1 LEARNING OBJECTIVES

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

- discuss the concept of a complex function and its derivative
- explain the meaning of an analytic function and learn about Cauchy-Riemann equations
- evaluate harmonic function and its relationship with analytic function
- solve complex problems involving standard complex functions.

12.2 INTRODUCTION

In the previous unit, the theory of Fourier transforms was introduced as an important analytical tool for representing functions in terms of frequency components and for solving a variety of physical and mathematical problems. The study of Fourier transforms highlighted the usefulness of complex exponential functions and their role in transforming problems from one domain to another.

The present unit introduces the basic concepts of complex analysis with special emphasis on functions of a complex variable and their fundamental properties. The discussion begins with the definition and representation of complex functions, followed by some important elementary functions frequently used in complex-variable theory. The concepts of limit, continuity, and differentiability are then developed in the setting of the complex plane, extending familiar ideas from real-variable calculus to complex variables. The unit further examines the Cauchy–Riemann equations as the fundamental conditions for differentiability and analyticity. Finally, harmonic functions and their relationship with analytic functions are discussed, along with methods for constructing analytic functions when either the real or imaginary part is known.

12.3 COMPLEX FUNCTIONS

Let z be a complex variable, where $z = x + iy$, and x and y are real variables with $i = \sqrt{-1}$. A complex-valued function $f(z)$ defined on a set Smay be regarded as a mapping

that assigns to each $z \in S$ a unique complex number w, written as $w = f(z)$. If a function f is defined on a set Ssuch that to each $z \in S$ there corresponds a unique complex number $w = f(z)$, then the set S containing all such admissible values of z is called the **domain** of the function. The corresponding value w is known as the value of the function at the point z. If $w = u + iv$, then

$$f(z) = u(x, y) + iv(x, y),$$

where u and v are real-valued functions of the real variables $x$and y. In polar form, if $z = re^{i\theta}$, then the function may also be expressed in terms of the polar variables r and θ.

Example 1: Let $f(z) = \frac{1}{z}$. Then $f(z) = f(x + iy) = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}$. Hence $u(x, y) = \frac{x}{x^2+y^2}$ and $v(x, y) = -\frac{y}{x^2+y^2}$. In case of polar coordinates $f(z) = f\left(re^{i\theta}\right) = \frac{1}{re^{i\theta}} = \frac{1}{r}e^{-i\theta} = \frac{1}{r}(\cos\theta - i\sin\theta)$ consequently $u(r, \theta) = \frac{1}{r}\cos\theta$ and $v(r, \theta) = \frac{1}{r}\sin\theta$.

Example 2: Let $f(z) = z^2$ Then $f(z) = f(x + iy) = x^2 - y^2 + i2xy$. Hence $u(x, y) = x^2 - y^2$ and $v(x, y) = 2xy$.

A multi-valued function is one in which a single point of the domain may correspond to more than one functional value. Suppose $f(z) = z^{\frac{1}{2}}$ is a multiple valued function as this function assigns two distinct values to each $z \neq 0$.

Elementary Functions

1. The function $p(z) = a_0 + a_1z + a_2z^2 + \cdots + a_nz^n$, where the coefficients are complex numbers is called a polynomial function. The domain of $p(z)$ is the whole complex plane.
2. Rational function $(z) = \frac{p(z)}{q(z)}$, where p, q are polynomial functions. The function $r(z)$ is defined at every point of z such that $q(z) \neq 0$.
3. The exponential function e^z if $z = x + iy$ (using Euler's formula $e^{iy} = \cos y + i\sin y$) is defined by

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

4. Let $z = re^{i\theta}$ and suppose $\log z = \log r + i(\theta + 2\pi k), (k = 0, \pm1, \pm2, \dots)$, then we have $e^{\log z} = z, (z \neq 0)$ and this is a multi-valued logarithm function of a nonzero complex variable. The principal value of $\log$ function is given by $\log z = \log r + i\theta$ This is well defined and single-valued when $z \neq 0$.
5. The sine and cosine functions of a complex variable z are defined as follows:

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \qquad \cos z = \frac{e^{iz} + e^{-iz}}{2}.$$

6. The hyperbolic sine and hyperbolic cosine of a complex variable are defined by

$$\sinh z = \frac{e^z - e^{-z}}{2}, \qquad \cosh z = \frac{e^z + e^{-z}}{2}.$$

12.4 CAUCHY- RIEMANN EQUATIONS

To determine whether a complex function $f(z)$ is differentiable, it is necessary to examine the relationship between its real and imaginary components. These relationships are expressed through the **Cauchy–Riemann equations** (C-R for short) which form the basic conditions that a function of the complex variable z must satisfy for differentiability.

Theorem 1 (Necessary conditions for differentiability): Let $f(z) = u(x,y) + iv(x,y)$ and suppose $f'(z_0)$ exists for $z_0 = x_0 + iy_0$. Then all the first order partial derivatives of $u(x,y)$ and $v(x,y)$ exist at (x_0, y_0) and satisfy the equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}\ ,\ \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \qquad (12.1)$$

at that point.

Proof:

Let $f(z) = u(x,y) + iv(x,y)$ be differentiable at z_0. Take $z_0 = x_0 + iy_0, h = h_1 + ih_2$ and

$f(z+h) - f(z) = [u(x_0 + h_1, y_0 + h_2) - u(x_0, y_0)] + i[v(x_0 + h_1, y_0 + h_2) - v(x_0, y_0)]$

Since $f'(z_0)$ exists, we have

$$f'(z_0) = \lim_{(h_1,h_2)\to(0,0)} \left[\frac{u(x_0 + h_1, y_0 + h_2) - u(x_0, y_0)}{h_1 + ih_2} + i\,\frac{[v(x_0 + h_1, y_0 + h_2) - v(x_0, y_0)]}{h_1 + ih_2}\right] (12.2)$$

must exist and be unique. By taking h to be purely real i.e. $h = h_1$ and $h_1 \to 0$, from (12.2) we get

$$\begin{aligned} f'(z_0) &= \lim_{h_1\to 0}\left[\frac{u(x_0 + h_1, y_0) - u(x_0, y_0)}{h_1} + i\,\frac{v(x_0 + h_1, y_0) - v(x_0, y_0)}{h_1}\right] \\ &= \lim_{h_1\to 0}\left[\frac{u(x_0 + h_1, y_0) - u(x_0, y_0)}{h_1}\right] + i\lim_{h_1\to 0}\left[\frac{v(x_0 + h_1, y_0) - v(x_0, y_0)}{h_1}\right] \\ &= \frac{\partial u}{\partial x}(x_0, y_0) + i\,\frac{\partial v}{\partial x}(x_0, y_0) \\ &= u_x + iv_x \end{aligned}$$

Since $f'(z_0)$ exists so u_x and v_x exists at (x_0, y_0).

Next, by taking h to be purely imaginary i.e. $h = ih_2$ and $h_2 \to 0$, from (12.2) we get

$$f'(z_0) = \lim_{h_2 \to 0}\left[\frac{u(x_0, y_0 + h_2) - u(x_0, y_0)}{h_2}\right] + i\lim_{h_2 \to 0}\left[\frac{v(x_0, y_0 + h_2) - v(x_0, y_0)}{ih_2}\right]$$
$$= -i\frac{\partial u}{\partial y}(x_0, y_0) + \frac{\partial v}{\partial y}(x_0, y_0)$$
$$= -iu_y + v_y$$

Similarly, since $f'(z_0)$ exists,$u_y$and $v_y$also exist at (x_0, y_0).As we know that the limit should be unique and hence the two limits obtained above should be identical at (x_0, y_0) implies

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

Equating the real and imaginary parts, we have

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

These two equations are known as Cauchy-Riemann equations, and they may be written as

$$u_x = v_y, \qquad u_y = -v_x$$

NOTE: The Cauchy–Riemann equations $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$ provide necessary conditions for a function $f(z) = u(x, y) + iv(x, y)$ to be analytic. However, the satisfaction of these equations alone is not sufficient to ensure analyticity. For $f(z)$ to be analytic, it is further required that the first-order partial derivatives of u and v exist and are continuous in a neighbourhood of the point under consideration.These sufficient conditions are stated formally in Theorem 2 below.

Theorem 2 (Sufficient condition for differentiability): Let $f(z) = u(x, y) + iv(x, y)$ be a function defined on some domain containing the point $z_0 = x_0 + iy_0$. If the first partial derivatives of u and v exist and are continuous at (x_0, y_0) and satisfy Cauchy-Riemann equations at (x_0, y_0), then f is differentiable at z_0 and $f'(z_0) = u_x(x_0, y_0) + iv_x(x_0, y_0) = v_y(x_0, y_0) - iu_y(x_0, y_0)$.

Definition: A function $f(z)$is said to be analytic at a point z_0 if it possesses a derivative at every point in some neighbourhood of z_0, however small that neighbourhood may be. An ϵ-neighbourhood of $z_0$is the set of all points z satisfying $|z - z_0| < \epsilon$.
A function $f(z)$is analytic in an open set S if it is differentiable at each point of that set. If $f(z)$ is analytic at every point of the complex plane $\mathbb{C}$, then the function is called **entire**.

Example 3: Let $f(z) = z^3$. If $z = x + iy$ then $f(x + iy) = (x + iy)^3 = (x^3 - 3xy^2) + i(3x^2y - y^3)$. Here $u = x^3 - 3xy^2$ and $v = 3x^2y - y^3$.

$$u_x = 3x^2 - 3y^2 v_x = 6xy$$
$$u_y = -6xy v_y = 3x^2 - 3y^2$$

Since $u_x = v_y$ and $u_y = -v_x$for all values of x and y, it follows that $f(z)$is analytic at every point of the complex plane. Hence, $f(z) = z^3$ is an entire function.

Cauchy–Riemann Equations in Polar Form

In many problems involving circular or cylindrical symmetry, it is convenient to express the Cauchy–Riemann equations in polar coordinates. Let

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

where, $r = \sqrt{x^2 + y^2},\ \theta = \tan^{-1}\left(\frac{y}{x}\right).$

Using the chain rule, we obtain

$$u_r = u_x\cos\theta + u_y\sin\theta\,,$$
$$u_\theta = -ru_x\sin\theta + ru_y\cos\theta\,.$$

Similarly,

$$v_r = v_x\cos\theta + v_y\sin\theta\,,$$
$$v_\theta = -rv_x\sin\theta + rv_y\cos\theta\,.$$

Using the Cartesian Cauchy–Riemann equations

$$u_x = v_y,\ u_y = -v_x,$$

we obtain

$$u_r = \frac{1}{r}v_\theta,$$

and $$v_r = -\frac{1}{r}u_\theta.$$

Hence, the Cauchy–Riemann equations in polar form are

$$\boxed{u_r = \frac{1}{r}\,v_\theta,\ v_r = -\frac{1}{r}\,u_\theta.}$$

These equations are known as the Cauchy–Riemann equations in polar form. They are particularly useful in establishing the analyticity of functions expressed naturally in polar coordinates, such as $f(z) = z^n$, where n is a positive integer. They also simplify the analysis of problems possessing circular or cylindrical symmetry.

The Cauchy–Riemann equations provide the mathematical criterion for testing whether a complex function is differentiable. We now proceed to study analytic functions in greater detail by examining the concepts of limits, continuity and differentiability of complex functions, which form the basis of analyticity.

12.5 ANALYTIC FUNCTIONS

Analytic functions constitute one of the fundamental concepts in complex analysis. To study such functions, it is necessary to understand how the operations of taking limits and derivatives, which are familiar for functions of a real variable x, can be extended to

functions of a complex variable z. Although the notation for functions of a complex variable and the associated operations appears similar to that used in real-variable calculus, the interpretation and analytical consequences in the complex plane are much broader and lead to important theoretical results. The main objective of this section is to introduce analytic functions and examine their basic properties.

Limits:

A function $f(z)$ is said to approach the limit w_0 as $z \to z_0$ if the difference $f(z) - w_0$can be made arbitrarily small whenever $z - z_0$is sufficiently small. In other words, for every positive number ϵ, there exists a positive number δ such that $| f(z) - w_0 | < \epsilon$whenever $0 < | z - z_0 | < \delta$.

The following result gives a connection between limit of a function of a complex variable and the limits of real valued functions of two real variables.

Theorem 3: Let $f(z) = u(x,y) + iv(x,y), z_0 = x_0 + iy_0$and $w_0 = u_0 + iv_0$.Then $\lim_{z \to z_0} f(z) = w_0$ if and only if $\lim_{(x,y) \to (x_0,y_0)} u(x,y) = u_0$ and $\lim_{(x,y) \to (x_0,y_0)} v(x,y) = v_0$.

Continuity:

A complex function $f(z)$ is said to be continuous at a point $z_0$if the function value at that point agrees with the limiting value of the function as $z \to z_0$. The function f is continuous in a domain if it is continuous at all points in that domain.

Example 3: Let $f(z) = \frac{z^2}{z^2+4}$. If $z \to i$, then $f(z) \to \frac{i^2}{i^2+4} = -\frac{1}{3} = f(i)$. Thus $f(z)$ is continuous at $z = i$. When $z = 2i$ or $z = -2i$ then $z^2 + 4 = 0$ and neither $f(2i)$ nor $f(-2i)$ exists. Hence $f(z) = \frac{z^2}{z^2+4}$ is discontinuous both at $2i$ and $-2i$.

Derivatives:

A complex function $f(z)$is differentiable at a point $z_0$if the limit

$$\lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

exists uniquely and remains independent of the path by which Δz approaches zero.

By taking $h = z - z_0$, we can write the definition as

$$f'(z_0) = \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

A point at which the derivative does not exist is called a singular point of the function.

Example 4: The function $f(z) = \frac{1}{z^2+4}$ has singular points where $z^2 + 4 = 0$, i.e. at $z = \pm 2i$.

Example 5: Suppose that $f(z) = z^2$. At any point a,

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{(a+h)^2 - a^2}{h} = \lim_{h \to 0} (2a + h) = 2a$$

So $f'(a) = 2a$. Since a be any arbitrary point, it follows that $f'(z) = 2z$.

12.6 HARMONIC FUNCTIONS

A real-valued function $u(x,y)$ is called harmonic in a domain if its first and second order partial derivatives exist and are continuous and it satisfies Laplace's equation throughout that domain.

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$$

Example 6: Let $u(x,y)=e^x\cos y$.

$$\frac{\partial u}{\partial x}=e^x\cos y,\frac{\partial^2 u}{\partial x^2}=e^x\cos y$$

And $\frac{\partial u}{\partial y}=-e^x\sin y,\frac{\partial^2 u}{\partial y^2}=-e^x\cos y$

Hence $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=e^x\cos y-e^x\cos y=0$

for all values x and y. So u is a harmonic function.

Theorem 4: When a function $f(z)=u+iv$ is analytic in a domain D, its real and imaginary parts u and v each satisfy Laplace's equation and hence are harmonic in D.

Definition: Suppose $u\,(x,y)$ and $v(x,y)$ are harmonic functions in a domain D whose first-order partial derivatives satisfy the Cauchy–Riemann equations (12.1) throughout D. Then v is called a harmonic conjugate of u.

Theorem 5: A function $f(z)=u(x,y)+iv(x,y)$ is analytic in a domain D if and only if v is a harmonic conjugate of u.

NOTE: If v is a harmonic conjugate of u in some domain, it is not in general true that u is a harmonic conjugate of v.

(The Milne–Thomson) method provides a direct technique for constructing an analytic function when either its real part or imaginary part is known.

Let $v(x,y)$ be the harmonic conjugate of $u(x,y)$ so $f(z)=u+iv$ is analytic. First, we find $f'(z)$ in terms of z

$$\begin{aligned}f'(z)&=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}\\&=\frac{\partial u}{\partial x}-i\frac{\partial u}{\partial y}\ (\text{by C}-\text{R equations})\\&=\frac{\partial u}{\partial x}(x,y)-i\frac{\partial u}{\partial y}(x,y)\\&=\frac{\partial u}{\partial x}(z,0)-i\frac{\partial u}{\partial y}(z,0)\end{aligned}$$

(by Milne-Thomson rule replace x by z and y by 0). By integrating, we get

$$f(z) = \int \left[\frac{\partial u}{\partial x}(z,0) - i\frac{\partial u}{\partial y}(z,0)\right] dz + C$$

Separating the real and imaginary parts of $f(z)$ we can find the imaginary part $v(x,y)$.
NOTE: If the imaginary part $v(x,y)$ is given, then

$$f(z) = \int \left[\frac{\partial v}{\partial y}(z,0) + i\frac{\partial v}{\partial x}(z,0)\right] dz + C$$

Separating the real and imaginary parts of $f(z)$ we can find the real part $u(x,y)$.

CHECK YOUR PROGRESS

Q1. Let $z = x + iy$ and suppose that $f(z) = \frac{1}{z}$. Write $f(z)$ in the form $u(x,y) + iv(x,y)$.

Q2. Is the function $f(z) = \frac{z}{\bar{z}}$ continuous at $z = 0$? Justify your answer.

Q3. Show that the function $f(z) = |z|^2$ is differentiable only at $z = 0$.

Q4. If a function $f(z)$ is differentiable at z_0 then show that $f(z)$ is continuous at z_0.

Q5. Find the points where the function $f(z) = \bar{z}$ is differentiable.

Q6. Find a harmonic conjugate $v(x,y)$ of a given harmonic function $u(x,y) = x^2 - y^2$ and hence find the corresponding analytic function.

Q7. Find the analytic function $f(z) = u + iv$ where $v(x,y) = 3x^2y - y^3$.

12.7 LET US SUM UP

- The concept of complex-valued functions was introduced and their representations in both Cartesian and polar forms were discussed.
- Important elementary functions of a complex variable, such as polynomial, rational, exponential, logarithmic, trigonometric and hyperbolic functions, were studied.
- The fundamental concepts of limit, continuity and differentiability for functions of a complex variable were developed in the complex plane.
- The Cauchy–Riemann equations were derived as necessary conditions for differentiability, and the sufficient conditions for analyticity were examined.
- The notion of analytic functions and entire functions was introduced with suitable examples.
- The relationship between analytic functions and harmonic functions was established through Laplace's equation.
- The concept of harmonic conjugates was discussed, and the Milne–Thomson method was introduced for constructing analytic functions when either the real part or imaginary part is known.

✓ 12. 8 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: $f(z) = \frac{x-iy}{(x+iy)(x-iy)} = \frac{x-iy}{x^2+y^2} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}$.

Answer to Q2: The function $f(z) = \frac{z}{\bar{z}}$ is not continuous at $z = 0$ as the $\lim_{z\to 0} f(z)$does not exist. When $z = x + i0$ is a nonzero point in the real axis, then

$$f(z) = \frac{x+i0}{x-i0} = 1.$$

So, by letting $z \to 0$ along the real axis, $\lim_{z\to 0} f(z) = 1$ and when $z = 0 + iy$ is a nonzero point in the imaginary axis, then

$$f(z) = \frac{0+iy}{0-iy} = -1$$

Similarly, by letting $z \to 0$ along the imaginary axis, $\lim_{z\to 0} f(z) = -1$, Since the limit is unique, we conclude that $\lim_{z\to 0} f(z) = \frac{z}{\bar{z}}$ does not exist.

Answer to Q3: Let $z = x + iy$ then $f(z) = x^2 + y^2$. Here $u(x, y) = x^2 + y^2$ and $v(x, y) = 0$. We have $\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}$, if $x = 0$; $\frac{\partial u}{\partial y} = 2y = -\frac{\partial v}{\partial x}$ if $y = 0$. Hence CR equations hold if $x = y = 0$. Therefore $f(z)$ is not differentiable at $z \neq 0$. For $z = 0, \lim_{z\to 0} \frac{f(z)-f(0)}{z-0} = \lim_{z\to 0} \frac{|z|^2}{z} = \lim_{z\to 0} \bar{z} = 0$. This proves the result.

Answer to Q4: Since $f'(z_0) = \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ it follows that

$$\lim_{z\to z_0} f(z) = \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}(z - z_0) + f(z_0) = f'(z_0) \times 0 + f(z_0).$$

Hence proved.

Answer to Q5: Here $f(z) = x - iy$ if $z = x + iy$ so $u = x$ and $v = -y$ so the Cauchy–Riemann equations (12.1) are not satisfied for any value of x and y. So nowhere differentiable.

Answer to Q6: Let, $u(x, y) = x^2 - y^2$, so $u_x = 2x$ and $u_y = -2y$. Since $v(x, y)$ is a harmonic conjugate to $u(x, y)$ we have

$$u_x = v_y . u_y = -v_x \qquad (12.3)$$

From the first equation we get $v_y(x, y) = 2x$. Integrating both side with respect to y, we find that

$$\int v_y(x,y)dy = \int 2xdy$$

$$v(x,y) = 2xy + \phi(x) \qquad (12.4)$$

where ϕ is an arbitrary function of x. Using second equation of (12.3), we have

$$-2y = -2y - \phi'(x)$$

or $\phi'(x) = 0$. By integrating we get $\phi(x) = C$, where C is an arbitrary real number. Substituting this in (12.4), then the function

$$v(x,y) = 2xy + C$$

is a harmonic conjugate of $u(x,y)$. The corresponding analytic function is

$$f(z) = x^2 - y^2 + i(2xy + C) = z^2 + iC$$

Answer to Q7: Here $v(x,y) = 3x^2y - y^3$. So $v_y(x,y) = 3x^2 - 3y^2$ and $v_x(x,y) = 6xy$. By Milne Thomas rule, we find

$$\begin{aligned} f(z) &= \int \left[v_y(z,0) + iv_x(z,0)\right] dz \\ &= \int 3z^2 dz \\ &= z^3 + C \end{aligned}$$

12.9 FURTHER READING

- Brown , J. V. and Churchill,R. V. (1999), Complex Variables and Applications, Mc Graw Hill.
- Ponnusamy, S. (1995), Foundations of Complex Analysis, Narosa
- Dass, H. K., & Verma, R. (2013), Mathematical physics. S. Chand.

12.10 MODEL QUESTIONS

1. Express $f(z) = \cos z$ in the form $u + iv$.
2. Show that $\lim_{z\to(1+i)}(z^2 - z + 2)$ exists and evaluate the limit.
3. Find the points where $f(z) = \frac{z}{z^2+9}$ is discontinuous.
4. Find the singular point of the function $f(z) = \frac{z^2}{z+2i}$.
5. Show that $f'(z)$ does not exist at any point if $f(z) = z - \bar{z}$.
6. Let $f(z) = e^z$. Show that the function is differentiable everywhere and find $f'(z)$.
7. Prove that $f(z) = z^n$ is analytic for all z and $f'(z) = nz^{n-1}$, here n is any natural number. (Hints: Use C-R equations in polar form).

8. Show that the function $f(z) = \sqrt{|xy|}$ if $z = x + iy$ is not analytic at the origin although the C-R equations are satisfied at origin.
9. State why the function $f(z) = 3z^3 + ze^z + e^{-z} + 1$ is entire.
10. Find all values of $\log(1 - i)$.
11. Write the Cauchy-Riemann equations in polar form.
12. Prove that the identity $\sin^2 z + \cos^2 z = 1$ holds for all $z \in \mathbb{C}$.
13. Find all possible values of z such that $e^z = -1$.
14. Find where the function $f(z) = \tan^2 z$ is analytic.
15. Let $f(z) = \sin z$. Find $f'(z)$.
16. Prove that there does not exist any non-constant analytic function f(z) whose real part is constant.
17. Verify whether the function $u(x, y) = x^3 - 3xy^2 + 3x^2 - 3y^2 + 1$ is harmonic.
18. Find a harmonic conjugate $v(x, y)$ of a given harmonic function $u(x, y) = 2x(1 - y)$ and hence find the corresponding analytic function.
19. Find the analytic function $f = u + iv$, where $u(x, y) = \frac{x}{x^2+y^2}$.
20. Find the analytic function $f = u + iv$ if $v(x, y) = e^x \sin y$. Hence find $u(x, y)$.

*** *** ***

UNIT 13: COMPLEX ANALYSIS-II

UNIT STRUCTURE

13.1 Learning Objectives
13.2 Introduction
13.3 Complex Integration
13.3.1 Cauchy Theorem and Integral Formula
13.4 Taylor and Laurent Series
13.5 Poles
13.6 Residues
13.7 Let Us Sum Up
13.8 Answers to Check Your Progress
13.9 Further Reading
13.10 Model Questions

13.1 LEARNING OBJECTIVES

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

- evaluate contour integrals involving complex functions
- apply integral theorems to evaluate integrals and derivatives of an analytic function
- write an analytic function near a non-singular point as an infinite Taylor series
- express an analytic function near a singular point as an infinite Laurent series
- classify the type of singularities of an analytic function
- find the residue at isolated singular points
- evaluate the integrals of analytic function having finitely many singular points.

13.2 INTRODUCTION

In the previous unit, the fundamental properties of functions of a complex variable were studied, including analyticity, differentiability and the Cauchy–Riemann equations. These ideas established the basic framework for understanding complex functions. Building on that foundation, the present unit extends the discussion to integration in the complex plane and introduces some important results associated with analytic functions. The theory of complex integration provides powerful techniques for evaluating contour integrals and offers useful methods for solving many difficult real integrals that arise in advanced mathematics and physics.

13.3 COMPLEX INTEGRATION

Integration plays a central role in complex analysis and provides a deeper understanding of analytic functions. In this section, the theory of integration along paths in the complex plane is developed, leading to important integral theorems and formulas that have wide applications in mathematics, engineering and physics.

We will start by introducing the meaning of a contour in the complex plane and the idea of contour integration involving complex functions.

CONTOURS:

A contour (or arc) Cin the complex plane may be represented parametrically as a set of points $z = z(t) = x(t) + iy(t), a \le t \le b$,where $x(t)$ and $y(t)$ are continuous real-valued functions of the parameter t. In this representation, each value of the parameter t determines a unique point on the curve C, which provides a convenient way to study the geometry of the path and to define integration of complex functions along the contour systematically.

An arc is called simple if it does not cross itself, i.e. if $z(t_1) = z(t_2)$ then $t_1 = t_2$. An arc is called closed if $z(a) = z(b)$. An arc which is simple and closed is known as a simple closed curve or a Jordan curve.

An arc is smooth if the derivatives $x'(t)$ and $y'(t)$ are continuous in $[a, b]$ and $z'(t) = x'(t) + iy'(t) \neq 0$ for all t. The length of a smooth arc is defined as

$$L = \int_a^b |z'(t)|\, dt = \int_a^b [(x'(t))^2 + (y'(t))^2]^{1/2}\, dt.$$

A contour or piecewise smooth arc is a continuous chain of a finite number of smooth arcs. For a contour $z'(t)$ is piecewise continuous.
A contour is called simple closed if its first and last points are the same and it does not cross itself.

Example 1: Let C be a circle with center at $z_0 = (x_0, y_0)$ and radius r. In terms of radius r and angle θ, a simple calculation gives $z = z_0 + r(\cos\theta + i\sin\theta), 0 \le \theta \le 2\pi$. Hence $z - z_0 = re^{i\theta}, 0 \le \theta \le 2\pi$ and $|z - z_0| = r$ is the parametric equation of the circle C. For a circle centre at origin and radius r, the parametric equation of a circle takes the form $z = re^{i\theta}, 0 \le \theta \le 2\pi$ and $|z| = r$.

Example 2: Let C be a simple contour consisting of the line segment from 0 to $1 + 2i$, followed by the line segment from $1 + 2i$ to $2 + 2i$. Then the parametric equation is given by

$$z(t) = \begin{cases} t + i2t\,, 0 \le t \le 1, \\ t + i2\ \,, 1 \le t \le 2. \end{cases}$$

CONTOUR INTEGRALS

A contour (or arc) in the complex plane is a continuous curve that can be described by a parameter. If C denotes a contour, then every point on C may be represented in the parametric form

$$z = z(t) = x(t) + iy(t),\ a \le t \le b$$

where $x(t)$ and $y(t)$ are continuous real-valued functions of the parameter t. As $t$varies from a to b, the point $z(t)$traces the contour from its initial point to its terminal point. This parametrization provides a convenient way to describe curves in the complex plane and forms the basis for defining contour integrals.

Let C be a contour joining the points z_1 and z_2 in the complex plane and let $f(z)$be a continuous function defined on C. The contour integral of $f(z)$ along C is defined as follows:

$$\int_C f(z)dz = \int_C (u+iv)(dx+idy)$$
$$= \int_C (u\,dx - v\,dy) + i\int_C (u\,dy + v\,dx)$$

Remark 1. Suppose that $z(t) = x(t) + iy(t), (a \le t \le b)$ be the parametric equation of the contour C, extending from a point $z_1 = z(a)$ to a point $z_2 = z(b)$. Then we find

$$\int_C f(z)dz = \int_a^b f(z(t))\frac{dz(t)}{dt}dt.$$

2. In defining the contour integral with respect to a closed contour, we say that the contour should be described in the positive sense, i.e., anticlockwise direction.

Example 3: Find the value of the integral $\int_C (z)dz$, where $f(z) = z^2$ and the contour C is the upper half of the unit circle from $z = 1$ to $z = -1$. Suppose $z = e^{i\theta}, 0 \le \theta \le \pi$. So $f(z(\theta)) = e^{i2\theta}, dz = ie^{i\theta}d\theta$.Therefore,

$$\int_C f(z)dz = \int_0^\pi e^{i2\theta}e^{i\theta}\,d\theta = \int_0^\pi e^{i3\theta}\,d\theta = -2/3$$

Some properties of integration are mentioned below:

1. $\int_C \alpha f(z)dz = \alpha\int_C f(z)\,dz,$ where$\alpha$is a complex constant.

2. $\int_C (f(z)+g(z))dz = \int_C f(z)dz + \int_C g(z)\,dz$

3. $\int_{-C} f(z)\,dz = -\int_C f(z)\,dz$

4. If $C = C_1 + C_2$then $\int_C f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz$

5. If $|f(z)| \le M$ for all $z \in C$ then $\left|\int_C f(z)dz\right|$
$\le ML$ where L is the length of the contour C.

This inequality is called the Darboux's inequality. It is very useful particularly in the evaluation of contour integrals.

13.3.1 Cauchy Theorem and Integral Formula

After introducing contour integration, we now consider some fundamental results that govern the behaviour of analytic functions under integration. These results form the basis of many important developments in complex analysis.

A domain or region D in the complex plane is said to be simply connected if every simple closed curve contained entirely in D can be continuously contracted to a single point without leaving the region D. Typical examples of simply connected regions include the interior of a circle, square, or ellipse. A region that does not satisfy this property is called multiply connected. For example, an annular region bounded by two concentric circles is a multiply connected domain.

One of the most important results in complex integration is Cauchy's integral theorem, which establishes a fundamental property of analytic functions integrated over closed contours.

Theorem 1: (Cauchy's Integral Theorem): Let $f(z)$ be analytic in a simply connected domain D and suppose that its derivative $f'(z)$ is continuous throughout D. Then, for every closed contour C lying entirely within D,

$$\oint_C f(z)\,dz = 0$$

A refinement of Cauchy's theorem was later established by Goursat, showing that the continuity of derivative of the function $f(z)$ is not an essential requirement for the validity of the theorem. So, the revised version of Cauchy's theorem, sometimes referred to as Cauchy-Goursat theorem, is stated in the following way:

Theorem 2: (Cauchy-Goursat Theorem): Let $f(z)$ be analytic throughout a simply connected domain D. If C is any simple closed contour contained entirely within D, then

$$\oint_C f(z)\,dz = 0$$

Cauchy's Integral Formula:
Although Cauchy's theorem applies to analytic integrands over closed contours, situations often arise where the integrand contains a singularity at an interior point. In such cases, Cauchy's integral formula provides a useful method for evaluating the integral.
Theorem 3: Let $f(z)$ be analytic in a domain D, and let C be any simple closed contour in D.If a is any point interior to C, then

$$f(a) = \frac{1}{2\pi i}\int_C \frac{f(z)dz}{z-a}$$

Example 4: For the integral $\int_C \frac{dz}{z-1}$, where C is any simple closed contour such the point $z = 1$ is interior to C, if we take $f(z) = 1$ then by Cauchy's integral formula

$$\int_C \frac{dz}{z-1} = 2\pi i f(1) = 2\pi i$$

An important consequence of Cauchy's integral formula is that it provides explicit expressions for the derivatives of an analytic function $f(z)$. This is given in the following results:

Theorem 4: If $f(z)$ is analytic at a point $z = a$, then all its derivatives exist and are analytic at that point, Also

$$f^{(n)}(a) = \frac{n!}{2\pi i}\int_C \frac{f(z)dz}{(z-a)^n + 1}$$

From the above result If a function f is analytic at a point, its derivatives of all orders $f', f'', f''', \ldots$ are guaranteed. Moreover, the derivatives of f are analytic at that point.

13.4 TAYLOR AND LAURENT SERIES

Sequences and Series of Complex Numbers

An infinite sequence $\{z_n\}$ of complex numbers forms the basis for studying series representations of analytic functions. A sequence $\{z_n\}$ is said to converge to a limit z_0 if for every positive number ϵ, there exists a positive integer N such that $|z_n - z_0| < \epsilon$ for all $n \geq N$. When such a limit exists, we write

$$\lim_{n\to\infty} z_n = z_0$$

These ideas naturally lead to infinite series expansions, which are useful in analyzing the local behaviour of analytic functions near ordinary as well as singular points.

Example 5: It is easy to verify that the sequence $z_n = \frac{1}{n} + i$ converges to i, the sequence $z_n = z^n$ converges to 0 if $|z| < 1$ and diverges for $|z| > 1$.

Definition 1 Let $\{z_n\}$ be a sequence of complex numbers. Then the infinite sum $z_1 + z_2 + z_3 + z_4 + \ldots = \sum_{n=1}^{\infty} z_n$ is called a series.

Let $S_N(z)$ denote the sum of the first N terms of the series and $\{S_N\}$ is known to be sequence of partial sums of the series. A series is said to be convergent if $S_N(z)$ tends to a finite limit say S as $n \to \infty$ for all z in a domain D, we then write $\sum_{n=1}^{\infty} z_n = S$. A series which is not convergent is said to be divergent.

Example 6: We can verify that $\sum_{n=0}^{\infty} z_n = \frac{1}{1-z}$, valid for $|z| < 1$. This follows from identity $1 + z + z^2 + \cdots + z^n = \frac{1-z^{n+1}}{1-z}$ for $z \neq 1$ and by taking the partial sum $S_N(z) = 1 + z + z^2 + \ldots + z^{N-1}$. We have if $|z| < 1$ then $\lim_{n\to\infty} S_N(z) = \frac{1}{1-z}$. Thus, we verified.

One of the most important applications of infinite series in complex analysis is the representation of analytic functions by power series. Such representations form the basis of the Taylor and Maclaurin series expansions.

Taylor Series Expansion

Suppose that a function $f(z)$is analytic in a neighbourhood of a point z_0. Then the function can be expressed as a power series in powers of $(z - z_0)$. This representation is known as the Taylor series expansion of $f(z)$ about the point z_0. The coefficients of the series are determined by the derivatives of the function evaluated at $z = z_0$, and the expansion is valid within its radius of convergence.

Theorem 5: Suppose that a function $f(z)$ is analytic at all points within a circle C with center at z_0 and radius R_0. Then for every z inside $C, f(z)$ can be represented by an infinite series of the form

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

where $a_n = \frac{f^{(n)}(z_0)}{n!} n = 0,1,2,3\ldots$. Here $f^{(0)}(z_0) = f(z_0)$. In other words the series converge to $f(z)$ for every z for which $|z - z_0| < R_0$,

$$\boxed{f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n, \ \ | \ z - z_0 \ | < R_0.}$$

This expansion of $f(z)$ is called the Taylor series expansion about the point z_0. Thus, every analytic function can be represented locally by a Taylor series about the point z_0.

Maclaurin Series Expansion

A Maclaurin series is a special case of the Taylor series obtained by choosing the expansion point as the origin, that is, $z_0 = 0$.

When $z_0 = 0$, the Taylor series reduces to

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n, |z| < R_0$$

and is referred to as Maclaurin series expansion. For instance, the Maclaurin series expansion of $f(z) = e^z$ about $z = 0$ is given by

$$e^z \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \quad \text{for} |z| < \infty \tag{13.1}$$

$(\text{Since } f(z) = f'(z) = f''(z) = \ldots = e^z, \text{we have } f(0) = f'(0) = f''(0) = \ldots = 1.)$

Laurent Series Expansion:

Consider the annular domain between two concentric circles C_1 and C_2 centred at z_0. Suppose that the function $f(z)$ is analytic on the circles and in the annular domain and let C be any simple closed contour around z_0 and lying in the annular domain. Then at each point z in the domain, $f(z)$ is represented by a series of positive and negative. In many situations, a function $f(z)$may not be analytic at a point z_0, but it may remain analytic in an annular region surrounding that point, say

$$r_1 <| z - z_0 |< r_2.$$

In such a case, the Taylor series expansion about z_0 cannot be applied because analyticity at the centre z_0 is required. To represent $f(z)$ in such a region, a more general series expansion involving both positive and negative powers of $(z - z_0)$is used. This expansion is called the Laurent series expansion of $f(z)$ about z_0.

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} \qquad (13.2)$$

where

$$a_n = \frac{1}{2\pi i} \int_C \frac{f(z)dz}{(z - z_0)^{n+1}}, \quad n = 0,1,2, \ldots$$

and $$b_n = \frac{1}{2\pi i} \int_C \frac{f(z)dz}{(z - z_0)^{n+1}}, \quad n = 1,2, \ldots$$

This expansion can also be written as following:

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n,$$

where

$$c_n = \frac{1}{2\pi i} \int_C \frac{f(z)dz}{(z - z_0)^{n+1}}, \quad n = 0, \pm 1, \pm 2, \ldots$$

Both the forms are called a Laurent series expansion of $f(z)$.

Remark 2 In practice, the coefficients of a Laurent series are often obtained by algebraic manipulation or series expansion techniques rather than by directly evaluating their integral representations.

Example 7: If we replace z by $1/z$ in the Maclaurin series expansion given in (1), we have

$$e^{1/z} \sum_{n=0}^{\infty} \frac{1}{n!\, z^n} = 1 + \frac{1}{1!\, z} + \frac{1}{2!\, z^1} + \frac{1}{3!\, z^3} + \cdots \quad \text{for } 0 < |z| < \infty. \qquad (13.3)$$

This is the Laurent series expansion of the function $e^{1/z}$ valid for $0 < |z| < \infty$.

13.5 POLES

We will start this section with the definition of singular point.

SINGULARITIES:

A point $z = a$ is called a **singular point** of an analytic function $f(z)$if the function $f(z)$ fails to be analytic at $z = a$. Such a singular point is said to be an **isolated singularity** if there exists a positive number δ such that $f(z)$remains analytic for every point satisfying

$$0 <| z - a |< \delta$$

Thus, the function is analytic in a deleted neighbourhood of the point $z = a$.

For example, $f(z) = \frac{1}{z+1}$ has an isolated singularity at $z = -1$, while $f(z) = \frac{1}{\sin(\pi z)}$has isolated singularities at $z = 0, \pm 1, \pm 2, \ldots$."

CLASSIFICATION OF SINGULARITIES:

If z_0 is an isolated singularity of a function $f(z)$, then there exists a positive number r such that $f(z)$ is analytic for all points in the deleted neighbourhood

$$0 <| z - z_0 |< r.$$

Hence, in this neighbourhood, the function $f(z)$can be expanded in a Laurent series about z_0 of the form

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n.$$

The part of the series containing the negative powers of $(z-z_0)$, namely

$$\sum_{n=1}^{\infty} \frac{a_{-n}}{(z - z_0)^n},$$

is called the **principal part** of the Laurent expansion of $f(z)$. Depending on the form of this principal part, the singularity at z_0 can be classified into different types.

We can classify the singularities as following:

1. $z = z_0$ is a removable singularity if all $b_n = 0$
2. $z = z_0$ is a pole of order m, if $b_m \neq 0$ but $b_n = 0$ for $n > m$. If m = 1 then it is called a simple pole.
3. $z = z_0$ is an essential singularity iff $b_n \neq 0$ for infinitely many negative integer n.

Example 8: The function $f(z) = \frac{\sin z}{z}$ has a removable singularity at $z = 0$.

Example 9: $f(z) = \frac{z+2}{z-1}$ has a simple pole at $z = 1$ and $f(z) = \frac{z^2+5}{(z-2)^2}$ has a pole of order 2 at $z = 2$. Also $f(z) = e^{1/z}$ has an essential singularity at $z = 0$.

A function $f(z)$is called **meromorphic** if it is analytic at every point in the finite complex plane except at a finite number of isolated poles.

Example 10 $f(z) = \frac{z^3}{(z-1)^2((z-2))}$ is a meromorphic function as it has only two poles a double pole at $z = 1$ and a simple pole at $z = 2$.

Branch Point

Not all complex functions are single-valued. Functions such as $\sqrt{z}$, $\log z$ can assume more than one value for the same complex number z. Such functions are called **multivalued functions**.

A **branch point** is a point in the complex plane around which the value of a multivalued function changes after one complete circuit. Thus, the function does not return to its original value after encircling the point once. Unlike poles and isolated singularities, a branch point is **not an isolated singularity.**

Branch Cut

To make a multivalued function single-valued, a curve is introduced in the complex plane across which the function is not allowed to cross. This curve is called a **branch cut.**

A branch cut removes the ambiguity associated with multivalued functions and defines a single-valued branch of the function.

Example 11: $\boldsymbol{f(z) = \sqrt{z}}$

The function $f(z) = \sqrt{z}$ has a branch point at $z = 0$.

If the point $z$moves once around the origin, the argument of $z$increases by 2π, giving

$$\sqrt{z} = \sqrt{r}\, e^{i\theta/2}.$$

After one complete revolution, $\theta \to \theta + 2\pi$,

So $$\sqrt{z} \to -\sqrt{z}$$

Hence the function changes its value after one circuit around the origin. Therefore,

$z = 0$ is a **branch point**. Usually, the negative real axis is chosen as the branch cut.

Example 12: $\boldsymbol{f(z) = \log z}$

The logarithmic function is defined as

$$\log z = \ln r + i(\theta + 2n\pi),$$

where n is any integer.

Since infinitely many values are possible, $\log z$ is a multivalued function.

When $z$completes one revolution about the origin,

$$\theta \to \theta + 2\pi,$$

and therefore $$\log z \to \log z + 2\pi i.$$

Hence, $z = 0$ is again a branch point.

The principal branch is usually obtained by taking $-\pi < \theta \leq \pi$, with the branch cut along the negative real axis.

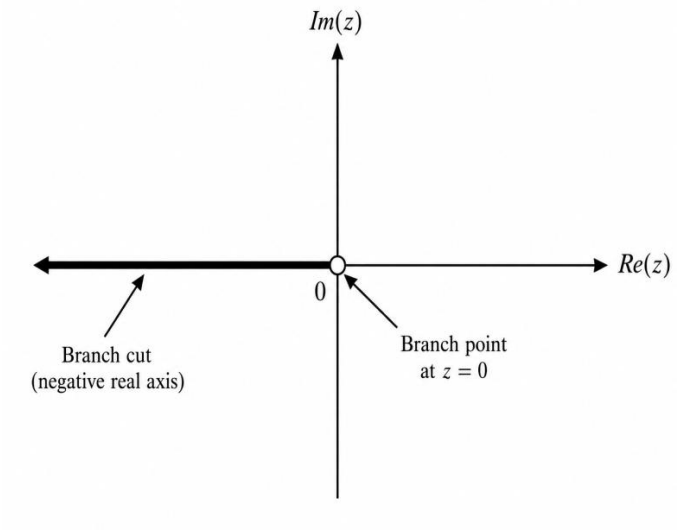


Figure 13.1: *Branch point at $z = 0$and branch cut along the negative real axis in the complex plane.*

Remark: A **pole** is an isolated singularity where the function becomes unbounded, whereas a **branch point** is a non-isolated singularity arising from the multivalued nature of the function. Branch cuts are introduced to define a single-valued branch of such functions.

13.6 RESIDUES

Suppose a function $f(z)$ has an isolated singularity at $z = z_0$. Then there exists $r > 0$such that $f(z)$ is analytic in the deleted neighbourhood $0 < |z - z_0| < r$, and hence $f(z)$ admits a Laurent series expansion about $z_0$of the form

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n.$$

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n},$$

then the coefficient b_1 is called the residue of $f(z)$ at $z = z_0$. Hence the residue of $f(z)$ at $z = z_0$ is also given by

$$[\mathrm{Res} f(z)]_{z=z_0} = \frac{1}{2\pi i} \int_C f(z) dz$$

where C is any closed contour around z_0 such that $f(z)$ is analytic within and on C except at $z = z_0$.

Theorem 6 (Cauchy's Residue Theorem) If a function $f(z)$ analytic inside and on a simple closed contour C, except for a finite many singular points $z_1, z_2, z_3 \ldots, z_k$ inside C, then

$$\int_C f(z)\,dz = 2\pi i \sum_{k=1}^{n} [Res f(z)]_{z=z_k}$$

Jordan's Lemma: Let $f(z)$ be a function that is continuous on the semicircular contour C_R in the upper half-plane and satisfies

$$| f(z) | \leq \frac{M}{R^{1+\delta}}$$

where M and $\delta > 0$ are constants. Then

$$\lim_{R\to\infty} \int_{C_R} e^{iaz} f(z)\,dz = 0,$$

provided $a > 0$.Thus, under suitable conditions, the contribution from the large semicircular arc vanishes as its radius tends to infinity.

Importance: Jordan's Lemma is a useful result in contour integration because it enables the evaluation of many improper real integrals by showing that the integral over the large semicircular arc approaches zero as the contour radius becomes infinitely large. Consequently, only the integral along the real axis and the residues of the enclosed poles contribute to the result.

Remark: Jordan's Lemma is frequently used together with the Cauchy Residue Theorem to evaluate definite integrals involving trigonometric or exponential functions, particularly integrals of the forms

$$\int_{-\infty}^{\infty} e^{iax} f(x)\,dx, \ \int_{-\infty}^{\infty} \cos(ax) f(x)\,dx, \ \int_{-\infty}^{\infty} \sin(ax) f(x)\,dx.$$

Example 13: Let $f(z) = ze^{1/z}$. The value of the integral $\int_C f(z)\,dz = \pi i$, where C is the unit circle $|z| = 1$ can be obtained as following:

The function $f(z) = ze^{1/z}$ has an essential singularity at the point $z = 0$. Using equation (13.3), $[Res f(z)]_{z=0} = 1/2$. Hence by Cauchy's residue theorem, we have

$$\int_C f(z)\,dz = \int_C ze^{1/z} = 2\pi i [Res f(z)]_{z=0} = \pi i.$$

$$\oint_C z\, e^{1/z}\,dz = 2\pi i\, [\text{Res}\, f(z)]_{z=0} = \pi i.$$

Example 14: Evaluate $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$ using a semicircular contour and the Residue Theorem.

Solution: Consider the complex function

$$f(z) = \frac{1}{1+z^2} = \frac{1}{(z-i)(z+i)}.$$

The poles of $f(z)$are $z = i,\ z = -i$.

To evaluate the given real integral, we consider a **closed contour** C consisting of two parts:

(i) the line segment along the real axis from $-R$ to R, and

(ii) the upper semicircular arc of radius R, joining R to $-R$.

Thus, $C = C_1 + C_2$, where C_1denotes the line segment on the real axis and C_2 denotes the upper semicircular arc.

By the Residue Theorem,

$$\oint_C \frac{dz}{1+z^2} = 2\pi i \ \text{Res}\ \left(\frac{1}{1+z^2},\ z = i\right).$$

Since $z = i$ is a simple pole,

$$\text{Res}\ \left(\frac{1}{1+z^2},\ z = i\right) = \lim_{z\to i} \frac{z-i}{(z-i)(z+i)} = \frac{1}{2i}$$

Hence, $$\oint_C \frac{dz}{1+z^2} = 2\pi i\left(\frac{1}{2i}\right) = \pi$$

Now, $$\oint_C \frac{dz}{1+z^2} = \int_{-R}^{R} \frac{dx}{1+x^2} + \int_{\Gamma_R} \frac{dz}{1+z^2}$$

We now examine the contribution from the semicircular arc C_2. To estimate this integral, we use **Darboux's inequality (ML inequality)** discussed earlier in this unit. According to this inequality, the modulus of a contour integral does not exceed the product of the maximum value of the integrand on the contour and the length of the contour.

$$\left|\int_{C_2} \frac{dz}{1+z^2}\right| \le \frac{\pi R}{R^2-1}$$

Since $\lim_{R\to\infty} \frac{\pi R}{R^2-1} = 0$, the integral over the semicircular arc tends to zero as $R \to \infty$.

Since $\lim_{R\to\infty} \frac{\pi R}{R^2-1} = 0$, it follows that $\lim_{R\to\infty} \int_{C_2} \frac{dz}{1+z^2} = 0.$

Therefore, $$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi.$$

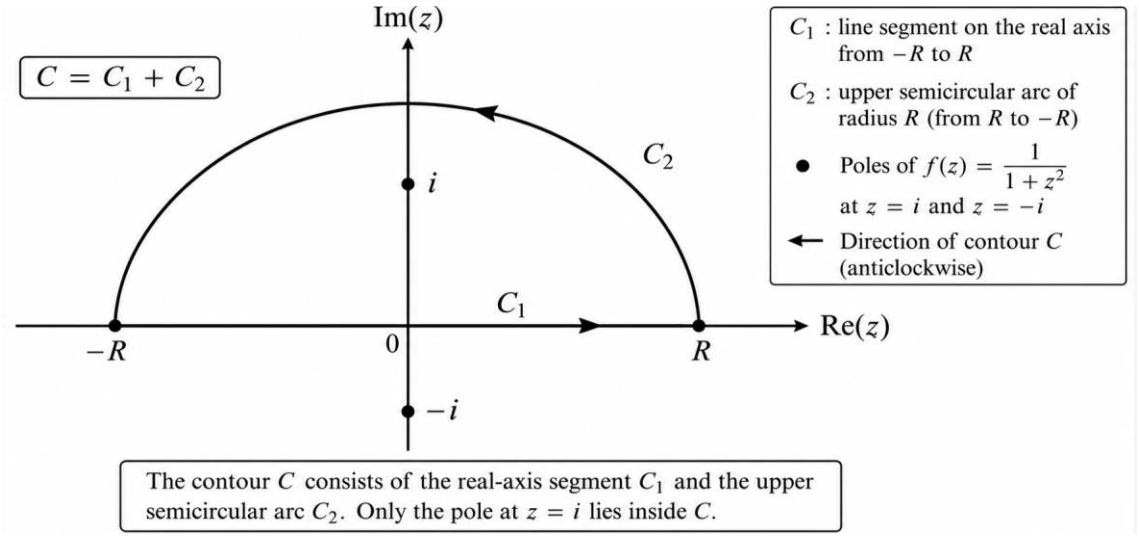


Figure 13.2: *Contour $C = C_1 + C_2$ used to evaluate $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$ using the Cauchy Residue Theorem. Here, C_1 is the real-axis segment from $-R$ to R, C_2 is the upper semi- circular arc, and only the pole at $z = i$ lies inside the contour*.

RESIDUES AT POLES:

The following result provides convenient formulas for evaluating residues at poles of different orders.

Theorem 7 If a function $f(z)$ has an isolated singularity $z = z_0$, then $z = z_0$is a pole of order m if and only if $f(z)$ can be written in the form $f(z) = \frac{g(z)}{(z-z_0)^m}$, where $g(z)$ is analytic and $g(z_0) \neq 0$. Also, for simple pole$[\text{Res } f(z)]_{z=z_0} = \lim_{z\to z_0}(z - z_0)\, f(z)$ and for pole of order $m \geq 2$,

$$[\text{Res } f(z)]_{z=z_0} = \lim_{z\to z_0}\left[\frac{1}{(m-1)!}\frac{d^{m-1}}{dz^{m-1}}((z - z_0)^m f(z))\right]$$

Example 15: Let $f(z) = \frac{1}{z-1}$. Then residue of $f(z)$ at $z = 1$ is given by $[\text{Res } f(z)]_{z=0} = 1$ and if $f(z) = \frac{1}{z(z-2)^4}$ then $f(z)$ has a simple pole at $z = 0$ with $[\text{Res } f(z)]_{z=0} = \frac{1}{16}$and pole of order 4 with $[\text{Res } f(z)]_{z=2} = -\frac{1}{4}$

Remark 3: If a function $f(z) = \frac{p(z)}{q(z)}$, where $p(z)$ and $q(z)$ are analytic at z_0 then

1. $f(z)$ is singular at z_0 if and only if

$q(z_0) = 0$ and $p(z_0) \neq 0$, then f(z) is singular at z_0

2. If $f(z)$ has a simple pole at $z = z_0$ and if $q(z_0) = 0$, $q'(z_0) \neq 0$, and $p(z_0) \neq 0$, then the residue of $f(z)$at $z = z_0$is

$$\boxed{\text{Res}(f; z_0) = \frac{p(z_0)}{q'(z_0)}}.$$

CHECK YOUR PROGRESS

Q.1: Find the parametric representation of a simple closed curve defined by

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

Q.2: Evaluate $\int_0^{1+i}(x^2+iy)dz$ along the path $y=x$.

Q.3: Show that $\int_C \frac{z}{z^2+4}dz=0$, where C is the closed contour $|z|=1$.

Q.4: Evaluate the integral $\int_C \frac{dz}{z^2+1}$, where C is the circle $|z|=2$.

Q.5: Show that $\int_C \frac{\sin 2z}{z^4}dz=\frac{-8\pi i}{3}$, where C is the circle $|z|=1$.

Q.6: Find the Taylor series expansion of $f(z)=\sin z$ about the point $z=0$.

Q.7: Find the Laurent series expansion of $f(z)=\frac{1}{z^3(1-z)}$ valid in the domain $0<|z|<1$ and $1<|z|<\infty$

Q.8: Calculate the residue of $f(z)=\frac{z^2}{(z-1)^2(z+2)}$ at its poles.

13.7 LET US SUM UP

- The concept of contour and contour integration in the complex plane was introduced, and its basic properties were discussed.
- Cauchy's Integral Theorem, the Cauchy–Goursat Theorem and Cauchy's Integral Formula were studied as fundamental tools in complex integration.
- The existence of derivatives of analytic functions of all orders was established through Cauchy's Integral Formula.
- Taylor series expansion was introduced for representing analytic functions about ordinary points, with the Maclaurin series discussed as a special case of the Taylor series.
- Laurent series expansion was introduced for representing functions in annular regions around singular points.
- Isolated singularities were classified into removable singularities, poles and essential singularities.
- Branch points and branch cuts were introduced in connection with multivalued complex functions.
- The concept of residue and the Cauchy Residue Theorem were introduced for evaluating contour integrals involving isolated singularities.
- Jordan's Lemma and its importance in contour integration were discussed.
- The application of the Cauchy Residue Theorem to the evaluation of real integrals using a semicircular contour was illustrated.

✓13.8 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q.1: The parametric equation of an ellipse

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

Is $x(\theta)=a\cos\theta\,, y(\theta)=b\sin\theta \quad 0\le\theta\le 2\pi.$

Answer to Q.2: Along $y=x$, we have $dy=dx$ and $dz=dx+idy=dx+idx$ joining the point $z=0$ to $1+i$ implies x varies from 0 to 1. Hence

$$\int_0^{1+i}(x^2+iy)\,dz=\int_0^1(x^2+iy)(dx+idx)$$

$$=\int_0^1(x^2+ix)(1+i)dx$$

$$=(1+i)\int_0^1(x^2+ix)dx$$

$$=(1+i)\left[\frac{x^3}{3}+i\frac{x^2}{2}\right]_0^1$$

$$=(1+i)(1/3+i/2)$$

$$=\frac{-1+5i}{6}.$$

Answer to Q.3: The singularities of the function $f(z)=\frac{z}{z^2+4}$ are the points for which $z^2+4=0$ or $z=\pm 2i$ and these singular point lies outside the circle $|z|=1$. So the function is analytic inside and, on the contour, C, the unit circle $|z|=1$. Therefore, by Cauchy's integral theorem

$$\int_C\frac{z}{z^2+4}dz=0.$$

Answer to Q.4: We can rewrite

$$f(z)=\frac{1}{z^2+1}=\frac{1}{(z+i)(z-i)}=\frac{1}{2i}\left(\frac{1}{z-i}-\frac{1}{z+o}\right).$$

So

$$\int_C\frac{dz}{z^2+1}=\frac{1}{2i}\left(\int_C\frac{dz}{z-i}-\int_C\frac{dz}{z+i}\right).$$

By Cauchy's integral formula $\int_C\frac{dz}{z-i}=2\pi i$ and $\int_C\frac{dz}{z+i}=2\pi i$ Therefore,

$$\int_C\frac{dz}{z^2+1}=\frac{1}{2i}\,(2\pi i-2\pi i)=0.$$

Answer to Q.5: The point $z = 0$ lies inside the unit circle $|z| = 1$ and let $f(z) = \sin 2z$. By Cauchy's integral formula for derivative, we have

$$\int_C \frac{f(z)}{z^4} dz = \frac{2\pi i}{3!} f'''(0).$$

Therefore

$$\int_C \frac{\sin 2z}{z^4} dz = \frac{\pi i}{3} \frac{d^3}{dz^3} [\sin 2z]_{z=0}$$

$$= \frac{\pi i}{3} [-8 \cos 2z]_z = 0$$

$$= \frac{-8\pi i}{3}.$$

Answer to Q.6: Given $f(z) = \sin z$, we find

$$f(z) = \sin z \qquad f(0) = 0$$

$$f'(z) \cos z \qquad f'(z) = 1$$

$$f''(z) = -\sin z \qquad f''(0) = 0$$

$$f'''(z) = -\cos z \qquad f'''(0) = -1$$

And so on. Hence, the Taylor series of $f(z) = \sin z$ about 0 is

$$f(z) = \sin z = f(0) + f'(0)z + \frac{f''(0)z^2}{2!}\, 2! + \frac{f'''(0)z^3}{3!} + \cdots$$

$$= z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots$$

Answer to Q.7: The domain $0 < |z| < 1$ is the interior of the unit circle $|z| = 1$, excluding the point $z = 0$. Hence the Laurent series expansion is given as following:

$$f(z) = \frac{1}{z^3(1-z)} = \frac{1}{z^3}(1-z)^{-1}$$

$$= \frac{1}{z^3}(1 + z + z^2 + \cdots)$$

$$= \frac{1}{z^3} + \frac{1}{z^2} + \frac{1}{z} + 1 + z + z^2 + \cdots$$

which is valid in the domain $0 < |z| < 1$.

To obtain a series expansion on the domain $1 < |z| < \infty$, we expand it as follows:

$$f(z) = \frac{1}{z^3(1-z)} = \frac{-1}{z^4\left(1 - \frac{1}{z}\right)}$$

$$-\frac{1}{z^4}\left(1-\frac{1}{z}\right)^{-1}$$

$$=-\frac{1}{z^4}\left(1+\frac{1}{z}+\frac{1}{z^2}+\frac{1}{z^3}\ldots\right)$$

$$=\sum_{n=0}^{\infty}\frac{-1}{z^{n+4}}.$$

This is the required Laurent series valid for $1<|z|<\infty$.

Answer to Q.8: Given $(z)=\frac{z^2}{(z-1)^2(z+2)}$, here $z=1$ is a pole of order 2 and $z=-2$ is a simple pole. So

1. Residue at z =1

$$[Resf(z)]_{z=1}=\lim_{z\to 1}\frac{d}{dz}[(z-1)^2 f(z)]$$

$$=\lim_{z\to 1}\frac{d}{dz}\left[\frac{z^2}{z+2}\right]$$

$$=\lim_{z\to 1}\left[\frac{z^2+4z}{z+2^2}\right]$$

$$=\frac{1+4}{(1+2)^2}$$

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

2. Residue at z = -2

$$[Resf(z)]_{z=-2}=\lim_{z\to -2}(z+2)f(z)$$

$$=\lim_{z\to -2}\frac{z^2}{(z-1)^2}$$

$$=\frac{(-2)^2}{(-2-1)^2}$$

$$=\frac{4}{9}$$

13.9 FURTHER READING

- Brown, J. W., & Churchill, R. V. (1999). *Complex Variables and Applications*. McGraw-Hill.
- Conway, J. B. (1973). *Functions of One Complex Variable*. Springer-Verlag.
- Dass, H. K., & Verma, R. (2013). *Mathematical Physics*. S. Chand.
- Ponnusamy, S. (1995). *Foundations of Complex Analysis*. Narosa Publishing House.

⍰ 13.10 MODEL QUESTIONS

1. Show that $\int_C e^z dz = 0$, where C is any simple closed contour.
2. Let $f(z) = z^2$. Evaluate $\int_C f(z)$ along
   (i) the straight line from (2,0) to (0,2)
   (ii) the straight lines (2,0) to (2,2) and then from (2,2) to (0,2).

3. Prove that $\left|\int_C z^2\, dz\right| < 10$, where the contour C is a straight line joining the point $z = i$ to $z = 2 + i$.

4. Evaluate the integral $\int_C \bar{z}$, where C is the upper half of the unit circle from $z = 1$ to $z = -1$.
5. Determine the domain over which the function $f(z) = \tan z$ is analytic. Evaluate $\int_C \tan z\, dz$, where C is the closed contour $|z| = 1$.

6. Find the value of the integral $\int_C \sec z\, dz$,where C is the unit circle $|z| = 1$

7. Find the value of the integral $\int_C \frac{3z^2+z}{z^2-1} dz$, where C is the circle $|z| = 2$.

8. Evaluate $\int_C \frac{zdz}{(z-1)(z-2)^2}$, where C is the circle $|z-2| = 1/2$ (Hint: Use Cauchy's Integral Formula).
9. Let C be the simple closed contour $|z| = \frac{22}{7}$and $g(z_0) = \int_C \frac{3z^2-6}{z-z_0} dz$. Calculate $g(2)$.

10. Evaluate $\int_C \frac{e^z}{(z+1)^4}$, where C is the circle $|z| = 2$.

11. Evaluate $\int_C \frac{\sin z}{z^2} dz$, where C is the simple closed contour formed by the lines $x = \pm 1$ and $y = \pm 1$.

12. Find the Taylor series expansion of f(z) = sinz about z = π/4.

13. Find the Maclaurin series expansion of the function $f(z) = \cos z$.
14. Find the Maclaurin series expansion of the function $f(z) = \frac{z}{z^4+9}$.

15. Expand $f(z) = \frac{1}{z^2-3z+2}$ in Laurent's series valid in the region $1 < |z| < 2$.

16. Obtain the Laurent series expansion for $f(z) = \frac{z-1}{z^2}$ in the powers of $(z-1)$. Indicate the region of validity.

17. Write the principal part of the following functions at its isolated singular point and classify them in each case.

$$(\mathrm{a})f(z)=\frac{\cos z}{z}\,(\mathrm{b})f(z)=\frac{z^2}{1+z}\,(\mathrm{c})f(z)=\sin\frac{1}{z}\,(\mathrm{d})f(z)=\frac{1-e^{2z}}{z^2}.$$

18. Find the residues of the function $f(z)=\frac{\cos z}{z^2(z-\pi)^3}$ at singular points.

19. Find the residue of $f(z)=\frac{1-e^2z}{z^4}$ at $z=0$.

20. Let $f(z)=\frac{z+1}{z^2-2z}$. Show that $\int_C f(z)\,dz=2\pi i$, where C is the circle $|z|=3$ (Hint: Use Cauchy's Residue Theorem).

21. Using contour integration and the Cauchy Residue Theorem, evaluate

$$\int_{-\infty}^{\infty}\frac{dx}{x^2+4}$$

*** *** ***

UNIT 14: GROUP THEORY

UNIT STRUCTURE

14.1 Learning Objectives
14.2 Introduction
14.3 Basics of Group Theory
14.3.1 Definition
14.3.2 Examples
14.4 Discrete Groups
14.4.1 Cyclic
14.4.2 Dihedral
14.4.3 Permutation groups
14.5 Subgroups
14.6 Lagrange theorem
14.7 Cosets and Conjugacy Classes
14.8 Introduction to Group Representations
14.9 Rotational Groups
14.9.1 SO (2)
14.9.2 SO (3)
14.9.3 U (1)
14.9.4 SU (2)
14.10 Let Us Sum Up
14.11 Answers to Check Your Progress
14.12 Further Reading
14.13 Model Questions

14.1 LEARNING OBJECTIVES

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

- explain the basic concepts and properties of groups with suitable examples
- describe cyclic, dihedral, permutation groups, subgroups, and Lagrange's theorem
- explain cosets, conjugacy classes, and normal subgroups
- describe group representations and distinguish between reducible and irreducible representations
- explain the continuous groups SO(2), SO(3), U(1), and SU(2), and their significance in geometry and physics.

14.2 INTRODUCTION

Group theory is an important branch of abstract algebra that studies mathematical structures called groups. A group consists of a set of elements together with an operation that combines any two elements and satisfies four basic properties: closure, associativity, the presence of an identity element, and the existence of inverses. These

simple rules help us understand how different mathematical objects behave when combined.

The main idea of group theory is to study symmetry. Many patterns and systems in mathematics, science, and engineering show symmetry, and groups provide a clear and systematic way to describe these symmetries. For example, rotations, reflections, and rearrangements (permutations) of objects can all be treated as group operations.

Groups can be classified into different types. They may be finite or infinite, abelian (where the order of operation does not matter) or non-abelian (where order matters). Special kinds of groups, such as cyclic, dihedral, and permutation groups, arise frequently in practical problems. Group theory has many important applications. It is used in solving algebraic equations, studying crystal structures in chemistry, understanding symmetry in physics, and in modern fields such as quantum mechanics, cryptography, and coding theory. Because of its wide range of uses and strong theoretical foundation, group theory is considered a fundamental and powerful tool in modern mathematics.

14.3 BASICS OF GROUP THEORY

Before proceeding further, we adopt simple and standard notations for commonly used number systems. Throughout this chapter, we denote the set of natural numbers by N, whole numbers by W, rational numbers by Q, real numbers by R, and complex numbers by C. These symbols will be used consistently in the discussions that follow.

Additionally, we introduce the following symbols, which will appear frequently and carry the meanings specified below

- ∈: belongs to or is an element of. Example: a \in A means a is an element of the set A.
- ∉: does not belong to. Example: $a \notin A$
- ∀: for all or for every. Example: ∀ $x \in A$
- ∃: there exists. Example: ∃ $x \in A$such that a given property holds.
- ∄: there does not exist.
- ⊆: subset of (possibly equal to). Example: $A \subseteq B$.
- ⊂: proper subset.
- ⊇: superset of.
- ∪: union of sets. Example: $A \cup B$
- ∩: intersection of sets. Example: $A \cap B = \{x \mid x \in A \; and \; x \in B\}$.
- ∅: empty (null) set.
- ⇒: implies.
- ⇔: if and only if (iff).
- |A|: cardinality (number of elements) of the set A.

14.3.1 Definitions

Let S be an any nonempty set. A function $f: s \times s \to s$ is known as binary operation on S. A binary operation $*$ on S is said to be binary if $a * b \in S$; $\forall\ a, b \in s$.

For example-

a) Let N be a set of natural numbers and $+$ defined on s then $(+)$ is a binary operation on N because addition of natural number gives another natural number.
   i.e. $a + b \in N, \quad \forall\ \ a, b \in N$.
b) If we define subtraction $(-)$ on the set of natural no. N then subtraction is not binary operation on N because subtraction of two natural no. is not natural number N.
   i.e. $a - b \notin N$, $\forall a, b \in N$. (e. g. $2,3 \in N$, but $2 - 3 = -1$, $-1 \notin N$).
c) Usual product $(*)$ is a binary operation on N, W, Z, Q, R & C. Similarly, usual addition $(+)$ is also binary operation on N, W, Z, Q, R & C. But $(-)$ is binary on Z, Q, R & C. Usual division is not a binary operation on $\mathrm{Q, R, C}$ because divison by zero is undefined.

Definition of Group
Let G be any non-empty set and $*$ be a binary operation defined on G. Set G together with the operation $*$, is called a group if itsatisfies following conditions:
1) Closure property: For any $a, b \in G$.

$$a * b \in G\ ; \qquad \forall\ \ a, b \in G$$

G is closed with respect to binary operation $*$

2) Associativity: For any $a, b, c \in G$

$$(a * b) * c = a * (b * c)$$

Binary operation $*$ is associative in group G.

3) Existence of identity: For any $a \in G$. If there exists an element $e \in G$, such that

$$a * e \ = \ a = e * a$$

Then, $'e'$ is known as an identity element of G.

4) Existence of inverses: For any $a \in G$ if there exists an element $b \in G$ such that

$$a * b \ = e = b * a$$

Then, $'b'$ is known as inverse element of a. It is denoted by $b = \ a^{-1}$.

Then, G together with binary operation $(*)$ is known as group. It is denoted by $(G,*)$ is a group or G is a group with respect to binary operation $*$.

Abelian group: A group $(G, \ *)$ is called abelian group if it satisfies the commutative property i.e. A group $(G, \ *)$ is called abelian group if

$$\forall\ \ a, b \in G, \ \ a * b = b * a.$$

e.g., $(R, +)$ is abelian group, $(Z, +)$ is abelian group.

($a * b = b * a$; $\forall \ a, b \in G$ is known as commutative property)

14.3.2 Examples

Example 1: Show that set of integers Z is a group w.r.t to binary operation $+$

Solution:

Let Z be the set of integers and usual addition is a binary operation define on Z.

To prove that Z is a group with respect to $'+'$

i) Closure property: We know that addition of two integers is an integer.Z is closed w.r.t $+$.

ii) Associativity: We know that addition of integers is associative.

$$\therefore \quad (a + b) + c = a + (b + c) \ ; \ \forall \ a, b, c \in Z$$

$\therefore$ z is associative w.r.t $'+'$.

iii) Existence of identity: There exist an element $0 \in Z$, such that

$$a + 0 = a = 0 + a; \ \forall \ a \in Z$$

0 is identity element of Z w.r.t $+$.

iv) Existence of inverse: For any $a \in Z$ there exists $a \in Z$ such that

$$a + (-a) = 0 = (-a) + a; \ \forall \ a \in Z.$$

This shows that $'-a'$ is inverse element of a in Z.

All these group axioms are satisfied by the set Z. Hence, $(Z, +)$is a group.

v) Commutative property: We know that addition of integers is commutative. i.e.

$$a + b = b + a \ ; \forall \ a, b \in z$$

14.4 DISCRETE GROUPS

A group is called a discrete group if its elements are distinct and isolated from one another, rather than depending continuously on one or more parameters. Finite groups are important examples of discrete groups. Some infinite groups, such as the additive group of integers $(\mathbb{Z}, +)$, are also discrete. Discrete groups are different from continuous groups, or Lie groups, whose elements depend continuously on one or more parameters. For example, the rotation groups SO (2) and SO (3) are continuous groups because their elements depend continuously on parameters such as the angle of rotation. Some important examples of discrete groups discussed in this unit are:

1. Cyclic groups
2. Dihedral groups
3. Permutation groups

14.4.1 Cyclic

A group $G$with respect to the operation multiplication is cyclic if there exists an element $q \in G$ such that every element in Gcan be written as:

$$G =< g >= \{g^n : n \in Z\}.$$

It means every element of the group G is of the form q^n i.e., if $a \in G$ then a can be written as $a = q^m$, where $m$is some integer.

- The element $q$is called a generator of the group G.
- If the group is finite of order, n, then
$$G = \{q^0, q^1, q^2, \ldots q^{n-1}\}$$
- In terms of additive group, the cyclic group $G$is defined as
$$G = \{nq : q \in G, n \in Z\}$$

Note: Any group G is said to be cyclic if there exists an element $a \in G$ that has order equal to the order of that group. Such element $'a'$ will called the generator of that group.

Key Properties:

1) Every cyclic group is abelian (commutative).
2) Every subgroup of a cyclic group is also cyclic.
3) A cyclic group of order *n* has $\phi(n)$ generators, where ϕ is Euler's totient function.

Some properties of cyclic group:

Properties 1: Every subgroup of a cyclic group is abelian group.

Proof: Let G be a cyclic group. Then there exists an element $g \in G$ such that:

$$G = \{q^n | n \in Z\}$$

Take any two elements $a, b \in G$. Since G is cyclic, there exist integers m and n such that

$$a = q^m , b = q^n$$

We want to show that $ab = ba.$
Now,

$$ab = q^m q^n = q^{m+n}$$

Similarly,

$$ba = q^n q^m = q^{n+m}$$

Since addition is commutative so,

$$q^{m+n} = q^{n+m}$$

This gives,

$$ab = ba$$

Therefore, G is abelian.
Properties 2: If g is the generator of a cyclic group, then its inverse is also a generator of the group.

Proof: Let us consider a group G is a cyclic group.

Therefore,

$$G = < g > = \{g^n : n \in Z\}$$

So, every element of G is of the form g^n for some integer n. Now consider the element g^{-1}. We want to show that $< g^{-1} >= G$. That is, the powers of g^{-1} also generate all elements of G.

Consider any element $g^n \in G$. This gives us

$$g^n = (g^{-1})^{-n}$$

So, every $g^n \in G$ is also in $<g^{-1}>$. In same way, the powers of g^{-1} are of the form

$$(g^{-1})^n = g^{-n}$$

But $g^{-n} \in G$ because Gcontains all powers of g, including negative ones. Therefore,

$$<g^{-1}> = \{(g^{-1})^n = \{g^{-n} : n \in Z\} = \{g^m : g^m\} = G.$$

Thus, g^{-1} is also a generator & G.

14.4.2 Dihedral

The dihedral group is an important concept in group theory. It describes all the symmetries of a regular polygon. These symmetries include both rotations and reflections.

The dihedral group of order $2n$, denoted by D_n, is the group of symmetries of a regular $n-$sided polygon. It consists of n rotations and n reflections, making a total of $2n$ elements.

Elements of D_n:

- **Rotations:** These are rotations about the centre of the polygon by multiples of $\frac{360°}{n}$i.e

  $$0°, \frac{360°}{n}, \frac{2 . 360°}{n}, \ldots$$

  There are n such rotations.

- **Reflections:** These are mirror reflections about symmetry axes passing through vertices or edges. There are also n reflections.

Group Properties

- The operation is composition of symmetries (doing one symmetry after another).
- The identity element e is the rotation by $0°$ (doing nothing).
- For $n \geq 3n$ the group is non-abelian, meaning

$$ab \neq ba$$

The order of operations matters.

Algebraic Presentation:

The group D_n can be generated by two elements r and s with relations

$$r^n = e, \qquad s^2 = e, \qquad srs = r^{-1}$$

where:

- r represents a rotation by $\frac{360^\circ}{n}$,
- s represents a reflection,
- e is the identity element.

Example: D_4 (Symmetries of a Square)

Consider a square, it has 4 sides. The dihedral group D_4 has$2 \times 4 = 8$ elements

- Rotations (4):
    1) $r^0 = e$ (identity)
    2) r^1 (90° rotation)
    3) r^2 (180° rotation)
    4) r^3 (270° rotation)
- Reflections (4): (reflections about different axes)
    1) s
    2) sr
    3) sr^2
    4) sr^3

If we first reflect and then rotate, we usually get a different result than rotating first and then reflecting. This shows that D_4 is non-abelian.

14.4.3 Permutation Groups

A permutation group is a set of permutations that forms a group under the operation of composition of mappings (performing one permutation after another). In other words, a permutation group is a collection of bijections of a set that satisfies the axioms of a group.

Thus, in a permutation group:

- Each mapping rearranges the elements of the set without repetition or omission.
- The composition of two permutations is again a permutation.
- The set of permutations satisfies all the group axioms.

Group Properties of Permutation Groups

For a set of permutations to form a group, the following properties must hold-

1. **Closure:** The composition of any two permutations in the group is also a permutation in the group.
2. **Associativity:** For any permutations $f, g,$ and h,
$$(fg)h = f(gh).$$
3. **Identity:** There exists an identity permutation III that leaves every element unchanged, that is,

$$I(x) = x \quad \text{for all } x$$

4. **Inverse:** For every permutation f, there exists an inverse permutation f^{-1}such that

$$ff^{-1} = f^{-1}f = I$$

Example: Consider the set

$$A = \{1, 2, 3\}$$

Some permutations of this set are:

- $(1,2,3) \rightarrow (2,3,1)$
- $(1,2,3) \rightarrow (3,1,2)$
- The identity permutation $(1,2,3) \rightarrow (1,2,3)$

The set of all permutations of A forms a group called the symmetric group of degree 3, denoted by S_3.

The Symmetric Group

The set S_n of all permutations on n symbols form a finite group of order $n!$ under the operation of composition of mappings. For $n \leq 2$, the group S_n is abelian, whereas for $n > 2$ it is non-abelian.

Proof. Let

$$X = \{a_1, a_2, \dots, a_n\}$$

be a finite set.

Define

$$S_n = \{f : X \rightarrow X \mid f \text{ is one} - \text{to} - \text{one and onto}\}.$$

Thus S_n is the set of all bijections from X to X. Since every bijection corresponds to a rearrangement of the n elements of X, the number of such permutations is

$$|S_n| = n!$$

We now show that S_n forms a group under composition of mappings.

1. Closure: Let $f, g \in S_n$. Since both f and g are bijections from X to X, their composition fg is also a bijection from X to X. Therefore,

$$fg \in S_n.$$

Hence S_n is closed under composition.

2. Associativity: Composition of mappings is associative in general. Thus, for all $f, g, h \in S_n$,

$$(fg)h = f(gh).$$

3. Identity: Define the mapping $I : X \rightarrow X$ by

$$I(x) = x \quad \text{for all } x \in X.$$

Then I is the identity permutation and satisfies

$$fI = If$$

for all $f \in S_n$. Hence S_n contains the identity element.

4. Inverse: Let $f \in S_n$. Since f is a bijection, it has an inverse mapping f^{-1}. The inverse mapping is also a bijection from X to X. Hence

$$f^{-1} \in S_n$$

and

$$ff^{-1} = f^{-1}f = I.$$

Therefore, every element of S_n has an inverse in S_n.

Thus S_n satisfies closure, associativity, identity, and inverse properties. Hence S_n is a group under composition. Since there are $n!$ permutations of n elements, the order of the group is $n!$.

Abelian and Non-Abelian Nature of S_n

If $n = 1$, then S_1 contains only one element, so it is trivially abelian.

If $n = 2$, then

$$S_2 = \{I, (12)\}$$

which has two elements. Every group of order 2 is abelian. For $n > 2$, the symmetric group S_n is non-abelian.
For example, in S_3,
let

$$f = (12), \qquad g = (23).$$

Then

$$fg = (123), \qquad gf = (132).$$

Since

$$fg \neq gf,$$

the group S_3 is non-abelian. Hence for all $n > 2$, S_n is non-abelian.

14.5 SUBGROUPS

A subgroup is a subset of a group that itself forms a group under the same binary operation as the parent group. To qualify as a subgroup, the subset must satisfy the group axioms:

1. Closure.
2. Associativity.
3. Identity.
4. Inverses.

We can define a Subgroup as: Let $(G,*)$ be a group and $S \subset G$. S is called a subgroup of G if, $(S,*)$ is itself a group. If this happens, we write $S < G$.

If G be a group and $S \subset G$. Then $S < G$ if and only if, for each $a, b \in S, ab^{-1} \in S$. This can be proof as below-

Conversely, suppose that $S$is a non-empty subset of G such that

$$ab^{-1} \in S, \forall\, a, b \in S.$$

We show that S is a subgroup of G.

1. Existence of Identity: Since S is non-empty, let $a \in S$. Taking $b = a$, we have

$$aa^{-1} = e \in S.$$

Hence, the identity element $e$belongs to S.

2. Existence of Inverse: Since $e \in S$, for any $a \in S$, applying the given condition to e and a, we obtain

$$ea^{-1} = a^{-1} \in S.$$

Thus, the inverse of every element of S also belongs to S.

3. Closure Property: Let $a, b \in S$. From the preceding step,

$$b^{-1} \in S$$

Applying the given condition to $a$and b^{-1}, we get

$$a(b^{-1})^{-1} \in S.$$

Since $(b^{-1})^{-1} = b$,

$$ab \in S.$$

Hence, S is closed under the group operation.

4. Associativity: Since the operation in S is the same as that in G, and G is a group, associativity is automatically satisfied in S.

Therefore, S satisfies all the group axioms under the operation inherited from G. Hence, $S \leq G$.

Theorem: Every subgroup of a cyclic group is cyclic.

Proof. Let $G = \langle a \rangle$ be a cyclic group generated by a, and let $H \leq G$ be a subgroup of G.

Define

$$S = \{\, n \in Z \mid an \in H \,\} \subseteq Z.$$

Since e\in H and $a^0 = e$, it follows that

$$0 \in S$$

Let $m, n \in S$. Then $a^m, a^n \in H$. Because H is a subgroup of G, it is closed under multiplication and inverses. Hence

$$a^{n-m} = a^n(a^m)^{-1} \in H.$$

Therefore

$$n - m \in S.$$

Thus S is a subgroup of Z. Now every subgroup of Z is of the form kZ there exists an integer k such that

$$S = kZ$$

Hence

$$a^n \in H \quad \text{if and only if} \quad n \in kZa$$

or equivalently,

$$k \mid n .$$

Therefore,

$$H = \langle a^k \rangle$$

Hence H is cyclic.

14.6 LAGRANGE THEOREM

In group theory, one of the foundational results is Lagrange's Theorem, named after the mathematician Joseph-Louis Lagrange. It explores a beautiful relationship between a group and its subgroups — specifically, how the size (or order) of a subgroup relates to the size of the entire group.

In essence, Lagrange's Theorem tells us that in a finite group, the number of elements in any subgroup divides evenly into the number of elements in the group.

This result is powerful because:

- It restricts the possible sizes of subgroups.
- It helps us understand the structure of a group.
- It leads to deeper results in algebra, such as Cauchy's Theorem and the classification of finite groups.

Theorem (Statement of Lagrange's theorem): The order of each subgroup of a finite group is divisor of the order of that group.

Proof:

Let Gbe a finite group of order $n$and H be any subgroup of Gsuch that order of $H$is m. Suppose $h_1, h_2, \ldots, h_m$ are the $m$members of H. Let $a \in G$. Then H_a is a right coset of H in G and we have,

$$Ha = \{h_1 a, h_2 a, \ldots, h_m a\}.$$

Here, $H_a$has m distinct members, since $h_i a = h_j a \Rightarrow h_i = h_j$(But we have given all

h_i's are different).Therefore, each right coset of Hin G has $m$distinct members. Any two distinct right cosets of H in $G$are disjoint, it means they have no common element. Since $G$is a finite group then the number of distinct right coset of H in Gwill be finite, say equal to k. The union of these k distinct right coset of Hin G is equal to G. Thus if, $H_{a_1}, H_{a_2}, \ldots, H_{a_k}$ are the $k$distinct right coset of H in G, then $G = H_{a_1} \cup H_{a_2} \cup \ldots \cup H_{a_k}$.

The number of elements of in $G =$ The number of elements in $+H_{a_1}$ The number of elements in $+ H_{a_2}$ …+. The number of elements in H_{a_k} (Because we know that two distinct right cosets are mutually disjoint).

$$| G | = km \Longrightarrow \quad n = km$$

$$\Longrightarrow k = \frac{n}{m} \Longrightarrow m \text{ is a divisor of } n.$$

$|H|$ is divisor of $| G |$. Hence theorem proved.

14.7 COSETS AND CONJUGACY CLASSES

Cosets are an important concept in group theory and play a crucial role in understanding the structure of groups. They are closely related to subgroups and are fundamental in proving many important results, such as Lagrange's Theorem.

Let G be a group and let H be a subgroup of G. Using the elements of G, we can construct new sets by multiplying elements of G with elements of H. These sets are called cosets of H in G.

Let G be a group and H be a subgroup of G. For any element $a \in G$:

- The left coset of H determined by a is
$$aH = \{ah \mid h \in H\}$$
- The right coset of H determined by a is
$$Ha = \{ha \mid h \in H\}.$$

Thus, a coset is obtained by multiplying every element of the subgroup by a fixed element of the group.

Consider the group

$$G = Z = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

under addition, and let

$$H = 2Z = \{\dots, -4, -2, 0, 2, 4, \dots\}$$

Then the left cosets of H in G are

$$0 + H = H = \{\dots, -4, -2, 0, 2, 4, \dots\}$$

$$1 + H = \{\dots, -3, -1, 1, 3, 5, \dots\}.$$

Thus G is partitioned into two cosets of H.

Conjugacy Classes:
Another important concept in group theory is that of conjugacy. Conjugacy describes how elements of a group can be related through internal symmetries of the group.

Let G be a group and let $a, b \in G$. The element b is said to be conjugate to a in G if there exists an element $g \in G$ such that

$$b = gag^{-1}.$$

The set of all elements conjugate to a in G is called the conjugacy class of a.

Thus, the conjugacy class of a is

$$Cl(a) = \{ gag^{-1} \mid g \in G \}.$$

Example

Consider the symmetric group S_3. The elements of S_3 can be written as

$$S_3 = \{ e, (12), (13), (23), (123), (132)\}.$$

The conjugacy classes of S_3 are:

$$\{e\}$$

$$\{(12), (13), (23)\}$$

$$\{(123), (132)\}.$$

Thus, elements with similar cycle structure belong to the same conjugacy class.

Normal Subgroup: A subgroup H of a group G is called a **normal subgroup** of G if

$$gHg^{-1} = H, \forall\, g \in G.$$

A normal subgroup is denoted by

$$H \trianglelefteq G.$$

Equivalently, H is normal in Gif every left coset of H is equal to the corresponding right coset, that is,

$$gH = Hg, \forall\, g \in G.$$

Thus, a normal subgroup is a subgroup that remains invariant under conjugation by the elements of the group.

14.8 INTRODUCTION TO GROUP REPRESENTATIONS

A group representation provides a way to represent the elements of a group as matrices acting on a vector space. This approach makes it possible to study groups using the tools of linear algebra.

Let G be a group and V be a vector space over a field F. A representation of the group G on V is a homomorphism i.e.

$$\rho : G \rightarrow GL(V)$$

where $GL(V)$ denotes the general linear group of all invertible linear transformations on V.

If V has finite dimension$n$, then $GL(V)$ can be identified with the group

$$GL(n, F),$$

which consists of all invertible $n \times n$ matrices with entries in the field F. Thus, a representation assigns to each element $g \in G$ an invertible matrix $\rho(g)$such that

$$\rho(gh) = \rho(g)\rho(h)$$

for all $g, h \in G$.

Matrix Representation

When the vector space V has dimension n, the representation can be written in terms of matrices. In this case, the representation is a mapping

$$\rho: G \to GL(n, F).$$

Each element of the group is represented by an $n \times n$ invertible matrix.

Note: The degree (or dimension) of a representation is the dimension of the vector space V. If V has dimension n, then the representation is said to be an n-dimensional representation.

14.9 ROTATIONAL GROUPS

A rotation is a transformation that turns a geometric object around a fixed point (in two dimensions) or around a fixed axis (in three dimensions) while preserving distances and angles. When all such rotations that leave an object unchanged are collected together, they form a rotational group.

Thus, rotational groups describe the symmetries of objects that can be obtained by rotations.

Let G be a set of rotations of a geometric object such that:

1. The composition of any two rotations is also a rotation in the set.
2. The identity transformation (no rotation) belongs to the set.
3. Every rotation has an inverse rotation.
4. Composition of rotations is associative.

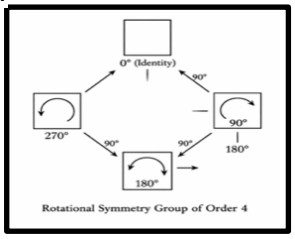
Then the set G forms a rotational group under the operation of composition of rotations.

Rotational Symmetry
A geometric object is said to possess rotational symmetry if it can be rotated about a point or an axis and still appear unchanged. The set of all such rotations forms the rotational symmetry group of the object.

Example 1: Rotations of a Square

A square has the following rotational symmetries about its centre:

1. $0°$ (identity)
2. $90°$
3. $180°$
4. $270°$



These four rotations form a group of order 4 under composition.

In many physical problems, rotations are not restricted to a few discrete angles. Instead, an object may be rotated by any possible angle. Such collections of rotations form continuous groups, often called Lie groups. Important examples of continuous rotational groups include

1. SO (2)
2. SO (3)
3. U (1)
4. SU (2)

14.9.1 SO (2)

The group $SO(2)$ is called the special orthogonal group in two dimensions. It describes all rotations about the origin in the two–dimensional plane. The group $SO(2)$ consists of all 2×2 real matrices A satisfying

$$A^T A = I$$

and. $$\det(A) = 1$$

Each element of $SO(2)$ represents a rotation through an angle θ and has the form

$$R(\theta) = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$$

Here θ represents the angle of rotation.

Physical Applications of SO (2)

The group SO(2) describes rotations in a two-dimensional plane. It is widely used in classical mechanics and geometry to study systems that remain unchanged under planar rotations. Examples include the rotation of a rigid body in a plane, circular motion, and coordinate transformations between two-dimensional reference frames. In physics, SO(2) symmetry helps identify quantities that remain invariant under rotations.

14.9.2 SO (3)

The group SO (3) is called the special orthogonal group in three dimensions**.** It describes all rotations about a fixed point in three-dimensional space and is one of the most important rotational groups in mathematical physics. The group SO (3) consists of all 3×3 real matrices R satisfying

$$R^T R = I$$

and $$\det(R) = 1$$

where R^T denotes the transpose of the matrix R, and I is the identity matrix. The orthogonality condition ensures that lengths and angles remain unchanged under rotation, while the determinant condition distinguishes proper rotations from reflections.

In three dimensions, every rotation can be specified by an axis of rotation and an angle of rotation. Thus, an element of SO (3) represents a rotation through an angle θabout some fixed axis.

For example, the matrix representing rotation about the z-axis through an angle $\theta$is given by

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, rotations about the x-axis and y-axis can also be represented by corresponding orthogonal matrices. The group operation in SO (3) is matrix multiplication. If R_1 and R_2 are two rotation matrices in SO (3), then their product R_1R_2 is also a rotation matrix in SO (3). The identity element is the identity matrix, and each rotation has an inverse given by its transpose.

The group SO (3) is a continuous group (Lie group) because the rotation angle can take any real value continuously. It plays an important role in classical mechanics, quantum mechanics, and crystallography, where rotational symmetry is fundamental. In physics, the generators of SO (3) are closely related to angular momentum operators. Thus, SO (3) provides the mathematical framework for describing rotational symmetry in three-dimensional space.

Physical Applications of SO (3):
The group SO (3) describes rotations in three-dimensional space and is fundamental to classical mechanics. It is used to study the rotational motion of rigid bodies, molecular rotations, and orbital angular momentum. Many physical systems possess SO (3) rotational symmetry, making this group one of the most important symmetry groups in physics.

14.9.3 U (1)

Another important group closely related to rotations is the group $U(1)$. The group $U(1)$ consists of all complex numbers whose absolute value is equal to 1

$$U(1) = \{z \in C \mid \mid z \mid = 1\}.$$

Every element of $U(1)$ can be written in the form

$$z = e^{i\theta}$$

where θ is a real number. Geometrically, these numbers lie on the unit circle in the complex plane.

Physical Applications of U (1):
The group U (1) represents transformations involving complex phase factors of the form $e^{i\theta}$. In quantum mechanics, multiplying a wave function by an overall phase factor does not change any physical observable. The U (1) symmetry therefore plays an important role in quantum theory. It also forms the mathematical basis of the gauge symmetry associated with electromagnetism.

14.9.4 SU (2)

The group $SU(2)$ called the special unitary group of degree 2, is an important group in modern mathematics and physics. The group $SU(2)$ consists of all 2×2 complex matrices U such that

$$U^{\dagger}U = I$$

and

$$det(U) = 1$$

Here $U^{\dagger}U$ denotes the conjugate transpose of the matrix U. A general element of $SU(2)$ can be written as

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^{*} & \alpha^{*} \end{pmatrix}$$

where the complex numbers α and β satisfy

Physical Applications of SU (2):
The group SU (2) plays a fundamental role in quantum mechanics. It describes the symmetry associated with the intrinsic angular momentum (spin) of spin-$\frac{1}{2}$ particles such as electrons. The generators of SU (2) are closely related to the Pauli matrices, which are widely used to represent spin operators. Consequently, SU (2) is essential in the study of atomic physics, quantum mechanics, and many areas of modern theoretical physics.

$$|\alpha|^2 + |\beta^2| = 1.$$

CHECK YOUR PROGRESS

Q 1. Define a Group. Give an example**.**
Q 2. Show that set of natural number N is not group with respect to addition $(+)$.
Q 3. Define a cyclic group with an example.
Q 4. Define a Dihedral group with an example.
Q 5. Show that every subgroup of an abelian group$G$ is a normal subgroup of G.
Q 6. Prove that set of integers is subgroup of set of real number with respect to the operation addition.
Q 7. Is it possible for an abelian group to contain a subgroup that is non-abelian?
Q 8. Let G be a finite group with order 10. Using Lagrange's Theorem, determine all possible orders of the subgroups of G. Provide proper justification.
Q 9. Determine whether the group G of order 20 can be cyclic. If it can be cyclic, give an example of such a group and state how many generators it has.
Q 10. Define the rotational group. Also differentiate between the groups $SO(2), U(1)$ and $SU(2)$.

⏱14.10 LET US SUM UP

- Group theory provides the mathematical framework for studying symmetry through groups, which satisfy the closure, associativity, identity, and inverse axioms.
- Discrete groups, including cyclic, dihedral, and permutation groups, describe important classes of finite symmetries occurring in mathematics and physics.
- Subgroups are subsets that themselves form groups, and the subgroup criterion provides a convenient test for identifying them.
- Lagrange's theorem establishes that the order of every subgroup of a finite group divides the order of the parent group, providing important information about group structure.
- Cosets**,** conjugacy classes, and normal subgroups help explain the internal structure and symmetry properties of groups.
- Group representations express abstract group elements as matrices, enabling the study of groups using linear algebra.
- Rotational groups such as SO(2), SO(3), U(1), and SU(2) describe continuous symmetries in geometry and physics, while generators and Lie algebras provide the mathematical foundation for analysing continuous symmetry groups.

✓ 14.11 ANSWERS TO CHECK YOUR PROGRESS

Answer to Q1: A **group** is a non-empty set G together with a binary operation ($*$) satisfying:

1. Closure: $a * b \in G$for all $a, b \in G$.
2. Associativity: $(a * b) * c = a * (b * c)$.
3. Identity: There exists $e \in G$such that $e * a = a * e = a$.
4. Inverse: For every $a \in G$, there exists $a^{-1} \in G$such that $a * a^{-1} = a^{-1} * a = e$.

For example: The set of integers $(\mathbb{Z}, +)$forms a group under addition. The identity element is $0$and the inverse of $a$is $-a$.

Answer to Q 2: A group must have an element e such that$a + e = e + a = a$. For addition, the identity element is 0. However, $0 \notin N$Therefore, N has no identity element under addition.

Answer to Q3: A **cyclic group** is a group in which every element can be generated by repeatedly applying the group operation to a single element, called the **generator** of the group. If $G$is a cyclic group generated by an element a, then

$$G = \langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}.$$

example: The group $(\mathbb{Z}, +)$of integers under addition is a cyclic group generated by 1(or -1), since every integer can be obtained by repeated addition of 1.

$$\mathbb{Z} = \langle 1 \rangle.$$

Thus, $(\mathbb{Z}, +)$is a cyclic group.

Answer to Q4: A **dihedral group** is the group of all symmetries of a regular polygon, including both rotations and reflections. The dihedral group of a regular n-sided polygon is denoted by $D_n$and contains $2n$ elements. It consists of $n$rotations and n reflections.

example: The symmetry group of a square is the dihedral group D_4. It has 4 rotations $(0°, 90°, 180°, 270°)$ and 4 reflections

Hence, | D_4 |$= 8$.Since the order of operations matters, D_4 is non-abelian.

Answer to Q5: Let G be an abelian group and H be any subgroup of G.
Since $G$is abelian, $gh = hg \forall\ g \in G,\ h \in H$
Multiplying by g^{-1}, we obtain $ghg^{-1} = h$.

Since $h \in H$, it follows that $ghg^{-1} \in H$

Thus, $gHg^{-1} = H \forall\ g \in G$.

Therefore, H is a normal subgroup of G.
Hence, every subgroup of an abelian group is a normal subgroup.

Answer to Q 6. To show Z is a subgroup of R, we verify the subgroup conditions.

1. **Closure:** If $a, b \in Z$, then $a + b \in Z$.
2. **Identity:** The additive identity is 0, and $0 \in Z$.
3. **Inverse:** For every $a \in Z$, the inverse $-a$ is also an integer. Since all subgroup conditions are satisfied, $Z \leq R$. Hence, the set of integers is a subgroup of the real numbers under addition.

Answer to Q7: No. Every subgroup of an abelian group is itself abelian. Since the group operation in the subgroup is inherited from the parent group, the commutative property also holds in the subgroup. Therefore, an abelian group cannot contain a non-abelian subgroup.

Answer to Q 8. Possible orders of subgroups of G: 1,2,5,10

Answer to Q 9. Yes, G can be cyclic. Example: the cyclic group Z_{20} . Number of generators: $\varphi(20) = 8$

Answer to Q10: A rotational group is a set of rotations of a geometric object that forms a group under the operation of composition of rotations. It satisfies closure, associativity, identity, and inverse properties. Difference Between SO (2), U(1), and SU(2)

Group	Meaning	Elements	Application
SO(2)	Special Orthogonal Group in 2D	*2 x 2* real orthogonal matrices with determinant 1	Rotations in a plane
U(1)	Unitary Group of Degree 1	Complex numbers of the form $e^{i\theta}$with modulus 1	Phase transformations, electromagnetism
SU(2)	Special Unitary Group of Degree 2	*2 x 2*complex unitary matrices with determinant 1	Quantum mechanics, spin systems

14.12 FURTHER READING

- Joshi, A. W. (2017). *Elements of group theory for physicists* (5th ed.). New Age International Publishers.
- Bhattacharya, P.B. Jain S.K., Nagpaul S.R. (1994),*Basic Abstract Algebra*, Cambridge Press, 1994.
- Tung, W. K. (1985). *Group theory in physics*. World Scientific Publishing.
- Cornwell, J. F. (1997). *Group theory in physics: An introduction*. Academic Press.
- Khanna V. K. and Bhambri S. K. (2021), *A Course in Abstract Algebra* (5th Edition), Vikas Publication House.

14.13 MODEL QUESTIONS

1. Prove that the collection of all n^{th} roots of unity form a finite abelian group G of order n under the operation of ordinary multiplication.
2. Consider the set of natural numbers N with the operation $a * b = LCM(a, b)$. Examine whether this structure forms a group.
3. Let $(G,*)$ be a non-empty semigroup. Show that it becomes a group if, for every pair of elements $a, b \in G$, the equations
   $$a * x = b \text{ and } y * a = b$$
   have unique solutions $x, y \in G$.
4. Demonstrate that the family of all subsets of a given set, equipped with the operation of set union, forms a semigroup.
5. Prove that a non-empty subset H of a group $(G,*)$ is a subgroup if and only if the following conditions hold
   I. For any $a, b \in H$, the element $a * b$ also belongs to H (closure under the group operation).
   II. For every $a \in H$, the inverse element a^{-1} is also contained in H.
6. Prove that, in any group, the order of an element is the same as the order of its inverse.
7. Show that the set S_3, consisting of all permutations of the symbols {1,2,3} forms a finite non-abelian group of order 3! under permutation composition.
8. State Lagrange's Theorem.
9. Show that the elements of $U(1)$ can be written in the form $e^{i\theta}$.
10. Write the general matrix form of a rotation in $SO(2)$.
11. Describe the continuous symmetry groups SO(2), SO(3), U(1), and SU(2). Discuss one important physical application of each.

*** ***** ***