POST-GRADUATE DEGREE PROGRAMME (CBCS)

M.SC. IN MATHEMATICS

SEMESTER-III

Paper Code: COR 3.1
(Pure and Applied Streams)

Linear Algebra

Special Functions

Integral Equations and Integral Transforms

Self-Learning Material

DIRECTORATE OF OPEN AND DISTANCE LEARNING

UNIVERSITY OF KALYANI

Kalyani, Nadia

West Bengal

Course Content Writers	
Course Name	**Writer**
Block-I : Linear Algebra	Ms. Audrija Choudhury Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani
Block-II : Special Functions	Dr. Chinmay Ghosh Assistant Professor Department of Mathematics Kazi Nazrul University
Block-III : Integral Equations and Integral Transforms	Dr. Meghlal Mallik, Assistant Professor in Mathematics Raiganj Surendranath Mahavidyalaya

August, 2023

Directorate of Open and Distance Learning, University of Kalyani
Published by the Directorate of Open and Distance Learning
University of Kalyani, 741235, West Bengal
Printed by East India Photo Composing Centre, 209A, Bidhan Sarani, Kolkata-700006

Director's Message

Satisfying the varied needs of distance learners, overcoming the obstacle of Distance and reaching the unreached students are the three fold functions catered by Open and Distance Learning (ODL) systems. The onus lies on writers, editors, production professionals and other personnel involved in the process to overcome the challenges inherent to curriculum design and production of relevant Self Learning Materials (SLMs). At the University of Kalyani a dedicated team under the able guidance of the Hon'ble Vice-Chancellorhas invested its best efforts, professionally and in keeping with the demands of Post Graduate CBCS Programmes in Distance Mode to devise a self-sufficient curriculum for each course offered by the Directorate of Open and Distance Learning (DODL), University of Kalyani.

Development of printed SLMs for students admitted to the DODL within a limited time to cater to the academic requirements of the Course as per standards set by Distance Education Bureau of the University Grants Commission, New Delhi, India under Open and Distance Mode UGC Regulations, 2020 had been our endeavor. We are happy to have achieved our goal.

Utmost care and precision have been ensured in the development of the SLMs, making them useful to the learners, besides avoiding errors as far as practicable. Further suggestions from the stakeholders in this would be welcome.

During the production-process of the SLMs, the team continuously received positive stimulations and feedback from Professor **(Dr.) Amalendu Bhunia, Hon'ble Vice-Chancellor, University of Kalyani**, who kindly accorded directions, encouragements and suggestions, offered constructive criticism to develop it with in proper requirements. We gracefully, acknowledge his inspiration and guidance.

Sincere gratitude is due to the respective chairpersons as well as each and every member of PGBOS (DODL), University of Kalyani. Heartfelt thanks are also due to the Course Writers-faculty members at the DODL, subject-experts serving at University Post Graduate departments and also to the authors and academicians whose academic contributions have enriched the SLMs. We humbly acknowledge their valuable academic contributions. I wouldespecially like to convey gratitude to all other University dignitaries and personnel involved either at the conceptual or operational level of the DODL of University of Kalyani.

Their persistent and coordinated efforts have resulted in the compilation of comprehensive, learner-friendly, flexible texts that meet the curriculum requirements of the Post Graduate Programme through Distance Mode.

All the Self Learning Materials are self writing and collected from e-book, journals and websites.

Professor (Dr.) Sanjib Kumar Datta
Director
Directorate of Open and Distance Learning
University of Kalyani

Post Graduate Board of Studies (PGBOS) Members of Department of Mathematics, Directorate of Open and Distance Learning (DODL), University of Kalyani.

Sl. No.	Name & Designation	Role
1	Prof. (Dr.) Samares Pal, Professor & Head, Department of Mathematics, University of Kalyani.	Chairperson
2	Prof. (Dr.) Pulak Sahoo, Professor, Department of Mathematics, University of Kalyani.	Member
3	Dr. Sahidul Islam, Associate Professor, Department of Mathematics, University of Kalyani.	Member
4	Prof. (Dr.) Sushanta Kumar Mohanta, Professor, Department of Mathematics, West Bengal State University.	External Nominated Member
5	Ms. Audrija Choudhury, Assistant Professor, Department of Mathematics, DODL, University of Kalyani.	Member
6	Prof. (Dr.) Sanjib Kumar Datta, Director, DODL, University of Kalyani.	Convener

Core Paper

PURE & APPLIED STREAMS

COR 3.1

Marks : 100 (SEE : 80; IA : 20); Credits : 6

Linear Algebra [Marks 40 : (SEE: 30; IA: 10)]
Special Functions [Marks : 25 (SEE: 20; IA: 05)]
Integral Equations and Integral Transforms [Marks : 35 (SEE: 30; IA: 05)]

Syllabus

Block I

- **Unit 1:** Matrices over a field: Matric polynomial, eigen values and eigen vectors, minimal polynomial.
- **Unit 2:** Linear Transformation (L.T.): Brief overview of L.T., Rank and Nullity of L.T.
- **Unit 3:** Dual space, dual basis, Representation of L.T. by matrices, Change of basis.
- **Unit 4:** Normal forms of matrices: Triangular forms, diagonalization of matrices.
- **Unit 5:** Smith's normal form, Invariant factors and elementary divisors.
- **Unit 6:** Jordan canonical form, Rational (or Natural Normal) form.
- **Unit 7:** Inner Product Spaces: Inner product and Norms. Adjoint of a linear operator, Normal, self adjoint, unitary, orthogonal operators and their matrices.
- **Unit 8:** Bilinear and Quadratic forms: Bilinear forms, quadratic forms, Reduction and classification of quadratic forms, Sylvester's law of Inertia.

Block II

- **Unit 9:** Legendre Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Schlafli's integral formula. Laplace's first and second integral formula. Construction of Legendre differential equation.
- **Unit 10:** Bessel's function: Generating function, Recurrence relation, Representation for the indices 1/2, -1/2, 3/2 and -3/2. Bessel's integral equation. Bessel's function of second kind.

- **Unit 11:** Hermite Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Construction and solution of Hermite differential equation.
- **Unit 12:** Laguerre Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Construction and solution of Laguerre differential equation.
- **Unit 13:** Chebyshev Polynomial: Definition, Series representation, Recurrence relations, Orthogonal property. Construction and solution of Chebyshev differential equation.

Block III

- **Unit 14:** Integral Equation: Symmetric, separable, iterated and resolvent kernel, Fredholm and Voltera integral equation & their classification, integral equation of convolution type, eigen value & eigen function, method of converting an initial value problem (IVP) into a Voltera integral equation, method of converting a boundary value problem (BVP) into a Fredholm integral equation.
- **Unit 15:** Homogeneous Fredholm integral equation of the second kind with separable or degenerate kernel; classical Fredholm theory- Fredholm alternative, Fredholm theorem.
- **Unit 16:** Method of successive approximations: Solution of Fredholm and Voltera integral equation of the second kind by successive substitutions & Iterative method (Fredholm integral equation only), reciprocal function, determination of resolvent kernel and solution of Fredholm integral equation.
- **Unit 17:** Hilbert-Schmidt theory: Orthonormal system of function, fundamental properties of eigen value and function for symmetric kernel, Hilbert theorem, Hilbert-Schmidt theorem.
- **Unit 18:** Integral Transform: Laplace transforms of elementary functions & their derivatives and Diracdelta function, Laplace integral, Lerch's theorem (statement only), property of differentiation, integration and convolution, inverse transform, application to the solution of ordinary differential equation, integral equation and BVP.
- **Unit 19:** Fourier Transform: Fourier transform of some elementary functions and their derivatives, inverse Fourier transform, convolution theorem & Parseval's relation and their application, Fourier sine and cosine transform.
- **Unit 20:** Hankel Transform, inversion formula and Finite Hankel transform, solution of two-dimensional Laplace and one-dimensional diffusion & wave equation by integral transform.

Contents

Director's Message

1 **1**
- 1.1 Introduction . . . 1
- 1.2 Matric Polynomials . . . 1
  - 1.2.1 Characteristic Polynomials . . . 2

2 **7**
- 2.1 Introduction . . . 7
- 2.2 Linear Transformations . . . 7
- 2.3 Algebra Of Linear Transformations . . . 11

3 **16**
- 3.1 Introduction . . . 16
- 3.2 Dual Spaces . . . 16
  - 3.2.1 Double Dual . . . 19
- 3.3 Matrix Representation of Linear Transformations . . . 20

4 **23**
- 4.1 Introduction . . . 23
- 4.2 Diagonalizability . . . 24
  - 4.2.1 Minimal Polynomials and Diagonalizability . . . 30

5 **33**
- 5.1 Introduction . . . 33
- 5.2 Smith's Normal Form . . . 33
- 5.3 Jordan Canonical Forms . . . 34

6 **37**
- 6.1 Introduction . . . 37
- 6.2 Invariant Subspaces . . . 38
- 6.3 Invariant Factors . . . 47
- 6.4 Rational Forms . . . 49

7 **51**
- 7.1 Introduction . . . 51
- 7.2 Inner Product Space . . . 52
- 7.3 Adjoint of a Linear Operator . . . 55
- 7.4 Unitary Operators . . . 57

7.5 Normal Operators . . . 58

8 **60**
8.1 Introduction . . . 60
8.2 Bilinear Forms . . . 61
8.2.1 Symmetric Bilinear Forms . . . 64
8.3 Quadratic Forms . . . 66
8.3.1 Definiteness of a 2 Variable Quadratic Form . . . 66
8.3.2 Definiteness of a 3 Variable Quadratic Form . . . 67
8.3.3 Definiteness and Eigen Values . . . 69
8.4 Sylvester's Law of Inertia . . . 69

9 **70**
9.1 Introduction . . . 70
9.2 Legendre's Equations and Legendre's Polynomials . . . 71
9.3 Rodrigue's Formula for Legendre's Polynomials . . . 77
9.4 Recurrence Relations of Legendre Polynomials . . . 77
9.5 Orthogonality Properties of Legendre's Polynomials . . . 80
9.5.1 Expansion of a given function in a series of Legendre polynomials . . . 82

10 **84**
10.1 Bessel's Equation and Bessel's Functions . . . 84
10.2 Recurrence Relations of Bessel's Equations . . . 90
10.3 Generating Function for Bessel's Functions . . . 92
10.4 Orthogonality Properties of Bessel's Polynomials . . . 93
10.5 Integrals of Bessel's Functions . . . 95

11 **103**
11.1 Introduction . . . 103
11.2 Hermite's Equation and Hermite's Polynomials . . . 104
11.3 Generating Function for Hermite's Polynomials . . . 107
11.4 Rodrigue's Formula for Hermite's Polynomials . . . 108
11.5 Recurrence Relations of Hermite's Polynomials . . . 108
11.6 Orthogonality Properties of Hermite's Polynomials . . . 109

12 **112**
12.1 Laguerre's Equation and Laguerre's Polynomials . . . 112
12.2 Generating Function for Laguerre's Polynomials . . . 113
12.3 Rodrigue's Formula for Laguerre's polynomials . . . 114
12.4 Recurrence Relations of Laguerre's Polynomials . . . 114
12.5 Orthogonality Properties of Laguerre's Polynomials . . . 115

13 **119**
13.1 Introduction . . . 119
13.2 Chebyshev's Equation and Chebyshev's Polynomials . . . 120
13.3 Generating Functions for Chebyshev's Polynomials . . . 121
13.4 Rodrigue's Formula for Chebyshev's Polynomials . . . 122
13.5 Recurrence Relations of Chebyshev's Polynomials . . . 122
13.6 Orthogonality Properties of Chebyshev's Polynomials . . . 124

13.7 Minimax Property 125

14 **127**
14.1 Introduction 127
14.2 Integral Equation 127
14.3 Special kinds of kernels 129
14.4 Integral equations of the convolution type 131
14.5 Eigenvalues and Eigen functions 131
14.6 Leibnitz's rule of differentiation under integral sign 132
14.7 Multiple integral into a single integral: Conversion formula 132
14.8 Method of converting an initial value problem into a volterra integral equation 134
14.8.1 Initial Value Problem 138
14.8.2 Boundary Value Problem 139
14.9 Volterra Integral Equations 141
14.9.1 Solution of Volterra integral equations 141
14.9.2 Determination of Resolvent kernel for Volterra integral equations 141

15 **143**
15.1 Introduction 143
15.2 Classical Fredholm Theory 144
15.2.1 Fredholm's First Fundamental Theorem 144
15.2.2 Alternative Procedure of calculating $\mathbf{B_m(x, t)}$ and $\mathbf{C_m}$ 145
15.2.3 Fredholm Second Fundamental Theorem 147
15.2.4 Fredholm Third Fundamental Theorem 148

16 **149**
16.1 Solution of Fredholm integral equations 149
16.1.1 Method of Successive approximations : 149
16.1.2 Resolvent kernel : 150
16.1.3 Solution in terms of resolvent kernel : 152
16.1.4 Iterative scheme for Fredholm integral equations 153
16.1.5 Solution of Homogeneous Fredholm Integral Equation of the second kind with separable (or Degenerate) kernel 155
16.1.6 Some Problems on Fredholm Integral Equation of the Second kind with separable kernels 159
16.2 Solution of Volterra integral equations 161
16.2.1 Solution of Volterra integral equation in terms of resolvent kernel 161
16.2.2 Method of Successive approximations for solving Volterra integral equation 163

17 **165**
17.1 Hilbert-Schmidt Theory 165
17.1.1 Symmetric Kernels 165
17.1.2 Orthogonal system of functions 165
17.1.3 Fundamental properties of eigenvalues and eigenfunctions of symmetric kernels 166
17.1.4 Hilbert-Schmidt Theorem 169

18 **172**
18.1 Introduction . . . 172
18.1.1 Laplace Transform of Derivatives . . . 175
18.1.2 The Dirac Delta Function . . . 175
18.1.3 Laplace Transform of Integral . . . 175
18.1.4 Multiplication by Powers of t . . . 175
18.1.5 Division by t . . . 176
18.1.6 Inverse Laplace transform . . . 176
18.1.7 Some special types of integral equations . . . 179
18.1.8 Application to solve ordinary differential equations . . . 180

19 **183**
19.1 Introduction . . . 183
19.1.1 The Infinite Fourier Transform . . . 183
19.1.2 Relationship between Fourier transform and Laplace transform . . . 184
19.1.3 Some theorems . . . 184
19.1.4 Problems related to Integral Equations . . . 187
19.1.5 The finite Fourier Transform . . . 188
19.1.6 Problems related to finite Fourier Sine and Cosine transform: . . . 192

20 **194**
20.1 Introduction . . . 194
20.1.1 Definition: Infinite Hankel Transform . . . 194
20.1.2 Definition: Inverse Hankel Transform . . . 194
20.1.3 Some Important Results on Bessel functions . . . 195
20.1.4 Finite Hankel Transform . . . 199
20.2 Laplace transforms and Fourier transforms to solve some partial differential equations . . . 203
20.2.1 Solution of two dimensional Laplace Equation using Finite Fourier Transform (FFT) . 204
20.2.2 Application to Heat Conduction and Wave Equations . . . 207

Unit 1

Course Structure

- Matrices over a field: Matric polynomial, eigen values and eigen vectors, minimal polynomial.

1.1 Introduction

You are already aware of matrices and its various properties such as determinants, characteristic polynomials, eigen values and eigen vectors. We will revisit them in this unit and learn about the minimal polynomial of matrices and read about the characteristic polynomial, the eigen values and eigen vectors using the information of the minimal polynomial.

Objectives

After reading this unit, you will be able to

- find the characteristic polynomial of a matrix
- find the eigen values and eigen vectors of a matrix
- learn the various properties of a matrix associated with its eigen vectors and eigen values and also its characteristic polynomial
- find the minimal polynomial of a matrix
- learn the relationship between minimal and characteristic polynomials of a matrix.

1.2 Matric Polynomials

Let F be a field and A be a matrix with entries from the field F. In this chapter, we are concerned mainly with the matrix polynomials, viz., the characteristic and minimal polynomials. Here, we will consider the underlying field to be either $\mathbb{R}$ or $\mathbb{C}$. Let A be an $n \times n$ matrix over the field $\mathbb{R}$. Then, a matrix polynomial for the matrix A is a polynomial with real coefficients and the variables as the matrix A, that is, if

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

is a real polynomial, then the matrix polynomial evaluated at A is given as

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n$$

where, I is the n-th order identity matrix. Next we will move on to the definition of the characteristic polynomials.

1.2.1 Characteristic Polynomials

Before stating the definition of characteristic polynomials, we will first define the eigen values and eigen vectors of a matrix.

Definition 1.2.1. Let A be an $n \times n$ matrix over the field $\mathbb{R}$. Then, a real number λ is said to be an eigen value of the matrix if there exists a non-zero vector $v \in \mathbb{R}^n$ such that

$$Av = \lambda v \tag{1.2.1}$$

holds. Then the non-zero vector v is said to be the eigen vector corresponding to the eigen value λ.

The equation (1.2.1) reduces to

$$(A - \lambda I)v = 0$$

which is an n-th order linear equation in n variables. This equation has non-trivial solution if

$$\det(A - \lambda I) = 0$$

The above equation is called the characteristic equation (polynomial) for the matrix A. The roots of the characteristic polynomials give us the eigen values of the matrix. It should be noted that the characteristic polynomial is a monic polynomial which has exactly degree n.

Consider the matrix

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

For any real number λ, the equation $\det(A - \lambda I) = 0$ gives

$$\begin{aligned}
\begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} &= 0, \\
\text{or, } \lambda^2 + 1 &= 0, \\
\text{or, } \lambda &= \quad i.
\end{aligned}$$

So, the characteristic equation is $\lambda^2 + 1 = 0$ which has no roots in the real field, but has roots $\pm i$ in the complex field. So, A has eigen values in the complex field but no eigen value in the real field.

Eigen values can also be defined as

Definition 1.2.2. If A is an $n \times n$ matrix over a field F, then $c \in F$ is called an eigen value of A in F if the matrix $(A - cI)$ is singular.

Eigen values are often called characteristic roots, latent roots, eigenvalues, proper values, or spectral values in several roots. We shall call them eigen values throughout. We will now discuss certain properties of characteristic polynomials.

Definition 1.2.3. Let A and B be two $n \times n$ matrices. Then A and B are said to be similar if there exists an invertible matrix P of order n such that

$$A = P^{-1}BP.$$

Theorem 1.2.4. Similar matrices have the same characteristic polynomial.

Proof. Let A and B be two $n \times n$ similar matrices. Then there exists an invertible matrix P such that

$$A = P^{-1}BP.$$

Then,

$$\begin{aligned}
\det(A - \lambda I) &= \det(P^{-1}BP - \lambda I) \\
&= \det(P^{-1}BP - \lambda P^{-1}IP) \\
&= \det(P^{-1}(B - \lambda I)P) \\
&= \det P^{-1}. \det(B - \lambda I). \det P \\
&= \det(B - \lambda I).
\end{aligned}$$

□

We will now move on to define the minimal polynomial of a matrix. Let us start with the following example.

Consider the following matrix

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

Then the characteristic polynomial for A is

$$\begin{vmatrix} 3-\lambda & 1 & -1 \\ 2 & 2-\lambda & -1 \\ 2 & 2 & -\lambda \end{vmatrix} = 0$$

which gives $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 = 0$. Thus, 1 and 2 are the eigen values of A. Find the corresponding eigen vectors!

So the characteristic polynomial for A is $f(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 = 0$. It is obvious that for any other polynomial $g(x)$ in $\mathbb{R}[x]$ (since in this case the underlying field is $\mathbb{R}$. Otherwise we would have taken $F[x]$.), we would have

$$h(x) = g(x)f(x) = 0,$$

or, writing it as

$$h(A) = g(A)f(A) = 0,$$

we can say that the polynomial $h(x)$ annihilates A. All such polynomials $h(x) \in \mathbb{R}[x]$ for which $h(A) = 0$ are called the annihilating polynomial of A. We formally define annihilating polynomial as follows.

Definition 1.2.5. Let A be an $n \times n$ matrix over a field F. Then a polynomial $f(x) \in F[x]$ is called an **Annihilating Polynomial** of A if $f(A) = 0$. By the definition, we can at once say that the characteristic polynomial of A is an annihilating polynomial of A.

We can check a simple fact that the set of all annihilating polynomials of a matrix A forms an ideal I of the polynomial ring $F[x]$ (verify). Now, since F is a field, so the ideal I is necessarily a principal ideal of $F[x]$. It means that there exists a polynomial $m(x) \in I$ such that $I = \langle m(x) \rangle$, that is I is generated by $m(x)$, that is, each element $f(x)$ of I can be written in the form $f(x) = p(x)m(x)$, where, $p(x) \in F[x]$. This $m(x)$ is called the **minimal polynomial** of the matrix A. We formally define the minimal polynomial of a matrix as follows.

Definition 1.2.6. Let A be an $n \times n$ matrix over a field F. Then the minimal polynomial $m(x)$ of A is the unique monic generator of the ideal of all polynomials over F which annihilate A.

Thus, we arrive at the following theorem.

Theorem 1.2.7. Let A be an $n \times n$ matrix over a field F and $m(x)$ be the minimal polynomial of A. Then, $m(x)$ divides each of the annihilating polynomial of A.

Theorem 1.2.8. Let A be an $n \times n$ matrix over a field F. Then the characteristic and minimal polynomials for A have the same roots, except for multiplicities.

Proof. Let m be the minimal polynomial for A. Let c be a scalar. We want to show that $m(c) = 0$ if and only if c is an eigen value. First suppose that $m(c) = 0$. Then

$$p(x) = (x - c)q(x),$$

where, q is a polynomial in F such that $\deg q < \deg p$. By the definition of minimal polynomial, we can say that $q(A) = 0$. Now, choose a vector β such that $q(A)\beta = 0$. Let $\alpha = q(A)\beta$. Then,

$$\begin{aligned} 0 &= m(A)\beta \\ &= (A - cI)q(A)\beta \\ &= (A - cI)\alpha \end{aligned}$$

and thus, α is an eigen value of A.

Now, suppose that c is an eigen value of A, say $A\alpha = c\alpha$ for some $\alpha = 0$. So, by the properties of matrices, we can say that

$$m(A)\alpha = m(c)\alpha. \qquad \square$$

Since $m(A) = 0$ and $\alpha = 0$, we have, $m(c) = 0$. Hence c is a root of the minimal polynomial of A. Thus the theorem.

Example 1.2.9. Consider the matrix of the previous example.

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

We have seen that the characteristic polynomial of the matrix is

$$f(x) = (x - 1)(x - 2)^2$$

Now, since minimal polynomial divides characteristic polynomial and both have same roots (excepting multiplicities), so the most probable candidates for the minimal polynomial are

1. $m(x) = (x - 1)(x - 2)^2$, or,

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

One may check whether $(A - I)(A - 2I) = 0$. If yes, then the second option is our required minimal polynomial. If not, then the characteristic polynomial and minimal polynomials coincide in this case.

There are various ways to find the minimal polynomial of a matrix (by finding the eigen vectors, rank, etc. of the matrix). We will deal with it in details in the upcoming units.

Exercise 1.2.10. 1. Find a 3×3 matrix whose minimal polynomial is x^2.

2. Find the minimal polynomial and eigen values of the following matrix.

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

3. Let a, b, c be elements of a filed F, and let A be the following 3×3 matrix over F:

$$A = \begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}.$$

Prove that the characteristic polynomial for A is $x^3 - ax^2 - bx - c$ and that this is also the minimal polynomial for A.

Unit 2

Course Structure

- Linear Transformation (L.T.): Brief overview of L.T., Rank and Nullity of L.T.

2.1 Introduction

We are already familiar with the idea of linear transformations from our undergraduate times. This unit helps to recapitulate those earlier notions and introduces certain new ideas on the algebra of linear transformations and the ideas of dual spaces of a vector space. We will learn of these things in detail. We will start with formally defining linear transformations, giving a few examples and stating the old theorems with their applications.

Objectives

After reading this unit, you will be able to

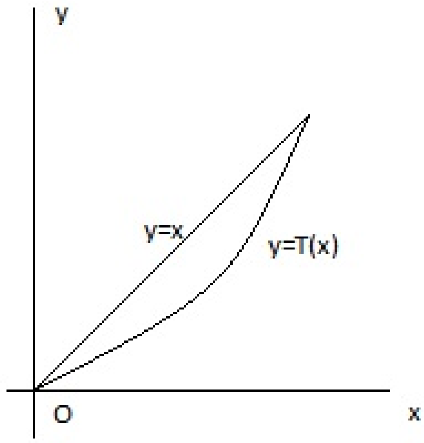
- recapitulate the basic notions of a linear transformation on a vector space
- solve the basic problems related to the representation of a linear transformation (LT) by matrices and change them with basis changes
- solve sums based on the Rank-Nullity theorem
- form the idea of the space of linear transformations from a particular vector space to another and discuss its properties
- form an idea about singular and non-singular linear transformations and conditions for being one

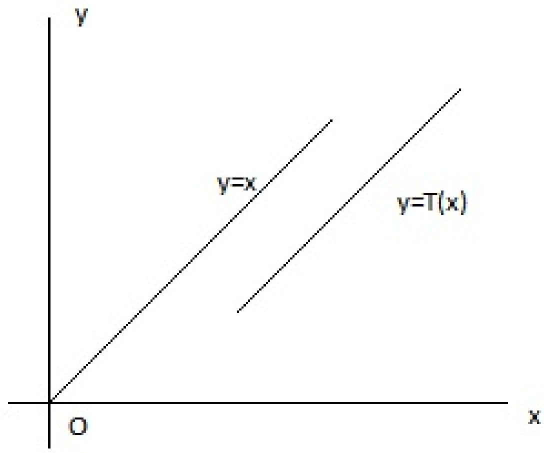
2.2 Linear Transformations

Definition 2.2.1. Let V and W be two vector spaces over the same field F. A linear transformation from V to W is a function that satisfies the following condition

$$T(ca + db) = cT(a) + dT(b)$$

for all c and $d \in F$ and a, b in V.





A simple calculation yields that $T(0) = 0$ always (can you show it?). Thus, for a simple intuitive example, if we consider the vector space $\mathbb{R}^2$ over the field $\mathbb{R}$, then we can say that any function T from $\mathbb{R}$ to itself is a LT if it takes a line passing through the origin to a line passing through the origin. Let us see the following examples.

Example 2.2.2. 1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a function defined as $T(v) = v^2$. Then clearly, T takes the line $y = x$ onto the curve $y = x^2$. Hence, T is not a linear transformation on $\mathbb{R}^2$.

2. Consider another example of T on the same vector space $\mathbb{R}^2$ where T is defined as

$$T(v) = v + \alpha$$

where α is a non-zero element of $\mathbb{R}^2$. Thus, we can see that T takes straight lines onto straight lines but does not take origin to itself. Hence, T is not a LT in this case too.

The above example illustrates a few examples of functions which are not LT. Below given are certain standard examples of a LT which are frequently used.

Example 2.2.3. 1. If V is any vector space, the identity transformation I, defined as $I(v) = v$, is a linear transformation from V into V.

2. The zero transformation 0 on a vector space V, defined as $0(v) = 0$ is also a linear transformation.

Certain other examples include

Example 2.2.4. 1. Let V be the vector space consisting of all continuous functions on the set of real numbers, over the field of reals. Then the integral operator defined as

$$(T(f))(x) = \int_0^x f(t)dt, \qquad f \in V,$$

is a LT on V.

2. Let V be the vector space consisting of all polynomials on the set of real numbers, over the field of reals. Then the differential operator defined as

$$\begin{aligned}(Df)(x) &= c_1 + 2c_2x + \cdots + kc_kx^{k-1} \\ \text{where,} \quad f(x) &= c_0 + c_1x + \cdots + c_kx^k \in V\end{aligned}$$

is a LT on V.

3. Let V be the vector space consisting of all convergent real sequences over the field of reals. Then the limit operator defined as

$$L(x) = \lim_{n\to\infty} x_n, \qquad x = \{x_n\} \in V,$$

is a LT on V.

Theorem 2.2.5. Let V be a finite dimensional vector space and $\{a_1, a_2, \ldots a_n\}$ be a basis of V and $\{b_1, b_2, \ldots b_n\}$ be any set of vectors (not necessarily distinct) in another vector space W under the same field F. Then, there exists a unique LT T from V into W such that

$$T(a_i) = b_i, \quad i = 1(1)n.$$

Proof. Since $\{a_1, a_2, \ldots a_n\}$ is a basis of V, so for any $v \in V$, there exists unique scalars $c_1, c_2, \ldots c_n$ of F such that

$$v = c_1a_1 + \cdots + c_na_n.$$

Then we define T as

$$T(v) = c_1b_1 + \cdots + c_nb_n.$$

Then T is a well-defined rule for associating with each vector v of V to a vector $T(v)$ in W. From the definition, we easily get

$$T(a_i) = b_i, \quad i = 1(1)n.$$

To see that T is linear, let us consider another vector w of V as

$$w = d_1a_1 + \cdots + d_na_n$$

and two other scalars x and y in F. Now,

$$\begin{aligned}xv + yw &= xc_1a_1 + \cdots + xc_na_n + yd_1a_1 + \cdots + yd_na_n \\ &= (xc_1 + yd_1)a_1 + \cdots (xc_n + yd_n)a_n.\end{aligned}$$

Then,

$$\begin{aligned}T(xv + yw) &= (xc_1 + yd_1)b_1 + \cdots (xc_n + yd_n)b_n \\ &= xc_1b_1 + \cdots + xc_nb_n + yd_1b_1 + \cdots + yd_nb_n \\ &= x(c_1b_1 + \cdots + c_nb_n) + y(d_1b_1 + \cdots + d_nb_n) \\ &= xT(v) + yT(w).\end{aligned}$$

Hence, T is linear. Now, let U be another LT from V into W such that $U(a_i) = b_i$, $i = 1(1)n$, then for any vector $v = \sum_{i=1}^n x_ia_i$, we have

$$\begin{aligned}U(v) &= U\left(\sum_{i=1}^n x_ia_i\right) \\ &= \sum_{i=1}^n x_iU(a_i) \\ &= \sum_{i=1}^n x_ib_i.\end{aligned}$$

so that U is exactly the same as the rule as T is defined. Hence, T is unique. $\square$

Example 2.2.6. The vectors $u = (1,2)$, $v = (3,4)$ are linearly independent and therefore form a basis for $\mathbb{R}^2$. Then, by the previous theorem, there exists a LT T from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that $T(u) = (3,2,1)$ and $T(v) = (6,5,4)$. Then, we must be able to find $T(1,0)$ such that

$$(1,0) = cu + dv = c(1,2) + d(3,4)$$

which gives $c = -2$ and $d = 1$. Thus,

$$\begin{aligned} T(1,0) &= -2(3,2,1) + (6,5,4) \\ &= (0,1,2). \end{aligned}$$

There are other interesting subspaces associated with a LT as we will define now.

Definition 2.2.7. Let V and W be vector spaces over the field and let T be a LT from V into W. Then the **Null Space** of T is the set of all vectors v in V such that $T(v) = 0$. This is clearly a subset of V because

1. $T(0) = 0$, so that N is non-empty;
2. if $T(v) = T(w) = 0$, then

$$T(cv + dw) = cT(v) + dT(w) = c0 + 0 = 0$$

so that $cv + dw$ also belongs to the null space. The dimension of the null space of T is called the **Nullity** of T.

Definition 2.2.8. The range of T is a subspace of the space W because if a, b in the range of T, then there exists vectors u and v in V such that $T(u) = a$ and $T(v) = b$. Then for the scalars x and y, $T(xu + yv) = xT(u) + yT(v) = xa + yb$. Hence, $xa + yb$ is also in the range T. The dimension of the range of T is called the **Rank** of T.

Theorem 2.2.9. A LT T is injective if and only if $N = \{0\}$.

Proof. The proof is trivial and has been left as an exercise. $\square$

We have the celebrated Rank-Nullity Theorem for Linear Transformations as follows:

Theorem 2.2.10. Let V and W be vector spaces over the field F and let T be a LT from V into W. Suppose that V is finite-dimensional. Then

$$\text{Rank}(T) + \text{Nullity}(T) = \text{Dim}V.$$

Proof. Let $\{v_1, v_2, \ldots, v_k\}$ be a basis of N, the null space of T. Then, the above basis can be extended to a basis $\{v_1, v_2, \ldots, v_n\}$ of V. We shall now prove that $\{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis for the range of T. The vectors $T(v_1), T(v_2), \ldots, T(v_n)$ certainly span the range of T, and since $T(v_j) = 0$ for $j \leq k$, we see that $T(v_{k+1}), \ldots, T(v_n)$ span the range of T. To check their independence, suppose that there are scalars c_i such that

$$\sum_{i=k+1}^{n} c_i T(v_i) = 0,$$

which gives

$$T\left(\sum_{i=k+1}^{n} c_i v_i\right) = 0$$

and hence, the vector $v = \sum_{i=k+1}^{n} c_i v_i$ is in the null space of T. Since $\{v_1, v_2, \ldots, v_k\}$ is a basis of N, so v can be represented as a finite linear combination of them, that is,

$$\sum_{i=k+1}^{n} c_i v_i = \sum_{i=1}^{k} b_i v_i$$

and hence

$$\sum_{i=1}^{k} b_i v_i - \sum_{i=k+1}^{n} c_i v_i = 0.$$

Since $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, so we have, $b_1 = b_2 = \cdots = b_k = c_{k+1} = \cdots = c_n = 0$. Thus, we have proved the linear independence of $T(v_{k+1}), \ldots, T(v_n)$ and hence it is a basis of the range of T. Thus, when nullity is k, the rank of T is $n - k$, thus giving us the required result. $\square$

Note 2.2.11. We know that any set of vectors with the zero element is always linearly dependent. So, the basis of the null space of T never contains the zero element. Thus, if N does not contain any element other than the zero element, then the nullity of T is zero.

The above theorem has huge applications.

Corollary 2.2.12. A LT T is surjective if and only if Rank$T = \dim V$.

Proof. Left as exercise. $\square$

Exercise 2.2.13. 1. Find the rank and nullity of the following linear transformations:

a. $T(x, y, z) = (x - y, y - z, z - x)$.

b. $T(x, y, z) = (2x, y, 0)$.

c. $T(x, y, z) = (2x + 3z, 4z, 5y - z)$.

2. Let T be a vector space and T a linear transformation from V to V. Prove that the following two statements are equivalent.

   a. The intersection of the range of T and the null space of T is the zero subspace of V.

   b. If $T(T(v)) = 0$, then $T(v) = 0$.

3. Describe explicitly a LT from $\mathbb{R}^3$ to $\mathbb{R}^2$ for which the range space is spanned by the vectors $(1, 0, -1)$ and $(1, 2, 2)$.

2.3 Algebra Of Linear Transformations

In the study of linear transformations from V to W, it is of fundamental importance that the set of these transformations inherits a natural vector space structure. The set of linear transformations from a space V into itself has even more algebraic structure, because ordinary composition of functions provide a "multiplication" of such transformations. Let us see.

Theorem 2.3.1. Let V and W be vector spaces over the field F. Let T and U be linear transformations from V into W. The function $T+U$ defined by

$$(T+U)(v) = T(v) + U(v)$$

is a linear transformation from V into W. If c is any scalar, then the function cT defined by

$$(cT)(v) = cT(v)$$

is a linear transformation from V into W. The set of all linear transformations from V into W, together with the addition and scalar multiplication defined above, is a vector space over the field F.

Proof. Suppose T and U are linear transformations from V into W and $T+U$ is defined as given. Then we first show that $T+U$ is linear. Let $c, d \in F$. Then

$$\begin{aligned}
(T+U)(cu+dv) &= T(cu+dv) + U(cu+dv) \\
&= cT(u) + dT(v) + cU(u) + dU(v) \\
&= c(T(u) + U(u)) + d(T(v) + U(v)) \\
&= c(T+U)(u) + d(T+U)(v).
\end{aligned}$$

Similarly, we can show that for scalar $c \in F$ and some additional scalars x, y, we have

$$\begin{aligned}
(cT)(xu+yv) &= c(T(xu+yv)) \\
&= c(xT(u) + yT(v)) \\
&= cxT(u) + cyT(v) \\
&= x(cT(u)) + y(cT(v)) \\
&= x((cT)(u)) + y((cT)(v))
\end{aligned}$$

This shows that cT is linear. The zero transformation from V into W is also linear. It is a routine exercise to check that the other properties of vector space are satisfied similarly. Hence the result. $\square$

The vector space thus formed, is denoted by the symbol $L(V,W)$. We note that $L(V,W)$ is defined only when V and W are defined over the same field.

Theorem 2.3.2. Let V be an n-dimensional vector space over the field F, and let W be an m-dimensional vector space over the field F. Then the space $L(V,W)$ is finite-dimensional and has dimension mn.

Proof. Let

$$\mathcal{B} = \{v_1, v_2, \ldots, v_n\} \qquad \mathcal{C} = \{w_1, w_2, \ldots, w_m\}$$

be ordered bases for V and W, respectively. For each integers (p,q) with $1 \leq p \leq m$ and $1 \leq q \leq n$, we define a linear transformation $E^{p,q}$ from V into W by

$$\begin{aligned}
E^{p,q}(v_i) &= 0, \quad \text{when } i \neq q \\
&= w_p, \quad \text{when } i = q
\end{aligned}$$

or,

$$E^{p,q}(v_i) = \delta_{iq} w_p.$$

According to our first theorem, there exists a unique linear transformation from V into W satisfying these conditions. The claim is that, these mn transformations $E^{p,q}$ form a basis for $L(V,W)$. Let T be a linear transformation from V into W. For each j, $1 \leq j \leq n$, let $A_{ij}, \ldots, A_{mj}$ be the coordinates of the vector $T(v_j)$ in the ordered basis $\mathcal{C}$, that is,

$$T(v_j) = \sum_{p=1}^{m} A_{pj} w_p.$$

We wish to show that

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq} E^{p,q}. \tag{2.3.1}$$

Let U be the linear transformation in the right hand member of the above equation. Then for each j,

$$\begin{aligned} U(v_j) &= \sum_p \sum_q A_{pq} E^{p,q}(v_j) \\ &= \sum_p \sum_q A_{pq} \delta_{jp} w_p \\ &= \sum_{p=1}^{m} A_{pj} w_p \\ &= T(v_j). \end{aligned}$$

and consequently $U = T$. Now, (2.3.1) shows that $E^{p,q}$ spans $L(V, W)$. We must prove that they are independent. But this is clear from what we did above; for, if the transformation

$$U = \sum_p \sum_q A_{pq} E^{p,q}$$

is the zero transformation, then $U(v_j) = 0$ for each j, so

$$\sum_{p=1}^{m} A_{pj} w_p = 0$$

and the independence if w_p implies that $A_{pj} = 0$ for every p and j. Hence the proof. □

Theorem 2.3.3. Let V, W and Z be vector spaces over the field F. Let $T : V \to W$ and $U : W \to Z$ be linear transformations. Then the composition function UT defined by $UT(v) = U(T(v))$ is a linear transformation from V into Z.

Proof. Left as exercise. □

Definition 2.3.4. If V is a vector space over a field F, then a linear operator on V is a linear transformation from V into V.

In the previous theorem, when $V = W = Z$, and U and T are linear operators on the space V, we see that the composition UT is again a linear operator on V. The space $L(V, V)$ "has a multiplication" defined on it by composition. In this case the operator TU is also defined, and one should note that in general $UT \neq TU$, that is, $UT - TU \neq 0$. We should take special note of the fact that if T is a linear operator on V then we can compose T with T. We shall use the notation $T^2 = TT$, and in general, $T^n = TT \cdots T (n$ factors) for $n = 1, 2, \ldots$. We define $T^0 = I$ if $T \neq 0$.

Theorem 2.3.5. Let V be a vector space over the field F; let U and T_1 and T_2 be linear operators on V and let $c \in F$. Then

1. $IU = UI = U$;
2. $U(T_1 + T_2) = UT_1 + UT_2$; $(T_1 + T_2)U = T_1U + T_2U$;
3. $c(UT_1) = (cU)T_1 = U(cT_1)$.

In everything we have so far discussed, we have left out the invertibility of linear operators. Under what conditions, does a linear operator admit of an inverse, that is, there exists a linear operator T^{-1} for which $TT^{-1} = T^{-1}T = I$?

Definition 2.3.6. A LT T from a space V to another space W is said to be invertible if there exists a LT U such that $TU = UT = I$. Such function U, if it exists, is unique.

We note that the by the theory of functions, we know that a function is invertible if it is bijective. Thus, by the rank-nullity theorem, we can say that the dimensions of both the spaces V and W must be the same. Let us see the following theorem.

Theorem 2.3.7. Let V and W be vector spaces over the field F and let T be a LT from V into W. If T is invertible, then the function T^{-1} is also a LT from W onto V.

Proof. When T is bijective, there exists a uniquely determined function T^{-1} which maps W onto V. To prove the linearity of T^{-1}, let us take two vectors b_1 and b_2 in W and two scalars x and y. Let $a_i = T^{-1}(b_i)$, $i = 1, 2$. Then, we have $T(a_i) = b_i$ for all i. Now, since T is linear,

$$\begin{aligned} T(xa_1 + ya_2) &= xT(a_1) + yT(a_2) \\ &= xb_1 + yb_2 \end{aligned}$$

Thus, $xa_1 + ya_2$ is the unique vector in V such that $T(xa_1 + ya_2) = xb_1 + yb_2$ which means that

$$T^{-1}(xb_1 + yb_2) = xa_1 + ya_2 = xT^{-1}(b_1) + yT^{-1}(b_2)$$

which shows that T^{-1} is linear. □

Definition 2.3.8. A linear transformation T is said to be non-singular if $T(v) = 0$ implies that $v = 0$, that is, if the null space comprises of only the singleton set $\{0\}$. Otherwise, T is said to be singular.

Theorem 2.3.9. Let T be a LT from V into W. Then T is non-singular if and only if T carries each linearly independent subset of V into a linearly independent subset of W.

Proof. First suppose that T is non-singular. Let S be a linearly independent subset of V. If $S = \{v_1, v_2, \ldots, v_k\}$, then $T(v_1), T(v_2), \ldots, T(v_k)$ are linearly independent, for if

$$c_1T(v_1) + c_2T(v_2) + \cdots + c_kT(v_k) = 0$$

and then

$$T(c_1v_1 + c_2v_2 + \cdots + c_kv_k) = 0$$

and since T is non-singular,

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$$

from which it follows that each $c_i = 0$ because S is linearly independent set. This shows that the image of S under T is independent.

Suppose that T carries linearly independent set into linearly independent set. Let a be a non-zero vector in V. Then the set S consisting of the one vector a is independent. The image of S is the set consisting of the one vector $T(a)$, and this set is independent. Thus, $T(a) \neq 0$, because the set consisting of the zero vector alone is independent. This shows that the null space of T is the zero subspace, that is, T is non-singular. □

Theorem 2.3.10. Let V and W be finite-dimensional vector spaces over the field F such that $\dim V = \dim W$. If T is a LT from V into W, the following are equivalent:

1. T is invertible.
2. T is non-singular.
3. T is onto.
4. If $\{v_1, v_2, \ldots, v_n\}$ is a basis for V, then $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W.
5. There is some basis $\{v_1, v_2, \ldots, v_n\}$ for V such that $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W.

Proof. Left as exercise □

Few Probable Questions

1. Show that there exists a unique linear transformation from a finite-dimensional vector space V into another vector space W over the same field sending the basis elements $\{v_1, v_2, \ldots, v_n\}$ to another set of arbitrary vectors $\{w_1, w_2, \ldots, w_n\}$, not necessarily distinct.
2. State and prove the Rank-Nullity Theorem.
3. Show that the space of linear transformations $L(V, W)$ from an n-dimensional space V into an m-dimensional space W is of dimension mn.
4. Show that a non-singular linear transformation takes a basis to a basis.
5. Define non-singular linear transformations. Show that the inverse of a non-singular linear transformation is also so.

Unit 3

Course Structure

- Dual space, dual basis, Representation of L.T. by matrices, Change of basis.

3.1 Introduction

In the previous unit we got the idea of the space $L(V, W)$, where V and W are two vector spaces over the same field F. Now, one may wonder, what if we replace W by the field F. It is easy to see that F is a 1-dimensional vector space over itself. So, if we consider $L(V, F)$, then we can study a new kind of linear transformations that are called functionals and the particular space is called the dual space of V. We will discuss this space in details. Also, We have seen simple representation of a linear transformation by means of its definition. However, as you may recall, linear transformations can also be represented by means of matrices depending upon the basis of the underlying vector spaces. In this unit, we will have a quick recapitulation of this idea.

Objectives

After reading this unit, you will be able to:

- form an idea about the linear functionals on a vector space V
- define the dual basis on a vector space V
- find the dual basis for the corresponding dual space
- define double dual for a vector space and form the corresponding basis
- recapitulate the idea of matrices representing linear transformations

3.2 Dual Spaces

Definition 3.2.1. If V is a vector space over the field F, a linear transformation f from V into the scalar field F is called a linear functional on V.

The concept of linear functional is important in the study of finite-dimensional spaces because it helps to organize and clarify the discussion of subspaces, linear equations, and coordinates.

Example 3.2.2. Let n be a positive integer and F a field. If A is an $n \times n$ matrix with entries in F, then the trace of A is a scalar

$$\mathrm{tr}A = A_{11} + A_{22} + \cdots + A_{nn}.$$

Then it is a linear functional on the matrix space $F^{n\times n}$ (verify!)

Example 3.2.3. Let $[a, b]$ be a closed interval on the real line and let $C([a, b])$ be the space of continuous real-valued functions on $[a, b]$. Then

$$L(g) = \int_a^b g(t)dt$$

defines a linear functional on $C([a, b])$.

Definition 3.2.4. If V is a vector space, then the collection of all linear functionals on V forms a vector space $L(V, F)$ and it is called the **Dual Space** of V. It is also denoted by V^*.

From the knowledge of the dimension of the space $L(V, W)$, we can say that

$$\dim V = \dim V^*.$$

Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then, by the first theorem of this unit, there exists a unique linear functional f_i on V such that

$$f_i(v_j) = \delta_{ij}.$$

In this way, we can obtain from $\mathcal{B}$, a set of n distinct linear functionals $f_1, f_2, \ldots, f_n$ on V. These functionals are also linearly independent. For, suppose

$$f = \sum_{i=1}^{n} c_i f_i.$$

Then,

$$\begin{aligned} f(v_j) &= \sum_{i=1}^{n} c_i f_i(v_j) \\ &= \sum_{i=1}^{n} c_i \delta_{ij} \\ &= c_j. \end{aligned}$$

In particular, if f is the zero functional, $f(v_j) = 0$ for each j and hence the scalars c_j are all 0. Now, $f_1, f_2, \ldots, f_n$ are n linearly independent functionals, and since we know that V^* has dimension n, it must be that $\mathcal{B}^* = \{f_1, f_2, \ldots, f_n\}$ is a basis for V^*. This is called the dual basis of $\mathcal{B}$.

Theorem 3.2.5. Let V be a finite-dimensional vector space over the field F, and let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then there is a unique dual basis $\mathcal{B}^* = \{f_1, f_2, \ldots, f_n\}$ for V^* such that $f_i(v_j) = \delta_{ij}$. For each linear functional f on V we have

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

and for each vector v in V we have

$$v = \sum_{i=1}^{n} f_i(v) v_i.$$

Proof. The above discussion shows that there i s a unique basis which is dual to the basis $\mathcal{B}$. If f is a linear functional on V, then f is some linear combination of f_i as

$$f = \sum_{i=1}^{n} c_i f_i.$$

Also we have observed that the scalars c_j must be given by $c_j = f(v_j)$. Similarly, if

$$v = \sum_{i=1}^{n} x_i v_i$$

is a vector in V, then

$$\begin{aligned} f_j(v) &= \sum_{i=1}^{n} x_i f_j(v_i) \\ &= \sum_{i=1}^{n} x_i \delta_{ij} \\ &= x_j. \end{aligned}$$

So that the unique expression for v as a linear combination of the v_i is

$$v = \sum_{i=1}^{n} f_i(v) v_i.$$

$\square$

Example 3.2.6. Consider $\mathbb{R}^2$ with the basis $\mathcal{B} = \{(2,1),(3,1)\}$. Let us find the dual basis $\mathcal{B}^*$.

Let $v_1 = (2,1)$ and $v_2 = (3,1)$. By definition, $f_i(v_j) = \delta_{ij}$, $i,j = 1,2$. Therefore,

$$f_1(v_1) = \delta_{11} = 1 \Leftrightarrow f_1(2,1) = 1 \Leftrightarrow f_1\left[2(1,0) + 1(0,1)\right] = 1 \Leftrightarrow 2_1(1,0) + f_1(0,1) = 1$$
$$f_1(v_2) = \delta_{12} = 0 \Leftrightarrow f_1(3,1) = 0 \Leftrightarrow f_1\left[3(1,0) + 1(0,1)\right] = 0 \Leftrightarrow 3_1(1,0) + f_1(0,1) = 0$$

Solving the system for $f_1(1,0)$ and $f_1(0,1)$, we get

$$f_1(1,0) = -1 \quad \text{and} \quad f_1(0,1) = 3.$$

Therefore,

$$f_1(x,y) = xf_1(1,0) + yf_1(0,1) = -x + 3y.$$

We can similarly show that

$$f_2(x,y) = xf_2(1,0) + yf_2(0,1) = x - 2y.$$

Therefore, the dual basis is given by $\mathcal{B}^* = \{-x+3y, x-2y\}$.

Now, let us try to check whether the functional $f(x,y) = 8x - 7y$ can be represented in terms of the dual basis just found. According to the preceding theorem, it should follow the equation below.

$$f(x,y) = f(v_1)f_1 + f(v_2)f_2.$$

Now, $f(v_1) = f(2,1) = 8 \cdot 2 - 7 \cdot 1 = 9$ and $f(v_2) = f(3,1) = 8 \cdot 3 - 7 \cdot 1 = 17$. Putting the values in the right hand side of the above equation, we get

$$\begin{aligned} f(v_1)f_1 + f(v_2)f_2 &= 9(-x+3y) + 17(x-2y) \\ &= 8x - 7y \end{aligned}$$

which is the same as $f(x,y)$.

Exercise 3.2.7. Let V be the vector space of all polynomial functions p from $\mathbb{R}$ to $\mathbb{R}$ having degree ≤ 2. Define three linear functionals on V as follows

$$f_1(p) = \int_0^1 p(x)dx, \quad f_2(p) = \int_0^2 p(x)dx, \quad f_3(p) = \int_0^{-1} p(x)dx.$$

Show that $\{f_1, f_2, f_3\}$ is a basis of V^* by exhibiting the basis of V of which it is the dual.

3.2.1 Double Dual

We saw that V^* is a vector space since it satisfies all the axioms of a vector space. So, it is quite tempting to think whether we can further find the dual of the space V^* itself and further continue finding duals beyond that. The set $(V^*)^* = V^{**}$ is known as the double dual of V. In other words, V^{**} is the set of all linear transformations $\phi : V^* \to F$.

Let v be a vector of the vector space V. Then v induces a linear functional L_v on V^* defined by

$$L_v(f) = f(v), \ \ f \in V^*.$$

Then L_v is a linear transformation (verify!). Further, if V is finite-dimensional and $v \neq 0$, then $L_v \neq 0$. In other words, there exists a linear functional f such that $f(v) \neq 0$ (prove it).

Theorem 3.2.8. Let V be a finite-dimensional vector space over the field F. For each vector $v \in V$, define

$$L_v(f) = f(v), \ \ f \in V^*.$$

Then the mapping $v \mapsto L_v$ is an isomorphism of V onto V^{**}.

Proof. We have left it as an exercise to show that for each v, L_v is linear. Now suppose v_1 and v_2 are in V and $c \in F$. Let $w = cv_1 + v_2$. Then for each $f \in V^*$,

$$\begin{aligned} L_w(f) &= f(w) \\ &= f(cv_1 + v_2) \\ &= cf(v_1) + f(v_2) \\ &= cL_{v_1}(f) + L_{v_2}(f) \end{aligned}$$

and so

$$L_w = cL_{v_1} + L_{v_2}.$$

This shows that $v \mapsto L_v$ is a linear transformation from V into V^{**}. This transformation is non-singular; for, according to the remarks above, $L_v = 0$ if and only if $v = 0$. Now, $v \mapsto L_v$ is a non-singular linear transformation from V into V^{**} and since

$$\dim V^{**} = \dim V^* = \dim V.$$

Thus, by theorem 2.3.10, this transformation is invertible and therefore an isomorphism. □

3.3 Matrix Representation of Linear Transformations

As we have mentioned earlier, linear transformation can be represented by matrices. In fact, a single linear transformation can give rise to different matrices which are similar. To each matrix, there is a linear transformation, but there may be many matrices corresponding to a single linear transformation, varying with the change in basis. Let us have an illustration.

Illustration 3.3.1. Let T be a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined as

$$T(x, y) = (x - y, y).$$

Consider the standard **ordered** basis $\mathcal{B} = \{(1, 0), (0, 1)\}$ of $\mathbb{R}^2$. Suppose we are to represent T with respect to the basis $\mathcal{B}$ on both sides. Then the resulting matrix is represented as $[T]_{\mathcal{B}}$. We find it as follows:

$$\begin{aligned} T(1,0) &= (1,0) = \mathbf{1}(1,0) + \mathbf{0}(0,1) \\ T(0,1) &= (-1,1) = -\mathbf{1}(1,0) + \mathbf{1}(0,1) \end{aligned}$$

and the resulting matrix becomes

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Again, if we consider another **ordered** basis $\mathcal{C} = \{(1, 1), (1, 0)\}$ of $\mathbb{R}^2$ as the domain set and the basis $\mathcal{B}$ of the range set. Then we have

$$\begin{aligned} T(1,1) &= (0,1) = \mathbf{0}(1,0) + \mathbf{1}(0,1) \\ T(1,0) &= (1,0) = \mathbf{1}(1,0) + \mathbf{0}(0,1) \end{aligned}$$

and the resulting matrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ is given by

$$[T]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We have certain theorems in connection to these.

Theorem 3.3.1. Let V and W be finite-dimensional vector spaces with ordered bases $\mathcal{B}$ and $\mathcal{C}$ respectively, and let $T : V \to W$ a be linear transformation. Then for each $v \in V$, we have

$$[T(v)]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[v]_{\mathcal{B}}.$$

Theorem 3.3.2. Let V and W be finite-dimensional vector spaces with ordered bases $\mathcal{B}$ and $\mathcal{C}$ respectively, and let $T, U : V \to W$ be linear transformations. Then

1. $[T + U]_{\mathcal{B}}^{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}} + [U]_{\mathcal{B}}^{\mathcal{C}}$.

2. $[aT]_{\mathcal{B}}^{\mathcal{C}} = a[T]_{\mathcal{B}}^{\mathcal{C}}$ for all scalars a.

Theorem 3.3.3. Let U, V, W be finite-dimensional vector spaces with ordered bases $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ respectively. Let $T : U \to V$ and $S : V \to W$ be linear transformations. Then

$$[ST]_{\mathcal{A}}^{\mathcal{C}} = [S]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{A}}^{\mathcal{B}}.$$

The purpose of matrix representation for a linear transformation T is to enable us to analyse T by working with the matrix, say M. If M is easy to work with, we have gained an advantage; if not, we have no advantage. Since different bases lead to different matrices, the "right" choice of basis to obtain a simple matrix M, such as a diagonal matrix, is important. Diagonal matrices are the easiest to work with. For now, we will restrict our attention to the cases when $v = W$. But, before going into details, let us check the following.

Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ and $\mathcal{C} = \{w_1, w_2, \ldots, w_n\}$ be two bases of a vector space V. Then, for each i, we have certain scalars p_{ij} such that

$$\begin{aligned} v_1 &= p_{11}w_1 + p_{12}w_2 + \cdots + p_{1n}w_n \\ v_2 &= p_{21}w_1 + p_{22}w_2 + \cdots + p_{2n}w_n \\ &\vdots \\ v_n &= p_{n1}w_1 + p_{n2}w_2 + \cdots + p_{nn}w_n \end{aligned}$$

which gives

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

Let

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

Then P is called the Transition matrix from the basis $\mathcal{B}$ to $\mathcal{C}$. This transition matrix is invertible. In fact, if Q is the transition matrix from the basis $\mathcal{C}$ to $\mathcal{B}$, then

$$Q = P^{-1}.$$

Now, let us come back to our discussion. We have seen that a linear transformation can have various matrix representations depending upon the choice of basis. Now, what strikes us is that whether there is certain relationship between these matrices. We have the following theorem in this direction.

Theorem 3.3.4. Let $T : V \to V$ be a linear transformation. Then, any two matrices representing T are similar.

Exercise 3.3.5. 1. Find the matrix representation of the following linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $T(x, y) = (x + 6y, 3x + 4y)$. Also find the matrix representation of T with respect to the basis $\{(2, -1), (1, 1)\}$.

2. Find the matrix representation of the rotation transformation by an angle $\pi/4$ radians counter-clockwise with respect to the standard basis and the basis $\{(1, 1), (1, 2)\}$.

3. Find the matrix representation of $T(x, y, z) = (x + 2y, x + y + z, z)$ with respect to the standard basis and the basis $\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$.

Few Probable Questions

1. Find a basis for the dual space of an n-dimensional space V.

Unit 4

Course Structure

- Diagonalization of matrices

4.1 Introduction

As we have already mentioned in the previous unit, diagonal matrices are the easiest to deal with. And we have also seen that different bases give rise to different matrices for a linear transformation, so our main aim is to find a particular basis $\mathbb{B}$ for a vector space V, for which a particular linear transformation (or rather, a linear operator) T, defined on V can be represented as a diagonal matrix. It is not always the case that there always exists such a basis for which T can be represented as a diagonal matrix. We will study mainly the cases and circumstances, under which this is possible. And if such basis does not exist, then what are the simplest possible type of matrix by which we can represent T. These are the various issues that will be addressed in this unit.

Objectives

After reading this unit, you will be able to

- define the characteristic values and vectors of a linear transformation
- recapitulate the basic notions about minimal and characteristic polynomials of a transformation
- define algebraic and geometric multiplicities of a particular eigen value
- define the eigen spaces of a transformation
- determine the cases when a transformation is diagonalizable
- determine the cases when a transformation is not diagonalizable
- find the necessary and sufficient condition for diagonalizability of a transformation
- learn about the Smith's Normal form

4.2 Diagonalizability

As we have already mentioned before, diagonalizability is something related to the matrix of a LT being diagonal. But, before going into the definition of diagonalizability, let us recollect the general notions of eigen values and eigen vectors of a matrix.

When we operate the matrix over a vector (v_1, v_2) of $\mathbb{R}^2$, and equate it to a constant multiple of (v_1, v_2), we get the system

$$\begin{aligned} v_1 + 2v_2 &= cv_1 \\ 3v_2 &= cv_2 \end{aligned}$$

Geometrically speaking, when we take a particular vector (v_1, v_2) of the xy-plane and operate the matrix on it, we the resulting vector is a scalar multiple of the original one. That is, the resulting vector is either a contracted or expanded form of the original vector depending on the value of c. For example, if we take the vector $(1, 1)$, then the resulting vector will be $(3, 3) = 3(1, 1)$. That is, the particular vector is expanding to thrice its original value.

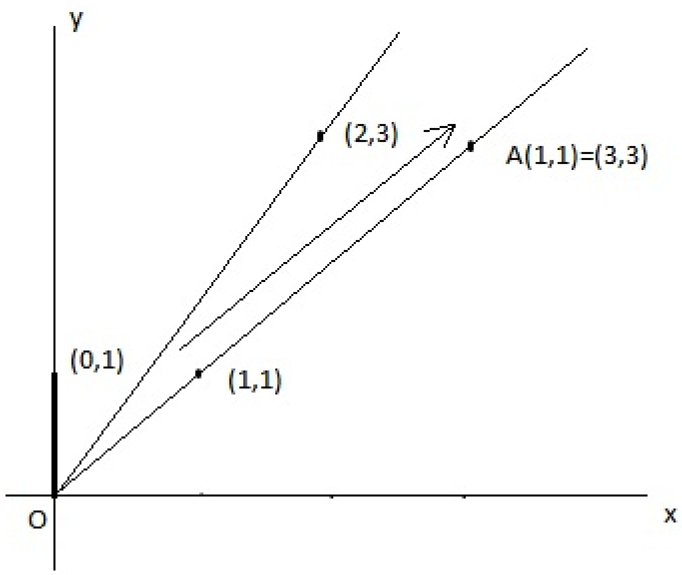


Figure 4.2.1: Eigen Values and Eigen Vectors Geometrically

On the other hand, if we operate the matrix over the vector $(0, 1)$, then the resulting vector $(2, 3)$ is not on the line joining $(0, 1)$ and $(2, 3)$. The vector $(1, 1)$ is called an eigen vector and 3 is the corresponding eigen value. $(0, 1)$ is not an eigen vector.

To summerize, we say that any matrix corresponds to a particular LT and those vectors which do not change their direction on the application of the LT are called its eigen vectors and the factor by which it contracts or expands, is called the corresponding eigen value. We are now in a position to formally define eigen values and eigen vectors of a matrix.

Definition 4.2.1. Let $T : V \to V$ be a LT over vector spaces on the field F. Then a **non-zero** vector $v \in V$ is said to be an eigen value of T if $T(v) = cv$ for some $c \in F$. This c is called the corresponding eigen value of T.

To find eigen value and eigen vectors of a LT, we generally find so for the corresponding matrix representations of T. It is independent of the bases since similar matrices have same eigen values.

It is important to note that T may not have any eigen value in the first place. And if V is finite-dimensional, say having dimension n, then T can have atmost n eigen values. And the eigen vectors can also be seen as the null space of the transformation $T - cI$ (of course ignoring the zero vector).

Now, our main concern is to check whether a given LT can be represented as a diagonal matrix or not. So, we are in search of that particular basis of V for which it can be done. If there exists certain basis for which T can be represented as a diagonal matrix, then T is said to be diagonalizable, otherwise T is non-diagonalizable.

Definition 4.2.2. A linear transformation $T : V \to V$, where V is a **finite-dimensional** vector space, is said to be diagonalizable if there exists a basis $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ for which the corresponding matrix is a diagonal matrix.

A diagonal matrix is of the form

$$\begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}.$$

An identity matrix is the most common example of a diagonal matrix. So, if we consider the eigen values and vectors of a LT T, that is the vectors v_i satisfying $Tv_i = c_i I v_i$, or the non-zero vectors of the null space $T - c_i I$. The intuitive idea is to break the matrix into diagonal blocks of the form

$$\begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_1 \end{bmatrix},$$

the above block being the diagonal block corresponding to the eigen value c_1. Thus, if the sum of the size of the blocks equals the dimension of V, then T stands diagonalized and the corresponding diagonal matrix is

$$\begin{bmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_1 & 0 & \cdots & 0 \\ 0 & 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & c_n \end{bmatrix}.$$

The size of each block is determined by the "size" of the null spaces, that is, dimension of the null spaces, that is the number of linearly independent eigen vectors spanning each null space. In this way, we come to another equivalent definition of diagonalizability.

Definition 4.2.3. A linear transformation $T : V \to V$, where V is a **finite-dimensional** vector space, is said to be diagonalizable if there exists a basis $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ comprising of the eigen vectors of T.

Let us illustrate the process.

Illustration 4.2.1. Let A be an $n \times n$ matrix over a filed F. We first find the eigen values using the "traditional" ways by finding the characteristic polynomial. Let $c_1, c_2, \dots, c_k \in F$ be the eigen values of A. We find the rank of each of the matrices $A - c_iI$ and then find out the nullity, that is, dimension of the null space of $A - c_iI$ using the Rank-Nullity theorem, for each i, $1 \leq i \leq k$. If $\sum_{i=1}^{k} \dim(A - c_iI) = n$, then the matrix A is diagonalizable otherwise, if $\sum_{i=1}^{k} \dim(A - c_iI) < n$, A is non-diagonalizable. For A, there exists a corresponding linear operator from the vector space F^n to F^n.

Example 4.2.4. Let A be a real 3×3 matrix

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}.$$

Then the characteristic polynomial of A is

$$\begin{vmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{vmatrix} = (x-1)(x-2)^2.$$

Then the eigen values of A are 1 and 2. Suppose that T is the linear operator on $\mathbb{R}^3$ which is represented by A in the standard basis. We will find the rank of the matrices $A - I$ and $A - 2I$. Now,

$$A - I = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}$$

has clearly rank equals to 2 and hence nullity equals to $3 - 2 = 1$. Also, the matrix

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

has rank 2 and hence nullity $3 - 2 = 1$. When we sum up the nullities of these two matrices, we get $1 + 1 = 2 \neq 3$. Thus, A is not diagonalizable. The nullities of the matrices $A - I$ and $A - 2I$ together tell us that the null space of the above matrices are spanned by one vector space each, that is, there are a maximum of two distinct linearly independent eigen vectors of A and hence we are unable to find a basis of A containing the eigen vectors.

Definition 4.2.5. Let T be a linear operator over a finite dimensional vector space V and let $c \in F$ be an eigen value of T. Then the null space of the linear operator $T - cI$ is called the **eigen space** of the corresponding eigen value c and the dimension of the eigen space, that is, the nullity of the operator $T - cI$ is called the **geometric multiplicity** of c.

It is a routine exercise to check that the eigen spaces form vector subspaces of V and has been left as an exercise.

Definition 4.2.6. For an eigen value c of a particular operator T, the power to which the factor $(x - c)$ is raised in the corresponding characteristic polynomial of the matrix representation of T is called the **algebraic multiplicity** of c.

Thus, in the previous example, the algebraic multiplicity of 1 and 2 are 1 and 2 respectively and their corresponding geometric multiplicities are equal to 1 each. We can say that the algebraic multiplicity of an eigen value if always greater than or equals to its geometric multiplicity. Also, the algebraic multiplicities of all the eigen values add up to the dimension of the parent vector space and is less than or equal to the dimension if we consider the geometric multiplicities. When the sum of the geometric multiplicities add up to the dimension of the vector space, we call the operator to be diagonalizable.

Example 4.2.7. Let T be a linear operator on $\mathbb{R}^3$ which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

Let us find the characteristic polynomial of A as

$$\begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} = (x-2)^2(x-1).$$

So, 2 and 1 are the eigen values of A with algebraic multiplicities 2 and 1 respectively. We will now find the algebraic and geometric multiplicities of the eigen values. The two matrices

$$A - I = \begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}$$

and

$$A - 2I = \begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}.$$

We know that $A - I$ is singular and obviously $\operatorname{rank}(A - I) \geq 2$ (by Rank-Nullity theorem). Therefore, $\operatorname{rank}(A - I) = 2$. It is evident that $\operatorname{rank}(A - 2I) = 1$. So, the nullity of the matrices are 1 and 2 respectively which sum up to 3. Hence, A is diagonalizable and the corresponding diagonal matrix is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Lemma 4.2.8. Suppose that $T(v) = cv$. If f is any polynomial, then $f(T)(v) = f(c)v$.

Proof. The proof of the lemma is based on the fact that

$$T^2(v) = T(T(v)) = T(cv) = cT(v) = c^2T(v).$$

We can prove by the principle of mathematical induction that

$$T^n(v) = c^n v.$$

Hence $f(T)(v) = f(c)v$, for any polynomial in T. $\square$

Lemma 4.2.9. Let T be a linear operator on the finite-dimensional space V. Let $c_1, c_2, \ldots, c_k$ be the distinct characteristic values of T and let W_i be the corresponding eigen spaces. If $W = W_1 + W_2 + \cdots + W_k$, then

$$\dim W = \dim W_1 + \dim W_2 + \cdots + \dim W_k.$$

In fact, if $\mathcal{B}_i$ is an ordered basis of W_i, then $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\}$ is an ordered basis for W.

Proof. The space $W = W_1 + W_2 + \cdots + W_k$ is the subspace spanned by all of the eigen vectors of T. Usually when one forms the sum W of subspaces W_i, one expects that $\dim W < \dim W_1 + \dim W_2 + \cdots + \dim W_k$ because of linear relations which may exist between vectors in the various spaces. This lemma states that the characteristic spaces associated with different characteristic values are independent of one another.

Suppose that (for each i) we have a vector b_i in W_i, and assume that

$$b_1 + b_2 + \cdots + b_k = 0.$$

We shall show that $b_i = 0$ for each i. Let f be any polynomial. Since $T(b_i) = c_i b_i$, the preceding lemma tells us that

$$\begin{aligned} 0 &= f(T)(0) \\ &= f(T)(b_1) + f(T)b_2 + \cdots + f(T)b_k \\ &= f(c_1)b_1 + f(c_2)b_2 + \cdots + f(c_k)b_k. \end{aligned}$$

Choose the polynomials $f_1, f_2, \ldots, f_k$ such that

$$\begin{aligned} f_i(c_j) = \delta_{ij} &= 1, \; i = j \\ &= 0, \; i \neq j. \end{aligned}$$

Then

$$0 = f_i(T)(0) = \sum_j \delta_{ij} b_j = b_i.$$

Now, let $\mathcal{B}_i$ be an ordered basis for W_i, and let $\mathcal{B}$ be the sequence $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\}$. Then $\mathcal{B}$ spans the subspace $W = W_1 + W_2 + \cdots + W_k$. Also, $\mathcal{B}$ is a linearly independent sequence of vectors, for the following reason. Any linear relation between the vectors in $\mathcal{B}$ will have the form $b_1 + b_2 + \cdots + b_k = 0$, where b_i is some linear combination of the vectors in $\mathcal{B}_i$. From what we just did, we know that $b_i = 0$ for each i. Since each $\mathcal{B}_i$ is linearly independent, we see that we have only the trivial linear relation between the vectors in $\mathcal{B}$. □

In the course of proving the above lemma, we have proved the following theorem.

Theorem 4.2.10. Eigen vectors corresponding to distinct eigen values are linearly independent.

Can you prove the theorem independently?

Thus, we arrive at the following theorem.

Theorem 4.2.11. Let T be a linear operator on a finite-dimensional space V. Let $c_1, c_2, \ldots, c_k$ be distinct eigen values of T and let W_i be the eigen space of c_i. Then the following are equivalent:

1. T is diagonalizable.
2. The characteristic polynomial for T is

$$f(x) = (x - c_1)^{d_1} \ldots (x - c_k)^{d_k},$$

where $\dim W_i = d_i$, $i = 1(1)k$.

3. $\dim W_1 + \dim W_2 + \cdots + \dim W_k = \dim V$.

Proof. We have observed that 1 implies 2. If the characteristic polynomial f is the product of linear factors, as in 2, then $d_1 + d_2 + \cdots + d_k = \dim V$. For, the sum of the $d_i's$ is the degree of the characteristic polynomial, and that degree is $\dim V$. Thus, 2 implies 3. Now suppose that 3 holds. Then by the previous lemma, we must have $V = W_1 + W_2 + \cdots + W_k$, that is, the eien vectors of T span V. $\square$

Let us summerize whatever we have learnt so far.

Let T be a linear operator on an n-dimensional vector space V. If T has n distinct eigen values then it has n linearly independent eigen vectors which form a basis of V and in that case, T is diagonalizable. If it has less number of eigen values, then we have to check that whether they fulfil the deficiency by having multiple eigen vector for a single eigen value so that the number of linearly independent eigen vectors are still n. In either case, we need to check whether the given operator has n linearly independent eigen vectors or not. We can also say that T is diagonalizable if and only if the geometric multiplicity and algebraic multiplicity for a given eigen value coincides.

Exercise 4.2.12. 1. Check whether the following matrices are diagonalizable. If yes, then find its diagonal form.

i.
$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$$

ii.
$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$$

2. Let T be a linear operator on the n-dimensional vector space V, and suppose that T has n distinct eigen values. Prove that T is diagonalizable.

3. Let V be the vector space of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ and let T be the linear operator on V defined as
$$T(f(x)) = \int_0^x f(t)dt.$$
Prove that T has no eigen values.

4. Let $\mathbb{P}_2$ denote the vector space of all polynomials of degree 2 or less, and let $T : \mathbb{P}_2 \to \mathbb{P}_2$ be a linear operator defined by
$$T(ax^2 + bx + c) = 2ax + b.$$
Check whether T is diagonalizable. If so, find the diagonal matrix.

5. Consider the matrix
$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$
where a and b are real numbers and $b \neq 0$. Fince all eigen values of A and determine the corresponding eigen spaces. Hence check whether A is diagonalizable.

6. Check whether the given matrix is diagonalizable. If yes, find the diagonalized matrix.
$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

4.2.1 Minimal Polynomials and Diagonalizability

We have seen in the previous units that minimal polynomials and characteristic polynomials of a matrix (or, linear operator) has same roots.

So, if T is a diagonalizable linear operator and $c_1, c_2, \ldots c_k$ are the distinct eigen values of T. Then it is easy to see that the minimal polynomial for T is the polynomial

$$m(x) = (x - c_1)(x - c_2)\ldots(x - c_k).$$

If v is an eigen vector, then one of the operators $T - c_1I, \ldots, T - c_kI$ sends v into 0. Hence

$$(T - c_1I)\ldots(T - c_kI)(v) = 0,$$

for every eigen vector v. There is a basis for the underlying space which consists of eigen vectors of T; hence

$$m(T) = (T - c_1I)\ldots(T - c_kI) = 0.$$

What we have concluded is this. If T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors. As we shall soon see, that property characterizes diagonalizable operators.

Theorem 4.2.13. Let V be a finite dimensional vector space over the field F and let T be a linear operator on V. Then T is diagonalizable if and only if the minimal polynomial of T is the product of distinct linear factors, that is, of the form

$$m(x) = (x - c_1)(x - c_2)\ldots(x - c_k),$$

where, $c_1, c_2, \ldots, c_k \in F$ are distinct.

Proof. We have noted earlier that, if T is diagonalizable, its minimal polynomial is a product of distinct linear factors. To prove the converse, let W be the subspace spanned by all of the eigen vectors of T, and suppose that $W \neq V$. By a previous lemma, there is a vector v not in W and an eigen value c_j of T such that the vector

$$b = (T - c_jI)(v)$$

lies in W. Since $b \in W$,

$$b = b_1 + b_2 + \cdots + b_k$$

where $T(b_i) = c_ib_i$, $1 \leq i \leq k$, and therefore the vector

$$h(T)(b) = h(c_1)(b_1) + \cdots + h(c_k)(b_k)$$

is in W, for every polynomial h. Now,

$$m(x) = (x - c_j)q(x),$$

for some polynomial q. Also,

$$q - q(c_j) = (x - c_j)h.$$

But we have

$$q(T)(v) - q(c_j)(v) = h(T)(T - c_jI)(v) = h(T)(b).$$

But, $h(T)(b) \in W$ and since

$$0 = m(T)(v) = (T - c_jI)q(T)(v)$$

the vector $q(T)(v)$ is in W. Hence $q(c_j)(v)$ is in W. Since v is not in W, we have $q(c_j) = 0$. This contradicts the fact that m has distinct roots. Hence the theorem. □

Example 4.2.14. Let A be a 4×4 matrix

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

The powers of A are easy to compute

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

$$A^3 = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{bmatrix}$$

Thus, $A^3 = 4A$, that is, $f(x) = x^3 - 4x = x(x+2)(x-2)$, then $m(A) = 0$. The minimal polynomial of A must divide f. Minimal polynomial is not of degree 1 since in that case, A would have been a scalar multiple of I, which is not true. Hence the candidates of minimal polynomial polynomial are f, $x(x+2)$, $x(x-2)$, $x^2 - 4$. The three quadratic polynomials can be eliminated since at a glance, we can see that $A^2 \neq 2A$, $A^2 \neq -2A$, and $A^2 \neq 4I$. Hence f is the minimal polynomial for A and since f is the product of distinct linear factors, so A is diagonalizable. Now, we can clearly see that the rank of A is 2 and hence its nullity is also $4-2=2$, which means that the eigen space of $A - 0I$ has dimension 2 and thus its algebraic multiplicity will be 2. Thus, the characteristic polynomial is $x^2(x^2-4)$. And the matrix A is similar to the diagonal form

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

Exercise 4.2.15.

1. Every matrix A such that $A^2 = A$ is similar to a diagonal matrix.
2. Using diagonalizability, compute A^n, $n \in \mathbb{N}$ for
$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$
3. Is every diagonalizable matrix invertible? Justify.
4. Let A be an $n \times n$ diagonalizable matrix whose characteristic polynomial is given by
$$f(x) = x^3(x-1)^2(x-2)^5(x+2)^4.$$
   i. Find the size of the matrix A.
   ii. Find the minimal polynomial of A.
   iii. Find the dimension of the eigen space for the eigen value 2.
   iv. Find the rank of the matrix.

Few Probable Questions

1. Show that the eigen vectors corresponding to distinct eigen values are linearly independent.

2. State a necessary and sufficient condition for diagonalizability. Check the diagonalizability of the following matrix
$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

3. Let f be the characteristic polynomial of a matrix A over the field $\mathbb{R}$ as
$$f(x) = x^2(x-3)(x+4)^5.$$
Also, let A be diagonalizable. Then
   (a) Find the minimal polynomial of A.
   (b) Find the eigen values along with their algebraic and geometric multiplicities.
   (c) Find the diagonalized form of A.

4. Let A be a matrix over the field $\mathbb{R}$ whose minimal polynomial is of the form
$$f(x) = (x^2-1)(x^2+1).$$
   (a) Is A diagonalizable over $\mathbb{R}$? Justify.
   (b) Is A diagonalizable over the field $\mathbb{C}$? Justify.

   Find the eigen values in each case.

5. Let P be a linear operator over $\mathbb{R}^2$ defined as
$$P((x,y)) = (x,0).$$
Show that P is linear. Find the matrix representation of P with respect to the standard basis of $\mathbb{R}^2$. What is the minimal polynomial of P? Is P diagonalizable?

Unit 5

Course Structure

- Smith's normal form, Jordan Canonical forms.

5.1 Introduction

We have seen in the preceding unit that we want to write the matrix of a linear operator in its simplest possible form, which is possible since the matrix representation of a single linear operator under various bases are similar. And we have also seen that the diagonal matrix is the simplest possible matrix to work with. We are always in search of a basis of the underlying vector space for which the corresponding matrix of the linear operator is diagonal. If such a basis exists, then we are happy and the operator is said to be diagonalizable. We have seen various circumstances under which an operator is diagonalizable. We are okay with them. But, what happens if a given operator is not diagonalizable. Can't we express the operator in a simpler form then? That is where the other canonical forms come into play. We can certainly express the operators in a simpler form, which is "almost" a diagonal matrix. One of them is the Jordan Canonical forms, which we shall come through in this unit. Another way of simplifying a matrix is by changing it into its Smith's normal form.

Objectives

After reading this unit, you will be able to

- learn about the Smith's normal form of a matrix
- learn about the Jordan forms and find those for any given matrix or linear operator

5.2 Smith's Normal Form

The Smith normal form is a normal form that can be defined for any matrix (not necessarily square) with entries in a principal ideal domain (PID). The Smith normal form of a matrix is diagonal, and can be obtained from the original matrix by multiplying on the left and right by invertible square matrices. In particular, the integers are a PID, so one can always calculate the Smith normal form of an integer matrix. We will talk particularly about the PID $\mathcal{Z}$.

Definition 5.2.1. Let A be an $m \times n$ matrix over $\mathbb{Z}$. We say that A is in Smith Normal form if there are non-zero $a_1, a_2, \dots, a_k \in \mathbb{Z}$ such that a_i divides a_{i+1} for $i < k$ such that

$$A = \begin{bmatrix} a_1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & \cdots & a_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Theorem 5.2.2. If A is a matrix with entries in $\mathbb{Z}$, then there are invertible matrices P and Q such that PAQ is in Smith normal form.

Theorem 5.2.3. Every matrix over $\mathbb{Z}$ has Smith Normal form.

In order to find the Smith Normal form of a matrix, we are allowed to use the following operations

1. interchange two rows and columns,
2. multiply a row or column by ± 1(which are the invertible elements in $\mathbb{Z}$)
3. add an integer multiple of a row (or column) to another row (or column)

Exercise 5.2.4. Obtain the Smith normal form and rank for

$$A = \begin{bmatrix} 0 & 2 & -1 \\ -3 & 8 & 3 \\ 2 & -4 & -1 \end{bmatrix}$$

over $\mathbb{Z}$.

5.3 Jordan Canonical Forms

We have seen that the diagonal matrices are "easiest" matrix to handle. So we are always in search of a basis for which a particular linear operator is diagonalizable. But this is not always possible. So we are in search of the next simplest matrix in which the operator can be represented. And the next "easiest" matrix to deal with are the triangular matrices. So we come to the Jordan canonical forms, or simply the Jordan forms. The Jordan Canonical Form is an upper triangular matrix of a particular form called a Jordan matrix representing a linear operator on a finite-dimensional vector space with respect to some basis. Such a matrix has each non-zero off-diagonal entry equal to 1, immediately above the main diagonal (on the superdiagonal), and with identical diagonal entries to the left and below them. Let us check for ourselves. Let A be a matrix as given

$$A = \begin{bmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{bmatrix}.$$

The eigen values of A are $1, 2, 4, 4$ and the dimensions of the eigen space corresponding to each eigen values are $1, 1, 1$ which does not sum up to 4, so A is not-diagonalizable. But A is similar to the matrix below

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

The matrix J is "almost" diagonal and is called the Jordan form of A.

Definition 5.3.1. Let A be an $n \times n$ matrix and c be an eigen value of A of algebraic multiplicity, say k. Then the elementary Jordan block of A corresponding to c, of size k is given by

$$\begin{bmatrix} c & 1 & 0 & \cdots & 0 \\ 0 & c & 1 & \cdots & 0 \\ 0 & 0 & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & c \end{bmatrix}.$$

Then the parent matrix is composed of the elementary Jordan blocks

$$A = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix}.$$

The Jordan form of a matrix has the following properties:

1. Given an eigen value c_j, the number of elementary Jordan blocks corresponding to c_j is equal to the geometric multiplicity of c_j.
2. The sum of the sizes of the Jordan blocks corresponding to an eigen value c_j is equal to its algebraic multiplicity.
3. The maximum size of a Jordan block corresponding to an eigen value c_j is equal to its multiplicity in the minimal polynomial of the parent matrix and there has to be a Jordan block with the maximum size for c_j.

Illustration 5.3.1. 1. Let us be given a matrix

$$A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}.$$

First of all, we calculate the eigen values of A which are 5 and 3. Then find the rank of the matrices $A - 5I$ and $A - 3I$ which happen to be 2 and 1 respectively and hence the nullity of the corresponding matrices are 1 and 2 respectively summing up to 3, the dimension of $\mathbb{R}^3$. Hence the minimal polynomial of A is $(x-3)(x-5)$ and the Jordan form for A is

$$J = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Here there are precisely three Jordan blocks, [5], [3], [3].

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

Then A has only one eigen value, which is 1 and the rank of $A - I$ is 1, which means that it has nullity equal to 2 which does not sum up to 3. Since the nullity, that is the geometric multiplicity of 1 is 2, so there will be two Jordan blocks for 1 and also the maximum size of the Jordan block should be 2. Thus, the Jordan form for A is
$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercise 5.3.2.

1. Put the matrix
$$A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{bmatrix}$$
into Jordan form.

2. Let A be a 5×5 matrix with characteristic polynomial $f(x) = (x-2)^3(x+7)^2$ and minimal polynomial $m = (x-2)^2(x+7)$. What is the Jordan form for A?

3. How many possible ,Jordan forms are there for a 6×6 complex matrix with characteristic polynomial $(x+2)^4(x-1)^2$?

4. The differentiation operator on the space of polynomials of degree less than or equal to 3 is represented in the 'natural' ordered basis by the matrix
$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$
What is the Jordan form of this matrix?

Few Probable Questions

1. Find the Jordan form of the matrix
$$A = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$
Show detailed steps.

Unit 6

Course Structure

- Invariant factors and elementary divisors, Rational (or Natural Normal) form.

6.1 Introduction

There are certain subspaces which remain invariant under a linear operator, that is, the linear operator sends each element of the subspace to itself. Such subspaces are of primary importance as we shall see that we can analyse many properties of the linear operator by finding out the various invariant subspaces of the operator. The primary purpose of this section is to prove that if T is any linear operator on a finite-dimensional space V, then there exist vectors $v_1, \ldots, v_k$ in V such that

$$V = Z(v_1; T) \oplus \cdots \oplus Z(v_k; T).$$

This will show that T is the direct sum of a finite number of linear operators, each of which has a cyclic vector. The cyclic decomposition theorem is closely related to the following question. Which T-invariant subspaces W have the property that there exists a T-invariant subspace W' such that $V = W \oplus W'$? In fact, there are many subspaces W' for which $V = W \oplus W'$ but we can't say whether they are invariant or not. This unit is dedicated to the study of the invariant factors and elementary divisors of a linear operator and certain canonical forms of it.

Objectives

After reading this unit, you will be able to

- define the invariant subspaces and see certain examples
- learn about the independent subspaces of a vector space
- learn about the direct-sum decomposition of a vector space into independent subspaces of it
- learn about the invariant direct sum decomposition of a vector space
- define the cyclic vectors of a vector space

- define the smallest invariant subspace containing a vector
- define T-admissible subspaces of a vector space
- learn the cyclic decomposition theorem for a finite-dimensional vector space with respect to a linear operator T
- learn the generalized Cayley-Hamilton theorem for a linear operator on a finite-dimensional vector space
- define the invariant factors of a matrix
- learn to find the rational canonical form for a matrix

6.2 Invariant Subspaces

Definition 6.2.1. Let V be a vector space and T, a linear operator on V. If W is a subspace of V, we say that W is invariant under T if for each $w \in W$, the vector $T(w)$ is also in W.

Example 6.2.2. If T is any linear operator on V, then V is invariant under T as is the zero subspace. The range of T and the null space of T are also invariant under T.

Example 6.2.3. Let F be a field and D be the differentiation operator on the space $F[x]$ of polynomials over F. Let n be a positive integer and W be a subspace of polynomials of degree not greater than n. Then W is invariant under T.

Example 6.2.4. Let T be the linear operator on $\mathbb{R}^2$ which is represented in the standard basis by the matrix

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

Then the only subspaces of $\mathbb{R}^2$ which are invariant under T are $\mathbb{R}^2$ and the zero subspace. Any other invariant subspace would necessarily have dimension 1. But, if W is the subspace spanned by some non-zero vector v, the fact that W is invariant under T means that v is an eigen vector, but A has no eigen value.

When the subspace W is invariant under the operator T, then T induces a linear operator T_W on the space W. The linear operator T_W is defined by $T_W(v) = T(v)$, for $v \in W$. Now we turn to an investigation of the simplest possible nontrivial invariant subspaces : invariant subspaces with dimension 1. How does an operator behave on an invariant subspace of dimension 1? Subspaces of a vector space V of dimension 1 are easy to describe. Take any non-zero vector $u \in V$ and let U equals the set of all scalar multiples of u, that is

$$U = \{au : \ a \in F\}.$$

where, F is the underlying field. The U is a one-dimensional subspace of V, and every one-dimensional subspace of V is of this form. If $u \in V$ and the subspace defined as above is invariant under T, then $T(u)$ must be in U, which means that there must exist a scalar $c \in F$ such that $T(u) = cu \in U$. Conversely, if u is a non-zero vector in V such that $T(u) = cu$ for some scalar c, then the subspace U defined above is a one-dimensional subspace of V invariant under T. The equation $T(u) = cu$ is same as $(T - cI)u = 0$, so that c is an eigen value and u is an eigen vector of T. Thus, we can see that the one dimensional invariant subspace of an operator T is precisely the eigen space of the operator. But the converse is not true always, that is, any eigen space of T need not be one-dimensional though it is invariant under T (can you think of such an example?).

When V is finite-dimensional, the invariance of a subspace W under the linear operator T has a simple matrix interpretation. Suppose we choose an ordered basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis of V and $\mathcal{B}' = \{v_1, \ldots, v_r\}$ of $W(r = \dim W)$. Let $A = [T]_{\mathcal{B}}$ so that

$$T(v_j) = \sum_{i=1}^{n} A_{ij} v_i.$$

Since W is invariant under T, the vector $T(v_j)$ belongs to W for $j \leq r$. This means that

$$T(v_j) = \sum_{i=1}^{r} A_{ij} v_i, \quad j \leq r.$$

In other words, $A_{ij} = 0$ if $j \leq r$ and $i > r$. Schematically A has the block form

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

where B is an $r \times r$ matrix, C is an $r \times (n-r)$ matrix, and D is an $(n-r) \times (n-r)$ matrix.

Direct-Sum Decompositions

Definition 6.2.5. The subspaces $W_1, W_2, \ldots, W_k$ of a vector space V are said to be independent if

$$w_1 + w_2 + \cdots + w_k = 0, \qquad w_i \in W_i$$

implies that each w_i is zero.

For $k = 2$, we can say that independence means that $W_1 \cap W_2 = \{0\}$. If $k > 2$, it says that each W_j intersects the sum of the other subspaces only at the zero vector.

The independence can be understood as this: If $W = W_1 + W_2 + \cdots + W_k$ be the subspace spanned by $W_1, W_2, \ldots, W_k$, then each vector $w \in W$ can be uniquely expressed as the sum of the vectors in W_j, that is,

$$w = w_1 + w_2 + \cdots + w_k, \quad w_i \in W_i$$

If w has another representation as

$$w = u_1 + u_2 + \cdots + u_k, \quad u_i \in W_i$$

then subtracting, we get

$$0 = (w_1 - u_1) + \cdots + (w_k - u_k), \quad w_k - u_k = 0$$

and the definition of independence implies that $w_j - u_j = 0$ for $1 \leq j \leq k$. Thus, when $W_1, W_2, \ldots, W_k$ are independent, we can operate with the vectors in W as k-tuples.

Lemma 6.2.6. Let V be a finite-dimensional vector space and let $W_1, W_2, \ldots, W_k$ be subspaces of V and let $W = W_1 + W_2 + \cdots + W_k$. Then the following are equivalent

1. $W_1, W_2, \ldots, W_k$ are independent.

2. For each j, $2 \leq j \leq k$, we have
$$W_j \cap (W_1 + \cdots + W_{j-1}) = \{0\}.$$

3. If $\mathcal{B}_i$ is an ordered basis for W_i, for each i, then the sequence $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\}$ is an ordered basis for W.

If the above conditions hold, we say that the sum $W = W_1 + W_2 + \cdots + W_k$ is direct or that W is the direct sum of $W_1, W_2, \ldots, W_k$ and we write it as
$$W = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Example 6.2.7. Let V be a finite-dimensional vector space over the field F and let $\{v_1, v_2, \ldots, l_n\}$ be a basis for V. If W_i be the one-dimensional subspace spanned by v_i, then
$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_n.$$

Example 6.2.8. Let T be any linear operator on a finite-dimensional space V. Let $c_1, c_2, \ldots, c_k$ be the distinct eigen values of T, and let W_i be the space of eigen vectors associated with the eigen value c_i. Then $W_1, W_2, \ldots, W_k$. And if T is diagonalizable, then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$.

Definition 6.2.9. If V is a vector space, a projection of V is a linear operator E on V such that $E^2 = E$.

Suppose E is a projection. Let R be the range of E and let N be the null space of E. We establish that $V = R \oplus N$. Because $w \in R$ if and only if $w = E(w)$, since $w = E(v)$ implies $E(w) = E(E(v)) = E^2(v) = E(v) = w$. Conversely, if $w = E(w)$, the obviously $w \in R$. The unique representation of v as the sum of vectors in R and N is $v = E(v) + (v - E(v))$.

If R and N are subspaces of V such that $V = R \oplus N$, there is a unique projection operator E which has range R and null space N. The operator is called the projection on R along N.

Projections are clearly diagonalizable since for any projection E, we always have $E^2 = E$ and since the minimal polynomial divides any annihilating polynomial of an operator, so the minimal polynomial can be either $x = 0$, or $x - 1 = 0$ or $x(x-1) = 0$ which is the product of distinct linear factors in all the cases.

Projections can be used to describe direct-sum decompositions of the space V.

Theorem 6.2.10. Let $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, then there exist k linear operators $E_1, E_2, \ldots, E_k$ on V such that

1. each E_i is a projection,

2. $E_iE_j = 0$, if $i \neq j$,

3. $I = E_1 + E_2 + \cdots + E_k$,

4. the range of E_i is W_i

Conversely, if $E_1, E_2, \ldots, E_k$ are k linear operators on V satisfying conditions 1-3, and if W_i is the range of E_i, then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$

Proof. Suppose $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Then for each j, we define an operator E_j on V. Let $v \in V$ and let $v = v_1 + v_2 + \cdots + c_k$ with $v_i \in W_i$. Then we define E_j as $E_j(v) = v_j$. Then E_j is well-defined and it is easy to check that it is linear and that, the range of E_j is W_j and that $E_j^2 = E_j$. The null space of E_j is the subspace

$$W_1 + W_2 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k$$

for, the statement that $E_j(v) = 0$ simply means $v_j = 0$, that is, v is actually a sum of vectors from the spaces W_i, with $i \neq j$. In terms of the projections E_j, we have

$$v = E_1(v) + \cdots + E_k(v)$$

for each $v \in V$. So, the identity operator on V can be written as

$$I = E_1 + E_2 + \cdots + E_k.$$

Also, if $i \neq j$, then we see that $E_i E_j = 0$ since the range of E_j is the subspace W_j which lies in the null space of E_j.

Conversely, suppose $E_1, E_2, \ldots, E_k$ are k linear operators on V satisfying conditions 1-4. Then certainly we must have

$$V = W_1 + W_2 + \cdots + W_k.$$

since by condition 3, we have

$$v = E_1(v) + \cdots + E_k(v)$$

for every $v \in V$, and $E_i(v) \in W_i$. This expression for v is unique, because if

$$v = v_1 + \cdots v_k, \quad v_i \in W_i,$$

say $v_i = E_i(w_i)$, then using 1 and 2, we have

$$E_j(v) = \sum_{i=1}^{k} E_j v_i = \sum_{i=1}^{k} E_j E_i w_i = E_j^2(w_j) = E_j(w_j) = v_j.$$

This shows that V is the direct sum of the W_i. $\square$

Invariant Direct Sums

We are primarily interested in direct-sum decompositions of V where each subspace if invariant under some linear operator T. Given such a decomposition of V, T induces a linear operator T_i on each W_i by restriction. Thus, if $v \in V$, then we have the unique representation

$$v = v_1 + \cdots + v_k, \quad v_i \in W_i$$

where, each W_i is an invariant subspace of V into which V decomposes. Then

$$T(v) = T_1(v_1) + \cdots + T_k(v_k)$$

We can say that T is the direct-sum of the operators $T_1, \cdots, T_k$. The fact that $V = W_1 \oplus \cdots \oplus W_k$, enables us to associate a unique k-tuple for each $v \in V$ (which is $(v_1, \ldots, v_k)$), in such a way that we can carry out the linear operations in V by working in the individual subspaces W_i. The fact, that each W_i is invariant under T enables us to view T as independent action of T_i on the subspaces W_i.

The above situation can be interpreted in terms of matrices. Suppose we select an ordered basis $\mathcal{B}_i$ of W_i and let $\mathcal{B}$ be the ordered basis for V consisting of the union of the $\mathcal{B}_i$, arranged in the order $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k$. Let $A = [T]_{\mathcal{B}}$ and let $A_i = [T]_{\mathcal{B}_\rangle}$, then A has the block form

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}$$

Each A_i is a $d_i \times d_i$ matrix, where $d_i = \dim W_i$, and 0's are symbols for rectangular blocks of scalars 0's of various sizes.

Theorem 6.2.11. Let T be a linear operator on the space V, and let $W_1, \ldots, W_k$ and $E_1, \ldots, E_k$ be the projections as in the previous theorem. Then a necessary and sufficient condition that each subspace W_i be invariant under T is that T commute with each of the projections E_i, that is

$$TE_i = E_iT, \quad i = 1(1)k.$$

We shall now describe a diagonalizable operator T in the language of invariant direct sum decompositions (projections which commute with T). This will be a great help to us in understanding some deeper decomposition theorems later.

Theorem 6.2.12. Let T be a linear operator on a finite-dimensional space V. If T is diagonalizable and $c_1, \ldots, c_k$ are the distinct eigen values of T, then there exist linear operators $E_1, \ldots, E_k$ on V such that

1. $T = c_1E_1 + \cdots + c_kE_k$;
2. $I = E_1 + \cdots + E_k$;
3. $E_iE_j = 0, i \neq j$;
4. $E_i^2 = E_i$;
5. the range of E_i is the eigen space for T associated with c_i.

Conversely, if there exist k distinct scalars $c_1, \ldots, c_k$ and k non-zero linear operators $E_1, \ldots, E_k$ satisfying conditions 1-3, then T is diagonalizable and conditions 4 and 5 are also satisfied.

Proof. Suppose that T is diagonalizable, with distinct eigen values $c_1, \ldots, c_k$. Let W_i be the eigen spaces of V. We know that,

$$V = W_1 \oplus \cdots \oplus W_k$$

Let $E_1, \ldots, E_k$ be the projections associated with this decomposition, as we have done before. Then 2-5 are satisfied. To verify 1, let $v \in V$ and we have

$$v = E_1(v) + \cdots + E_k(v)$$

So,

$$T(v) = TE_1(v) + \cdots + TE_k(v) = c_1E_1(v) + \cdots + c_kE_k(v).$$

Thus,

$$T = c_1E_1 + \cdots + c_kE_k.$$

Now suppose that we are given a linear operator T along with distinct scalars c_i and non-zero operators E_i which satisfy 1-3. Since $E_iE_j = 0$, for $i \neq j$, we multiply both sides of $I = E_1 + \cdots + E_k$ by E_i, and obtain immediately $E_i^2 = E_i$. Multiplying $T = c_1E_1 + \cdots + c_kE_k$ by E_i, we get $TE_i = c_iE_i$, which shows that any vector in the range of E_i, is in the null space of $T - c_iI$. Since we have assumed that $E_i \neq 0$, this proves that there is a non-zero vector in the null space of $T - c_iI$, that is, c_i is an eigen value of T. Furthermore, c_i are all of the eigen values of T; for if c is any scalar, then

$$T - cI = (c_1 - c)E_i + \cdots + (c_k - c)E_k$$

so that, if $(T - cI)(v) = 0$, we must have $(c_i - c)E_i(v) = 0$. If v is not the zero vector, then $E_i(v) \neq 0$ for some i, so that for this i, we have $c_i - c = 0$.

Certainly T is diagonalizable, since we have shown that every non-zero vector in the range of E_i is an eigen vector of T, and the fact that $I = E_1 + \cdots + E_k$ shows that these characteristic vectors span V. All that remains to be demonstrated is that the null space of $T - c_iI$ is exactly the range of E_i. But this is clear since if $T(v) = c_iv$, then

$$\sum_{j=1}^{k}(c_j - c_i)E_j(v) = 0, \quad \text{for each } j$$

and then

$$E_j(v) = 0, \quad i \neq i.$$

Since $v = E_1(v) + \cdots + E_k(v)$, and $E_j(v) = 0$ for $j \neq i$, we have $v = E_j(v)$, which proves that v is in the range of E_i. $\square$

Primary Decomposition Theorem

We studying a linear operator T on the finite-dimensional space V, by decomposing it into a direct sum of operators which are in some sense elementary. We can do this through the eigen values and vectors of T in certain special cases, i.e.,when T is diagonalizable, or, when the minimal polynomial for T factors over the scalar field F into a product of distinct monic polynomials of degree 1 . What can we do with the general T? While studying T using eigen values, we are confronted with two problems. First, T may not have a single eigen value ; this is really a deficiency in the scalar field, namely, that it is not algebraically closed, and we have nothing to do in that case. Second, even if the characteristic polynomial factors completely over F into a product of polynomials of degree 1, there may not be enough eigen vectors for T to span the space V; this is clearly a deficiency in T. The second situation is illustrated by the operator T on F^3, where F is any field represented in the standard basis by

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

The characteristic polynomial for A is $(x - 2)^2(x + 1)$ and this is also the minimal polynomial for A, and thus, for T. Hence, T is not diagonalizable and this happens since the nullity of $T - 2I$ is 1. On the other hand, the null space of $T + I$ and $(T - 2I)^2$ span V. From here, we get the motivation for our further work. Suppose we are given that

$$m = (x - c_1)^{r_1} \ldots (x - c_k)^{r_k}$$

where $c_1, \ldots, c_k \in F$, then we will show that V is the direct sum of the null spaces of $(T - c_iI)^{r_i}$, $i = 1(1)k$.

Theorem 6.2.13. Let T be a linear operator on the finite-dimensional vector space V over the field F. Let m be the minimal polynomial for T as

$$m = m_1^{r_1} \ldots m_k^{r_k}$$

where the m_i are distinct irreducible monic polynomials over F and r_i are positive integers. Let W_i be the null space of $m_i(T)^{r_i}$, $i = 1(1)k$. Then

1. $V = W_1 \oplus \cdots \oplus W_k$;

2. each W_i in invariant under T;

3. if T_i is the operator induced on W_i by T, then the minimal polynomial for T_i is $m_i^{r_i}$.

Proof. Let
$$f_i = \frac{m}{m_i^{r_i}} = \Pi_{j \neq i} m_j^{r_j}.$$
Since m_i are distinct polynomials, the polynomials f_i are relatively prime which implies that there are polynomials $g_1, \dots, g_k$ such that
$$\sum_{i=1}^{n} f_i g_i = 1.$$
Also, if $i \neq j$, then $f_i f_j$ is divisible by the polynomial m, since $f_i f_j$ contains each $m_l^{r_l}$ as factor. We shall show that the polynomials $h_i = f_i g_i$ such that $h_i(T)$ is the identity on W_i and is zero on the other W_j such that $h_1(T) + \cdots + h_k(T) = I$.

Let $E_i = h_i(T) = f_i(T)g_i(T)$. Since $h_1 + \cdots + h_k = 1$ and p divides $f_i f_j$ for $i \neq j$, we have
$$E_1 + \cdots + E_k = I, \quad E_i E_j = 0, \quad \text{if } i \neq j.$$
Thus, E_i are the projections which correspond to some direct-sum decomposition V. We will show that the range of E_i is exactly W_i. It is clear that each vector in the range of E_i is in W_i, since if $v \in E_i$, then $v = E_i(v)$, and so
$$m_i(T)(v) = m_i(T)^{r_i} E_i(v) = m_i(T)^{r_i} f_i(T) g_i(T)(v) = 0$$
since m divides $m_i^{r_i} f_i g_i$. Conversely, suppose that v is in the null space of $m_i(T)^{r_i}$. If $j \neq i$, then $f_j g_j$ is divisible by $m_i^{r_i}$ and so $f_j(T)g_j(T)(v) = 0$, that us $E_j(v) = 0$ for $j \neq i$. But this is immediate that $E_i(v) = v$, that is v is in the range of E_i. This completes the proof of 1.

Also, it is evident that W_i are invariant under T. If T_i is the operator induced on W_i by T, then obviously $m_i(T)^{r_i} = 0$, because by definition, $m_i(T)^{r_i}$ is zero on W_i. This shows that the minimal polynomial for T_i divides $m_i^{r_i}$. Conversely, let g be any polynomial such that $g(T_i) = 0$. Then $g(T)f_i(T) = 0$. Thus gf_i is divisible by the minimal polynomial of T, that is, $m_i^{r_i}$ divides gf_i. It is easily seen that $m_i^{r_i}$ divides g. Hence the minimal polynomial for T_i is $m_i^{r_i}$. □

Exercise 6.2.14. 1. Let T be a linear operator on a finite-dimensional vector space V. Let R be the range of T and let N be the null space of T. Prove that R and N are independent if and only if $V = R \oplus N$.

2. Let T be a linear operator on V. Suppose $V = W_1 \oplus \cdots \oplus W_k$, where each W_i is invariant under T. Let T_i be the induced operator on W_i. Then show that the characteristic polynomial f of T is the product of those of T_i.

3. Let T be a linear operator on V which commutes with every projection operator on V. What can you say about T?

4. Let T be a linear operator on the finite-dimensional space V with characteristic polynomial
$$f = (x - c_1)^{d_1} \dots (x - c_k)^{d_k}$$

and minimal polynomial

$$m = (x - c_1)^{r_1} \dots (x - c_k)^{r_k}.$$

Let W_i be the null space of $(T - c_i I)^{r_i}$. Then show that W_i is the set of all vectors $v \in V$ such that $(T - c_i I)^m(v) = 0$ for some positive integer m (which may depend on v).

Cyclic Subspaces and Annihilators

If V is a finite-dimensional vector space over a field F and T is a fixed linear operator on V. If v is any vector in V, there is a smallest subspace of V which is invariant under T and contains v. This subspace can be defined as the intersection of all T-invariant subspaces which contain v. If W is any subspace of V which is invariant under T and contains v, then W must also contain $T(v)$ and hence must contain $T^2(v)$, $T^3(v)$, and so on. In other words, W must contain $g(T)(v)$ for every polynomial g over F. This is clearly the smallest subspace which contains the vector v and invariant under T.

Definition 6.2.15. If v is any vector in V, the T-cyclic subspace generated by v is the subspace $Z(v;T)$ of all vectors of the form $g(T)(v)$, g in $F[x]$. If $Z(v;T) = V$, then v is called a cyclic vector for T.

In other words, $Z(v;T)$ is the subspace $\{v, T(v), T^2(v), \dots\}$ and v is a cyclic vector if and only if these vectors span V. Every arbitrary operator need not have cyclic vectors.

Example 6.2.16. For any operator T, the T-cyclic subspace generated by the zero vector is the zero subspace. The space $Z(v;T)$ is one-dimensional if and only if v is an eigen vector for T. For the identity operator, every non-zero vector generates a one-dimensional cyclic subspace; thus, if $\dim V > 1$, the identity operator has no cyclic vector.

For any operator T and vector v, we are interested in the linear relations

$$c_0 + c_1 T(v) + \dots + c_k T^k(v) = 0$$

between the vectors $T^i(v)$, or, we shall be interested in the polynomials $g = c_0 + c_1 x + \dots + c_k x^k$ such that $g(T)(v) = 0$. The set of all g satisfying the property in $F[x]$ is clearly a non-zero ideal since it contains the minimal polynomial m of the operator T.

Definition 6.2.17. If v is any vector in V, the T-annihilator of v is the ideal $M(v;T)$ in $F[x]$ consisting of all polynomials g over F such that $g(T)(v) = 0$. Then the unique monic polynomial m_v which generates this ideal will also be called the T-annihilator of v.

We note that the degree of m_v should be greater than zero unless v is the zero vector.

Theorem 6.2.18. Let v be any non-zero vector in V and m_v be the T-annihilator of v. Then

1. the degree of m_v is equal to the dimension of the cyclic subspace $Z(v;T)$;
2. if the degree of m_v is k, then the vectors $v, T(v), T^2(v), \dots, T^{k-1}(v)$ form a basis for $Z(v;T)$
3. if U is the linear operator on $Z(v;T)$ induced by T, then the minimal polynomial for U is m_v.

If v is a cyclic vector for T, then the minimal polynomial for T must have degree equal to the dimension of the space V ; hence, the Cayley-Hamilton theorem tells us that the minimal polynomial for T is the characteristic polynomial for T.

Our plan is to study the general T by using operators which have a cyclic vector. So, let us take a look at a linear operator U on a space W of dimension k which has a cyclic vector v. By the above theorem, the vectors $v, \dots, U^{k-1}(v)$ forms a basis for the space W, and the annihilator m_v of v is the minimal polynomial for U (and hence also the characteristic polynomial for U). If we let $v_i = U^{i-1}(v)$, $i = 1(1)k$, then the action of U on the ordered basis $\mathcal{B} = \{v_1, \dots, v_k\}$ is

$$\begin{aligned} U(v_i) &= v_{i+1}, \quad i = 1(1)k-1 \\ U(v_k) &= -c_0v_1 - c_1v_2 - \cdots - c_{k-1}v_k \end{aligned}$$

where, $m_v = c_0 + c_1x + \cdots + x^k$. The expression for $U(v_k)$ follows from the fact that $m_v(U)(v) = 0$, that is

$$U^k(v) + c_{k+1}U^{k-1}(v) + \cdots + c_1U(v) + c_0v = 0.$$

This says that the matrix of U in the ordered basis $\mathcal{B}$ is

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -c_{k-1} \end{bmatrix}.$$

The matrix is called the companion matrix of the monic polynomial m_v.

Theorem 6.2.19. If U is a linear operator on the finite-dimensional space W, then U has a cyclic vector if and only if there is some ordered basis for W in which U is represented by the companion matrix of the minimal polynomial for U.

Proof. If U has a cyclic vector, then there is such an ordered basis for W. Conversely, if we have some ordered basis $\{v_1, \dots, v_k\}$ for W in which U is represented by the companion matrix of its polynomial, it is obvious that v_1 is a cyclic vector for U. □

Corollary 6.2.20. If A is the companion matrix of a monic polynomial m, then m is both the minimal and the characteristic polynomial of A.

If T is any linear operator on the space V and v is any vector in V, then the operator U which T induces on the cyclic subspace $Z(v; T)$ has a cyclic vector, namely v. Thus, $Z(v; T)$ has an ordered basis in which U is represented by the companion matrix of m_v, the T-annihilator of v.

Exercise 6.2.21.

1. Show that $Z(v; T)$ is one dimensional if and only if v is an eigen vector of T.
2. Let T be the linear operator on $\mathbb{R}^3$ which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

   Prove that T has no cyclic vector. What is the T-cyclic subspace generated by the vector $(1, -1, 3)$?
3. Let V be an n-dimensional vector space, and let T be a linear operator on V. Suppose that T is diagonalizable. If T has a cyclic vector, show that T has n distinct eigen values.

6.3 Invariant Factors

Definition 6.3.1. Let T be a linear operator on a vector space V. A subspace W of V is said to be T -admissible if

1. W is T-invariant;
2. if $f(T)(v)$ is in W, there exists a vector w in W such that $f(T)(v) = f(T)(w)$.

Note that, from the discussion we had done in the introduction of this unit, if V is decomposed as $V = W \oplus W'$, where both W and W' are invariant, then any vector $v \in V$ has a unique representation $v = w + w'$, where $w \in W$ and $w' \in W'$. If f is any polynomial over the scalar field, then $f(T)(v) = f(T)(w) + f(T)(w')$. Since W and W' are T-invariant, the vectors $f(T)(w)$ and $f(T)(w')$ lies in W and W' respectively. Thus, $f(T)(v)$ is in W if and only if $f(T)(w') = 0$. Hence, we can say that for such a case, W is admissible.

Let W be a proper T-invariant subspace. Let us try to find a non-zero vector v such that

$$W \cap Z(v; T) = \{0\}.$$

We can choose a vector w' which is not in W. Consider the T-conductor $S(w'; W)$, which consists of all polynomials g such that $g(T)(w')$ is in W. Recall that the monic polynomial f which generates the ideal $S(w'; W)$ is also called the T-conductor of w' into W. The vector $f(T)(w')$ is in W. Now, if W is T-admissible, there is a w'' in W with $f(T)(w') = f(T)(w'')$. Let $w = w' - w''$ and let g be any polynomial. Since $w' - w$ is in W, $g(T)(w')$ will be in W if and only if $g(T)(w)$ is in W; in other words, $S(w; W) = S(w'; W)$. Thus, the polynomial f is also the T-conductor of w into W. But $f(T)(w) = 0$ which tells us that $g(T)(w)$ is in W if and only if $g(T)(w) = 0$, that is, the subspaces $Z(v; T)$ and W are independent and f is the T-annihilator of v.

Theorem 6.3.2. (Cyclic Decomposition Theorem) Let T be a linear operator on a finite-dimensional vector space V and let W_0 be a proper T-admissible subspace of V. There exist non-zero vectors $v_1, \ldots, v_k$ in V with respective T-annihilators $m_1, \ldots, m_k$ such that

1. $V = W_0 \oplus Z(v_1; T) \oplus \cdots \oplus Z(v_k; T)$;
2. m_r divides m_{r-1}, $r = 2, \ldots, k$.

Furthermore, the integer k and the annihilators $m_1, \ldots, m_k$ are uniquely determined by 1 and 2 and the fact that no v_r is 0.

The proof is rather lengthy and has been omitted for general good.

Our next corollary gives us the answer to our primary question which we asked at the beginning of this unit regarding the existence of a T-invariant subspace W' which forms a **complementary** for a T-invariant subspace W of V.

Corollary 6.3.3. If T is a linear operator on a finite-dimensional vector space, every T-admissible subspace has a complementary subspace which is also invariant under T.

Proof. Let W be an admissible subspace of V. If $W = V$, the required complement is $\{0\}$. If W is proper, then we apply the Cyclic decomposition theorem and let

$$W' = Z(v_1; T) \oplus \cdots \oplus Z(v_k; T).$$

Then W' is invariant under T and $V = W \oplus W'$. □

Corollary 6.3.4. Let T be a linear operator on a finite-dimensional vector space V.

1. There exists a vector v in V such that the T-annihilator of v is the minimal polynomial for T.
2. T has a cyclic vector if and only if the characteristic and minimal polynomials for T are identical.

Proof. If $V = \{0\}$, the results are trivially true. If $V \neq \{0\}$, let

$$V = Z(v_1;T) \oplus \cdots \oplus Z(v_k;T)$$

where the T-annihilators $m_1, \ldots, m_k$ are such that m_{r+1} divides m_r, $1 \leq r \leq k-1$. As we noted in the previous theorem, it follows easily that m_1 is the minimal polynomial for T, that is, the T-conductor of V into $\{0\}$.

We saw in the previous unit that if T has a cyclic vector, the minimal polynomial for T coincides with the characteristic polynomial. Choose any vector v as in 1. If the degree of the minimal polynomial is $\dim V$, then $V = Z(v;T)$. □

Theorem 6.3.5. (Generalized Cayley-Hamilton Theorem) Let T be a linear operator on a finite-dimensional vector space V. Let m and f be the minimal and characteristic polynomials for T, respectively. Then

1. m divides f;
2. m and f have the same prime factors, except for multiplicities;
3. If $m = f_1^{r_1} \ldots f_k^{r_k}$ is a prime factorization of m, then $f = f_1^{d_1} \ldots f_k^{d_k}$, where d_i is the nullity of $f_i(T)^{r_i}$ divided by the degree of f_i.

Proof. If $V = \{0\}$, then the case is trivial. To prove 1 and 2, consider a cyclic decomposition of V. As in the proof of the above corollary, $m_1 = m$. Let U_i be the restriction of T to $Z(v_i;T)$. Then U_i has a cyclic vector and so m_i is both the minimal as well as characteristic polynomial for U_i. Hence, the characteristic polynomial f is the product $f = m_1 \ldots m_r$. Clearly, $m_1 = m$ divides f and this proves 1. Obviously any prime divisor of m is a prime divisor of f. Conversely, a prime divisor of $f = m_1 \ldots m_r$ must divide one of the factors m_i, which is turn divides m_1.

Let the given factorization in the statement of the theorem be the prime factorization of m. We use the primary decomposition theorem which tells us that, if V is the null space of $f_i(T)^{r_i}$, then

$$V = V_1 \oplus \cdots \oplus V_k$$

and $f_i^{r_i}$ is the minimal polynomial of the operator T_i, obtained by restricting T to the subspace V_i. Apply part 2 of the present theorem to the operator T_i. Since its minimal polynomial is a power of the prime f_i, the characteristic polynomial for T_i has the form $f_i^{d_i}$, where $d_i \geq r_i$. Obviously

$$d_i = \frac{\dim V}{\deg f_i}$$

and (almost by definition) $\dim V_i = \text{nullity} f_i(T)^{r_i}$. Since T is the direct sum of the operators $T_1, \ldots, T_k$, the characteristic polynomial f is the product

$$f = f_1^{d_1} \ldots f_k^{d_k}.$$

□

The polynomials $m_1, \ldots, m_r$ are called the invariant factors for a matrix B.

6.4 Rational Forms

Let us try to understand the cyclic-decomposition theorem for matrices. If we have the operator T and the direct-sum decomposition and $\mathcal{B}_i$ be the cyclic ordered basis $\{v_i, T(v_i), \ldots, T^{k_i-1}(v_i)\}$ for $Z(v_i;T)$. Here, k_i denotes the dimension of $Z(v_i;T)$, that is, the degree of the annihilator m_i. The matrix of the induced operator T_i in the ordered basis $\mathcal{B}_i$ is the companion matrix of the polynomial m_i. Thus, if we let $\mathcal{B}$ be the ordered basis for V which is the union of the $\mathcal{B}_i$ arranged in the order $\mathcal{B}_1, \ldots, \mathcal{B}_r$, then the matrix of T in the ordered basis $\mathcal{B}$ will be

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix}$$

where A_i is the $k_i \times k_i$ companion matrix of m_i. An $n \times n$ matrix A, which is the direct-sum of companion matrices of non-scalar monic polynomials $m_1, \ldots, m_r$ such that m_{i+1} divides m_i for $i = 1, \ldots, r-1$, will be said to be in rational form.

Theorem 6.4.1. Let F be a field and let B be an $n \times n$ matrix over F. Then B is similar over the field F to unique matrix which is in rational form.

We have seen a simpler form for non-diagonalizable matrices, that is the Jordan form. We have a theorem for triangular matrices which states that

Theorem 6.4.2. An $n \times n$ is triangulable, that is, similar to a triangular matrix if and only if its minimal polynomial is the product of linear factors (not necessarily distinct).

Now, the Jordan form is a triangular matrix and we know that the triangular matrices are the next "simplest" matrices to deal with, right after diagonal ones and we have also seen with certain examples that the Jordan form was deducible for a matrix when its minimal polynomial, or we can also say that its characteristic polynomial was the product of linear factors. But this is not always the case. For example, consider the matrix over the real field

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

The characteristic polynomial of the above matrix is $f(x) = x^2 + 1$. Since the minimal polynomial of a matrix divides its characteristic polynomial, and since the characteristic polynomial is irreducible, so the minimal polynomial of the matrix is also $m(x) = x^2 + 1$. These are the cases when the rational forms come into play. We will illustrate how we find the rational form for a matrix.

Illustration 6.4.1. 1. Consider the real matrix

$$A = \begin{bmatrix} -2 & 0 & 0 \\ -1 & -4 & -1 \\ 2 & 4 & 0 \end{bmatrix}.$$

Then the characteristic polynomial of the matrix can be calculated and is equal to $f(x) = x^3 + 6x^2 + 12x + 8 = (x+2)^3$. We have, $A + 2I \neq 0$, but $(A+2I)^2 = 0$. Thus, the minimal polynomial of the matrix is $(x+2)^2$. We know that the largest invariant factor is simply the minimal polynomial. Furthermore, we know that the size of our canonical form matrix must be 3×3, and that our invariant factors must divide the minimal polynomial. Thus, there are two invariant factors $(x+2)^2 = x^2+4x+4$ and $x+2$. Therefore, the rational canonical form of the matrix is

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & 1 & -4 \end{bmatrix}.$$

Note that the minimal polynomial of A is the product of linear factors and hence we can find the Jordan form for A. (Find it)

Exercise 6.4.3. 1. Find the minimal polynomials and the rational form for the following matrices

$$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} c & 0 & -1 \\ 0 & c & 1 \\ -1 & 1 & c \end{bmatrix}.$$

2. Find the rational form of the matrix

$$\begin{bmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ -3 & -3 & -5 \end{bmatrix}.$$

Few Probable Questions

1. Show that for a direct-sum decomposition of a finite-dimensional vector space V, $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, there exists k projection operators E_i such that the range of each E_i is W_i and $I = E_1 + \cdots + E_k$.
2. State and prove the primary decomposition theorem.
3. State and prove the Generalized Cayley-Hamilton theorem.
4. Find the minimal polynomial, invariant factors and the rational form of the following matrix

$$\begin{bmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{bmatrix}.$$

Unit 7

Course Structure

- Inner Product Spaces: Inner product and Norms. Adjoint of a linear operator, Normal, self adjoint, unitary, orthogonal operators and their matrices,

---

7.1 Introduction

The inner product or the dot product on $\mathbb{R}^n$ is defined by

$$(u, v) = u_1v_1 + u_2v_2 + \ldots + u_nv_n,$$

where $u = (u_1, \ldots, u_n),\ v = (v_1, \ldots, v_n) \in \mathbb{R}^n$. It satisfies the following properties

1. Linearity : $(au + bv, w) = a(u, w) + b(v, w)$, for real constants $a,\ b$;
2. Symmetry : $(u, v) = (v, u)$;
3. Positive definite : $(u, u) \geq 0$ and $(u, u) = 0$ if and only if $u = 0$,

for all $u, v, w \in \mathbb{R}^n$. With the dot product we have geometric concepts such as the length of a vector, the angle between two vectors, orthogonality, etc. We shall try to push these concepts to abstract vector spaces so that geometric concepts can be applied to describe abstract vectors.

Objectives

After reading this unit, you will be able to

- define inner product and norm on a vector space
- define and discuss the basic properties of inner product space
- recall the idea of orthogonal and orthonormal set and their basic properties
- define adjoint of a linear operator and discuss its properties
- define unitary operator and discuss its properties
- define normal operators and discuss some of its properties

7.2 Inner Product Space

Let us begin with the definition of inner product on a vector space V.

Definition 7.2.1. Let F be the field of real numbers or the field of complex numbers, and V a vector space over F. An inner product on V is a function $\langle , \rangle : V \times V \to F$ such that it satisfies the following axioms.

1. Linearity : $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$, for real constants $a,\ b$;
2. Symmetry : $\langle u, v\rangle = \overline{\langle v, u\rangle}$, the bar denoting complex conjugation;
3. Positive definite : $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0$ if and only if $u = 0$

for every vector $u, v, w \in V$ and scalar $a, b \in F$.

Let us look into a few common examples.

Example 7.2.2. The standard dot product on $\mathbb{R}^n$ defined in the introduction is one of the most common examples of the inner product.

Example 7.2.3. Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in $\mathbb{R}^2$. Define

$$\langle x, y\rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + 5x_2y_2.$$

Then $\langle , \rangle$ is an inner product on $\mathbb{R}^2$. t is easy to see the linearity and the symmetric property. As for the positive definite property, note that

$$\begin{aligned} \langle x, x\rangle &= 2x_1^2 - 2x_1x_2 + 5x_2^2 \\ &= (x_1 + x_2)^2 + (x_1 - 2x_2)^2 \geq 0. \end{aligned}$$

Moreover, $\langle x, x\rangle = 0$ if and only if

$$x_1 + x_2 = 0, \quad x_1 - 2x_2 = 0,$$

which implies $x_1 = x_2 = 0$.

Example 7.2.4. Let V be the vector space of all continuous complex-valued functions on the unit interval, $0 \leq t \leq 1$. Let

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt.$$

Example 7.2.5. The vector space $M_{m,n}$ of all $m \times n$ real matrices can be made into an inner product space under the inner product

$$\langle A, B\rangle = \text{tr}\left(B^T A\right),$$

where $A,\ B \in M_{m,n}$.

For instance, when $m = 3$, $n = 2$, and for

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix},$$

we have

$$B^T A = \begin{bmatrix} b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} & b_{11}a_{12} + b_{21}a_{22} + b_{31}a_{32} \\ b_{12}a_{11} + b_{22}a_{21} + b_{32}a_{31} & b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32} \end{bmatrix}.$$

Thus,

$$\begin{aligned}\langle A, B\rangle &= b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} + b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32} \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}b_{ij}.\end{aligned}$$

Definition 7.2.6. Let V be a vector space with an inner product $\langle , \rangle$. Let $v \in V$. The norm (also called the length) of v is defined as the number

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

If $\|v\| = 1$, then v is called a unit vector and v is said to be normalized. For any nonzero vector $u \neq 0$, we have the unit vector

$$\hat{u} = \frac{1}{\|u\|}u.$$

This process is called normalizing u.

Also, we can think of a matrix for an inner product $\langle , \rangle$ with respect to some basis of the underlying vector. This means the matrix of an inner product varies with the underlying basis. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis of an n-dimensional inner product space V . For vectors $u,\ v \in V$, we have

$$\begin{aligned}u &= x_1v_1 + x_2v_2 + \ldots + x_nv_n \\ v &= y_1v_1 + y_2v_2 + \ldots + y_nv_n.\end{aligned}$$

The linearity implies

$$\begin{aligned}\langle u, v\rangle &= \left\langle \sum_{i=1}^{n} x_iv_i \sum_{j=1}^{n} y_jv_j \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j\langle v_i, v_j\rangle.\end{aligned}$$

We call the $n \times n$ matrix

$$A = \begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \cdots & \langle v_1, v_n\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \cdots & \langle v_2, v_n\rangle \\ \vdots & \vdots & & \vdots \\ \langle v_n, v_1\rangle & \langle v_n, v_2\rangle & \cdots & \langle v_n, v_n\rangle \end{bmatrix}$$

is the matrix of the inner product $\langle , \rangle$ relative to the basis $\mathcal{B}$. The inner product can be written as

$$\langle u, v\rangle = [u]_{\mathcal{B}^T} A[v]_{\mathcal{B}}.$$

Thus, we have arrived at the following theorem.

Theorem 7.2.7. Let V be an n-dimensional vector space with an inner product $\langle , \rangle$ and let A be the matrix of $\langle , \rangle$ relative to a basis $\mathcal{B}$. Then for any vectors $u, v \in V$,

$$\langle u, v\rangle = x^T Ay$$

where, x and y are the coordinate vectors of u and v respectively with respect to the basis $\mathcal{B}$.

Example 7.2.8. For the inner product

$$\langle x, y\rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + 5x_2y_2.$$

on $\mathbb{R}^2$, the matrix with respect to the standard basis is given by

$$A = \begin{bmatrix} \langle e_1, e_1\rangle & \langle e_1, e_2\rangle \\ \langle e_2, e_1\rangle & \langle e_2, e_2\rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 5 \end{bmatrix}.$$

Having discussed about the inner product and norm, we are in a position to define the inner product space.

Definition 7.2.9. An inner product space is a real or complex vector space, together with a specified inner product on that space.

A finite-dimensional real inner product space is often called a Euclidean space. A complex inner product space is often referred to as a unitary space. Certain basic properties of inner product spaces are listed in the theorem below.

Theorem 7.2.10. If V is an inner product space, then for any vectors $u, v \in V$ and any scalar c

1. $\|cu\| = |c|\|u\|$;
2. $\|u\| > 0$ for $u \neq 0$;
3. $|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$ (Cauchy-Schwarz inequality);
4. $\|u + v\| \leq \|u\| + \|v\|$.

Definition 7.2.11. Let $u,\ v$ be vectors in an inner product space V. Then u is orthogonal to v if $\langle u, v\rangle = 0$. We can also say that u and v are orthogonal. If S is a set of vectors in V, S is called an orthogonal set provided all pairs of distinct vectors in S are orthogonal. An orthonormal set is an orthogonal set S with the additional property that $\|u\| = 1$ for every u in S.

The zero vector is orthogonal to every vector in V and is the only vector with this property. It is appropriate to think of an orthonormal set as a set of mutually perpendicular vectors, each having length 1.

Example 7.2.12. The standard basis of either $\mathbb{R}^n$ or $\mathbb{C}^n$ is an orthonormal set with respect to the standard inner product.

Orthogonal sets have certain basic properties.

Theorem 7.2.13. An orthogonal set of non-zero vectors is linearly independent.

Proof. Let S be a finite or infinite orthogonal set of non-zero vectors in a given inner product space. Suppose $v_1, v_2, \ldots, v_n$ are distinct vectors in S and that

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n.$$

Then

$$\begin{aligned} \langle v, v_i\rangle &= \left\langle \sum_{j=1}^{n} c_jv_j, v_i \right\rangle \\ &= \sum_{j=1}^{n} c_j\langle v_i, v_j\rangle \\ &= c_i\langle v_i, v_i\rangle. \end{aligned}$$

Since $\langle v_i, v_i\rangle \neq 0$, it follows that

$$c_i = \frac{\langle v, v_i\rangle}{\|v_i\|^2}, \quad 1 \leq i \leq n.$$

Thus, when $v = 0$, each c_i is zero. Hence, S is linearly independent. □

7.3 Adjoint of a Linear Operator

the adjoint of an operator is a generalization of the notion of the Hermitian conjugate of a complex matrix to linear operators on complex Hilbert spaces.

Definition 7.3.1. Let $T : V \to V$ be a linear transformation on an inner product space V. The adjoint of T is a transformation $T^* : V \to V$ satisfying

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle$$

for all $x, y \in V$.

It must come to mind that the matrix representation of T^* with respect to any orthonormal basis $\mathcal{B}$ is the complex conjugate of $[T]_{\mathcal{B}}$. Also, it will be observed that the adjoint of T depends not only on T but on the inner product as well. To see the existence of the "adjoint" of a linear operator T on V , we begin with linear functionals on an inner product space and their relation to the inner product.

The basic result is that any linear functional f on a finite-dimensional inner product space is "inner product with a fixed vector in the space," i.e., that such an f has the form

$$f(u) = \langle u, v\rangle,$$

for some fixed $v \in V$.

Theorem 7.3.2. Let V be a finite-dimensional inner product space, and f a linear functional on V. Then there exists a unique vector $v \in V$ such that

$$f(u) = \langle u, v\rangle, \;\; \forall u \in V.$$

Proof. Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis for V. Put

$$v = \sum_{j=1}^{n} \overline{f(v_j)} v_j$$

and let f_v be a linear functional defined by

$$f_v(u) = \langle u, v\rangle.$$

Then

$$f_v(v_i) = \left\langle v_i, \sum_{j=1}^{n} \overline{f(v_j)} v_j \right\rangle = f(v_i).$$

Since this is true for each v_i, it follows that $f = f_v$. Now, for the uniqueness suppose that $w \in V$ such that $\langle u, v\rangle = \langle u, w\rangle$ for all $u \in V$. Then $\langle v - w, v - w\rangle = 0$ which implies $v = w$. Thus there is exactly one vector v determining the linear functional f in the stated manner. □

We use this result to prove the existence of the "adjoint" of a linear operator T on V.

Theorem 7.3.3. For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T^* on V such that

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle$$

for all $u, v \in V$.

Proof. Let $v \in V$. Then $\mapsto \langle T(u), v\rangle$ is a linear functional on V. By the previous theorem, there is a unique $v' \in V$ such that $\langle T(u), v\rangle = \langle u, v'\rangle$ for every $u \in V$. Let T^* denote the mapping $v \mapsto v'$

$$v' = T^*(v),$$

which implies that $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$. We will verify the linearity of T^*. Let $v, w \in V$ and c be a scalar. Then for any u,

$$\begin{aligned}\langle u, T^*(cv + w)\rangle &= \langle T(u), cv + w\rangle \\ &= \langle T(u), cv\rangle + \langle T(u), w\rangle \\ &= \bar{c}\langle T(u), v\rangle + \langle T(u), w\rangle \\ &= \bar{c}\langle u, T^*(v)\rangle + \langle u, T^*(w)\rangle \\ &= \langle u, cT^*(v)\rangle + \langle u, T^*(w)\rangle \\ &= \langle u, cT^*(v) + T^*(w)\rangle.\end{aligned}$$

Thus, $T^*(cv + w) = cT^*(v) + T^*(w)$ and T^* is linear.

The uniqueness of T^* is clear. For any $v \in V$, the vector $T^*(v)$ is uniquely determined as the vector v' such that $\langle T(u), v\rangle = \langle u, v'\rangle$ for every u. □

Theorem 7.3.4. Let V be a finite-dimensional inner product space and let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be an (ordered) orthonormal basis for V. Let T be a linear operator on V and let A be the matrix of T in the ordered basis $\mathcal{B}$. Then $a_{kj} = \langle T(v_j), v_k\rangle$.

Proof. Since $\mathcal{B}$ is an orthonormal basis, we have

$$v = \sum_{k=1}^{n} \langle v, v_k\rangle v_k.$$

The matrix A is defined by

$$T(v_j) = \sum_{k=1}^{n} a_{kj} v_k$$

and since

$$T(v_j) = \sum_{k=1}^{n} \langle T(v_j), v_k\rangle v_k$$

we have, $a_{kj} = \langle T(v_j), v_k\rangle$. □

Corollary 7.3.5. Let V be a finite-dimensional inner product space, and let T be a linear operator on V. In any orthonormal basis for V , the matrix of T^* is the conjugate transpose of the matrix of T.

Proof. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for V, let $A = [T]_{\mathcal{B}}$ and $B = [T^*]_{\mathcal{B}}$. By the previous theorem,

$$a_{kj} = \langle T(v_j), v_k\rangle, \quad b_{kj} = \langle T^*(v_j), v_k\rangle. \tag{7.3.1}$$

By the definition of T^*, we then have,

$$\begin{aligned}b_{kj} &= \langle T^*(v_j), v_k\rangle \\ &= \overline{\langle v_k, T^*(v_j)\rangle} \\ &= \overline{\langle T(v_k), v_j\rangle} \\ &= \overline{a_{kj}}.\end{aligned}$$

□

By Theorem 7.3.3, every linear operator on a finite-dimensional inner product space V has an adjoint on V. In the infinite-dimensional case this is not always true. But in any ease there is at most one such operator T^*.

Let us now discuss a few properties of adjoint of linear operators.

Theorem 7.3.6. Let V be a finite-dimensional inner product space. If T and U are linear operators on V and c be a scalar,

1. $(T+U)^* = T^* + U^*$;
2. $(cT)^* = \overline{c}T^*$;
3. $(TU)^* = U^*T^*$;
4. $(T^*)^* = T$.

Proof. Left as exercise. □

Definition 7.3.7. A linear operator T is called self-adjoint if $T = T^*$.

If $\mathcal{B}$ is an orthonormal basis for V, then $[T^*]_{\mathcal{B}} = [T]^*_{\mathcal{B}}$. Hence, T is self-adjoint if and only if its matrix in every orthonormal basis is a self-adjoint matrix.

7.4 Unitary Operators

Before defining unitary operators, let us first define the notion of linear operators which preserve the underlying inner products.

Definition 7.4.1. Let V and W be inner product spaces over the same field, and let T be a linear transformation from V into W. We say that T preserves inner products if $\langle T(u), T(v)\rangle = \langle u, v\rangle$. An isomorphism of V onto W is a vector space isomorphism T of V onto W which also preserves inner products.

If T preserves inner products, then $\|T(u)\| = \|u\|$ and so T is necessarily non-singular.

Theorem 7.4.2. Let V and W be finite-dimensional inner product spaces over the same field, having the same dimension. If T is a linear transformation from V into W, the following are equivalent.

1. T preserves inner products.
2. T is an (inner product space) isomorphism.
3. T carries every orthonormal basis for V onto an orthonormal basis for W.
4. T carries some orthonormal basis for V onto an orthonormal basis for W.

Theorem 7.4.3. Let V and W be inner product spaces over the same field, and let T be a linear transformation from V into W. Then T preserves inner products if and only if $\|T(u)\| = \|u\|$ for every u in V.

Definition 7.4.4. A unitary operator on an inner product space is an isomorphism of the space onto itself.

The product of two unitary operators is unitary. For if U_1 and U_2 are two unitary operators, then U_2U_1 is invertible and $\|U_2U_1(u)\| = \|U_1(u)\| = \|u\|$ for each u. Also, the inverse of a unitary operator is unitary (verify). Since the identity operator is clearly unitary, we see that the set of all unitary operators on an inner product space is a group, under the operation of composition.

Theorem 7.4.5. Let U be a linear operator on an inner product space V. Then U is unitary if and only if the adjoint U^* of U exists and $UU^* = U^*U = I$.

Proof. Suppose U is unitary. Then U is invertible and

$$\langle U(v), w\rangle = \langle U(v), UU^{-1}(w)\rangle = \langle u, U^{-1}(w)\rangle$$

for all v, w. Hence, U^{-1} is the adjoint of U.

Conversely, suppose U^* exists and $UU^* = U^*U = I$. Then U is invertible and $U^{-1} = U^*$. We only need to show that U preserves inner products. We have,

$$\langle U(v), U(w)\rangle = \langle v, U^*U(w)\rangle = \langle v, I(w)\rangle = \langle v, w\rangle$$

for all v, w. □

Definition 7.4.6. A complex $n \times n$ matrix is called unitary if $A^*A = I$.

Theorem 7.4.7. Let V be a finite-dimensional inner product space and let U be a linear operator on V. Then U is unitary if and only if the matrix of U in some (or every) ordered orthonormal basis is a unitary matrix.

Proof. Left as exercise. □

Definition 7.4.8. A real or complex $n \times n$ matrix A is said to be orthogonal if $A^tA = I$.

A real orthogonal matrix i s unitary ; and, a unitary matrix is orthogonal if and only if each of its entries is real.

7.5 Normal Operators

Definition 7.5.1. Let V be a finite-dimensional inner product space and T a linear operator on V. We say that T is normal if it commutes with its adjoint i.e., $TT^* = T^*T$.

Any self-adjoint operator is normal, as is any unitary operator. Any scalar multiple of a normal operator is normal; however, sums and products of normal operators are not generally normal. Although it is by no means necessary, we shall begin our study of normal operators by considering self-adjoint operators.

Theorem 7.5.2. Let V be an inner product space and T a self-adjoint linear operator on V. Then each characteristic value of T is real, and characteristic vectors of T associated with distinct characteristic values are orthogonal.

Proof. Suppose c is a characteristic value of T, i.e., $T(v) = cv$ for some non-zero vector v. Then

$$\begin{aligned} c\langle v, v\rangle &= \langle cv, v\rangle \\ &= \langle T(v), v\rangle \\ &= \langle v, T(v)\rangle \\ &= \langle v, cv\rangle \\ &= \overline{c}\langle v, v\rangle. \end{aligned}$$

Since $\langle v, v\rangle \neq 0$, we must have $c = \overline{c}$. Suppose that $T(w) = dw$, $w \neq 0$. Then

$$\begin{aligned} c\langle v, w\rangle &= \langle T(v), w\rangle \\ &= \langle v, T(w)\rangle \\ &= \langle v, dw\rangle \\ &= \overline{d}\langle v, w\rangle \\ &= d\langle v, w\rangle. \end{aligned}$$

If $c \neq d$, then $\langle v, w\rangle = 0$. □

Theorem 7.5.3. On a finite-dimensional inner product space of positive dimension, every self-adjoint operator has a (non-zero) characteristic vector.

Proof. Let V be an inner product space of dimension n, where $n > 0$, and let T be a self-adjoint operator on V. Choose an orthonormal basis $\mathcal{B}$ for V and let $A = [T]_{\mathcal{B}}$. Since $T = T^*$, we have $A = A^*$. Now let W be the space of $n \times 1$ matrices over $\mathbb{C}$, with inner product $\langle X, Y\rangle = Y * X$. Then $U(X) = AX$ defines a self-adj oint linear operator U on W. The characteristic polynomial, $\det(xI - A)$, is a polynomial of degree n over the complex numbers ; every polynomial over $\mathbb{C}$ of positive degree has a root. Thus, there is a complex number c such that $\det(cI - A) = O$. This means that $A - cI$ is singular, or that there exists a non-zero X such that $AX = cX$. Since the operator U (multiplication by A) is self-adjoint, it follows from the previous theorem that c is real. If V is a real vector space, we may choose X to have real entries. For then A and $A-cI$ have real entries, and since $A - cI$ is singular, the system $(A - cI)X = 0$ has a non-zero real solution X. It follows that there is a non-zero vector α: in V such that $T\alpha = c\alpha$. □

Few Probable Questions

1. Define adjoint of a linear operator. Show that the adjoint of a linear operator over a finite dimensional vector space exists.
2. What can you say about the matrix of the adjoint of a linear operator defined on a finite dimensional vector space V? Give complete justification.
3. Show that U is a unitary operator if and only if $UU^* = U^*U = I$ for the adjoint U^* of U.

Unit 8

Course Structure

- Bilinear and Quadratic forms: Bilinear forms, quadratic forms, Reduction and classification of quadratic forms, Sylvester's law of Inertia.

8.1 Introduction

A bilinear form on a real vector space V is a function f which assigns a number to each pair of elements of V, a scalar from the underlying field, satisfying certain properties. We can begin with an example of a map from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, where $\mathbb{R}$ is the underlying field, defined by

$$\langle X, Y \rangle = X^T.Y = x_1y_1 + \cdots + x_ny_n.$$

This is the most common dot product, that we are familiar with. The property of the dot product which we will use to generalize to bilinear forms is bilinearity: the dot product is a linear function from V to F, where F is the underlying field, if one of the elements is fixed. Bilinear forms are meant to be a generalization of the dot product on $\mathbb{R}^n$.

Objectives

After reading this unit, you will be able to

- define bilinear forms and see certain examples of it
- learn properties related to them
- define quadratic forms and associated matrices
- define definiteness of a form and its associated matrices
- learn about the equivalent definitions of definiteness of a matrix and form
- solve problems related to the definiteness of matrices

8.2 Bilinear Forms

Definition 8.2.1. Let V be a vector space over F. We define a bilinear form to be a function $f : V \times V \to F$ such that

$$\begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w), \ \ v_1, v_2, w \in V \\ f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2), \ \ v, w_1, w_2 \in V \\ f(cv, w) &= cf(v, w) = f(v, cw), \ \ v, w \in W, \ c \in F \end{aligned}$$

We will often use the notation $\langle v, w\rangle$ for $f(v, w)$.

The zero function from $V \times V$ into F is clearly a bilinear form. It is also true that any linear combination of bilinear forms on V is again a bilinear form(check it). All this may be summarized by saying that the set of all bilinear forms on V is a subspace of the space of all functions from $V \times V$ into F. We denote the space of bilinear forms on V by $L(V, V, F)$.

Example 8.2.2. Let V be a vector space over the field F and let L_1 and L_2 be linear functions on V. Define f by

$$f(u, v) = L_1(u)L_2(v).$$

If we fix v and regard f as a function of u, then we simply have a scalar multiple of the functional L_1. And fixing u, f is a scalar multiple of L_2. Hence f is a bilinear form on V.

Example 8.2.3. Let m and n be positive integers and F a field. Let V be the vector space of $m \times n$ matrices over F. Let A be a fixed $m \times n$ over F. Define

$$f_A(X, Y) = \text{tr}(X^T AY).$$

Then f_A is a bilinear form on V. If X, Y, Z are $m \times n$ matrices over F, then

$$\begin{aligned} f_A(cX + Z, Y) &= \text{tr}[(cX + Z)^T AY] \\ &= \text{tr}(cX^T AY) + \text{tr}(Z^T AY) = cf_A(X, Y) + f_A(Z, Y). \end{aligned}$$

Of course, we have used the fact that the transpose operation and the trace function are linear. It is even easier to show that f_A is linear as a function of its second argument. In the special case, $n = 1$, the matrix $X^T AY$ is 1×1 matrix, that is, a scalar, and the bilinear form is simply

$$f_A(X, Y) = \sum_{i,j} A_{ij} x_i y_j.$$

Example 8.2.4. Let F be a field. Let us find all bilinear forms on the space F^2. Suppose f is such a bilinear form. If $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are in F^2, then

$$\begin{aligned} f(x, y) &= f(x_1 e_1 + x_2 e_2, y) \\ &= x_1 f(e_1, y) + x_2 f(e_2, y) \\ &= x_1 f(e_1, y_1 e_1 + y_2 e_2) + x_2 f(e_2, y_1 e_1 + y_2 e_2) \\ &= x_1 y_1 f(e_1, e_1) + x_1 y_2 f(e_1, e_2) + x_2 y_1 f(e_2, e_1) + x_2 y_2 f(e_2, e_2). \end{aligned}$$

Hence, f is completely determined by the four scalars $A_{ij} = f(e_i, e_j) = \langle e_i, e_j\rangle$ by

$$\begin{aligned} f(x, y) &= A_{11} x_1 y_1 + A_{12} x_1 y_2 + A_{21} x_2 y_1 + A_{22} x_2 y_2 \\ &= \sum_{i,j} A_{ij} x_i y_j. \end{aligned}$$

Thus, if X and Y are the coordinate matrices of x and y, and if A is the above matrix, then

$$f(x,y) = X^T AY.$$

This can be generalized for any finite-dimensional vector spaces.

Definition 8.2.5. (**Bilinear forms on** $\mathbb{R}^n$) Every bilinear form on $\mathbb{R}^n$ has the form

$$\langle x, y\rangle = x^T Ay = \sum_{i,j} a_{ij}x_i y_j, \quad x, y \in \mathbb{R}^n$$

for some $n \times n$ matrix A and we also have $a_{ij} = \langle e_i, e_j\rangle$ for all i, j. e_i is the n tuple of real numbers whose ith entry is 1 and all other entries are 0.

Definition 8.2.6. Let V be a finite-dimensional vector space, and let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis for V. If f is a bilinear form on V, the matrix of f in the ordered basis $\mathcal{B}$ is the $n \times n$ matrix A with entries $A_{ij} = f(v_i, v_j)$. We shall denote this matrix by $[f]_{\mathcal{B}}$.

Theorem 8.2.7. Let V be a finite-dimensional vector space over the field F. For each ordered basis $\mathcal{B}$ of V, the function which associates with each bilinear form on V, its matrix in the ordered basis $\mathcal{B}$ is an isomorphism of the space $L(V, V, F)$ onto the space of $n \times n$ matrices over the field F.

Proof. We have seen that $f \to [f]_{\mathcal{B}}$ is a one-one correspondence between the set of bilinear forms on V and the set of all $n \times n$ matrices over F. That this is a linear transformation is easy to see, because

$$(cf + g)(v_i, v_j) = cf(v_i, v_j) + g(v_i, v_j)$$

for each i and j. This simply says that

$$[cf + g]_{\mathcal{B}} = c[f]_{\mathcal{B}} + [g]_{\mathcal{B}}.$$

□

Corollary 8.2.8. If $\mathcal{B} = \{v_1, \ldots, v_n\}$ is an ordered basis for V, and $\mathcal{B}^* = \{L_1, \ldots, L_n\}$ be an ordered basis for V^*, then the n^2 bilinear forms

$$f_{ij}(x, y) = L_i(x)L_j(y), \quad 1 \neq i \neq n, \; 1 \leq j \leq n$$

form a basis for $L(V, V, F)$. In particular, the dimension of $L(V, V, F)$ is n^2.

The concept of the matrix of a bilinear form in an ordered basis is similar to that of the matrix of a lineal' operator in an ordered basis. Just as for linear operators, we shall be interested in what happens to the matrix representing a bilinear form, as we change from one ordered basis to another. So, suppose $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{v_1', \ldots, v_n'\}$ two ordered bases for V and that f is a bilinear form on V. How are the matrices $[f]_{\mathcal{B}}$ and $[f]_{\mathcal{B}'}$ related? Well, let P be the (invertible) $n \times n$ matrix such that

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$$

for all $v \in V$. In other words, define P by

$$v_j' = \sum_{i=1}^{n} P_{ij} v_i.$$

For any vectors $v, w \in V$,

$$\begin{aligned} f(v,w) &= [v]_{\mathcal{B}}^T[f]_{\mathcal{B}}[v]_{\mathcal{B}} \\ &= (P[v]_{\mathcal{B}'})^T[f]_{\mathcal{B}}P[w]_{\mathcal{B}'} \\ &= [v]_{\mathcal{B}'}^T(P^T[f]_{\mathcal{B}}P)[w]_{\mathcal{B}'}. \end{aligned}$$

By the definition and uniqueness of the matrix representing f in the ordered basis $\mathcal{B}^{'}$, we must have

$$[f]_{\mathcal{B}'} = P^T[f]_{\mathcal{B}}P.$$

One consequence of the change of basis formula is the following: If A and B are $n \times n$ matrices which represent the same bilinear form on V in (possibly) different ordered bases, then A and B have the same rank. For, if P is an invertible $n \times n$ matrix and $B = P^TAP$, it is evident that A and B have the same rank. This makes it possible to define the rank of a bilinear form on V as the rank of any matrix which represents the form in an ordered basis for V.

It is desirable to give a more intrinsic definition of the rank of a bilinear form. This can be done as follows : Suppose f is a bilinear form on the vector space V. If we fix a vector v in V, then $f(v,w)$ is linear as a function of w. If we fix a vector $v \in V$, then $f(v,w)$ is linear as a function of w. In this way, each fixed v determines a linear functional on V; let us denote this linear functional by $L_f(v)$. To repeat, if v is a vector in V, then $L_f(v)$ is the linear functional on V whose value on any vector w is $f(v,w)$. This gives us a transformation $v \to L_f(v)$ from V into the dual space V^*. Since

$$f(cv_1 + v_2, w) = cf(v_1, w) + f(v_2, w)$$

we see that

$$L_f(cv_1 + v_2) = cL_f(v_1) + L_f(v_2)$$

that is, L_f is a linear transformation from V into V^*.

In a similar manner, f determines a linear transformation R_f from V into V^*. For each fixed $w \in V$, $f(v,w)$ is linear as a function of v. We define $R_f(w)$ to be the linear functional on V whose value on the vector v is $f(v,w)$.

Theorem 8.2.9. Let f be a bilinear form on the finite-dimensional vector space V. Let L_f and R_f be the linear transformations from V into V^* defined by $(L_f(v))(w) = f(v,w) = (R_f(w))(v)$. Then rank$(L_f)$ =rank(R_f).

Definition 8.2.10. If f is a bilinear form on the finite-dimensional space V, the rank of f is the integer r =rank(L_f) =rank(R_f).

Corollary 8.2.11. The rank of a bilinear form is equal to the rank of the matrix of the form in any ordered basis.

Corollary 8.2.12. If f is a bilinear form on the n-dimensional vector space V, the following are equivalent:

1. rank$(f) = n$;
2. For each non-zero $v \in V$, there is a vector $w \in V$ such that $f(v,w) \neq 0$;
3. For each non-zero $w \in V$, there is a vector $v \in V$ such that $f(v,w) \neq 0$.

Definition 8.2.13. A bilinear form f on a vector space V is called non-degenerate (or non-singular) if it satisfies conditions 2 and 3 of the above corollary.

If V is finite-dimensional, then f is non-degenerate provided f satisfies any one of the three conditions of the above corollary. In particular, f is non-degenerate (non-singular) if and only if its matrix in some (every) ordered basis for V is a non-singular matrix.

Example 8.2.14. Let $V = \mathbb{R}^n$, and let f be the bilinear form defined on $v = (x_1, \ldots, x_n)$ and $w = (y_1, \ldots, y_n)$ by

$$f(v, w) = x_1 y_1 + \cdots + x_n y_n.$$

Then f is a non-degenerate bilinear form on $\mathbb{R}^n$. The matrix of f in the standard ordered basis is the $n \times n$ identity matrix

$$f(X, Y) = X^T Y.$$

Example 8.2.15. Let $V = \mathbb{P}_2$ denote the space of real polynomials of degree at most 2. We can define a bilinear form on V by

$$\langle f, g \rangle = \int_0^1 f(x) g(x) dx, \quad f, g \in V.$$

By definition, the matrix of the form is given by

$$a_{ij} = \langle x^{i-1}, x^{j-1} \rangle = \int_0^1 x^{i+j-2} dx = \frac{1}{i+j+2}.$$

Thus, the matrix of the form with respect to the standard basis is

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

8.2.1 Symmetric Bilinear Forms

The main purpose of this section is to answer the following question : If f is a bilinear form on the finite-dimensional vector space V, when is there an ordered basis $\mathcal{B}$ for V in which f is represented by a diagonal matrix? We prove that this is possible if and only if f is a symmetric bilinear form, that is, $f(v, w) = f(w, v)$. The theorem is proved only when the scalar field has characteristic zero, that is, that if n is a positive integer the sum $1 + \cdots + 1$ (n times) in F is not 0.

Definition 8.2.16. Let f be a bilinear form on the vector space V. We say that f is symmetric if $f(v, w) = f(w, v)$ for all $v, w \in V$.

If V is a finite-dimensional, the bilinear form f is symmetric if and only if its matrix A in some (or every) ordered basis is symmetric, $A^T = A$. To see this, one inquires when the bilinear form

$$f(X, Y) = X^T A Y$$

is symmetric. This happens if and only if $X^T A Y = Y^T A X$, for all column matrices X and Y. Since $X^T A Y$ is a 1×1 matrix, we have $X^T A Y = Y^T A^T X$. Thus f is symmetric if and only if $Y^T A^T X = Y^T A X$ for all X, Y. Clearly this just means that $A^T = A$. In particular, one should note that if there is an ordered basis for V in which f is represented by a diagonal matrix, then f is symmetric, for any diagonal matrix is a symmetric matrix.

Definition 8.2.17. 1. (**Positive definite**) A bilinear form f on a real vector space V is positive definite, if

$$\langle v, v \rangle = f(v, v) > 0, \quad v \neq 0.$$

A real $n \times n$ matrix A is positive definite if $x^T A x > 0$ for all $x \neq 0$.

2. (**Negative definite**) A bilinear form f on a real vector space V is negative definite, if

$$\langle v, v\rangle = f(v, v) < 0, \quad v \neq 0.$$

A real $n \times n$ matrix A is positive definite if $x^T Ax < 0$ for all $x \neq 0$.

3. (**Positive Semi-definite**) A bilinear form f on a real vector space V is positive semi-definite, if

$$\langle v, v\rangle = f(v, v) \geq 0, \quad v \in V.$$

A real $n \times n$ matrix A is positive semi-definite if $x^T Ax \geq 0$ for all x.

4. (**Negative Semi-definite**) A bilinear form f on a real vector space V is negative semi-definite, if

$$\langle v, v\rangle = f(v, v) \leq 0, \quad v \in V.$$

A real $n \times n$ matrix A is negative semi-definite if $x^T Ax \leq 0$ for all x.

5. (**Indefinite**) A bilinear form f on a real vector space V is indefinite, if

$$\langle v, v\rangle = f(v, v) > 0, \quad \text{for some } v \in V$$

and

$$\langle v, v\rangle = f(v, v) < 0, \quad \text{for some } v \in V.$$

Example 8.2.18. 1. Consider the bilinear form on $\mathbb{R}^2$ defined by

$$f(x, y) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2.$$

To check whether it is positive definite, we find

$$\begin{aligned} f(x, x) &= x_1^2 - 4x_1x_2 + 5x_2^2 \\ &= (x_1 - 2x_2)^2 + x_2^2 > 0 \end{aligned}$$

for $x = (x_1, x_2) \neq (0, 0)$. Thus, it is positive definite.

2. Consider the bilinear form on $\mathbb{R}^2$ defined by

$$f(x, y) = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 3x_2y_2.$$

Again,

$$\begin{aligned} f(x, x) &= x_1^2 + 4x_1x_2 + 3x_2^2 \\ &= (x_1 + 2x_2)^2 - x_2^2 < 0 \end{aligned}$$

when $x_1 = -2x_2$ and $x_2 \neq 0$.

8.3 Quadratic Forms

Definition 8.3.1. If f is a symmetric bilinear form, the quadratic form associated with f is the function q from V into F defined by

$$q(v) = f(v, v).$$

Theorem 8.3.2. Any quadratic form can be represented by symmetric matrix.

Indeed, if $a_{ij} \neq a_{ji}$, we replace them by new $a_{ij}^{'} = a_{ji}^{'} = \frac{a_{ij}+a_{ji}}{2}$, this does not change the corresponding quadratic form.

Example 8.3.3.

1. The quadratic form $f(x, y) = x^2 + y^2$ is positive for all nonzero (x, y). Hence f is positive definite.
2. The quadratic form $f(x, y) = -x^2 - y^2$ is negative for all nonzero (x, y). Hence f is negative definite.
3. The quadratic form $f(x, y) = (x - y)^2$ is non-negative. This means that f is either zero or positive for all (x, y). Hence f is positive semi-definite.
4. The quadratic form $f(x, y) = -(x - y)^2$ is non-positive. This means that f is either zero or negative for all (x, y). Hence f is negative semi-definite.
5. The quadratic form $f(x, y) = x^2 - y^2$ is indefinite since it can take both positive as well as negative for example, $f(3, 1) = 9 - 1 = 8 > 0$ and $f(1, 3) = 1 - 9 = -8 < 0$.

8.3.1 Definiteness of a 2 Variable Quadratic Form

Let $f(x, y) = ax^2 + 2bxy + cy^2$ which is equal to

$$f(x, y) = \begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} a & b \\ b & a \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}.$$

Here,

$$A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$$

is the symmetric matrix of the quadratic form. The determinant

$$\begin{vmatrix} a & b \\ b & a \end{vmatrix} = ac - b^2$$

is called the discriminant of f. It can be easily seen that

$$ax^2 + 2bxy + cy^2 = a\left(ax + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2.$$

Let us use the notation $D_1 = a$, $D_2 = ac - b^2$. Actually D_1 and D_2 are leading principal minors of A. Note that there exists one more principal (non leading) minor (of degree 1) $D_1^{'} = c$. Then

$$f(x, y) = D_1\left(ax + \frac{b}{a}y\right)^2 + \frac{D_2}{D_1}y^2.$$

From this expression we obtain:

1. If $D_1 > 0$ and $D_2 > 0$, then the form $x^2 + y^2$ type, so it is positive definite;
2. If $D_1 < 0$ and $D_2 > 0$, then the form $-x^2 - y^2$ type, so it is negative definite;
3. If $D_1 > 0$ and $D_2 < 0$, then the form $x^2 - y^2$ type, so it is indefinite; If $D_1 < 0$ and $D_2 > 0$, then the form $-x^2 + y^2$ type, so it is also indefinite.

Thus, if $D_2 < 0$, then the form is indefinite.

Semidefiniteness depends not only on leading principal minors D_1, D_2 but also on all principal minors, in this case on $D_1^{'} = c$ too.

4. If $D_1 \geq 0$, $D_1^{'} \geq 0$ and $D_2 \geq 0$, then the form is positive semidefinite.

   Note that the condition $D_1^{'} \geq 0$ is necessary since the form $f(x, y) = -y^2$ with $a = 0$, $b = 0$ and $c = -1$ for which $D_1 = a \geq 0$, $D_2 = ac - b^2 \geq 0$, nevertheless the form is not positive semidiefinite.

5. If $D_1 \leq 0$, $D_1^{'} \leq 0$ and $D_2 \geq 0$, then the form is negative semidefinite.

   Note that the condition $D_1^{'} \leq 0$ is necessary since the form $f(x, y) = y^2$ with $a = 0$, $b = 0$ and $c = 1$ for which $D_1 = a \leq 0$, $D_2 = ac - b^2 \geq 0$, nevertheless the form is not negative semidiefinite.

8.3.2 Definiteness of a 3 Variable Quadratic Form

Let us start with the following example.

Example 8.3.4. Let $f(x, y, z) = x^2 + 2y^2 - 7z^2 - 4xy + 8xz$. The symmetric matrix of this quadratic form is

$$\begin{bmatrix} 1 & -2 & 4 \\ -2 & 2 & 0 \\ 4 & 0 & -7 \end{bmatrix}.$$

The leading principal minors of this matrix are

$$\left|D_1\right| = 1, \quad \left|D_2\right| = \begin{vmatrix} 1 & -2 \\ -2 & 2 \end{vmatrix} = -2, \quad \left|D_3\right| = \begin{vmatrix} 1 & -2 & 4 \\ -2 & 2 & 0 \\ 4 & 0 & -7 \end{vmatrix} = -18.$$

Also, on simplification, we get

$$f(x, y, z) = x^2 + 2y^2 - 7z^2 - 4xy + 8xz = |D_1| l_1^2 + \frac{D_2}{D_1} l_2^2 + \frac{D_3}{D_3} l_3^2,$$

where

$$\begin{aligned} l_1 &= x - 2y + 4z, \\ l_2 &= y - 4x, \\ l_3 &= z \end{aligned}$$

That is, (l_1, l_2, l_3) are linear combinations of (x, y, z). More precisely,

$$\begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

where

$$P = \begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$$

is a non-singular matrix (changing variables).

In general if

$$f(x,y,z) = \begin{bmatrix} x & y & z \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

The following three determinants

$$\left|D_1\right| = \left|a_{11}\right|, \quad \left|D_1\right| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \left|D_3\right| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

are leading principal minors. It is possible to show that, if $|D_1| \neq 0$, $|D_2| \neq 0$, then

$$f(x,y,z) = |D_1| l_1^2 + \frac{|D_2|}{|D_1|} l_2^2 + \frac{|D_3|}{|D_2|} l_3^2,$$

where l_1, l_2, l_3 are some linear combinations of x, y, z. This is called Lagrange's Reduction. This implies the following

1. The form is positive definite iff $|D_1| > 0$, $|D_2| > 0$, $|D_3| > 0$, that is all principal minors are positive.
2. The form is negative definite iff $|D_1| < 0$, $|D_2| > 0$, $|D_3| < 0$, that is all principal minors alternate in sign starting with negative one.

Example 8.3.5. Determine the definiteness of the form $f(x,y,z) = 3x^2 + 2y^2 + 3z^2 - 2xy - 2yz$.

The matrix of our form is

$$\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}.$$

The leading principal minors are

$$\left|D_1\right| = 3 > 0, \quad \left|D_1\right| = \begin{vmatrix} 3 & -1 \\ -1 & 2 \end{vmatrix} = 5 > 0, \quad \left|D_3\right| = \begin{vmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{vmatrix} = 18 > 0,$$

thus the form is positive definite.

The above process can be generalized for n variable, which we omit here. We arrive at the following theorems.

Theorem 8.3.6. 1. A quadratic form is positive definite if and only if

$$|D_1| > 0, \quad |D_2| > 0, \quad \cdots \quad , |D_n| > 0,$$

that is all principal minors are positive;

2. A quadratic form is negative definite if and only if

$$|D_1| < 0, \quad |D_2| > 0, \quad |D_3| < 0, \quad , |D_4| > 0, \quad \cdots ,$$

that is principal minors alternate in sign starting with negative one.

3. If some kth order leading principal minor is nonzero but does not fit either of the above two sign patterns, then the form is indefinite.

Theorem 8.3.7. 1. A quadratic form is positive semidefinite if and only if all principal minors are ≥ 0;

2. A quadratic form is negative semidefinite if and only if all principal minors of odd degree are ≤ 0, and all principal minors of even degree are ≥ 0.

8.3.3 Definiteness and Eigen Values

As we know a symmetric $n \times n$ matrix has n real eigenvalues (maybe some multiple).

Theorem 8.3.8. Given a quadratic form $f(x) = x^T Ax$ and let $c_1, \ldots, c_n$ be eigen values of A. Then f is

1. positive definite iff $c_i > 0$, $i = 1, \ldots, n$;
2. negative definite iff $c_i < 0$, $i = 1, \ldots, n$;
3. positive semidefinite iff $c_i \geq 0$, $i = 1, \ldots, n$;
4. negative semidefinite iff $c_i \leq 0$, $i = 1, \ldots, n$;

8.4 Sylvester's Law of Inertia

This section is all about the possible diagonal entries for the diagonalisation of a real quadratic form. This is stated as the Sylvester's Law of Inertia in the following.

Theorem 8.4.1. Suppose V is a finite-dimensional real vector space and Q is a quadratic form on V. Then the numbers of positive diagonal entries, zero diagonal entries, and negative diagonal entries in any diagonalization of Q is independent of the diagonalization.

The idea is to decompose V as a direct sum of three spaces, one on which Q acts as a positive-definite quadratic form, one on which Q acts as the zero map and one on which Q acts as a negative-definite quadratic form.

Few Probable Questions

1. Define bilinear forms. Determine the definiteness of the form $f(x, y) = x^2 + 2xy + y^2$.
2. Define quadratic forms. For which real numbers k is the quadratic form $f(x, y) = kx^2 - 6xy + ky^2$ positive-definite?
3. Determine the definiteness of the matrix

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Unit 9

Course Structure

- Legendre Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Schlafli's integral formula. Laplace's first and second integral formula. Construction of Legendre differential equation.

9.1 Introduction

We are familiar with the method of solving ordinary differential equations via series solutions. In particular, we have learnt to find solutions of ODE around a regular point and a regular singular point for the given ODE. We used to employ Frobenius Method to calculate the solution in the latter case. Here, we will study the solutions of certain standard and "difficult" ODE which have applications in various fields using the same method. We will start with Legendre polynomials and explore certain properties of them.

Objectives

After reading this unit, you will be able to

- find the solution of Legendre equations
- define Legendre polynomials
- represent the solutions in a standard manner for further use
- learn the orthogonal properties and Rodrigue's formula for Legendre polynomials

9.2 Legendre's Equations and Legendre's Polynomials

The differential equation of the form

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0 \tag{9.2.1}$$

where n is a constant is called **Legendre's equation**.The points $x = \pm 1$ are the regular singular points of the equation. But the point $x = 0$ is an ordinary point of (9.2.1).

To check whether $x = \infty$ is a regular singular point of (9.2.1), set $x = \frac{1}{t}$. Then

$$\frac{dx}{dt} = -\frac{1}{t^2}$$

and hence

$$\frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = -t^2\frac{dy}{dt}.$$

Also,

$$\frac{d^2y}{dx^2} = \frac{d}{dt}\left(\frac{dy}{dt}.\frac{dt}{dx}\right)\frac{dt}{dx} = t^4\frac{d^y}{dt^2} + 2t^3\frac{dy}{dt}.$$

Hence equation (9.2.1) becomes

$$t^2(t^2-1)\frac{d^2y}{dt^2} + 2t^3\frac{dy}{dt} + n(n+1)y = 0. \tag{9.2.2}$$

$t = 0$ is clearly a regular singular point of (9.2.2), which implies that $x = \infty$ is a singular point of (9.2.1). Now, check that

$$\lim_{t\to 0}\frac{2t^4}{t^2(t^2-1)} = 0 \quad \& \quad \lim_{t\to 0} t^2\frac{n(n+1)}{t^2(t^2-1)} = -n(n+1).$$

Hence $t = 0$ is a regular singular point of (9.2.2).

Consider the series solution of (9.2.2) about $t = 0$, and let

$$y = t^s\sum_{m=0}^{\infty} a_m t^m$$

be a solution of (9.2.2) such that $a_0 \neq 0$. Then

$$\frac{dy}{dt} = \sum_{m=0}^{\infty}(m+s)a_m t^{s+m-1} \quad \text{and} \quad \frac{d^2y}{dt^2} = \sum_{m=0}^{\infty}(m+s)(m+s-1)a_m t^{s+m-2}.$$

Then (9.2.2) becomes

$$\sum_{m=0}^{\infty}\{(m+s-2)(m+s-1)a_{m-2}-(m+s+n)(m+s-n-1)a_m\}t^m-(s+n)(s-n-1)a_0-(s+n+1)(s-n)a_1 t = 0.$$

Then the indicial equation is

$$-(s+n)(s-n-1)a_0 = 0 \implies s = -n, n+1, \text{ since } a_0 \neq 0.$$

When $s = -n$, $a_1 = 0$ and when $s = n+1$, $a_1 = 0$. Hence $a_1 = 0$ in all case and the general recurrence relation is

$$a_m = \frac{(m+s-2)(m+s-1)}{(m+s-n)(m+s-n-1)}, \quad m \geq 2.$$

Since $a_1 = 0$, so $a_3 = a_5 = \cdots = a_{2m+1} = \cdots = 0$.

Now,

$$\begin{aligned} a_2 &= \frac{s(s+1)}{(s+n+2)(s-n+1)}a_0 \\ a_4 &= \frac{s(s+1)(s+2)(s+3)}{(s+n+2)(s+n+4)(s-n+1)(s-n+3)}a_0 \\ &\vdots \end{aligned}$$

Let n be a positive integer. Taking $m = n+1$, we have

$$a_{n+1} = \frac{(n+s)(n+s-1)}{(2n+s+1)s}a_{n-1}, \quad a_{n+2} = \frac{(n+s)(n+s+1)}{(2n+s+2)(s+1)}a_n.$$

When $s = -n$,

$$a_2 = -\frac{n(n-1)}{2(2n-1)}a_0, \quad a_4 = \frac{n(n-1)(n-2)(n-3)}{2.4.(2n-1)(2n-3)}a_0, \ldots$$

and $a_{n+1} = a_{n+2} = 0$. Then

$$y = a_0\left(x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2.4.(2n-1)(2n-3)}x^{n-4} + \cdots\right). \tag{9.2.3}$$

Taking $s = n+1$, we have

$$y = a_0\left(x^{-n-1} - \frac{(n+1)(n+2)}{2(2n+3)}x^{-n-3} + \frac{(n+1)(n+2)(n+3)(n+4)}{2.4.(2n+3)(2n+5)}x^{-n-5} + \cdots\right). \tag{9.2.4}$$

When n is a positive integer, the roots of the **indicial equation** differ by $2n+1$, which is an integer. There could be problem in evaluating a_{2n+1} for $s = -n$. But, $a_{n+1} = a_{n+2} = \cdots = 0$, and hence we don't face that problem.

When $n = 1$, $y_1 = a_0 x$.

When $n = 2$, $y_1 = a_0\left(x^2 - \frac{1}{3}\right)$.

When $n = 3$, $y_1 = \left(x^3 - \frac{3}{5}x\right)$.

If we take

$$a_0 = \frac{1.3.5\ldots(2n-1)}{n!},$$

then the solution of (9.2.2) is called the **Legendre function of first kind** or **Legendre Polynomial of degree** n and is denoted by $P_n(x)$. Thus, $P_n(x)$ is a solution of (9.2.1). But even if n is a positive integer, solution (9.2.3) is an infinite series. In this case if we take

$$a_0 = \frac{n!}{1.3.5\ldots(2n+1)},$$

then solution (9.2.3) is denoted by $Q_n(x)$ and is called the Legendre function of second kind. $Q_n(x)$ is not a polynomial and it is linearly independent from $P_n(x)$ and we get the general solution of (9.2.1) as

$$y = AP_n(x) + BQ_n(x).$$

Definition 9.2.1. Legendre Polynomial of degree n is defined as

$$P_n(x) = \frac{1.3.5\ldots(2n-1)}{n!}\left(x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2.4.(2n-1)(2n-3)}x^{n-4} + \cdots\right) \tag{9.2.5}$$

The general term of this polynomial is

$$(-1)^r \frac{n(n-1)(n-2)\ldots(n-2r+1)}{2.4\ldots 2r(2n-1)(2n-3)\ldots(2n-2r+1)} \frac{1.3.5\ldots(2n-1)}{n!} x^{n-2r} \tag{9.2.6}$$

Now,

$$1.3.5\ldots(2n-1) = \frac{1.2.3.\ldots(2n)}{2.4.\ldots(2n)} = \frac{(2n)!}{2^n.n!}.$$

Also,

$$n(n-1)(n-2)\ldots(n-2r+1) = \frac{n!}{(n-2r)!}.$$

$$2.4.\ldots(2r) = 2^r.r!.$$

And

$$(2n-1)(2n-3)\ldots(2n-2r+1) = \frac{(2n)!(n-r)!}{2^r.n!(2n-2r)!}.$$

So, using these things, (9.2.6) becomes

$$(-1)^r \frac{(2n-2r)!}{2^n r!(n-2r)!(n-r)!} x^{n-2r}.$$

(9.2.5) is a polynomial of degree n. Hence $n-2r \geq 0$ or 1 according as n is even or odd, that is, $r \leq \left[\frac{n}{2}\right]$. Hence, Legendre polynomial of degree n is given by

$$P_n(x) = \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \frac{(2n-2r)!}{2^n r!(n-2r)!(n-r)!} x^{n-2r}.$$

Determination of few Legendre Polynomials

For $n = 0$, we have

$$P_0(x) = (-1)^0 \frac{(2.0-2.0)!}{2^0 0!0!0!} = 1.$$

Similarly, putting $n = 1, 2, 3, 4$ we get

$$\begin{aligned}
P_1(x) &= x \\
P_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} \\
P_3(x) &= \frac{5}{3}x^3 - \frac{3}{2}x \\
P_4(x) &= \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8} \\
P_5(x) &= \frac{63}{8}x^5 - \frac{35}{4}x^3 + \frac{15}{8}x \\
P_6(x) &= \frac{231}{16}x^6 - \frac{315}{16}x^4 + \frac{105}{16}x^2 - \frac{5}{16}.
\end{aligned}$$

Generating Function for Legendre Polynomial

We shall now show that $P_n(x)$ is the coefficient of z^n in the expansion of $(1-2xz+z^2)^{-1/2}$ in ascending powers of z.The function

$$w(x,z) = (1-2xz+z^2)^{-1/2}$$

is the **generating function for Legendre polynomials**

Theorem 9.2.2. Show that

$$w(x,z) = \sum_{n=0}^{\infty} P_n(x).z^n,$$

holds for sufficiently small values of $|z|$.

Proof. We know that,

$$P_n(x) = \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \frac{(2n-2r)!}{2^n r!(n-2r)!(n-r)!} x^{n-2r}.$$

When $\left|2xz - z^2\right| < 1$, the expression $(1-2xz+z^2)^{-1/2}$can be expanded in a series of ascending powers of $\left(2xz - z^2\right)$. Moreover if $|2xz| + \left|z^2\right| < 1$ the powers can be multiplied out and the resulting series can be rearranged in any manner. Thus using the expansion

$$\begin{aligned}
(1-a)^{-1/2} &= 1 + \frac{a}{2} + \frac{(-1/2)(-1/2-1)}{2!}a^2 + \frac{(-1/2)(-1/2-1)(-1/2-2)}{3!}a^3 + \cdots \\
&= 1 + \frac{a}{2} + \frac{1.3}{2^2.2!}a^2 + \frac{1.3.5}{2^3.3!}a^3 + \cdots \\
&= \sum_{k=0}^{\infty} \frac{1.3.5\ldots(2k-1)}{k!.2^k} a^k.
\end{aligned}$$

we get,

$$\begin{aligned}
(1-2xz+z^2)^{-1/2} &= \sum_{k=0}^{\infty} \frac{1.3.5\ldots(2k-1)}{k!.2^k}(2xz-z^2)^k \\
&= \sum_{k=0}^{\infty} \frac{1.2.3\ldots(2k-1).\,(2k)}{2.4.6...\,(2k)\,.\,(k!.2^k)}(2xz-z^2)^k \\
&= \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2}(2xz-z^2)^k \\
&= \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2} \sum_{s=0}^{k} \binom{k}{s} (2xz)^k(-z^2)^{k-s} \\
&= \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2} \sum_{s=0}^{k} \binom{k}{s} (2x)^k(-1)^{k-s} z^{2k-s} \\
&= \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} (-1)^{k-s} \frac{(2k)!}{2^{2k}(k!)^2} \frac{k!}{s!(k-s)!}(2x)^s z^{2k-s}.
\end{aligned}$$

Consider the portion $(k-s)!$, where s varies from 0 to k. If $s = k+1$, $(k-s)! = (-1)! = \infty$. Similarly, for other $s > k$, $(k-s)! \to \infty$ and so, the terms for $s > k$ becomes zero and the summation can be extended from k to ∞. Interchanging the summations, we get

$$w(x,z)=\sum_{s=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{k-s}\frac{(2k)!}{2^{2k}(k!)^2}\frac{k!}{s!(k-s)!}(2x)^s z^{2k-s}.$$

When $k=0,1,\ldots,(s-1)$, we get $(k-s)!=\infty$. And when $k=s$, $(k-2)!=0!=1$. So, we can effectively start the summation from $k=s$ instead of $k=0$ and the equation becomes

$$w(x,z)=\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}(-1)^{k-s}\frac{(2k)!}{2^{2k}(k!)^2}\frac{k!}{s!(k-s)!}(2x)^s z^{2k-s}.$$

Putting $k-s=p$, and eliminating k, we get

$$w(x,z)=\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}(-1)^p\frac{(2p+2s)!}{2^{2p+s}(s+p)!}\frac{1}{s!p!}x^s z^{2p+s}.$$

Put $2p+s=n$ and eliminate s. Then since p varies from 0 to ∞, s varies from 0 to ∞, n varies from 0 to ∞. Now, $s\geq 0$. So, $n-2p\geq 0$ which implies that $p\leq\left[\frac{n}{2}\right]$. Since p is an integer, $p\leq\left[\frac{n}{2}\right]$. So,

$$\begin{aligned}
w(x,z) &= \sum_{n=0}^{\infty}\sum_{p=0}^{\left[\frac{n}{2}\right]}(-1)^p\frac{(2n-2p)!}{2^n(n-p)!}\frac{1}{(n-2p)!p!}x^{n-2p}z^n\\
&= \sum_{n=0}^{\infty}z^n\sum_{p=0}^{\left[\frac{n}{2}\right]}(-1)^p\frac{(2n-2p)!}{2^n(n-p)!}\frac{1}{(n-2p)!p!}x^{n-2p}\\
&= \sum_{n=0}^{\infty}P_n(x)z^n.
\end{aligned}$$

□

Show that $i)$ $P_n(1)=1$, $ii)$ $P_n(-1)=(-1)^n$, $iii)$ $P_{2n+1}(0)=0$, $iv)$ $P_{2n}(0)=(-1)^n\,\frac{1.3.5...(2n-1)}{n!.2^n}$.

Solution. We use the relation

$$\sum_{n=0}^{\infty}P_n(x)z^n=(1-2xz+z^2)^{-1/2}$$

to prove the results.

$i)$ Put $x=1$ in the above relation. Then we get

$$\begin{aligned}
\sum_{n=0}^{\infty}P_n(1)z^n &= (1-2z+z^2)^{-1/2}\\
&= (1-z)^{-1}\\
&= 1+z+z^2+...+z^n+...\\
&= \sum_{n=0}^{\infty}z^n.
\end{aligned}$$

Therefore

$$P_n(1)=1.$$

$ii)$ Put $x = -1$ in the above relation. Then we get

$$\begin{aligned}\sum_{n=0}^{\infty} P_n(-1)z^n &= (1+2z+z^2)^{-1/2} \\ &= (1+z)^{-1} \\ &= 1 - z + z^2 - ... + (-1)^n z^n + ... \\ &= \sum_{n=0}^{\infty} (-1)^n z^n.\end{aligned}$$

Therefore

$$P_n(-1) = (-1)^n .$$

To prove $iii)$ and $iv)$ put $x = 0$ in the above relation. Then we get

$$\begin{aligned}\sum_{n=0}^{\infty} P_n(0)z^n &= (1+z^2)^{-1/2} \\ &= 1 - \frac{1}{2}z^2 + \frac{1.3}{2^2.2!}z^4 - \frac{1.3.5}{2^3.3!}z^6 + ... + (-1)^n \frac{1.3.5....(2n-1)}{2^n.n!}z^{2n} + ...\end{aligned}$$

Therefore the coefficients of the odd powers of z are 0 i.e.,

$$P_{2n+1}(0) = 0$$

and those for the even powers of z we get

$$P_{2n}(0) = (-1)^n \frac{1.3.5\ldots(2n-1)}{n!.2^n}.$$

■

Show that $P_n(x)$ is an even or odd function of x according as n is even or odd respectively.

Solution. We have

$$\sum_{n=0}^{\infty} P_n(x)z^n = (1-2xz+z^2)^{-1/2}.$$

Now

$$\begin{aligned}\sum_{n=0}^{\infty} P_n(-x)z^n &= (1+2xz+z^2)^{-1/2} \\ &= \left\{1-2x(-z)+(-z)^2\right\}^{-1/2} \\ &= \sum_{n=0}^{\infty} P_n(x)(-z)^n \\ &= \sum_{n=0}^{\infty} (-1)^n P_n(x)z^n.\end{aligned}$$

Therefore

$$P_n(-x) = (-1)^n P_n(x).$$

and hence the result follows. ■

9.3 Rodrigue's Formula for Legendre's Polynomials

For integral values of n, Legendre Polynomials $P_n(x)$ satisfy

$$\frac{1}{2^n n!}\frac{d^n y}{dx^n}(x^2-1)^n = P_n(x),$$

called the **Rodrigue's formula**.

$$\begin{aligned}
\frac{d^n}{dx^n}(x^2-1)^n &= \frac{d^n}{dx^n}\left[\sum_{r=0}^{n}\binom{n}{r}(-1)^r(x^2)^{n-r}\right] \\
&= \sum_{r=0}^{n}(-1)^r\binom{n}{r}\frac{d^n}{dx^n}x^{2n-2r} \\
&= \sum_{r=0}^{m}(-1)^r\binom{n}{r}\frac{(2n-2r)!}{(n-2r)!}x^{n-2r},
\end{aligned}$$

where $m=\frac{n}{2}$ or $\frac{n-1}{2}$ according as n is even or odd.
For $r>m$, the powers of x in x^{2n-2r} becomes less than n, and then $\frac{d^n}{dx^n}x^{2n-2r}$ vanishes.
For $r\le m$, $\frac{d^n}{dx^n}x^{2n-2r}$ does not vanish.
Therefore

$$\begin{aligned}
\frac{d^n}{dx^n}(x^2-1)^n &= \sum_{r=0}^{m}(-1)^r\frac{n!(2n-2r)!}{r!\,(n-r)!(n-2r)!}x^{n-2r} \\
\frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n &= \frac{1}{2^n n!}\sum_{r=0}^{m}(-1)^r\frac{n!(2n-2r)!}{r!\,(n-r)!(n-2r)!}x^{n-2r} \\
&= \sum_{r=0}^{[\frac{n}{2}]}(-1)^r\frac{(2n-2r)!}{2^n.r!\,(n-r)!(n-2r)!}x^{n-2r} \\
&= P_n(x).
\end{aligned}$$

9.4 Recurrence Relations of Legendre Polynomials

The Legendre polynomials $P_n(x)$ satisfies the following recurrence relations.
1) $(n+1)P_{n+1}(x)-(2n+1)xP_n(x)+nP_{n-1}(x)=0$.
Using generating function of $P_n(x)$ we get

$$w(x,z)=(1-2xz+z^2)^{-1/2}=\sum_{n=0}^{\infty}z^nP_n(x). \tag{9.4.1}$$

Differentiating equation (9.4.1) with respect to z, we get

$$\frac{\partial w}{\partial z}=-\frac{1}{2}(1-2xz+z^2)^{-3/2}(-2x+2z)^{-1/2}=\sum_{n=0}^{\infty}nz^{n-1}P_n(x)$$

i.e.,

$$\begin{aligned}(x-z)\,(1-2xz+z^2)^{-3/2} &= \sum_{n=0}^{\infty} nz^{n-1}P_n(x)\\ (x-z)\,(1-2xz+z^2)^{-1/2} &= \sum_{n=0}^{\infty} nz^{n-1}(1-2xz+z^2)P_n(x)\\ (x-z)\sum_{n=0}^{\infty} z^nP_n(x) &= (1-2xz+z^2)\sum_{n=0}^{\infty} nz^{n-1}P_n(x).\end{aligned}$$

Equating the coefficient of z^n from both sides of the above identity we get

$$(n+1)P_{n+1}(x)-(2n+1)xP_n(x)+nP_{n-1}(x)=0. \tag{9.4.2}$$

2) $nP_n(x)=xP_n'(x)-P_{n-1}'(x)$.
Differentiating equation (9.4.1) with respect to x, we get

$$\frac{\partial w}{\partial x}=-\frac{1}{2}(1-2xz+z^2)^{-3/2}(-2z)=\sum_{n=0}^{\infty} z^nP_n'(x)$$

i.e.,

$$\begin{aligned}z(1-2xz+z^2)^{-3/2} &= \sum_{n=0}^{\infty} z^nP_n'(x)\\ z\,(x-z)\,(1-2xz+z^2)^{-3/2} &= (x-z)\sum_{n=0}^{\infty} z^nP_n'(x)\end{aligned}$$

Equating the coefficients of z^n on both sides, we get

$$nP_n(x)=xP_n'(x)-P_{n-1}'(x). \tag{9.4.3}$$

3) $P_{n+1}'(x)=(2n+1)P_n(x)+P_{n-1}'(x)$.
Differentiating (9.4.2) with respect to x we get

$$(n+1)P_{n+1}'(x)-(2n+1)P_n(x)-(2n+1)xP_n'(x)+nP_{n-1}'(x)=0.$$

Using relation (9.4.3),

$$\begin{aligned}(n+1)P_{n+1}'(x)-(2n+1)P_n(x)-(2n+1)\left\{nP_n(x)+P_{n-1}'(x)\right\}+nP_{n-1}'(x) &= 0\\ (n+1)P_{n+1}'(x)-(2n+1)(n+1)P_n(x)-(n+1)P_{n-1}'(x) &= 0.\end{aligned}$$

Cancelling the factor $(n+1)$, we get

$$P_{n+1}'(x)=(2n+1)P_n(x)+P_{n-1}'(x). \tag{9.4.4}$$

4) $(n+1)\,P_n(x)=P_{n+1}'(x)-xP_n'(x)$
Adding relation (9.4.3) and (9.4.4) we get

$$(n+1)\,P_n(x)=P_{n+1}'(x)-xP_n'(x). \tag{9.4.5}$$

5) $(1-x^2)P_n^{'}(x) = n\left\{P_{n-1}(x) - xP_n(x)\right\}$
Replacing n by $n-1$ in (9.4.5) we get

$$nP_{n-1}(x) = P_n^{'}(x) - xP_{n-1}^{'}(x).$$

Using relation (9.4.3) we get

$$\begin{aligned} nP_{n-1}(x) &= P_n^{'}(x) - x\left\{xP_n^{'}(x) - nP_n(x)\right\} \\ &= \left(1-x^2\right)P_n^{'}(x) + xnP_n(x). \end{aligned}$$

Rearranging we get

$$(1-x^2)P_n^{'}(x) = n\left\{P_{n-1}(x) - xP_n(x)\right\} \tag{9.4.6}$$

6) $(1-x^2)P_n^{'}(x) = (n+1)\left\{xP_n(x) - P_{n+1}(x)\right\}$
Replacing n by $n+1$ in (9.4.3) we get

$$(n+1)\,P_{n+1}(x) = xP_{n+1}^{'}(x) - P_n^{'}(x).$$

Using relation (9.4.5) we get

$$\begin{aligned} (n+1)\,P_{n+1}(x) &= x\left\{(n+1)\,P_n(x) + xP_n^{'}(x)\right\} - P_n^{'}(x) \\ &= (n+1)\,xP_n(x) + x^2P_n^{'}(x) - P_n^{'}(x) \\ &= (n+1)\,xP_n(x) - (1-x^2)P_n^{'}(x). \end{aligned}$$

Rearranging we get

$$(1-x^2)P_n^{'}(x) = (n+1)\left\{xP_n(x) - P_{n+1}(x)\right\} \tag{9.4.7}$$

7) $(2n+1)\left(x^2-1\right)P_n^{'}(x) = n\,(n+1)\left\{P_{n+1}(x) - P_{n-1}(x)\right\}$
Multiplying (9.4.6) by $n+1$ and (9.4.7) by n and adding we get

$$\begin{aligned} (n+1)\,(1-x^2)P_n^{'}(x) + n(1-x^2)P_n^{'}(x) &= n\,(n+1)\{P_{n-1}(x) - xP_n(x)\} + n\,(n+1)\{xP_n(x) - P_{n+1}(x)\} \\ (2n+1)\left(1-x^2\right)P_n^{'}(x) &= n\,(n+1)\left\{P_{n-1}(x) - P_{n+1}(x)\right\}. \end{aligned}$$

Therefore

$$(2n+1)\left(x^2-1\right)P_n^{'}(x) = n\,(n+1)\left\{P_{n+1}(x) - P_{n-1}(x)\right\}. \tag{9.4.8}$$

The relation (9.4.8) is called **Beltrami's relation**.
8) $P_{n+1}^{'}(x) - 2xP_n(x) + P_{n-1}(x) - P_n(x) = 0.$
Taking logarithm on both sides of (9.4.1) and then differentiating with respect to x, we get

$$\begin{aligned} \frac{d}{dx}[\ln((1-2xz+z^2)^{-1/2})] &= \frac{d}{dx}\left[\ln\left(\sum_{n=0}^{\infty} z^nP_n(x)\right)\right] \\ \text{or,}\quad \frac{1}{2}\frac{2z}{1-2xz+z^2} &= \frac{\sum_{n=0}^{\infty} z^nP_n^{'}(x)}{\sum_{n=0}^{\infty} z^nP_n(x)} \\ \text{or,}\quad (1-2xz+z^2)\sum_{n=0}^{\infty} z^nP_n^{'}(x) &= z\sum_{n=0}^{\infty} z^nP_n(x) \end{aligned}$$

Equating the coefficients of z^n on both sides, we get

$$P_n^{'}(x) - 2xP_{n-1}^{'}(x) + P_{n-2}^{'}(x) = P_{n-1}(x).$$

Replacing n by $n+1$, we get,

$$P'_{n+1}(x) - 2xP_n(x) + P_{n-1}(x) - P_n(x) = 0. \tag{9.4.9}$$

8) $nP_n(x) = xP'_n(x) - P'_{n-1}(x)$.
Differentiating (9.4.8) with respect to z, we get,

$$\frac{x-z}{(1-2xz+z^2)^{3/2}} = \sum_{n=1}^{\infty} nP_n(x)z^{n-1} \tag{9.4.10}$$

Again, differentiating (9.4.8) with respect to x, we get,

$$\frac{z}{(1-2xz+z^2)^{3/2}} = \sum_{n=0}^{\infty} P'_n(x)z^n \tag{9.4.11}$$

Multilying (9.4.10) by z we get

$$\frac{z\,(x-z)}{(1-2xz+z^2)^{3/2}} = \sum_{n=1}^{\infty} nP_n(x)z^n. \tag{9.4.12}$$

Multilying (9.4.11) by $(x-z)$ we get

$$\frac{z\,(x-z)}{(1-2xz+z^2)^{3/2}} = \sum_{n=1}^{\infty} (x-z)\,P'_n(x)z^n. \tag{9.4.13}$$

Subtracting (9.4.13) from (9.4.12) we get

$$(x-z)\sum_{n=0}^{\infty} P'_n(x)z^n = \sum_{n=1}^{\infty} nP_n(x)z^n. \tag{9.4.14}$$

Equating the coefficients of z^n from both sides of (9.4.14) we get

$$nP_n(x) = xP'_n(x) - P'_{n-1}(x). \tag{9.4.15}$$

9.5 Orthogonality Properties of Legendre's Polynomials

Definition 9.5.1. A sequence of functions $\{\phi_n(x)\}$ is said to be orthogonal in the interval $[a, b]$ with the weight function $\omega(x)$ if
i) $\int_a^b \omega(x)\,\phi_m(x)\phi_n(x)dx = 0$ for $m \neq n$
ii) $\int_a^b \omega(x)\,\phi_m(x)\phi_n(x)dx \neq 0$ for $m = n$ and m, n are positive integers.

The orthogonal functions have the remarkable property that any integrable function over a given interval can be expanded in a series of orthogonal functions over that intervals.

We will show that the Legendre polynomials are orthogonal in the interval $[-1, 1]$ with the weight function $\omega(x) = 1$, i.e.,
i) $\int_{-1}^1 P_m(x)P_n(x)dx = 0$ for $m \neq n$
ii) $\int_{-1}^1 P_m(x)P_n(x)dx = \frac{2}{2m+1}$ for $m = n$ and m, n are positive integers.
First consider $m \neq n$.

Legendre equation of order m is

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

$P_m(x)$ is a solution of the above equation.

So,

$$(1-x^2)P_m''(x) - 2xP_m'(x) + m(m+1)P_m(x) = 0. \tag{9.5.1}$$

Also, $P_n(x)$ is a solution of Legendre equation of order n.

So,

$$(1-x^2)P_n''(x) - 2xP_n'(x) + m(m+1)P_n(x) = 0. \tag{9.5.2}$$

Multiplying (9.5.1) by $P_n(x)$ and (9.5.2) by $P_m(x)$ and subtracting, we get $(1-x^2)[P_m''(x)P_n(x) - P_n''(x)P_m(x)] - 2x[P_m'(x)P_n(x) - P_n'(x)P_m(x)] + [m(m+1) - n(n+1)]P_m(x)P_n(x) = 0$

i.e.,

$$\frac{dy}{dx}[(1-x^2)\{P_m'(x)P_n(x) - P_n'(x)P_m(x)\}] + [m(m+1) - n(n+1)]P_m(x)P_n(x) = 0$$

Integrating both sides with respect to x from -1 to 1, we get

$$\begin{aligned}
\int_{-1}^{1} \frac{d}{dx}[(1-x^2)\{P_m'(x)P_n(x) - P_n'(x)P_m(x)\}]dx &= (n-m)(n+m+1)\int_{-1}^{1} P_m(x)P_n(x)dx \\
\text{or, } (n-m)(n+m+1)\int_{-1}^{1} P_m(x)P_n(x)dx &= [(1-x^2)\{P_m'(x)P_n(x) - P_n'(x)P_m(x)\}]_{-1}^{1} \\
&= 0 \\
\text{or, } \int_{-1}^{1} P_m(x)P_n(x)dx &= 0.
\end{aligned}$$

Now consider $m = n$ and m, n are positive integers.

We have from the generating function that

$$(1-2xz+z^2)^{-1/2} = \sum_{n=0}^{\infty} z^n P_n(x). \tag{9.5.3}$$

Replacing n by m in (9.5.3), we have

$$(1-2xz+z^2)^{-1/2} = \sum_{m=0}^{\infty} z^m P_m(x). \tag{9.5.4}$$

Multiplying (9.5.3) and (9.5.4), we get

$$(1-2xz+z^2)^{-1} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} z^{n+m} P_n(x)P_m(x).$$

Integrating both sides with respect to x from -1 to 1, we get

$$\begin{aligned}
\int_{-1}^{1}(1-2xz+z^2)^{-1}dx &= \int_{-1}^{1}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} z^{n+m}P_n(x)P_m(x)dx \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\int_{-1}^{1} P_n(x)P_m(x)z^{n+m}dx \\
&= \sum_{n=0}^{\infty}\int_{-1}^{1} P_n(x)P_n(x)z^{2n}dx
\end{aligned}$$

Now,

$$\int_{-1}^{1}(1-2xz+z^2)^{-1}dx=\frac{1}{z}\ln\left(\frac{1+z}{1-z}\right).$$

Hence,

$$\sum_{n=0}^{\infty}\int_{-1}^{1}P_n(x)P_n(x)z^{2n}dx=\frac{2}{z}\left\{z+\frac{z^3}{3}+\frac{z^5}{5}+\cdots\right\}=\sum_{n=0}^{\infty}\frac{2z^{2n}}{2n+1}.$$

Hence, equating the like coefficients from both sides, we get

$$\int_{-1}^{1}P_n(x)P_n(x)z^{2n}dx=\frac{2}{2n+1}.$$

9.5.1 Expansion of a given function in a series of Legendre polynomials

Let $f(x)$ be any function piecewise continuous over the interval $[-1,1]$. We express $f(x)$ in a series of Legendre polynomials as

$$f(x)=\sum_{n=0}^{\infty}a_nP_n(x) \tag{9.5.5}$$

Multiplying (9.5.5) by $P_m(x)$ and integrating from -1 to $+1$ we get

$$\begin{aligned}\int_{-1}^{1}f(x)P_m(x)dx &= \sum_{n=0}^{\infty}a_n\int_{-1}^{1}P_m(x)P_n(x)dx\\ &= a_m\int_{-1}^{1}P_m^2(x)dx\\ &= \frac{2a_m}{2m+1}.\end{aligned}$$

Thus

$$a_m=\frac{2m+1}{2}\int_{-1}^{1}f(x)P_m(x)dx. \tag{9.5.6}$$

From (9.5.6) we can determine the coefficients a_m for $m=0,1,2,...$ in the equation (9.5.5).

In particular let $f(x)=b_0+b_1x+...+b_kx^k$ be a polynomial of degree k, then using the fact $\int_{-1}^{1}x^rP_m(x)dx=0$ if $r<m$ we get

$$\begin{aligned}\int_{-1}^{1}f(x)P_m(x)dx &= \int_{-1}^{1}\left(b_0+b_1x+...+b_kx^k\right)P_m(x)dx\\ &= 0,\ k<m.\end{aligned}$$

Thus

$$a_{k+1}=a_{k+2}=...=0.$$

Hence from (9.5.5) we can express $f(x)$ as a terminating series of Legendre polynomials

$$f(x)=\sum_{n=0}^{k}a_nP_n(x).$$

Therefore any polynomial $f(x)$ of degree k can be expanded by a series of Legendre polynomials $P_0(x)$, $P_1(x)$, ..., $P_k(x)$.

Exercise 9.5.2.

1. Show that $(2n+1)\,P_n^2 - (2n-1)\,P_{n-1}^2 = \frac{d}{dx}\left\{xP_n^2 + xP_{n-1}^2 - 2P_nP_{n-1}\right\}$
2. Show that $P_n(x)$ is a solution of Legendre equation of order n.
3. Show that $\int_{-1}^{1} P_n(x)dx = 2$, if $n = 0$ and $\int_{-1}^{1} P_n(x)dx = 0$, if $n \geq 1$.
4. Show that $\int_{-1}^{1} x^m P_n(x)dx = 0$ if $m < n$.
5. Evaluate $\int_{-1}^{1} x^m P_n(x)dx$ if $m \geq n$.
6. Expand $x^3 - 7x^2 + 3x + 2$ in a series of Legendre polynomials in $[-1, 1]$.
7. Expand $f(x)$ in a series of Legendre polynomials in $[-1, 1]$, where $f(x) = 0$ for $-1 \leq x \leq 0$ and $f(x) = x$ for $0 \leq x \leq 1$.
8. Show that all the roots of $P_n(x) = 0$ are real and distinct and lie between -1 to $+1$.

Few Probable Questions

1. Prove that the Legendre polynomials are orthogonal.
2. State and prove the Rodrigue's formula.
3. Prove that for any non-negative integer n, we have $P'_{n+1}(x) - 2xP_n(x) + P'_{n-1}(x) - P_n(x) = 0$.
4. Prove the following:
   (a) $P_n(1) = 1$, $P_n(-1) = (-1)^n$.
   (b) $P'_n(1) = \frac{n(n+1)}{2}$, and $P'_n(-1) = (-1)^{n-1}\frac{n(n+1)}{2}$
   (c) $P_n(-x) = (-1)^n P_n(x)$. Hence deduce that $P_n(-1) = (-1)^n$.

Unit 10

Course Structure

- Bessel's function: Generating function, Recurrence relation, Representation for the indices $1/2,\ -1/2,\ 3/2$ and $-3/2$. Bessel's integral equation. Bessel's function of second kind.

Objectives

After reading this section, you will be able to

- solve Bessel's equations by Frobenius method
- deduce the generating functions for integral index
- deduce the recurrence relations for Bessel's functions
- deduce the orthogonality condition for Bessel's functions
- solve related problems

10.1 Bessel's Equation and Bessel's Functions

The differential equation

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0 \tag{10.1.1}$$

where n is a constant, is called a Bessels's equation whose solution gives the Bessel's functions. These are an orthogonal sequence of functions that have many closely related definitions. This unit is dedicate to the study of Bessel's functions and its properties.

To solve the differential equation (10.1.1), we see that $x = 0$ is a regular singular point of (10.1.1) (Verify!). We will thus attempt to solve it about $x = 0$ by Frobenius method. The resulting series solution is valid in a neighborhood of $x = 0$. Let us assume the solution to be

$$y = \sum_{m=0}^{\infty} c_m x^{m+s}, \text{ where } c_0 \neq 0$$

and s is to be determined. Thus

$$\frac{dy}{dx}=\sum_{m=0}^{\infty}(m+s)c_m x^{m+s-1}, \quad \& \quad \frac{d^2y}{dx^2}=\sum_{m=0}^{\infty}(m+s)(m+s-1)c_m x^{m+s-2}$$

Thus, (10.1.1) becomes

$$\sum_{m=0}^{\infty}(m+s)(m+s-1)c_m x^{m+s}+\sum_{m=0}^{\infty}(m+s)c_m x^{m+s}+\sum_{m=0}^{\infty}c_m x^{m+s+2}-n^2\sum_{m=0}^{\infty}c_m x^{m+s}=0$$

$$\text{or, } \sum_{m=0}^{\infty}\left\{(m+s)(m+s-1)+(m+s)-n^2\right\}c_m x^{m+s}+\sum_{m=0}^{\infty}c_m x^{m+s+2}=0$$

$$\text{or, } \sum_{m=0}^{\infty}(m+s+n)(m+s-n)c_m x^m+\sum_{m=2}^{\infty}c_{m-2}x^m=0$$

$$\text{or, } \sum_{m=2}^{\infty}\left\{(m+s+n)(m+s-n)c_m+c_{m-2}\right\}x^m+(s+n)(s-n)c_0+(1+s+n)(1+s-n)c_1x=0.$$

The indicial equation is

$$(s+n)(s-n)c_0=0 \implies s=-n,n, \quad \text{since } c_0\neq 0$$

and the general recurrence relation is

$$c_m=-\frac{1}{(m+s+n)(m+s-n)}c_{m-2}, \quad m\geq 2.$$

When $s=n$, we have $(1+n-n)(1+n+n)c_1=0$ which implies $c_1=0$ if $n\neq -1/2$.
When $s=-n$, we have $(1-n-n)(1-n+n)c_1=0$ which implies $c_1=0$ if $n\neq 1/2$.

Case I: When $n\neq 1/2$, then $c_1=0$. Then

$$\begin{aligned}
c_2 &= -\frac{1}{(2+s+n)(2+s-n)}c_0\\
c_4 &= \frac{1}{(2+s+n)(4+s+n)(2+s-n)(4+s-n)}c_0\\
c_6 &= -\frac{1}{(2+s+n)(4+s+n)(6+s+n)(2+s-n)(4+s-n)(6+s-n)}c_0\\
&\vdots\\
c_{2m} &= (-1)^m\frac{1}{(2+s+n)(4+s+n)\ldots(2m+s+n)(2+s-n)(4+s-n)\ldots(2m+s-n)}c_0
\end{aligned}$$

and $c_1=c_3=c_5=\cdots=c_{2m+1}=\cdots=0$. Thus, we get the solution as

$$y=c_0x^s\left[1-\frac{1}{(2+s+n)(2+s-n)}x^2+\cdots\right]$$

Putting $s=n$, we get

$$y=y_1=c_0x^n\left[1-\frac{x^2}{4(n+1)}+\frac{x^4}{4.8.(n+1)(n+2)}-\cdots\right]. \tag{10.1.2}$$

Putting $s = -n$, we get

$$y = y_2 = c_0^{'} x^{-n}\left[1 - \frac{x^2}{4(1-n)} + \frac{x^4}{4.8.(1-n)(2-n)} - \cdots\right]. \tag{10.1.3}$$

The particular solution (10.1.2) of (10.1.1), taking

$$c_0 = \frac{1}{2^n\Gamma(n+1)}$$

is called the Bessel's function of first kind of order n and denoted by $J_n(x)$, that is,

$$J_n(x) = \frac{x^n}{2^n\Gamma(n+1)}\left[1 - \frac{x^2}{4(1-n)} + \frac{x^4}{4.8.(1-n)(2-n)} - \cdots\right].$$

The general term of $J_n(x)$ is, on simplification,

$$(-1)^r.\frac{1}{r!(n+1)\ldots(n+r).\Gamma(n+1)}\frac{x^{n+2r}}{2^{n+2r}}.$$

Now,

$$\Gamma(n+1) = n\Gamma(n).$$

Thus,

$$\Gamma(n+r+1) = (n+r)(n+r-1)\ldots(n+1)\Gamma(n+1).$$

Hence, the general term of the summation in $J_n(x)$ is

$$(-1)^r.\frac{x^{2r+n}}{2^{2r+n}\Gamma(r+1).\Gamma(n+r+1)}.$$

Thus, we can now formally define the Bessel's function as

Definition 10.1.1. The Bessel's function of first kind, of order n is defined as

$$J_n(x) = \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(n+r+1)}\left(\frac{x}{2}\right)^{2r+n}.$$

Similarly, the other solution is obtained by putting $-n$ for n, that is, the other solution, when n is not an integer is given by

$$J_{-n}(x) = \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(r-n+1)}\left(\frac{x}{2}\right)^{2r-n}.$$

Case II: When $n = \pm 1/2$, then the Bessel's equation becomes

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + \left(x^2 - \frac{1}{4}\right)y = 0.$$

Let

$$y = \sum_{m=0}^{\infty} c_m x^{m+s}$$

be a solution of the above equation, where $c_0 \neq 0$ and s is to be determined. Then the above equation becomes

$$\sum_{m=0}^{\infty}(m+s)(m+s-1)c_m x^{m+s} + \sum_{m=0}^{\infty}(m+s)c_m x^{m+s} + \sum_{m=0}^{\infty} c_m x^{m+s+2} - \frac{1}{4}\sum_{m=0}^{\infty} c_m x^{m+s} = 0$$

$$\text{or,}\quad \sum_{m=0}^{\infty}\left(m+s+\frac{1}{2}\right)\left(m+s-\frac{1}{2}\right)c_m x^m + \sum_{m=0}^{\infty} c_m x^{m+2} = 0$$

$$\text{or,}\quad \sum_{m=0}^{\infty}\left\{\left(m+s+\frac{1}{2}\right)\left(m+s-\frac{1}{2}\right)c_m + c_{m-2}\right\}x^m + \left(s+\frac{1}{2}\right)\left(s-\frac{1}{2}\right)c_0$$

$$+\left(s+\frac{3}{2}\right)\left(s+\frac{1}{2}\right)c_1 x = 0$$

Thus, the indicial equation is

$$\left(s+\frac{1}{2}\right)\left(s-\frac{1}{2}\right)c_0 = 0 \implies s = \pm\frac{1}{2}, \text{ since } c_0 \neq 0.$$

When $s = 1/2$, $c_1 = 0$ and when $s = -1/2$, c_1 is indeterminate. Take c_1 as constant. Now, the general recurrence relation is

$$c_m = -\frac{1}{(m+s+1/2)(m+s-1/2)}c_{m-2}, \;\; m \geq 2.$$

When $s = -1/2$, we have

$$c_m = -\frac{1}{m(m-1)}c_{m-2}, \;\; m \geq 2.$$

Thus

$$\begin{aligned} c_2 &= -\frac{1}{2!}c_0, \quad c_4 = \frac{1}{4!}c_0, \quad c_6 = -\frac{1}{6!}c_0, \ldots, \\ c_3 &= -\frac{1}{3!}c_1, \quad c_5 = \frac{1}{5!}c_1, \quad c_7 = -\frac{1}{7!}c_1, \ldots. \end{aligned}$$

Thus, the solution becomes

$$y = c_0\left[x^{-1/2} - \frac{x^{3/2}}{2!} + \frac{x^{7/2}}{4!} - \cdots\right] + c_1\left[x^{1/2} - \frac{x^{5/2}}{3!} + \frac{x^{9/2}}{5!} - \cdots\right] = 0.$$

We have,

$$J_n(x) = \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(n+r+1)}\left(\frac{x}{2}\right)^{2r+n}.$$

Thus,

$$J_{1/2}(x) = \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(1/2+r+1)}\left(\frac{x}{2}\right)^{2r+1/2}.$$

Simplification of the above equation yields the same result as we have got on solving the Bessel's equation for $n = \pm 1/2$.

The function $J_n(x)$ is one of the solutions of the Bessel's equation. Also, $J_{-n}(x)$ represents a solution of Bessel's equation which may or may not be independent of $J_n(x)$ always.

Theorem 10.1.2. If n is an integer, then

$$J_{-n}(x) = (-1)^n J_n(x).$$

This shows that $J_n(x)$ and $J_{-n}(x)$ do not provide with two independent solutions of Bessel's equations when n is an integer.

Proof. •

CaseI: When n is a positive integer, then

$$\begin{aligned} J_n(x) &= \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(n+r+1)}\left(\frac{x}{2}\right)^{2r+n}, \\ \text{and} \quad J_{-n}(x) &= \sum_{r=0}^{\infty}(-1)^r.\frac{1}{\Gamma(r+1)\Gamma(r-n+1)}\left(\frac{x}{2}\right)^{2r-n}. \end{aligned}$$

We know that $\Gamma(m) = \infty$ if $m = 0$ or a negative integer. Thus, $-n+r+1$ should be greater than zero, that is, $-n+r+1 \geq 1 \implies r \neq n$. So,

$$J_{-n}(x) = \sum_{r=n}^{\infty}(-1)^r.\frac{1}{r!\Gamma(r-n+1)}\left(\frac{x}{2}\right)^{2r-n}.$$

Put $m = r - n$ and eliminate r. Thus,

$$\begin{aligned} J_{-n}(x) &= \sum_{m=0}^{\infty}(-1)^{m+n}.\frac{1}{(m+n)!\Gamma(m+1)}\left(\frac{x}{2}\right)^{2m+n} \\ &= \sum_{m=0}^{\infty}(-1)^{m+n}.\frac{1}{\Gamma(m+n+1)\Gamma(m+1)}\left(\frac{x}{2}\right)^{2m+n} \\ &= (-1)^n\sum_{m=0}^{\infty}(-1)^{m}.\frac{1}{\Gamma(m+n+1)\Gamma(m+1)}\left(\frac{x}{2}\right)^{2m+n} = (-1)^n J_n(x). \end{aligned}$$

CaseII: When n is a negative integer. Let $n = -p$, where p is a positive integer. Then,

$$\begin{aligned} J_{-p}(x) &= (-1)^p J_p(x), \quad \text{[by CaseI]} \\ \text{or,} \quad J_n(x) &= (-1)^{-n} J_{-n}(x) \\ \text{or,} \quad J_{-n}(x) &= (-1)^n J_n(x). \end{aligned}$$

□

When n is an integer, $J_{-n}(x)$ is not independent of $J_n(x)$. Hence $y = AJ_n(x) + BJ_{-n}(x)$ is not a general solution of (10.1.1) when n is an integer. But when n is non-integral, then the general solution is given by $y = AJ_n(x)+BJ_{-n}(x)$, for arbitrary constants A and B. We will investigate the general solution for Bessel's equation for integral n. We have the theorem below in this direction.

Theorem 10.1.3. The two linearly independent solutions of (10.1.1) may be taken to be two functions taken as $y_1(x) = J_n(x)$ and

$$y_2(x) = \lim_{\nu \to n}\frac{\cos(\nu\pi)J_\nu(x) - J_{-\nu}(x)}{\sin(\nu\pi)} = Y_n(x).$$

Proof. •

CaseI: When n is not an integer. Since n is not an integer, so $\sin n\pi \neq 0$. Hence

$$Y_n(x) = \cot(n\pi)J_n(x) - \text{cosec}(n\pi)J_{-n}(x),$$

that is, Y_n is a linear combination of $J_n(x)$ and $J_{-n}(x)$. But we know that J_n and J_{-n} are independent solutions of Bessel's equation, when n is not an integer, that is,

$$W(J_n(x), J_{-n}(x)) = \begin{vmatrix} J_n & J_{-n} \\ J_n' & J_{-n}' \end{vmatrix} \neq 0.$$

Now, on simplifying, we get

$$W(J_n, Y_n) = \begin{vmatrix} J_n & Y_n \\ J_n' & Y_n' \end{vmatrix} = -\text{cosec}(n\pi)W(J_n(x), J_{-n}(x)) \neq 0.$$

Thus, J_n and Y_n are two independent solutions of Bessel's equation of order n.

CaseII: Let n be an integer. Then $\sin(n\pi) = 0$ and $\cos(n\pi) = (-1)^n$ and also $J_{-n}(x) = (-1)^n J_n(x)$. First, we deduce a simplified form of $Y_n(x)$.

$$\begin{aligned} Y_n(x) &= \lim_{\nu \to n} \frac{\cos(\nu\pi)J_\nu(x) - J_{-\nu}(x)}{\sin(\nu\pi)} \left(\frac{0}{0}\right) \\ &= \frac{-\pi\sin(n\pi)J_n(x) + \cos(n\pi)\left[\frac{\partial}{\partial\nu}J_\nu(x)\right]_{\nu=n} - \left[\frac{\partial}{\partial\nu}J_{-\nu}(x)\right]_{\nu=n}}{\pi\cos(n\pi)} \\ &= \frac{1}{\pi}\left[\frac{\partial}{\partial\nu}J_\nu(x) - (-1)^n\frac{\partial}{\partial\nu}J_{-\nu}(x)\right]_{\nu=n}. \end{aligned} \tag{10.1.4}$$

We now establish the following two results for $Y_n(x)$.

1. $Y_n(x)$ is a solution of Bessel's equation.

   Proof. $J_\nu(x)$ and $J_{-\nu}(x)$ are solutions of Bessel's equation of order ν. Thus,

$$x^2\frac{d^2}{dx^2}J_\nu + x\frac{d}{dx}J_\nu + (x^2-\nu^2)J_\nu = 0 \tag{10.1.5}$$

$$x^2\frac{d^2}{dx^2}J_{-\nu} + x\frac{d}{dx}J_{-\nu} + (x^2-\nu^2)J_{-\nu} = 0 \tag{10.1.6}$$

   Differentiating (10.1.5) and (10.1.6), with respect to ν we get,

$$x^2\frac{d^2}{dx^2}\left(\frac{\partial}{\partial\nu}J_\nu\right) + x\frac{d}{dx}\left(\frac{\partial}{\partial\nu}J_\nu\right) + (x^2-\nu^2)\left(\frac{\partial}{\partial\nu}J_\nu\right) - 2\nu J_\nu = 0 \tag{10.1.7}$$

$$x^2\frac{d^2}{dx^2}\left(\frac{\partial}{\partial\nu}J_{-\nu}\right) + x\frac{d}{dx}\left(\frac{\partial}{\partial\nu}J_{-\nu}\right) + (x^2-\nu^2)\left(\frac{\partial}{\partial\nu}J_{-\nu}\right) - 2\nu J_{-\nu} = 0 \tag{10.1.8}$$

   By (10.1.7) $-(-1)^\nu$(10.1.8), we get

$$x^2\frac{d^2}{dx^2}\left[\frac{\partial}{\partial\nu}J_\nu - (-1)^\nu\frac{\partial}{\partial\nu}J_{-\nu}\right] + x\frac{d}{dx}\left[\frac{\partial}{\partial\nu}J_\nu - (-1)^\nu\frac{\partial}{\partial\nu}J_{-\nu}\right] +$$

$$(x^2-\nu^2)\left[\frac{\partial}{\partial\nu}J_\nu-(-1)^\nu\frac{\partial}{\partial\nu}J_{-\nu}\right]-2\nu[J_\nu-(-1)^\nu J_{-\nu}]=0.$$

At $\nu=n$, the above equation becomes

$$\left[x^2\frac{d^2}{dx^2}+x\frac{d}{dx}+(x^2-n^2)\right]\left[\frac{\partial}{\partial\nu}J_\nu-(-1)^\nu\frac{\partial}{\partial\nu}J_{-\nu}\right]_{\nu=n}-2n(J_n-(-1)^nJ_{-n})=0$$

$$\text{or,}\quad \left[x^2\frac{d^2}{dx^2}+x\frac{d}{dx}+(x^2-n^2)\right]\left[\frac{\partial}{\partial\nu}J_\nu-(-1)^\nu\frac{\partial}{\partial\nu}J_{-\nu}\right]_{\nu=n}=0$$

$$\text{or,}\quad \left[x^2\frac{d^2}{dx^2}+x\frac{d}{dx}+(x^2-n^2)\right]Y_n(x)=0.$$

Thus, Y_n is a solution of Bessel's equation of order n. □

2. $Y_n(x)$ is independent of $J_n(x)$.

   This is evident from the structure of Y_n and J_n.

□

Definition 10.1.4. Bessel's function of second kind of order n, denoted by $Y_n(x)$ is defined as

$$\begin{aligned}Y_n(x) &= \lim_{\nu\to n}\frac{\cos(\nu\pi)J_\nu(x)-J_{-\nu}(x)}{\sin(\nu\pi)},\quad \text{when n is an integer}\\ &= \frac{J_n(x)\cos(n\pi)-J_{-n}(x)}{\sin(n\pi)},\quad \text{when n is not an integer.}\end{aligned}$$

Thus, the general solution of Bessel's equation is

$$y(x)=C_1J_n(x)+C_2Y_n(x)$$

where C_1 and C_2 are independent constants.

10.2 Recurrence Relations of Bessel's Equations

For integral n, the Bessel's function satisfies the following recurrence relations:

1. $xJ_n'(x)=nJ_n(x)-xJ_{n+1}(x)$.

   Proof. We have

$$J_n(x)=\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}.$$

   Differentiating with respect to x, we get

$$J_n'(x)=\frac{1}{2}\sum_{r=0}^{\infty}(-1)^r\frac{(n+2r)}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r-1}$$

$$\text{or,}\qquad xJ_n'(x)=\sum_{r=0}^{\infty}(-1)^r\frac{(n+2r)}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}$$

$$= n\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r} + 2\sum_{r=0}^{\infty}(-1)^r\frac{r}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}$$

$$= nJ_n(x) - 2\sum_{s=0}^{\infty}(-1)^s\frac{1}{s!\Gamma(n+s+2)}\left(\frac{x}{2}\right)^{n+2s+2} \quad \text{[Putting r-1=s and eliminating r]}$$

$$= nJ_n(x) - x\frac{1}{\Gamma(s+1)\Gamma(n+s+2)}\left(\frac{x}{2}\right)^{n+1+2s} = nJ_n(x) - xJ_{n+1}(x).$$

□

2. $xJ_n'(x) = xJ_{n-1}(x) - nJ_n(x)$.

Proof. We have

$$J_n(x) = \sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}.$$

Differentiating with respect to x, we get

$$J_n'(x) = \frac{1}{2}\sum_{r=0}^{\infty}(-1)^r\frac{(n+2r)}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r-1}$$

$$\text{or,}\quad 2J_n'(x) = \sum_{r=0}^{\infty}(-1)^r\frac{2(n+2r)}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r-1} - \sum_{r=0}^{\infty}(-1)^r\frac{n}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r-1}$$

$$= 2\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n)}\left(\frac{x}{2}\right)^{n-1+2r} - n\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r-1}$$

$$= 2J_{n-1}(x) - \frac{2n}{x}\sum_{r=0}^{\infty}(-1)^r\frac{1}{r!\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r} = 2J_{n-1}(x) - \frac{2n}{x}J_n(x).$$

Simplifying, we get the desired result. □

3. $\frac{2n}{x}J_n(x) = J_{n-1}(x) + J_{n+1}(x)$.

Proof. From 1, we have

$$xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x).$$

From 2, we have

$$xJ_n'(x) = xJ_{n-1}(x) - nJ_n(x).$$

Subtracting and simplifying, we get the desired result. □

4. $2J_n'(x) = J_{n-1}(x) - J_{n+1}(x)$.

Proof. From 1, we have

$$xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x).$$

From 2, we have

$$xJ_n'(x) = xJ_{n-1}(x) - nJ_n(x).$$

Adding and simplifying, we get the desired result. □

10.3 Generating Function for Bessel's Functions

For all values of x and for all values of z such that $0 < |z| < \infty$, the function

$$w(x,z) = e^{\frac{x}{2}\left\{z-\frac{1}{z}\right\}},$$

generates the Bessel's function of **integral** order n, that is,

$$e^{\frac{x}{2}\left\{z-\frac{1}{z}\right\}} = \sum_{n=-\infty}^{\infty} J_n(x) z^n.$$

Proof. We have,

$$\begin{aligned} e^{\frac{x}{2}\left\{z-\frac{1}{z}\right\}} &= e^{xz/2}.e^{-x/(2z)} \\ &= \left\{\sum_{r=0}^{\infty}\left(\frac{xz}{2}\right)^r \frac{1}{r!}\right\}\left\{\sum_{s=0}^{\infty}(-1)^s\left(\frac{x}{2z}\right)^s \frac{1}{s!}\right\} \end{aligned}$$

$$= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}(-1)^s \frac{1}{\Gamma(r+1)\Gamma(s+1)}\left(\frac{x}{2}\right)^{r+s} z^{r-s}. \tag{10.3.1}$$

CaseI: When $r - s \geq 0$, let $r - s = n$. Then $n \geq 0$, that is, n is an integer varying from 0 to ∞. Eliminating s, we get

$$\sum_{n=0}^{\infty}\sum_{r=n}^{\infty}(-1)^{r-n}\frac{1}{\Gamma(r+1)\Gamma(r-n+1)}\left(\frac{x}{2}\right)^{2r-n} z^n.$$

[For real values of $\Gamma(r+1)$ we must have $r+1 \geq 1$, that is, $r \geq 0$ and for real values of $\Gamma(r-n+1)$ we must similarly have $r \geq n$.] Now, let

$$w(x,z) = \sum_{n=0}^{\infty} f_n(x) z^n, \quad \text{where } f_n(x) = \sum_{r=n}^{\infty}(-1)^{r-n}\frac{1}{\Gamma(r+1)\Gamma(r-n+1)}\left(\frac{x}{2}\right)^{2r-n}.$$

Putting $r - n = p$ and eliminating r, we get

$$f_n(x) = \sum_{p=0}^{\infty}(-1)^p \frac{1}{\Gamma(p+1)\Gamma(p+n+1)}\left(\frac{x}{2}\right)^{2p+n} = J_n(x). \tag{10.3.2}$$

CaseII: When $r - s < 0$, let $r - s = -n_1$, where n_1 is a positive integer. So n_1 takes values from 1 to ∞. Eliminating s, we get,

$$w(x,z) = \sum_{n_1=1}^{\infty}\sum_{r=0}^{\infty}(-1)^{r+n_1}\frac{1}{\Gamma(r+1)\Gamma(r+n_1+1)}\left(\frac{x}{2}\right)^{2r+n_1} z^{-n_1}.$$

[For real values of $\Gamma(r+1)$ we must have $r+1 \geq 1$, that is, $r \geq 0$ and for real values of $\Gamma(r+n_1+1)$ we must similarly have $r \geq -n_1$. Both inequalities simultaneously give $r \geq 0$.] We assume

$$w(x,z) = \sum_{n_1=1}^{\infty} f_{n_1}(x) z^{-n_1},$$

where

$$f_{n_1}(x) = \sum_{r=0}^{\infty}(-1)^{r+n_1}\frac{1}{\Gamma(r+1)\Gamma(r+n_1+1)}\left(\frac{x}{2}\right)^{2r+n_1} = (-1)^{n_1}J_{n_1}(x).$$

Thus,

$$\begin{aligned} w(x,z) &= \sum_{n_1=1}^{\infty}(-1)^{n_1}J_{n_1}(x)z^{-n_1} \\ &= \sum_{n=-\infty}^{-1}(-1)^{-n}J_{-n}(x)z^n \\ &= \sum_{n=-\infty}^{-1}(-1)^{-n}(-1)^nJ_n(x)z^n = \sum_{n=-\infty}J_n(x)z^n. \end{aligned} \tag{10.3.3}$$

Combining (10.3.2) and (10.3.3), we get the desired result.

□

10.4 Orthogonality Properties of Bessel's Polynomials

If c_i and c_j are the roots of the equation $J_n(ca) = 0$, then

$$\begin{aligned} \int_0^a xJ_n(c_ix)J_n(c_jx)dx &= 0, \quad i \neq j \\ &= \frac{a^2}{a}J_{n+1}^2(c_ia), \quad i = j. \end{aligned}$$

To prove the above, we first write the Bessel's equation of order n.

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

The general solution is

$$y(x) = AJ_n(x) + BY_n(x).$$

We first show that $J_n(cx)$ satisfies the following equation, known as the **modified Bessel's equation**,

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

Put $z = cx$. Then we get

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = c\frac{dy}{dz}, \qquad \frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dz}.c\right) = \frac{d}{dz}\left(\frac{dy}{dz}.c\right)\frac{dz}{dx} = c^2\frac{d^2y}{dz^2}.$$

Using these in the modified Bessel's equation, we get

$$z^2\frac{d^2y}{dz^2} + z\frac{dy}{dz} + (z^2 - n^2)y = 0.$$

This is Bessel's equation in the variable z whose general solution is

$$y(z) = AJ_n(z) + BY_n(z).$$

Thus $J_n(z)$, that is, $J_n(cx)$ is the solution of the modified Bessel's equation. We now move on to prove the orthogonality condition.

Proof. •

CaseI: When $i \neq j$. c_i and c_j are the roots of $J_n(ca) = 0$, that is,

$$J_n(c_i a) = 0, \qquad J_n(c_j a) = 0.$$

We know that $J_n(cx)$ satisfies the modified Bessel's equation. Thus,

$$x^2 \frac{d^2}{dx^2} J_n(cx) + x \frac{d}{dx} J_n(cx) + (c^2x^2 - n^2) J_n(cx) = 0.$$

Thus,

$$x^2 \frac{d^2}{dx^2} J_n(c_i x) + x \frac{d}{dx} J_n(c_i x) + (c_i^2 x^2 - n^2) J_n(c_i x) = 0 \tag{10.4.1}$$

$$x^2 \frac{d^2}{dx^2} J_n(c_j x) + x \frac{d}{dx} J_n(c_j x) + (c_j^2 x^2 - n^2) J_n(c_j x) = 0 \tag{10.4.2}$$

Putting $u = J_n(c_i x)$ and $v = J_n(c_j x)$, we get from (10.4.1) and (10.4.2),

$$x^2 \frac{d^2 u}{dx^2} + x \frac{du}{dx} + (c_i^2 x^2 - n^2) u = 0 \tag{10.4.3}$$

$$x^2 \frac{d^2 v}{dx^2} + x \frac{dv}{dx} + (c_j^2 x^2 - n^2) v = 0 \tag{10.4.4}$$

Now, (10.4.3)$\times v -$(10.4.4)$\times u$ gives on simplification

$$\frac{d}{dx}\left\{ x \left(\frac{du}{dx} v - \frac{dv}{dx} u \right) \right\} + (c_i^2 - c_j^2) x u v = 0.$$

Integrating the above equation with respect to x from 0 to a, we get

$$\begin{aligned}
(c_j^2 - c_i^2) \int_0^a xuv\,dx &= \left[x \left(\frac{du}{dx} v - \frac{dv}{dx} u \right) \right]_0^a \\
&= \left[x \left(v \frac{d}{dx} u - u \frac{d}{dx} v \right) \right]_0^a \\
&= a J_n(c_j a) \left[\frac{d}{dx} J_n(c_i x) \right]_{x=a} - a J_n(c_i a) \left[\frac{d}{dx} J_n(c_j x) \right]_{x=a} = 0
\end{aligned}$$

[since c_i and c_j are the roots of the equation $J_n(ca) = 0$]. Hence the result.

CaseII: When $i = j$. Multiplying (10.4.3) by $2\frac{du}{dx}$, we get on simplifying,

$$\frac{d}{dx}\left[\left(x^2 \left(\frac{du}{dx} \right)^2 \right) - n^2 u^2 + c_i^2 x^2 u^2 \right] - 2 c_i^2 x u^2 = 0.$$

Integrating the above equation with respect to x from 0 to a, we get

$$\begin{aligned}
2c_i^2 \int_0^a x u^2\,dx &= \left[\left(x^2 \left(\frac{du}{dx} \right)^2 \right) - n^2 u^2 + c_i^2 x^2 u^2 \right]_0^a \\
&= \left[\left(x^2 \left(\frac{d}{dx} J_n(c_i x) \right)^2 \right) - n^2 [J_n(c_i x)]^2 + c_i^2 x^2 [J_n(c_i x)]^2 \right]_0^a \\
&= a^2 \left[\frac{d}{dx} J_n(c_i x) \right]_{x=a} - n^2 [J_n(c_i a)]^2 + c_i^2 a^2 [J_n(c_i a)]^2 + n^2 (J_n(0))^2 \\
&= a^2 \left[\frac{d}{dx} J_n(c_i x) \right]_{x=a} + n^2 (J_n(0))^2 = a^2 \left[\frac{d}{dx} J_n(c_i x) \right]_{x=a}.
\end{aligned} \tag{10.4.5}$$

Recurrence relation 1 gives

$$xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x).$$

Now, put $x = c_i y$. Then $dx = c_i dy$. Thus,

$$\begin{aligned} c_i y \frac{1}{c_i}\frac{d}{dy} J_n(c_i y) &= nJ_n(c_i y) - c_i y J_{n+1}(c_i y) \\ \text{or,} \quad \frac{d}{dy} J_n(c_i y) &= \frac{n}{y} J_n(c_i y) - c_i J_{n+1}(c_i y) \\ \text{or,} \quad \frac{d}{dx} J_n(c_i x) &= \frac{n}{x} J_n(c_i x) - c_i J_{n+1}(c_i x). \end{aligned}$$

Putting $x = a$ on both sides of the last equation, we get

$$\left[\frac{d}{dx} J_n(c_i x)\right]_{x=a} = \frac{n}{a} J_n(c_i a) - c_i J_{n+1}(c_i a) = -c_i J_{n+1}(c_i a).$$

Thus, (10.4.5) gives

$$2c_i^2 \int_0^a x[J_n(c_i x)]^2 dx = a^2 c_i^2 [J_{n+1}(c_i a)]^2.$$

Simplifying, we get the desired result.

□

10.5 Integrals of Bessel's Functions

1. We have for any values of n,

$$J_n(x) = \sum_{r=0}^{\infty} (-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)} \left(\frac{x}{2}\right)^{n+2r}.$$

Then,

$$\begin{aligned} \frac{d}{dx}\left(x^n J_n(x)\right) &= \frac{d}{dx}\left(\sum_{r=0}^{\infty} (-1)^r \frac{x^{2n+2r}}{2^{n+2r}\Gamma(r+1)\Gamma(r+n+1)}\right) \\ &= \sum_{r=0}^{\infty} (-1)^r \frac{(2n+2r)\, x^{2n+2r-1}}{2^{n+2r}\Gamma(r+1)\Gamma(r+n+1)} \\ &= x^n \sum_{r=0}^{\infty} (-1)^r \frac{x^{(n-1)+2r}}{2^{(n-1)+2r}\Gamma(r+1)\Gamma(r+(n-1)+1)} \\ &= x^n J_{n-1}(x) \qquad (10.5.1) \end{aligned}$$

Thus,

$$\int x^n J_{n-1}(x) dx = x^n J_n(x) + c \qquad (10.5.2)$$

In particular, for $n = 1$, we get

$$\int x J_0(x) dx = x J_1(x) + c \qquad (10.5.3)$$

2. We have for any values of n,

$$\frac{d}{dx}\left(x^{-n}J_n(x)\right) = -x^{-n}J_{n+1}(x) \qquad \text{(verify!)}$$

Thus

$$\int x^{-n}J_{n+1}(x)dx = -x^{-n}J_n(x) + c \tag{10.5.4}$$

In particular, for $n = 0$, we get

$$\int J_1(x)dx = -J_0(x) + c \tag{10.5.5}$$

3. Rewriting the recurrence relation 4, we get

$$J_{n+1}(x) = J_{n-1}(x) - 2J_n^{'}(x)$$

Integrating w.r.t. x we get

$$\int J_{n+1}(x)dx = \int J_{n-1}(x)dx - 2J_n(x). \tag{10.5.6}$$

Example 10.5.1. Find $J_0(x)$ and $J_1(x)$.
We have,

$$J_n(x) = \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}.$$

Putting $n = 0$, we get

$$\begin{aligned} J_0(x) &= \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+1)}\left(\frac{x}{2}\right)^{2r} = \sum_{r=0}^{\infty}(-1)^r \frac{1}{(r!)^2}\left(\frac{x}{2}\right)^{2r} \\ &= 1 - \frac{1}{(1!)^2}\left(\frac{x}{2}\right)^2 + \frac{1}{(2!)^2}\left(\frac{x}{2}\right)^4 - \frac{1}{(3!)^2}\left(\frac{x}{2}\right)^6 + \cdots \\ &= 1 - \frac{1}{4}x^2 + \frac{1}{64}x^4 - \frac{1}{2304}x^6 + \cdots. \end{aligned}$$

Again putting $n = 1$, we get

$$\begin{aligned} J_1(x) &= \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+2)}\left(\frac{x}{2}\right)^{1+2r} = \sum_{r=0}^{\infty}(-1)^r \frac{1}{r!(r+1)!}\left(\frac{x}{2}\right)^{1+2r} \\ &= \frac{x}{2}\left[1 - \frac{1}{1!2!}\left(\frac{x}{2}\right)^2 + \frac{1}{2!3!}\left(\frac{x}{2}\right)^4 - \frac{1}{3!4!}\left(\frac{x}{2}\right)^6 + \cdots\right] \\ &= \frac{x}{2} - \frac{1}{16}x^3 + \frac{1}{384}x^5 - \frac{1}{18432}x^7 + \cdots. \end{aligned}$$

Example 10.5.2. Show that $J_n(x)$ is an even function when n is even and odd function when n is odd.
Suppose n is even. Then

$$\begin{aligned} J_n(-x) &= \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{-x}{2}\right)^{n+2r} \\ &= (-1)^n \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r} \\ &= (-1)^n J_n(x) = J_n(x), \text{ since } n \text{ is even.} \end{aligned}$$

Next suppose n is odd. Then

$$\begin{aligned}
J_n(-x) &= \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{-x}{2}\right)^{n+2r} \\
&= (-1)^n \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r} \\
&= (-1)^n J_n(x) = -J_n(x), \text{ since } n \text{ is odd.}
\end{aligned}$$

Example 10.5.3. Express $J_6(x)$ in terms of $J_0(x)$ and $J_0(x)$.

Rewriting the recurrence relation 3 we get

$$J_{n+1}(x) = \frac{2n}{x} J_n(x) - J_{n-1}(x).$$

Putting $n = 1, 2, 3, 4, 5$ we get

$$J_2(x) = \frac{2}{x} J_1(x) - J_0(x)$$

$$J_3(x) = \frac{4}{x} J_2(x) - J_1(x) = \left(\frac{8}{x^2} - 1\right) J_1(x) - \frac{4}{x} J_0(x)$$

$$J_4(x) = \frac{6}{x} J_3(x) - J_2(x) = \left(\frac{48}{x^3} - \frac{8}{x}\right) J_1(x) + \left(1 - \frac{24}{x^2}\right) J_0(x)$$

$$J_5(x) = \frac{8}{x} J_4(x) - J_3(x) = \left(\frac{384}{x^4} - \frac{72}{x^2} - 1\right) J_1(x) + \left(\frac{12}{x} - \frac{192}{x^3}\right) J_0(x)$$

$$J_6(x) = \frac{10}{x} J_5(x) - J_4(x) = \left(\frac{3840}{x^5} - \frac{768}{x^3} - \frac{2}{x}\right) J_1(x) + \left(-1 + \frac{144}{x^2} - \frac{1920}{x^4}\right) J_0(x).$$

Example 10.5.4. Show that

$$J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cos x \quad \text{and} \quad J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x.$$

We have,

$$J_n(x) = \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+n+1)}\left(\frac{x}{2}\right)^{n+2r}.$$

Thus,

$$\begin{aligned}
J_{-1/2}(x) &= \sum_{r=0}^{\infty}(-1)^r \frac{1}{\Gamma(r+1)\Gamma(r+1/2)}\left(\frac{x}{2}\right)^{-1/2+2r} \\
&= \sqrt{\frac{2}{\pi x}}\left[1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right] = \sqrt{\frac{2}{\pi x}} \cos x.
\end{aligned}$$

We can similarly prove the other part by expanding the Bessel's function $J_{1/2}(x)$.

Example 10.5.5. Show that

$$\int_0^{\pi/2} \sqrt{\pi x} J_{1/2}(2x) dx = 1.$$

We have,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x.$$

Thus,

$$J_{1/2}(2x) = \sqrt{\frac{1}{\pi x}} \sin 2x.$$

Integrating the above with respect to x from 0 to $\pi/2$, we get the required result.

Example 10.5.6. Express $J_{\frac{7}{2}}(x)$ in terms of *sine* and *cosine* functions.

Rewriting the recurrence relation 3 we get

$$J_{n+1}(x) = \frac{2n}{x} J_n(x) - J_{n-1}(x).$$

Putting $n = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$, we get

$$J_{\frac{3}{2}}(x) = \frac{1}{x} J_{\frac{1}{2}}(x) - J_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \left(\frac{\sin x}{x} - \cos x \right)$$

$$\begin{aligned} J_{\frac{5}{2}}(x) &= \frac{3}{x} J_{\frac{3}{2}}(x) - J_{\frac{1}{2}}(x) = \frac{3}{x} \left[\sqrt{\frac{2}{\pi x}} \left(\frac{\sin x}{x} - \cos x \right) \right] - \sqrt{\frac{2}{\pi x}} \sin x \\ &= \sqrt{\frac{2}{\pi x}} \left[\left(\frac{3 - x^2}{x^2} \right) \sin x - \frac{3}{x} \cos x \right] \end{aligned}$$

$$\begin{aligned} J_{\frac{7}{2}}(x) &= \frac{5}{x} J_{\frac{5}{2}}(x) - J_{\frac{3}{2}}(x) \\ &= \frac{5}{x} \sqrt{\frac{2}{\pi x}} \left[\left(\frac{3 - x^2}{x^2} \right) \sin x - \frac{3}{x} \cos x \right] - \sqrt{\frac{2}{\pi x}} \left(\frac{\sin x}{x} - \cos x \right) \\ &= \sqrt{\frac{2}{\pi x}} \left[\left(\frac{15 - 6x^2}{x^3} \right) \sin x - \left(\frac{15 - x^2}{x^2} \right) \cos x \right]. \end{aligned}$$

Example 10.5.7. Show that $J_1''(x) = -J_1(x) + \frac{1}{x} J_2(x)$.

Rewriting the recurrence relation 2 we get

$$J_n'(x) = J_{n-1}(x) - \frac{n}{x} J_n(x) \tag{10.5.7}$$

Putting $n = 1$, in (10.5.7) we get

$$J_1'(x) = J_0(x) - \frac{1}{x} J_1(x) \tag{10.5.8}$$

Differentiating (10.5.8) w.r.t. x, we get

$$J_1''(x) = J_0'(x) + \frac{1}{x^2} J_1(x) - \frac{1}{x} J_1'(x). \tag{10.5.9}$$

Putting $n = 0$, in (10.5.7) we get

$$J_0'(x) = J_{-1}(x) = -J_1(x) \tag{10.5.10}$$

since, $J_{-n}(x) = (-1)^n J_n(x)$, when n is an integer.

From (10.5.8), (10.5.9) and (10.5.10) we get

$$J_1''(x) = -J_1(x) + \frac{1}{x^2}J_1(x) - \frac{1}{x}\left[J_0(x) - \frac{1}{x}J_1(x)\right] \tag{10.5.11}$$

Rewriting the recurrence relation 3 we get

$$J_{n+1}(x) = \frac{2n}{x}J_n(x) - J_{n-1}(x).$$

Putting $n = 1$, we get

$$J_2(x) = \frac{2}{x}J_1(x) - J_0(x) \tag{10.5.12}$$

Eliminating $J_0(x)$ from (10.5.11) and (10.5.12) we get

$$\begin{aligned} J_1''(x) &= -J_1(x) + \frac{1}{x^2}J_1(x) - \frac{1}{x}\left[\frac{2}{x}J_1(x) - J_2(x) - \frac{1}{x}J_1(x)\right] \\ &= -J_1(x) + \frac{1}{x}J_2(x). \end{aligned}$$

Example 10.5.8. Show that, $x^n J_n(x)$ is a solution of

$$x\frac{d^2y}{dx^2} + (1-2n)\frac{dy}{dx} + xy = 0. \tag{10.5.13}$$

Let $y = x^n J_n(x)$.
From (10.5.1), we have

$$\begin{aligned} \frac{d}{dx}\left[x^n J_n(x)dx\right] &= x^n J_{n-1}(x) \\ \text{or } \frac{dy}{dx} &= x^n J_{n-1}(x) \end{aligned}$$

Differentiaing w.r.t x, we get

$$\frac{d^2y}{dx^2} = x^n J_{n-1}'(x) + nx^{n-1}J_{n-1}(x).$$

Putting the value of y, $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ in the LHS of (10.5.13) we get

$$\begin{aligned} x\frac{d^2y}{dx^2} + (1-2n)\frac{dy}{dx} + xy &= x\left[x^n J_{n-1}'(x) + nx^{n-1}J_{n-1}(x)\right] + (1-2n)x^n J_{n-1}(x) + x^{n+1}J_n(x) \\ &= x^{n+1}J_{n-1}'(x) - (n-1)x^n J_{n-1}(x) + x^{n+1}J_n(x) \\ &= x^{n+1}\left[J_{n-1}'(x) - \frac{(n-1)}{x}J_{n-1}(x) + J_n(x)\right] \\ &= 0 \qquad \text{by recurrence relation 1.} \end{aligned}$$

Example 10.5.9. Evaluate $\int J_5(x)dx$.
From (10.5.6), we have

$$\int J_{n+1}(x)dx = \int J_{n-1}(x)dx - 2J_n(x).$$

Putting $n = 4$, we get

$$\int J_5(x)dx = \int J_3(x)dx - 2J_4(x).$$

Again putting $n = 2$, we get

$$\int J_3(x)dx = \int J_1(x)dx - 2J_2(x).$$

Also from (10.5.5), we know that

$$\int J_1(x)dx = -J_0(x) + c$$

From the above relations we get

$$\begin{aligned}\int J_5(x)dx &= \int J_1(x)dx - 2J_2(x) - 2J_4(x)\\ &= -J_0(x) - 2J_2(x) - 2J_4(x) + c\end{aligned}$$

Example 10.5.10. Evaluate $\int x^3 J_3(x)dx$.
From (10.5.4), we get, $\int \left\{x^{-2}J_3(x)\right\} dx = -x^{-2}J_2(x)$.
Integrating the given integral by parts w.r.t. x we get,

$$\begin{aligned}\int x^3 J_3(x)dx &= \int x^5 \left[x^{-2}J_3(x)\right] dx\\ &= -x^5x^{-2}J_2(x) + \int 5x^4x^{-2}J_2(x)dx\\ &= -x^3J_2(x) + 5\int x^2J_2(x)dx\end{aligned}$$

Again (10.5.4), we get, $\int \left\{x^{-1}J_2(x)\right\} dx = -x^{-1}J_1(x)$.
Integrating $\int x^2J_2(x)dx$ by parts w.r.t. x we get,

$$\begin{aligned}\int x^2 J_2(x)dx &= \int x^3 \left[x^{-1}J_2(x)\right] dx\\ &= -x^3x^{-1}J_1(x) + \int 3x^2x^{-1}J_1(x)dx\\ &= -x^2J_1(x) + 3\int xJ_1(x)dx\end{aligned}$$

Again,

$$\begin{aligned}\int xJ_1(x)dx &= -\int xJ_0'(x)dx\\ &= -\left[xJ_0(x) - \int J_0(x)dx\right]\end{aligned}$$

Substituting all the above values, we get

$$\begin{aligned}\int x^3J_3(x)dx &= -x^3J_2(x) + 5\int x^2J_2(x)dx\\ &= -x^3J_2(x) + 5\left[-x^2J_1(x) + 3\left\{-xJ_0(x) + \int J_0(x)dx\right\}\right]\\ &= -x^3J_2(x) - 5x^2J_1(x) - 15xJ_0(x) + 15\int J_0(x)dx\end{aligned}$$

Example 10.5.11. Prove that $J_{n+1}(x) = x\int_0^1 J_n(xy)y^{n+1}dy$.
Let $xy = t$. Then $xdy = dt$ and

$$\begin{aligned}
x\int_0^1 J_n(xy)y^{n+1}dy &= x\int_0^x J_n(t)\left(\frac{t}{x}\right)^{n+1}\frac{dt}{x} \\
&= x^{-n-1}\int_0^x t^{n+1}J_n(t)dt \\
&= x^{-n-1}\int_0^x x^{n+1}J_n(x)dx \\
&= x^{-n-1}x^{n+1}J_{n+1}(x) \qquad \text{from (10.5.2)} \\
&= J_{n+1}(x)
\end{aligned}$$

Exercise 10.5.12.

1. Prove that $2J_0'' = J_2 - J_0$.
2. Prove that for integral n, $4J_n^{"} = J_{n-2} - 2J_n + J_{n+2}$.
3. Show that, $\frac{d}{dx}\{xJ_nJ_{n+1}\} = x[J_n^2 - J_{n+1}^2]$
4. Prove that, $\frac{d}{dx}\{J_n^2 + J_{n+1}^2\} = 2[\frac{n}{x}J_n^2 - \frac{n+1}{x}J_{n+1}^2]$.
5. Prove that, $J_0^2 + 2(J_1^2 + J_2^2 + J_3^2 + \cdots) = 1$
6. Prove that

$$J_{-3/2}(x) = \sqrt{\frac{2}{\pi x}}\left(-\frac{\cos x}{x} - \sin x\right) \quad \text{and} \quad J_{3/2}(x) = \sqrt{\frac{2}{\pi x}}\left(\frac{\sin x}{x} - \cos x\right).$$

7. Express J_3 and J_4 in terms of J_0 and J_1.
8. Prove that $J_0^{'} = -J_1$.
9. For $n > 1$ show that

$$\int_0^x x^{n+1}J_n(x)dx = x^{n+1}J_{n+1}(x)dx.$$

10. Prove that

(a)

$$\frac{d}{dx}(xJ_1(x)) = xJ_0(x).$$

(b)

$$\int_0^b xJ_0(ax)dx = \frac{b}{a}J_1(ab).$$

11. Express $\int x^{-3}J_4(x)dx$ in terms of J_0 and J_1.
12. Show that, $x^{-n}J_n(x)$ is a solution of

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

13. Prove that

(a)
$$\int_0^x x^{-n} J_{n+1}(x)dx = \frac{1}{2^n\Gamma(n+1)} - x^{-n}J_n(x),\ n > 1.$$

(b)
$$\int_0^\infty x^{-n} J_{n+1}(x)dx = \frac{1}{2^n\Gamma(r+1)},\ n > -\frac{1}{2}.$$

14. Prove that
$$\int J_0(x)\sin x dx = xJ_0(x)\sin x - xJ_1(x)\cos x + c$$

Few Probable Questions

1. Deduce a generating function for Bessel's functions.
2. Deduce the orthogonality of Bessel's functions.
3. Define Bessel's function of second kind. Show that the Bessel's function of second kind is a solution of Bessel's equations.
4. Prove that $xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x)$. Hence show that
$$\frac{d}{dx}J_0(x) = -J_1(x) \tag{10.5.14}$$
and
$$\int_a^b J_0(x)J_1(x)dx = \frac{1}{2}[(J_0(a))^2 - (J_0(b))^2]. \tag{10.5.15}$$
5. If $a > 0$, prove that
$$\int_0^\infty e^{-ax}J_0(bx)dx = \frac{1}{\sqrt{a^2+b^2}}.$$

Unit 11

Course Structure

- Hermite Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Construction and solution of Hermite differential equation.

11.1 Introduction

In mathematics, the Hermite polynomials are a classical orthogonal polynomial sequence. These arise in probability, combinatorics, numerical analysis, systems theory, random matrix theory and many more. Hermite polynomials were defined by Pierre-Simon Laplace in 1810, though in scarcely recognizable form, and studied in detail by Pafnuty Chebyshev in 1859. Chebyshev's work was overlooked, and they were named later after Charles Hermite, who wrote on the polynomials in 1864, describing them as new. They were consequently not new, although Hermite was the first to define the multidimensional polynomials in his later 1865 publications. And the Laguerre polynomials arise in quantum mechanics, in the radial part of the solution of the Schrödinger equation for a one-electron atom. They also describe the static Wigner functions of oscillator systems in quantum mechanics in phase space. They further enter in the quantum mechanics of the Morse potential and of the 3D isotropic harmonic oscillator. The generalized Laguerre polynomials are related to the Hermite polynomials. This unit is dedicated to the study of Hermite as well as Laguerre polynomials.

Objectives

After reading this unit, you will be able to

- solve the Hermite's equation and find the general structure of Hermite's polynomial
- define a general Laguerre polynomial
- derive the Rodrigue's formula for both Hermite and Laguerre polynomials
- establish the orthogonality of Hermite and Laguerre polynomials
- find a generating function for Laguerre and Hermite's polynomials
- learn some recurrence relations relating to both
- solve certain problems relating to both

11.2 Hermite's Equation and Hermite's Polynomials

The Hermite's equation is

$$\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2ny = 0 \tag{11.2.1}$$

where, n is a constant.

The most important single application of the Hermite polynomials is to the theory of the linear harmonic oscillator in quantum mechanics. A differential equation that arises in this theory and is closely related to Hermite's equation (11.2.1) is

$$\frac{d^2w}{dx^2} + \left(2n + 1 - x^2\right) w = 0, \tag{11.2.2}$$

where n is a constant. Physicist are interested in sollutions of (11.2.2) that approaches zero as $|x| \to \infty$. To solve (11.2.2) directly by power series method is very tedious job. We use the transformation

$$w = ye^{-\frac{x^2}{2}}$$

and simplify equation (11.2.2) into (11.2.1). Then the desired solution of (11.2.2) will correspond to the solution of (11.2.1) that grow in magnitude less rapidly than $e^{\frac{x^2}{2}}$ as $|x| \to \infty$.

We solve (11.2.1) by Frobenius Method, about $x = 0$. Assume that

$$y = \sum_{m=0}^{\infty} a_m x^{s+m}$$

be the solution of (11.2.1), where $a_0 \neq 0$ and s is to be determined. Then

$$\begin{aligned} \frac{dy}{dx} &= \sum_{m=0}^{\infty} (s+m) a_m x^{s+m-1} \\ \frac{d^2}{dx^2} &= \sum_{m=0}^{\infty} (s+m)(s+m-1) a_m x^{s+m-2} \end{aligned}$$

Thus, equation (11.2.1) becomes

$$\begin{aligned} \sum_{m=0}^{\infty} (s+m)(s+m-1)a_m x^{s+m-2} - 2\sum_{m=0}^{\infty}(s+m)a_m x^{s+m} + 2n\sum_{m=0}^{\infty} a_m x^{s+m} &= 0 \\ \text{or, } \sum_{m=0}^{\infty} (s+m)(s+m-1)a_m x^{s+m-2} - 2\sum_{m=0}^{\infty}(s+m-n)a_m x^{s+m} &= 0 \\ \text{or, } \sum_{m=0}^{\infty} (s+m)(s+m-1)a_m x^{m} - 2\sum_{m=0}^{\infty}(s+m-n)a_m x^{m+2} &= 0 \\ \text{or, } \sum_{m=0}^{\infty} (s+m)(s+m-1)a_m x^{m} - 2\sum_{m=2}^{\infty}(s+m-n-2)a_{m-2} x^{m} &= 0 \\ \text{or, } \sum_{m=2}^{\infty} \{(s+m)(s+m-1)a_m - 2(s+m-n-2)a_{m-2}\}x^m + s(s-1)a_0 + s(s+1)a_1 x &= 0 \end{aligned}$$

The indicial equation is

$$s(s-1)a_0 = 0 \implies s = 0, 1.$$

When $s = 0$, a_1 is indeterminate. When $s = 1$, $a_1 = 0$. The general recurrance relation is

$$a_m = \frac{2(s+m-n-2)}{(s+m)(s+m-1)} a_{m-2}, \quad m \geq 2$$

For $s = 0$, we have

$$a_m = 2\frac{m-n-2}{m(m-1)} a_{m-2}, \quad m \geq 2.$$

Putting $m = 2, 4, \ldots, 2m, \ldots$, we get

$$\begin{aligned}
a_2 &= \frac{(-1)^1 2^1 . n a_0}{2!} \\
a_4 &= \frac{(-1)^2 . 2^2 n(n-2) a_0}{4!} \\
a_6 &= \frac{(-1)^3 . 2^3 n(n-2)(n-4) a_0}{6!} \\
&\vdots \\
a_{2m} &= \frac{(-1)^m . 2^m n(n-2)(n-4)\ldots(n-2m+2)}{(2m)!} a_0
\end{aligned}$$

Next, put $m = 3, 5, \ldots, 2m+1, \ldots$, we get

$$\begin{aligned}
a_3 &= \frac{(-1).2.(n-1)}{3!} a_1 \\
a_5 &= \frac{(-1)^2 . 2^2 (n-1)(n-3)}{5!} a_1 \\
&\vdots \\
a_{2m+1} &= \frac{(-1)^m . 2^m (n-1)(n-3)\ldots(n-2m+1)}{(2m+1)!} a_1
\end{aligned}$$

Hence, the series solution gives

$$\begin{aligned}
y \;=\; & a_0 \left[1 + \frac{(-1)^1 2^1 . n}{2!} x^2 + \cdots + \frac{(-1)^m . 2^m n(n-2)(n-4)\ldots(n-2m+2)}{(2m)!} x^{2m} + \cdots \right] + \\
& a_1 \left[x + \frac{(-1).2.(n-1)}{3!} x^3 + \cdots + \frac{(-1)^m . 2^m (n-1)(n-3)\ldots(n-2m+1)}{(2m+1)!} x^{2m+1} + \cdots \right]
\end{aligned}$$

which is of the form $a_0 y_1(x) + a_1 y_2(x)$. It is observed that in the case when the constant represents a positive integer, then one of the solutions y_1 or y_2 reduces to a polynomial according as n is even or odd.

When $n = 2$, $y_1 = 1 - 2x^2$.
When $n = 4$, $y_1 = 1 - 4x^2 + 4/3x^4$. and so on.
If $n = 1$, $y_2 = x$.
If $n = 3$, $y_2 = x - 2/3x^3$ and so on.

Thus, when n is a positive integer, one solution of Hermite's equation will be a polynomial and the other solution will be an infinite power series. We try to find the form of the polynomial solution which is as follows.

$$\begin{aligned}
y &= a_n x^n + a_{n-2} x^{n-2} + \cdots + a_1 x, \quad \text{when } n \text{ is odd} \\
&= a_n x^n + a_{n-2} x^{n-2} + \cdots + a_0, \quad \text{when } n \text{ is even}
\end{aligned} \tag{11.2.3}$$

Here, we are going to have series solution of Hermite's equation in decreasing powers of x. While solving Hermite's equation by Frobenius method, we got the recurrence relation when $s = 0$, as

$$a_m = 2\frac{m-n-2}{m(m-1)}a_{m-2}, \quad m \geq 2.$$

Here, we would express all the coefficients in terms of a_n instead of a_1 or a_0. We have from the previous equation,

$$a_{m-2} = \frac{m(m-1)}{2(m-n-2)}a_m.$$

Replacing m by $m+2$, we get

$$a_m = \frac{(m+1)(m+2)}{2(m-n)}a_{m+2}. \tag{11.2.4}$$

Put $m = n-2$. Then we get

$$a_{n-2} = (-1)\frac{n(n-1)}{2.2}a_n.$$

and putting $m = n-4$, we get

$$a_{n-4} = (-1)^2\frac{n(n-1)(n-2)(n-3)}{2.2.2.4}a_n.$$

Putting these in (11.2.3), we get,

$$y = a_n\left[x^n - \frac{n(n-1)}{2.2}x^{n-2} + \cdots + (-1)^r\frac{n(n-1)\ldots(n-2r+1)}{2^r.2.4\ldots 2r}x^{n-2r} + \cdots\right]$$

When n is even, $n-2r \geq 0$ which gives $r \leq n/2$. And when n is odd, $n-2r \geq 1$ which gives $r \leq (n-1)/2$. Thus, r vanishes from 0 to $n/2$ or $(n-1)/2$ according as n is even or odd, which implies that r varies from 0 to $[n/2]$. Hence, the general form of the polynomial solution to (11.2.1) is

$$y = a_n\sum_{r=0}^{[\frac{n}{2}]}(-1)^r\frac{n(n-1)\ldots(n-2r+1)}{2^{2r}.r!}x^{n-2r}.$$

Taking $a_n = 2^n$, and denoting the solution by $H_n(x)$, we obtain a standard solution to (11.2.1) known as the **Hermite's polynomial** of order n.

Definition 11.2.1. Hermite's polynomial $H_n(x)$ of order n is defined by

$$H_n(x) = \sum_{r=0}^{[\frac{n}{2}]}(-1)^r\frac{n!}{r!(n-2r)!}(2x)^{n-2r}.$$

The first few Hermite polynomials are

1. $H_0(x) = 1$
2. $H_1(x) = 2x$
3. $H_2(x) = 4x^2 - 2$
4. $H_3(x) = 8x^3 - 12x$
5. $H_4(x) = 16x^4 - 48x^2 + 12$
6. $H_5(x) = 32x^5 - 160x^3 + 120x$
7. $H_6(x) = 64x^6 - 480x^4 + 720x^2 - 120$
8. $H_7(x) = 128x^7 - 1344x^5 + 3360x^3 - 1680x.$

11.3 Generating Function for Hermite's Polynomials

The Hermite's polynomials $H_n(x)$ may also be defined by the power series expansion:

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x). \tag{11.3.1}$$

Each term in the expansion of e^{2tx-t^2} gives a Hermite's polynomial. In fact, coefficient of $t^n/n!$ gives a Hermite's polynomial of order n. Thus e^{2tx-t^2} is called the **generating function for Hermite's polynomial**. This definition has the advantage of efficiency for deducing properties of $H_n(x)$'s.

We have,

$$\begin{aligned} e^{2tx-t^2} &= e^{2tx}.e^{-t^2} \\ &= \sum_{s=0}^{\infty} \frac{(2tx)^s}{s!} \sum_{r=0}^{\infty} \frac{(-t^2)^r}{r!} \\ &= \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} (-1)^r \frac{t^{s+2r}}{r!s!} (2x)^s. \end{aligned}$$

Putting $s + 2r = n$ and eliminating s we get

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \frac{t^n}{r!(n-2r)!} (2x)^{n-2r}$$

since both r and s varies from 0 to ∞, so n varies from 0 to ∞ and $n - 2r \geq 0$ which gives $r \leq n/2$ and we arrive at the same conclusion as we had arrived in case of Legendre's polynomial. Thus, we have

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \left\{ \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \frac{n!}{r!(n-2r)!} (2x)^{n-2r} \right\} \frac{t^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x).$$

Show that

$$H_n(-x) = (-1)^n H_n(x).$$

Solution. From (11.2.1) we get

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x).$$

Change x to $-x$ and t to $-t$ in (11.3.1), then we get

$$e^{2(-t)(-x)-(-t)^2} = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} H_n(-x).$$

Since

$$e^{2(-t)(-x)-(-t)^2} = e^{2tx-t^2},$$

we get

$$\sum_{n=0}^{\infty} \frac{(-t)^n}{n!} H_n(-x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x)$$

i.e.,

$$\sum_{n=0}^{\infty} (-1)^n \frac{t^n}{n!} H_n(-x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x).$$

Equating the coefficient of t^n from both sides we get the result. ■

11.4 Rodrigue's Formula for Hermite's Polynomials

Hermite's polynomials satisfy the following formula which is known as the **Rodrigue's formula for Hermite's polynomials**

$$H_n(x) = (-1)^n e^{-x^2} \frac{d^n}{dx^n}(e^{-x^2}).$$

The function e^{2tx-t^2} is analytic in any neighbourhood of the point $t = 0$. Thus, for any fixed value of x it has a Taylor series expansion of the form

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[\frac{\partial^n}{\partial t^n} \left(e^{2tx-t^2} \right) \right]_{t=0}. \tag{11.4.1}$$

Now, we have

$$\frac{\partial^n}{\partial t^n} \left(e^{2tx-t^2} \right) = e^{x^2} \frac{\partial^n}{\partial t^n} \left(e^{-(x-t)^2} \right).$$

On calculation, we get

$$\frac{\partial}{\partial t} \left(e^{-(x-t)^2} \right) = -2(x-t)e^{-(x-t)^2} = -\frac{\partial}{\partial x} \left(e^{-(x-t)^2} \right).$$

By repeated use of this, we get

$$\frac{\partial^n}{\partial t^n} \left(e^{-(x-t)^2} \right) = (-1)^n \frac{\partial^n}{\partial x^n} \left(e^{-(x-t)^2} \right)$$

Thus, we get

$$\left[\frac{\partial^n}{\partial t^n} \left(e^{2tx-t^2} \right) \right]_{t=0} = e^{x^2} (-1)^n \frac{d^n}{dx^n} \left(e^{-x^2} \right).$$

Using this result in (11.4.1), we get

$$e^{2tx-t^2} = e^{x^2} \sum_{n=0}^{\infty} (-1)^n \frac{d^n}{dx^n} \left(e^{-x^2} \right) \frac{t^n}{n!}.$$

From the formula of generating function of Hermite's polynomial and equating the coefficients of $t^n/n!$ on both sides, we get the required result.

11.5 Recurrence Relations of Hermite's Polynomials

The Hermite's polynomials $H_n(x)$ satisfy some recurrence relations.

1. $H_n^{'}(x) = 2nH_{n-1}(x)$, for $n \geq 1$ and $H_0^{'}(0) = 0$.

   Proof. We have

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{n!}.$$

Differentiating both sides with respect to x, we have

$$\begin{aligned}\sum_{n=0}^{\infty} H_n'(x)\frac{t^n}{n!} = 2t.e^{2tx-t^2} &= 2t\sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty} 2H_n(x)\frac{t^{n+1}}{n!} \\ &= \sum_{n=1}^{\infty} 2nH_{n-1}(x)\frac{t^n}{n!}\end{aligned}$$

Equating the power of t^0 on both sides, we get $H_0'(0) = 0$ and equating the coefficients of $t^n/n!$ on both sides for $n \geq 1$, we get the desired result. □

2. $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$, for $n \geq 1$ and $H_1(x) = 2xH_0(x)$.

   Proof. We have
$$\sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!} = e^{2tx-t^2}.$$
   Differentiating both sides with respect to t, we get

$$\begin{aligned}2(x-t).e^{2tx-t^2} &= \sum_{n=0}^{\infty} nH_n(x)\frac{t^{n-1}}{n!} \\ \text{or,}\ \ 2(x-t)\sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!} &= \sum_{n=1}^{\infty} H_n(x)\frac{t^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} H_{n+1}(x)\frac{t^n}{n!} \\ \text{or,}\ \ 2x\sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!} &= 2\sum_{n=0}^{\infty} H_n(x)\frac{t^{n+1}}{n!} + \sum_{n=0}^{\infty} H_{n+1}(x)\frac{t^n}{n!}\end{aligned}$$

   Equating the coefficients of t^0 and $t^n/n!$ on both sides, we get the desired results. □

3. $H_n'(x) = 2xH_n(x) - H_{n+1}(x)$.

   Proof. Left as an exercise. □

4. $H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0$.

   Proof. Left as an exercise. □

11.6 Orthogonality Properties of Hermite's Polynomials

The Hermite's polynomials are orthogonal in the interval $(-\infty, \infty)$ i.e.,

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-x^2} H_m(x)H_n(x)dx &= 0,\ \ m \neq n \\ &= \sqrt{\pi}2^n.n!,\ \ m = n\end{aligned}$$

We have

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{n!}.$$

Replacing n by m and t by s and multiplying the resulting equation with the above equation we get,

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{H_n(x)H_m(x)}{n!.m!} t^n s^m = e^{2tx-t^2=2sx-s^2}.$$

Multiplying both sides by e^{-x^2} and integrating with respect to x from $-\infty$ to ∞, we get

$$\begin{aligned}
\int_{-\infty}^{\infty} \left(\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{e^{-x^2} H_n(x) H_m(x)}{n!.m!} \right) dx.t^n s^m &= \int_{-\infty}^{\infty} e^{-x^2+2x(t+s)-(t^2+s^2)} \\
\text{or, } \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left(\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x) dx \right) \frac{t^n.s^m}{n!.m!} &= e^{2ts} \int_{-\infty}^{\infty} e^{-(x-(t+s))^2} dx \qquad (11.6.1)
\end{aligned}$$

Putting $x - (t + s) = y$, we get, $dx = dy$. Thus,

$$\begin{aligned}
e^{2ts} \int_{-\infty}^{\infty} e^{-y^2} &= e^{2ts} \sqrt{\pi}, \quad \text{using gamma integral} \\
&= \sum_{n=0}^{\infty} \frac{2^n.t^n.s^n}{n!} \sqrt{\pi}.
\end{aligned}$$

Thus, (11.6.1) becomes

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left(\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x) dx \right) \frac{t^n.s^m}{n!.m!} = \sqrt{\pi} \sum_{n=0}^{\infty} \frac{2^n.t^n.s^n}{n!} \qquad (11.6.2)$$

We note that the powers of t and s are always equal in each term of the RHS of (11.6.2). So when $m \neq n$, equating the coefficients of $t^n s^m$ on both sides of (11.6.2), we have

$$\int_{-\infty}^{\infty} \frac{e^{-x^2} H_n(x) H_m(x)}{n!m!} dx = 0,$$

and when $m = n$, we have

$$\int_{-\infty}^{\infty} \frac{e^{-x^2} H_n(x) H_n(x)}{n!} dx = 2^n.n!\sqrt{\pi}.$$

Hence, we are done.

Exercise 11.6.1.

1. Convert the Hermite polynomial $2H_4(x) + 3H_3(x) - H_2(x) + 5H_1(x) + 6H_0(x)$ into an ordinary polynomial.
2. Convert the ordinary polynomial $64x^4 + 8x^3 - 32x^2 + 40x + 10$ into a Hermite polynomial.
3. Verify the fact that e^{2tx-t^2} is indeed the generating function for $H_n(x)$ by expanding the exponential function and showing that the coefficients of the individual terms $t^n/n!$ are indeed the Hermite's polynomials.
4. Verify Rodrigue's formula for first three Hermite's polynomials.

5. Find Hermite's polynomials upto order 6 by using Rodrigue's formula.
6. Prove that $H_{2n}(0) = (-1)^n \frac{(2n)!}{n!}$ and $H_{2n+1}(0) = 0$.
7. Prove that
$$\int_{-\infty}^{\infty} x^2 e^{-x^2} H_n(x) H_n(x) dx = \sqrt{\pi} 2^n . n! \left(n + \frac{1}{2} \right).$$

Few Probable Questions

1. Solve the Hermite's differential equation and deduce the structure of the Hermite's polynomial.
2. Establish the Rodrigue's polynomial for Hermite's polynomial.
3. Show that e^{2xt-t^2} is the generating function for Hermite's polynomial.
4. Establish the orthogonality of Hermite's polynomials.
5. Show that e^{2tx-t^2} is the generating function for the Hermite's polynomial. Hence show that for $m < n$,
$$\frac{d^m}{dx^m} H_m(x) = \frac{2^m n!}{(n-m)!} H_{n-m}(x).$$

Unit 12

Course Structure

- Laguerre Polynomial: Generating function, Recurrence relations, Rodrigue's formula, Orthogonal property. Construction and solution of Laguerre differential equation.

12.1 Laguerre's Equation and Laguerre's Polynomials

The Laguerre's equation is of the form

$$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + ny = 0 \tag{12.1.1}$$

where n is a constant. The point $x = 0$ is a regular singular point of the equation (12.1.1). We solve (12.1.1) by Frobenius Method, about $x = 0$. Assume that

$$y = \sum_{m=0}^{\infty} a_m x^{s+m}$$

be the solution of (12.1.1), where $a_0 \neq 0$ and s is to be determined. The indicial equation in this case will be

$$s^2 = 0,$$

which has repeated roots $s = 0, 0$.

Using this value of s, we get

$$a_{m+1} = -\frac{n-m}{(m+1)^2}a_m$$

for $m = 0, 1, 2, \ldots$. So all the coefficients are found in terms of a_0, and one of the solution of (12.1.1) will be

$$y(x) = a_0\left[1 - \frac{n}{1^2}x + \frac{n(n-1)}{1^2.2^2}x^2 - \ldots + (-1)^m\frac{n(n-1)\ldots(n-m+1)}{1^2.2^2\ldots m^2}x^m + \ldots\right]. \tag{12.1.2}$$

If n is not a positive integer or zero (12.1.2) is an infinite series. But if n is a positive integer or zero (12.1.2) will terminate to a finite series and becomes a ploynomial of degree n. Further if we choose $a_0 = 1$, the resultant expression defines the **Laguerre's polynomial**

$$\begin{aligned} L_n(x) &= 1 - \frac{n}{1^2}x + \frac{n(n-1)}{1^2.2^2}x^2 - ... + (-1)^n \frac{n(n-1)...1}{1^2.2^2...n^2}x^n \\ &= 1 - \frac{n}{1!.1!}x + \frac{n(n-1)}{2!.2!}x^2 - ... + (-1)^m \frac{n(n-1)...1}{m!.m!}x^n \\ &= \binom{n}{0}x^0 - \binom{n}{1}\frac{x^1}{1!} + \binom{n}{2}\frac{x^2}{2!} + ... + (-1)^n \binom{n}{n}\frac{x^n}{n!} \end{aligned}$$

i.e.,

$$L_n(x) = \sum_{r=0}^{n} \frac{(-1)^r}{r!}\binom{n}{r}x^r. \tag{12.1.3}$$

The first few Laguerre's polynomials are

1. $L_0(x) = 1$
2. $L_1(x) = 1 - x$
3. $L_2(x) = 1 - 2x + \frac{x^2}{2}$
4. $L_3(x) = 1 - 3x + \frac{3x^2}{2} - \frac{x^3}{6}$.

12.2 Generating Function for Laguerre's Polynomials

The generating function for the Laguerre's polynomials is

$$g(x,t) = \frac{e^{-\frac{xt}{1-t}}}{1-t}$$

i.e., we can show that

$$\frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^n L_n(x),$$

for $|t| < 1$.

For

$$\begin{aligned} \frac{e^{-\frac{xt}{1-t}}}{1-t} = \frac{1}{1-t}\sum_{r=0}^{\infty}\left(\frac{-xt}{1-t}\right)^r \frac{1}{r!} &= \sum_{r=0}^{\infty}\frac{(-1)^r}{r!}x^r t^r (1-t)^{-r-1} \\ &= \sum_{r=0}^{\infty}\frac{(-1)^r}{r!}x^r t^r \sum_{s=0}^{\infty}\frac{(r+s)!}{r!s!}t^s \\ &= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}(-1)^r\frac{(r+s)!}{(r!)^2 s!}x^r t^{r+s}. \end{aligned}$$

We put $r + s = n$, that is, $s = n - r$, where r is fixed.
Now, $s \geq 0$ implies $r \leq n$, giving all possible values of r.
Therefore the coefficients of t^n is given by

$$\begin{aligned} \sum_{r=0}^{n}(-1)^r\frac{n!}{(r!)^2(n-r)!}x^r &= \sum_{r=0}^{n}\frac{(-1)^r}{r!}\binom{n}{r}x^r \\ &= L_n(x). \end{aligned}$$

Hence the result follows.

12.3 Rodrigue's Formula for Laguerre's polynomials

The Rodrigue's representation for Laguerre's polynomials is

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x}).$$

We use the Leibnitz's theorem for successive differentiation

$$\begin{aligned} D(uv) &= \frac{d^n}{dx^n}(uv) \\ &= D^n u.v + \binom{n}{1} D^{n-1}u.Dv + \cdots + \binom{n}{r} D^{n-r}u.D^r v + \cdots + uD^n v. \end{aligned}$$

Then we get

$$\frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x}) = \frac{e^x}{n!}\sum_{r=0}^{n}\binom{n}{r} D^{n-r}x^n D^r e^{-x}.$$

Since $D^n x^m = \frac{m!}{(m-n)!}x^{m-n}$ and $D^n e^{ax} = a^n e^{ax}$, we get

$$\begin{aligned} \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x}) &= \frac{e^x}{n!}\sum_{r=0}^{n}\binom{n}{r}\frac{n!}{(n-(n-r))!}x^{n-(n-r)}(-1)^r e^{-x} \\ &= \sum_{r=0}^{n}\frac{e^x}{n!}\frac{n!}{r!(n-r)!}.\frac{n!}{r!}.x^r.(-1)^r e^{-x} \\ &= L_n(x). \end{aligned}$$

12.4 Recurrence Relations of Laguerre's Polynomials

The Laguerre's polynomials satisfy some recurrence relations.

$i)$ $(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$.

We have, $g(x,t) = \frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^n L_n(x)$. Differentiating both sides with respect to t, we get

$$\begin{aligned} \sum_{n=0}^{\infty} t^{n-1} n L_n(x) &= \frac{1}{(1-t)^2}e^{-\frac{xt}{1-t}} - \frac{1}{(1-t)}e^{-\frac{xt}{1-t}}.\frac{x}{(1-t)^2} \\ &= \frac{1}{1-t}\sum_{n=0}^{\infty} t^n L_n(x) - \frac{x}{(1-t)^2}\sum_{n=0}^{\infty} t^n L_n(x). \end{aligned}$$

Multiplying both sides by $(1-t^2)$ and simplifying, we obtain

$$\sum_{n=0}^{\infty} t^{n-1} nL_n(x) - 2\sum_{n=0}^{\infty} t^n nL_n(x) + \sum_{n=0}^{\infty} t^{n+1} nL_n(x) = \sum_{n=0}^{\infty} t^n L_n(x) - \sum_{n=0}^{\infty} t^{n+1} L_n(x) - x\sum_{n=0}^{\infty} t^n L_n(x).$$

Equating the coefficients of t^n on both sides, we get the desired result.

$ii)$ $xL_n^{'}(x) = nL_n(x) - nL_{n-1}(x)$.

We have,

$$g(x,t) = \frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^n L_n(x).$$

Differentiating both sides with respect to x, and using the generating function, we get

$$\sum_{n=0}^{\infty} t^n L_n^{'}(x) = \frac{1}{1-t} e^{-\frac{xt}{1-t}} . \frac{-t}{1-t} = \frac{-t}{1-t} \sum_{n=0}^{\infty} t^n L_n(x).$$

Multiplying both sides by $(1-t)$ and simplifying, we get

$$\sum_{n=0}^{\infty} t^n L_n^{'}(x) - \sum_{n=0}^{\infty} t^{n+1} L_n^{'}(x) = \sum_{n=0}^{\infty} t^{n+1} L_n(x).$$

Equating the coefficients of t^n on both sides, we get the desired result.
$iii)$ $L_n^{'}(x) = -\sum_{r=0}^{n-1} L_r(x)$.
We have,

$$g(x,t) = \frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^n L_n(x).$$

Differentiating both sides with respect to x, and using the generating function, we get

$$\sum_{n=0}^{\infty} t^n L_n^{'}(x) = \frac{1}{1-t} e^{-\frac{xt}{1-t}} . \frac{-t}{1-t} = \frac{-t}{1-t} \sum_{r=0}^{\infty} t^r L_r(x) = -t \sum_{s=0}^{\infty} t^s \sum_{r=0}^{\infty} t^r L_r(x).$$

Thus,

$$\sum_{n=0}^{\infty} t^n L_n^{'}(x) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} t^{r+s+1} L_r(x).$$

The coefficients of t^n on the LHS is clearly $L_n^{'}(x)$. We will find the coefficients of t^n on the RHS.
Let $r+s+1 = n$, so that $s = n-r-1$.
Hence, for a fixed value of r, the coefficient of t^n on the RHS of the above equation is $-L_r(x)$.
But, $s \geq 0$, which implies that $n-r-1 \geq 0 \implies r \leq n-1$, which gives all the values of r for which $-L_r(x)$ is the coefficient of t^n.
Hence the total coefficients of t^n on the RHS is given by $-\sum_{r=0}^{n-1} L_r(x)$ and equating the coefficients on both sides, we get the desired result.

12.5 Orthogonality Properties of Laguerre's Polynomials

If $L_m(x)$ and $L_n(x)$ are Laguerre's polynomials (m, n being positive integers), then

$$\begin{aligned} \int_0^{\infty} e^{-x} L_n(x) L_m(x) dx &= 0, \ m \neq n \\ &= 1, \ m = n. \end{aligned}$$

The generating function for Laguerre's polynomial gives

$$\begin{aligned} \frac{e^{-\frac{xt}{1-t}}}{1-t} &= \sum_{n=0}^{\infty} t^n L_n(x) \\ \&, \quad \frac{e^{-\frac{xs}{1-s}}}{1-s} &= \sum_{m=0}^{\infty} s^m L_m(x). \end{aligned}$$

Multiplying both the equations we get

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} s^m.t^n.L_m(x)L_n(x) = \frac{e^{-\frac{xs}{1-s}-\frac{xt}{1-t}}}{(1-s)(1-t)}$$

Multiplying both sides by e^{-x} and integrating with respect to x from 0 to ∞, we get

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\left\{\int_0^{\infty} e^{-x}L_m(x)L_n(x)dx\right\} s^m.t^n &= \frac{1}{(1-t)(1-s)}\int_0^{\infty} e^{-x(1+t/(1-t)+s(1-s)}dx \\ &= \frac{1}{(1-t)(1-s)}\left|\frac{e^{-x(1+t/(1-t)+s(1-s)}}{-(1+t/(1-t)+s(1-s)}\right|_0^{\infty} \\ &= \frac{1}{1-st}.\end{aligned}$$

Now, we have

$$(1-st)^{-1} = 1+st+s^2t^2+\cdots = \sum_{n=0}^{\infty} s^nt^n.$$

Using this in the previous equation, we get

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\left\{\int_0^{\infty} e^{-x}L_m(x)L_n(x)dx\right\} s^m.t^n = \sum_{n=0}^{\infty} s^nt^n.$$

Equating the coefficients of t^ns^m on both sides, we get the desired result.

Prove that $i)$ $L_n\left(0\right)=1$, $ii)$ $L_n'\left(0\right)=-n$, $iii)$ $L_n''\left(0\right)=\frac{n(n-1)}{2}$.

Solution. $i)$ Put $x=0$ in

$$\frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^nL_n(x).$$

Then

$$\frac{1}{1-t} = \sum_{n=0}^{\infty} t^nL_n(x)$$

i.e.,

$$\sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} t^nL_n(x).$$

Equating the coefficient of t^n from both sides we get the result.
$ii)$ Since $L_n(x)$ satisfies the Laguerre's equation

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

we have

$$x\frac{d^2L_n(x)}{dx^2}+(1-x)\frac{dL_n(x)}{dx}+nL_n(x)=0. \tag{12.5.1}$$

Put $x=0$ in (12.5.1). Then we get

$$\begin{aligned}L_n'(0)+nL_n(0) &= 0\\ L_n'(0)+n &= 0\end{aligned}$$

Therefore

$$L_n'(0) = -n.$$

$iii)$ The generating for Laguerre's polynomials is

$$\frac{e^{-\frac{xt}{1-t}}}{1-t} = \sum_{n=0}^{\infty} t^n L_n(x).$$

Differentiating it twice with respect to x,we get

$$\frac{e^{-\frac{xt}{1-t}}}{1-t}\left(\frac{-t}{1-t}\right)^2 = \sum_{n=0}^{\infty} t^n L_n''(x). \tag{12.5.2}$$

Put $x = 0$ in (12.5.2). Then we get

$$\begin{aligned}
\sum_{n=0}^{\infty} t^n L_n''(0) &= \frac{t^2}{(1-t)^3} \\
\sum_{n=0}^{\infty} t^n L_n''(0) &= t^2 (1-t)^{-3} \\
\sum_{n=0}^{\infty} t^n L_n''(0) &= t^2 \left\{1 + 3t + \frac{3.4}{2!}t^2 + \frac{3.4.5}{3!}t^3 + ... + \frac{3.4.5...n}{(n-2)!}t^{n-2} + ...\right\} \\
\sum_{n=0}^{\infty} t^n L_n''(0) &= t^2 + 3t^3 + \frac{3.4}{2!}t^4 + \frac{3.4.5}{3!}t^5 + ... + \frac{3.4.5...n}{(n-2)!}t^n + ...
\end{aligned}$$

Equating the coefficient of t^n from both sides we get

$$\begin{aligned}
\frac{3.4.5...n}{(n-2)!} &= L_n''(0) \\
\frac{n!}{2.(n-2)!} &= L_n''(0).
\end{aligned}$$

Therefore

$$L_n''(0) = \frac{n(n-1)}{2}.$$

■

Exercise 12.5.1. 1. Compute the first few Laguerre polynomials using the summation formula.

2. Prove that $\int_0^{\infty} e^{-st} L_n(t)dt = 1/s(1-1/s)^n$.

3. Verify the Rodrigue's formula for first four positive integers.

4. Prove that $\int_x^{\infty} e^{-y} L_n(y)dy = e^{-x}[L_n(x) - L_{n-1}(x)]$.

5. Prove that

   (a) $L_n'(x) = n[L_{n-1}'(x) - L_{n-1}(x)]$.

   (b) $xL_n(x) = nL_n(x) - n^2 L_{n-1}(x)$.

Few Probable Questions

1. Establish the Rodrige's polynomial for Laguerre polynomial.
2. Show that $e^{-xt/(1-t)}/(1-t)$ is the generating function for Laguerre polynomial.
3. Establish the orthogonality of Laguerre polynomials.

Unit 13

Course Structure

- Chebyshev Polynomial: Definition, Series representation, Recurrence relations, Orthogonal property. Construction and solution of Chebyshev differential equation.

13.1 Introduction

In mathematics the Chebyshev polynomials, named after Pafnuty Chebyshev, are a sequence of orthogonal polynomials which are related to de Moivre's formula and which can be defined recursively. Chebyshev polynomials are important in approximation theory because the roots of the Chebyshev polynomials of the first kind, which are also called Chebyshev nodes, are used as nodes in polynomial interpolation. The resulting interpolation polynomial minimizes the problem of Runge's phenomenon and provides an approximation that is close to the polynomial of best approximation to a continuous function under the maximum norm. This approximation leads directly to the method of Clenshaw–Curtis quadrature. We will study about the Chebyshev polynomials and its properties in this unit.

Objectives

After reading this section, you will be able to

- know the Chebyshev's equations
- define Chebyshev's polynomials
- learn the Rodrigue's formula for Chebyshev's polynomials
- deduce a generating function for Chebyshev's polynomials
- learn the orthogonality condition for Chebyshev's polynomials
- learn the recurrence relations concerning Chebyshev polynomials
- solve various problems related to the above topics

13.2 Chebyshev's Equation and Chebyshev's Polynomials

The Chebyshev differential equation is written as

$$(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + n^2y = 0, \tag{13.2.1}$$

where $|x| < 1$ and n is any real number. This equation can be converted to a simpler form using the substitution $x = \cos t$. Then we have

$$dx = -\sin t dt \implies \frac{dt}{dx} = -\frac{1}{\sin t}.$$

Hence,

$$\frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = -\frac{1}{\sin t}\frac{dy}{dt},$$

and

$$\begin{aligned}\frac{d^2y}{dx^2} &= \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dt}\frac{dt}{dx}\left(-\frac{1}{\sin t}\frac{dy}{dt}\right)\\ &= -\frac{1}{\sin t}\frac{d}{dt}\left(-\frac{1}{\sin t}\frac{dy}{dt}\right) = \frac{1}{\sin^2 t}\left[\left(-\frac{\cos t}{\sin t}\right)\frac{dy}{dt} + \frac{d^2y}{dt^2}\right].\end{aligned}$$

Substituting these in (13.2.1), and simplifying, we get

$$\frac{d^2y}{dt^2} + n^2y = 0,$$

whose general solution is given by

$$y(t) = C\cos(nt + a).$$

For simplicity, we set $a = 0$. Thus, the general solution of the equation (13.2.1) is given by

$$y(x) = C\cos(n \arccos x).$$

Now, if n is an integer, then the above function is the Chebyshev polynomial of first kind.

Definition 13.2.1. The Chebyshev polynomial of the first kind is called the function

$$T_n(x) = \cos(n \arccos x) = \frac{n}{2}\sum_{k=0}^{[n/2]}(-1)^k\frac{(n-k-1)!}{k!(n-2k)!}(2x)^{n-2k},$$

where $|x| < 1$ and $n = 0, 1, 2, \ldots$

$T_n(x)$ given by above is an nth degree polynomial in x. Here $T_n(x)$ is a solution of D.E (13.2.1) is known as the Chebyshev polynomial of the first kind.

In the expanded form Chebyshev polynomial is given by

$$T_n(x) = 2^{n-1}\left[x^n - \frac{n}{1!2^2}x^{n-2} + \frac{n\,(n-3)}{2!2^4}x^{n-4} - \frac{n\,(n-4)\,(n-5)}{3!2^6}x^{n-6} + \ldots\right].$$

The first few Chebyshev's p[olynomials are

1. $T_0(x) = 1$
2. $T_1(x) = x$

3. $T_2(x) = 2x^2 - 1$

4. $T_3(x) = 4x^3 - 3x$

5. $T_4(x) = 8x^4 - 8x^2 + 1.$

Show that $T_n(1) = 1$.

Proof. By definition we have

$$T_n(x) = \cos\left(n \cos^{-1} x\right).$$

Put $x = 1$,

$$T_n(1) = \cos\left(n \cos^{-1} 1\right) = \cos(n.0) = 1.$$

□

13.3 Generating Functions for Chebyshev's Polynomials

The function

$$w(x,t) = \frac{2 - xt}{1 - xt + t^2},$$

is the generating function for Chebyshev polynomials, that is,

$$\frac{2 - xt}{1 - xt + t^2} = \sum_{n=0}^{\infty} T_n(x) t^n.$$

We have,

$$\begin{aligned}
\frac{1}{1 - xt + t^2} &= \sum_{k=0}^{\infty} \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{k-l} x^l t^{2k-l} \\
&= \sum_{k,l} (-1)^{k-l} \binom{k}{k-l} x^l t^{2k-l}.
\end{aligned}$$

Put $m = 2k - l$ and $n = k - l$. Then $k = m - n$ and $l = m - 2n$. Then the above equation changes to

$$\sum_{m,n} (-1)^n \binom{m-n}{n} x^{m-2n} t^m.$$

Now,

$$\begin{aligned}
&(2-xt)\sum_{m,n}(-1)^n\binom{m-n}{n}x^{m-2n}t^m \\
= \;& 2\sum_{m,n}(-1)^n\binom{m-n}{n}x^{m-2n}t^m \\
& +\sum_{m,n}(-1)^{n+1}\binom{m-n}{n}x^{m-2n+1}t^{m+1} \\
= \;& 2\sum_{m,n}(-1)^n\binom{m-n}{n}x^{m-2n}t^m \\
& +\sum_{m,n}(-1)^n\binom{m-n-1}{n}x^{m-2n}t^m \\
= \;& \sum_{m,n}(-1)^n\left\{2\binom{m-n}{n}-\binom{m-n-1}{n}\right\}x^{m-2n}t^m \\
= \;& \sum_{m,n}(-1)^n\frac{m}{m-n}\binom{m-n}{n}x^{m-2n}t^m=\sum_m T_m(x)t^m.
\end{aligned}$$

Show that

$$\frac{1-t^2}{1-2xt+t^2}=T_0\left(x\right)+2\sum_{n=1}^{\infty}T_n(x)t^n$$

Hint: Put $x=\cos\theta=\frac{1}{2}\left(e^{i\theta}+e^{-i\theta}\right)$ in LHS we obtain

$$\sum_{k,l=0}^{\infty}t^{k+l}e^{i(k-l)\theta}-\sum_{k,l=0}^{\infty}t^{k+l+2}e^{i(k-l)\theta}.$$

Now, Coefficient of $t^0,(k=0,l=0):T_0\left(x\right).$
Coefficient of $t',(k=0,l=1):2x=2T_1\left(x\right),$
Coefficient of $t^n,(l=n-k):2\cos n\theta=2T_n\left(x\right).$

13.4 Rodrigue's Formula for Chebyshev's Polynomials

The Chebyshev's polynomial satisfy the following Rodrigue's formula

$$T_n(x)=\frac{(-2)^n n!}{(2n)!}\sqrt{1-x^2}\,\frac{d^n}{dx^n}(1-x^2)^{n-1/2}.$$

13.5 Recurrence Relations of Chebyshev's Polynomials

The Chebyshev's polynomial follows the following recurrence relations
$i)$ $2xT_n(x)=T_{n+1}(x)+T_{n-1}(x).$

1. *Proof.* Putting $x=\cos t$, we have

$$\begin{aligned}
T_{n-1}(t) &= \cos((n-1)\arccos x)=\cos((n-1)t) \\
T_{n+1}(t) &= \cos((n+1)\arccos x)=\cos((n+1)t).
\end{aligned}$$

Also, we have,

$$\begin{aligned} T_1(x) &= \cos(\arccos x) = \cos(\arccos(\cos t)) = \cos t = x \\ T_n(x) &= \cos(n \arccos x) = \cos(n \arccos(\cos t)) = \cos(nt). \end{aligned}$$

Further,

$$\begin{aligned} \cos((n-1)t) + \cos((n+1)t) &= 2\cos\frac{(n-1)t + (n+1)t}{2} \\ &= 2\cos\frac{2nt}{2}\cos\frac{-2t}{2} = 2\cos(nt)\cos t. \end{aligned}$$

Thus, we get

$$T_{n-1}(x) + T_{n+1} = 2T_n(x)T_1(x) = 2xT_n(x).$$

□

ii) $(1-x^2)T_n^{'}(x) = -nxT_n(x) + nT_{n-1}(x)$.

Proof. Differentiating bothside of $T_n(x) = \cos(n\theta)$ w.r.t. x we obtain

$$\begin{aligned} \frac{d}{dx}T_n(x) &= \frac{d}{dx}\cos(n\theta) = \frac{d}{dx}\cos\left(n\cos^{-1}x\right) \\ &= -\sin\left(n\cos^{-1}x\right).n.\left(-\frac{1}{\sqrt{1-x^2}}\right) \end{aligned}$$

i.e; $\sqrt{1-x^2}T_n^{'}(x) = n\sin(n\theta)$. □

Multiplying by $\sqrt{1-x^2} = \sin\theta$ we get

1. *Proof.*

$$\begin{aligned} (1-x^2)T_n^{'}(x) &= n.\sin\theta.\sin(n\theta) \\ &= \frac{1}{2}n.\left[\cos(n-1)\theta - \cos(n+1)\theta\right] \\ &= \frac{n}{2}\left[T_{n-1}(x) - T_{n+1}(x)\right]. \end{aligned}$$

Eliminate $T_{n+1}(x)$ using $2xT_n(x) = T_{n+1}(x) + T_{n-1}(x)$ we obtain

$$\begin{aligned} (1-x^2)T_n^{'}(x) &= \frac{n}{2}\left[T_{n-1}(x) - 2xT_n(x) + T_{n-1}(x)\right] \\ &= n\left[T_{n-1}(x) - xT_n(x)\right]. \end{aligned}$$

□

iii) For $-1 < x < 1$, we have $T_n^2(x) - T_{n-1}(x)T_{n+1}(x) = 1 - x^2$.

Proof. The proof is left to the reader. □

iv) $T_n(x) - 2xT_{n-1}(x) + T_{n-2}(x) = 0$.

Proof. We have

$$\begin{aligned} T_n(x) + T_{n-2}(x) &= \cos n\theta + \cos(n-2)\theta \\ &= \cos n\theta + \cos n\theta . \cos 2\theta + \sin n\theta . \sin 2\theta \\ &= \cos n\theta\,(1+\cos 2\theta) + 2\sin n\theta . \sin\theta\cos\theta \\ &= 2\cos\theta\,[\cos n\theta . \cos\theta + \sin n\theta . \sin\theta] \\ &= 2\cos\theta\cos(n-1)\theta = 2x.T_{n-1}(x). \end{aligned}$$

□

$v)$ $T_{m+n}(x) + T_{m-n}(x) = 2T_m(x).T_n(x)$

Proof. We have

$$\begin{aligned} T_{m+n}(x) + T_{m-n}(x) &= \cos(m+n)\theta + \cos(m-n)\theta \\ &= 2\cos m\theta\cos n\theta = 2T_m(x).T_n(x). \end{aligned}$$

□

$vi)$ $2\,[T_n(x)]^2 = 1 + T_{2n}(x)$

Proof. We have

$$\begin{aligned} 2\,[T_n(x)]^2 - T_{2n}(x) &= 2\cos^2 n\theta - \cos 2n\theta \\ &= 2\cos^2 n\theta - \cos^2 n\theta + \sin^2 n\theta = 1. \end{aligned}$$

□

13.6 Orthogonality Properties of Chebyshev's Polynomials

The Chebyshev's polynomials satisfy the following orthogonal property

$$\begin{aligned} \int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx &= 0, \;\; m \neq n \\ &= \pi/2, \;\; m = n \neq 0 \\ &= \pi, \;\; m = n = 0. \end{aligned}$$

Therefore, The Chebyshev's polynomials are orthogonal on $[-1, 1]$ with respect to the weight function $\frac{1}{\sqrt{1-x^2}}$.

To prove the orthogonality, consider

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx = \int_{-1}^{1} \frac{\cos(n\arccos x)\cos(m\arccos x)}{\sqrt{1-x^2}}dx.$$

Use the change of variable

$$\begin{aligned} x &= \cos\theta \\ dx &= -\sin\theta d\theta \\ &= -\sqrt{1-x^2}dx. \end{aligned}$$

Therefore

$$\begin{aligned}\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx &= -\int_{-\pi}^{0} \cos(n\theta)\cos(m\theta)\,d\theta \\ &= \int_{0}^{\pi} \cos(n\theta)\cos(m\theta)\,d\theta.\end{aligned}$$

Case 1: For $m \neq n$,

$$\begin{aligned}\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx &= \frac{1}{2}\int_{0}^{\pi} \{\cos(n+m)\theta + \cos(n-m)\theta\}\,d\theta \\ &= 0.\end{aligned}$$

Case 2: For $m = n \neq 0$,

$$\begin{aligned}\int_{-1}^{1} \frac{\{T_n(x)\}^2}{\sqrt{1-x^2}}dx &= \int_{0}^{\pi} \cos^2(n\theta)\,d\theta \\ &= \frac{1}{2}\int_{0}^{\pi} \{1+\cos 2n\theta\} \\ &= \pi/2.\end{aligned}$$

Case 3: For $m = n = 0$,

$$\begin{aligned}\int_{-1}^{1} \frac{\{T_n(x)\}^2}{\sqrt{1-x^2}}dx &= \int_{0}^{\pi} d\theta \\ &= \pi.\end{aligned}$$

Using the orthogonality properties of Chebyshev's polynomial an arbitrary function $f(x)$ can be expanded in **Chebyshev's series** as

$$f(x) = \sum_{i=0}^{\infty} a_i T_i(x),$$

where

$$a_0 = \frac{1}{\pi}\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}}dx$$

and

$$a_n = \frac{2}{\pi}\int_{-1}^{1} \frac{T_n(x)f(x)}{\sqrt{1-x^2}}dx$$

for $n > 1$.

13.7 Minimax Property

The Chebyshev's polynomials have the remarkable minimax property, which is stated without proof.

Statement: Among all polynomials $P(x)$ of degree $n > 0$ with leading coefficient 1, $2^{1-n}T_n(x)$ deviates least from zero in the interval $[-1, 1]$:

$$\max_{-1\leq x\leq 1} |P(x)| \geq \max_{-1\leq x\leq 1} \left|2^{1-n}T_n(x)\right| = 2^{1-n}.$$

Prove that

$$\int_{-1}^{1} \frac{x^m T_n(x)}{\sqrt{1-x^2}}dx = 0, \;\; m < n.$$

Solution: By Chebyshev's series we have

$$x^m = \sum_{i=0}^{m} a_i T_i(x).$$

When m is even, all the odd coefficients $a_1, a_3, ...$ will be zero and when m is odd, the even coefficients $a_0, a_2, a_4, ...$will be zero.

Now

$$\begin{aligned}
\int_{-1}^{1} \frac{x^m T_n(x)}{\sqrt{1-x^2}} dx &= \int_{-1}^{1} \frac{\left[\sum_{i=0}^{m} a_i T_i(x)\right] T_n(x)}{\sqrt{1-x^2}} dx \\
&= \sum_{i=0}^{m} a_i \int_{-1}^{1} \frac{T_i(x) T_n(x)}{\sqrt{1-x^2}} dx.
\end{aligned}$$

Since $i = 0, 1, 2, ..., m(< n)$, therefore using orthogonality of Chebyshev's polynomials the integral on the RHS vanishes for $i = 0, 1, 2, ..., m(< n)$.

Exercise 13.7.1.

1. Compute few Chebyshev's polynomials using the formula given above.
2. Compute few higher Chebyshev's polynomials using the first Recurrence relation and $T_1(x) = x$.
3. Prove that $T_n(-x) = (-1)^n T_n(x)$. Thus show that $T_n(-1) = (-1)^n$.
4. Show that
$$\begin{aligned}
T_n(0) &= (-1)^n, \text{ when } n \text{ is even} \\
&= 0, \text{ when } n \text{ is odd.}
\end{aligned}$$
5. Show that
$$[T_n(x)]^2 - T_{n+1}(x).T_{n-1}(x) = 1 - x^2.$$
Hint: LHS = $\cos^2 n\theta - \cos(n+1)\theta . \cos(n-1)\theta = \cos^2 n\theta - \left(\cos^2 n\theta - \sin^2\theta\right) = \sin^2\theta = 1 - x^2$ =RHS.
6. Verify Rodrigue's formula for first few Chebyshev's polynomials.

Few Probable Questions

1. Define Chebyshev polynomials. Solve Chebyshev's equation.
2. Deduce a generating function for Chebyshev polynomials.
3. Prove that $T_n(x)$ is a polynomial of nth degree in x.
4. Show that the leading coefficient of x^n in $T_n(x)$ is 2^{n-1}.
5. Show that Chebyshev polynomials are solutions of Chebyshev differential equation.

Hint:

$$y = T_n(x) = \cos\left(n\cos^{-1} x\right),$$

$$y' = \frac{n\sin\left(n\cos^{-1} x\right)}{\sqrt{1-x^2}}, y'' = \frac{nx}{(1-x^2)^{\frac{3}{2}}} \sin\left(n\cos^{-1} x\right),$$

so that $\left(1-x^2\right) y'' = xy' - n^2 y$.

Unit 14

Course Structure

- Integral Equation: Symmetric, separable, iterated and resolvent kernel, Fredholm and Voltera integral equation their classification, integral equation of convolution type, eigen value eigen function, method of converting an initial value problem (IVP) into a Voltera integral equation, method of converting a boundary value problem (BVP) into a Fredholm integral equation.

14.1 Introduction

Many physical problems of science and technology which were solved with the help of theory of ordinary and partial differential equations can be solved by better methods of theory of integral equations. For example, while searching for the representation formula for the solution of linear differential equation in such a manner so as to include boundary conditions or initial conditions explicitly, we arrive at an integral equation. The solution of the integral equation is much easier than the original boundary value or initial value problem. The theory of integral equations is very useful tool to deal with problems in applied mathematics, theoretical mechanics, and mathematical physics. Several situations of science lead to integral equations, e.g., neutron diffusion problem and radiation transfer problem etc.

14.2 Integral Equation

Definition 14.2.1. An integral equation is an equation is which an unknown function appears under one or more integral signs. For example, for $a \leq x \leq b,\ a \leq t \leq b$, the equations

$$\int_a^b K(x,t)y(t)dt = f(x) \tag{14.2.1}$$

$$y(x) - \lambda \int_a^b K(x,t)y(t)dt = f(x) \tag{14.2.2}$$

$$\text{and} \quad y(x) = \int_a^b K(x,t)[y(t)]^2 dt, \tag{14.2.3}$$

where the function $y(x)$, is the unknown function while the functions $f(x)$ and $K(x,t)$ are known functions and λ, a and b are constants, are all integral equations. The above mentioned functions may be complex-valued functions of the real variables x and t.

Definition 14.2.2. Linear and Non-linear Integral Equation : An integral equation is called *linear* if only linear operations are performed in it upon the unknown function. An integral equation which is not linear is known as a *non-linear integral equation*. By writing either

$$L(y) = \int_a^b K(x,t)\, y(t)\, dt \quad \text{or} \quad L(y) = y(x) - \lambda \int_a^b K(x,t) y(t) dt, \tag{14.2.4}$$

we can easily verify that L is a linear integral operator. In fact, for any constants c_1 and c_2, we have

$$L\{c_1 y_1(x) + c_2 y_2(x)\} = c_1 L\{y_1(x)\} + c_2 L\{y_2(x)\} \tag{14.2.5}$$

which is well known general criterion for a *linear operator*. In this block, we shall study only linear integral equations. The most general type of linear integral equation is of the form

$$g(x)y(x) = f(x) + \lambda \int_a K(x,t)y(t)dt, \tag{14.2.6}$$

where the upper limit may be either variable x or fixed. The functions f, g and K are known functions while y is to be determined; λ is a non-zero real or complex, *parameter*. The function $K(x,t)$ is known as the kernel of the integral equation.

Remark 14.2.3. The constant λ can be incorporated into the kernel $K(x,t)$ in Eq.(14.2.6). However, in many applications λ represents a significant parameter which may take on various values in a discussion being considered.

Remark 14.2.4. If $g(x) \neq 0$, Eq.(14.2.6) is known as linear integral equation of the third kind. When $g(x) = 0$, Eq.(14.2.6) reduces to

$$f(x) + \lambda \int_a K(x,t)y(t)dt, = 0 \tag{14.2.7}$$

which is known as *linear integral equation of the first kind*. Again, when $g(x) = 1$, Eq.(14.2.6) reduces to

$$y(x) = f(x) + \lambda \int_a K(x,t)y(t)dt. \tag{14.2.8}$$

which is known as linear integral equation of the second kind. In the present block, we shall study in details equations of the form (14.2.7) and (14.2.8) only.

Definition 14.2.5. Fredholm Integral Equation : A linear integral equation of the form

$$g(x)y(x) = f(x) + \lambda \int_a^b K(x,t)y(t)dt, \tag{14.2.9}$$

where a, b are both constants, $f(x)g(x)$ and $K(x,t)$ are known functions while $y(x)$ is unknown function and λ is a non-zero real or complex parameter, is called *Fredholm integral equation of third kind*. The function $K(x,t)$ is known as the kernel of the integral equation.

• Setting $g(x) = 0$ in Eq.(14.2.9), we have the *Fredholm integral equation of the first kind*.

• Setting $g(x) = 1$ in Eq.(14.2.9), we have the *Fredholm integral equation of the second kind*.

• Setting $g(x) = 1$ and $f(x) = 0$ in Eq.(14.2.9), we have the *Homogeneous Fredholm integral equation of the second kind*.

Definition 14.2.6. Volterra Integral Equation : A linear integral equation of the form

$$g(x)y(x) = f(x) + \lambda \int_a^x K(x,t)y(t)\,dt, \tag{14.2.10}$$

where a, b are both constants, $f(x),\ g(x)$ and $K(x,t)$ are known functions while $y(x)$ is unknown function and λ is a non-zero real or complex parameter, is called *Volterra integral equation of third kind*. The function $K(x,t)$ is known as the kernel of the integral equation.

• Setting $g(x) = 0$ in Eq. (14.2.10), we have the linear integral equation of the form

$$f(x) + \lambda \int_a^x K(x,t)\,y(t)\,dt = 0, \tag{14.2.11}$$

which is known as *Volterra integral equation of the first kind.*

• Setting $g(x) = 1$ in Eq. (14.2.10), we have a linear integral equation of the form

$$y(x) = f(x) + \lambda \int_a^x K(x,t)\,dt, \tag{14.2.12}$$

which is known as *Volterra integral equation of the second kind*.

• Setting $f(x) = 0$ in Eq. (14.2.10), we have the following integral equation of the form

$$y(x) = \lambda \int_a^x K(x,t)\,y(t)\,dt, \tag{14.2.13}$$

which is known as the *homogeneous Volterra integral equation of the second kind*.

14.3 Special kinds of kernels

The following special cases of the kernel of an integral equation are of main interest and we shall frequently come across with such kernels throughout the discussion of this unit.

(i) **Symmetric kernel:** A kernel $K(x,t)$ is symmetric (or complex symmetric or Hermitian) if

$$K(x,t) = \bar{K}(t,x)$$

where the bar donates the complex conjugate. A real kernel $K(x,t)$ is symmetric if

$$K(x,t) = K(t,x).$$

For example, $\sin(x+t), \log(xt), x^2t^2 + xt + 1$ etc. are all symmetric kernels. Again, $\sin(2x+3t)$ and $x^2t^3 + 1$ are not symmetric kernels. Again $i(x-t)$ is a symmetric kernel, since in this case, if $K(x,t) = i(x-t)$, then $k(t,x) = i(t-x)$ and so $\bar{K}(t,x) = -i(t-x) = i(x-t) = K(x,t)$. On the other hand, $i(x+t)$ is not a symmetric kernel, since in this case, if $K(x,t) = i(x+t)$, then $\bar{K}(t,x) = \overline{i(t+x)} = -i(t+x) =$ - $K(x,t)$ and so $K(x,t) \neq \bar{K}(x,t)$.

(ii) **Separable or degenerate kernel:** A kernel $K(x,t)$ is called separable if it can be expressed as the sum of a finite number of terms, each of which is the product of a function of x only and a function of t only, i.e.,

$$K(x,t) = \sum_{i=1}^{n} g_i(x)h_i(t). \tag{14.3.1}$$

Remark: The functions $g_i(x)$ can be regarded as linearly independent, otherwise the number of terms in relation (14.3.1) can be further reduced. Recall that the set of functions $g_i(x)$ is said to be linearly independent, if $c_1g_1(x) + c_2g_2(x) + \cdots + c_ng_n(x) = 0$, where $c_1, c_2, \ldots, c_n$ are arbitrary constants, then $c_1 = c_2 = \ldots = c_n = 0$.

(iii) **Iterated Kernels or functions:**

(a) Consider a Fredholm integral equation of the second kind

$$y(x) = f(x) + \lambda \int_a^b K(x,t)y(t)dt \tag{14.3.2}$$

Then the iterated kernels $K_n(x,t)$, $n = 1, 2, \ldots$ are defined as follows:

$$\begin{aligned} K_1(x,t) &= K(x,t) \\ \text{and} \quad K_n(x,t) &= \int_a^b K(x,z)K_{n-1}(z,t)dz, \quad n = 2, 3, \ldots \end{aligned} \tag{14.3.3}$$

(b) Consider a Volterra integral equation of the second kind

$$y(x) = f(x) + \lambda \int_a^x K(x,t)y(t)dt \tag{14.3.4}$$

Then the iterated kernels $K_n(x,t)$, $n = 1, 2, \ldots$ are defined as follows:

$$\begin{aligned} K_1(x,t) &= K(x,t) \\ \text{and} \quad K_n(x,t) &= \int_t^x K(x,z)K_{n-1}(z,t)dz, \quad n = 2, 3, \ldots \end{aligned} \tag{14.3.5}$$

(iv) **Resolvent Kernel or reciprocal kernel:** Suppose the solution of integral equations

$$y(x) = f(x) + \lambda \int_a^b K(x,t)y(t)dt \tag{14.3.6}$$

and

$$y(x) = f(x) + \lambda \int_a^x K(x,t)y(t)dt \tag{14.3.7}$$

be respectively

$$y(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) f(t) dt \tag{14.3.8}$$

and

$$y(x) = f(x) + \lambda \int_a^x \Gamma(x,t;\lambda) f(t) dt, \tag{14.3.9}$$

then $R(x,t;\lambda)$ and $\Gamma(x,t;\lambda)$ is called the resolvent kernel or reciprocal kernel of the given integral equation.

14.4 Integral equations of the convolution type

Consider an integral equation in which the kernel $K(x,t)$ is dependent solely on the difference $x-t$,i.e.,

$$K(x,t) = K(x-t), \tag{14.4.1}$$

where K is a certain function of one variable. Then integral equations

$$y(x) = f(x) + \lambda \int_a^x K(x-t)\, y(t)\, dt,$$

and

$$y(x) = f(x) + \lambda \int_a^b K(x-t)\, y(t)\, dt$$

are called integral equations of the convolution type. $K(x-t)$ is called difference kernel. Let $y_1(x)$ and $y_2(x)$ be two continuous functions defined for $x \geq 0$. Then the convolution or Faltung of y_1 and y_2 is denoted and defined by

$$y_1 * y_2 = \int_0^x y_1(x-t) y_2(t)\, dt = \int_0^x y_1(t) y_2(x-t)\, dt. \tag{14.4.2}$$

The integrals occurring in (14.4.2) are called the convolution integrals. Note that the convolution defined by relation (14.4.2) is a particular case of the standard convolution.

$$y_1 * y_2 = \int_{-\infty}^{\infty} y_1(x-t) y_2(t)\, dt = \int_{-\infty}^{\infty} y_1(t) y_2(x-t)\, dt. \tag{14.4.3}$$

By setting $y_1(t) = y_2(t) = 0$, for $t < 0$ and $t > x$, the integrals in (14.4.2) can be obtained from those in (14.4.3).

14.5 Eigenvalues and Eigen functions

Consider the homogeneous Fredholm integral equation

$$y(x) = \lambda \int_a^b K(x,t) y(t)\, dt. \tag{14.5.1}$$

Then (14.5.1) has the obvious solution $y(x) = 0$, which is called the zero or trivial solution of (14.5.1). The values of the parameter λ for which (14.5.1) has a non-zero solution $y(x) \neq 0$ are called eigenvalues of

(14.5.1) or of the kernel $K(x,t)$, and every non-zero solution of (14.5.1) is called on eigenfunction corresponding to the eigenvalue λ.

Remark 1. The number $\lambda = 0$ is not an eigenvalue since for $\lambda = 0$ it follows from (14.5.1) that $y(x) = 0$.
Remark 2. If $y(x)$ is an eigenfunction of (14.5.1), then $c\, y(x)$, where c is an arbitrary constant, is also an eigenfunction of (14.5.1), which corresponds to the same eigenvalue λ.
Remark 3. A homogeneous Fredholm integral equation of the second kind may, generally, have no eigenvalue and eigenfunction, or it may not have any real eigenvalue or eigenfunction.

14.6 Leibnitz's rule of differentiation under integral sign

Let $F(x,t)$ and $\partial F/\partial x$ be continuous functions of both x and t and let the first derivatives of $G(x)$ and $H(x)$ be continuous. Then

$$\frac{d}{dx}\int_{G(x)}^{H(x)} F(x,t)\,dt = \int_{G(x)}^{H(x)} \frac{\partial F}{\partial x}\,dt + F[x,H(x)]\frac{dH}{dx} - F[x,G(x)]\frac{dG}{dx} \tag{14.6.1}$$

Particular Case: If G and H are absolute constants, then (14.6.1) reduces to

$$\frac{d}{dx}\int_{G}^{H} F(x,t)dt = \int_{G}^{H} \frac{\partial F}{\partial x}\,dt.$$

14.7 Multiple integral into a single integral: Conversion formula

The multiple integral of order n can be convert to a ordinary integral of order one using the formula

$$\int_a^x y(t)\,dt^n = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!}y(t)\,dt.$$

Proof. Let

$$I_n(x) = \int_a^x (x-t)^{n-1}y(t)\,dt, \tag{14.7.1}$$

where n is a positive integer and a is constant. Differentiating (14.7.1) with respect to x and using Leibnitz's rule, we have

$$\frac{dI_n}{dx} = (n-1)\int_a^x (x-t)^{n-2}y(t)\,dt + (x-x)^{n-1}y(x)\frac{dx}{dx} - (x-0)^{n-1}y(0)\frac{d0}{dx}$$

$$\Rightarrow dI_n/dx = (n-1)\,I_{n-1},\ n > 1 \tag{14.7.2}$$

From (14.7.1),

$$I_1 = \int_a^x y(t)\,dt \quad \text{so that} \quad \frac{dI_1}{dx} = y(x) \tag{14.7.3}$$

Now, differentiating (14.7.2) with respect to x successively k times, we have

$$\frac{d^k I_n}{dx^k} = (n-1)(n-2)\cdots(n-k)\ I_{n-k}, \quad n > k \tag{14.7.4}$$

Using (14.7.4) for $k = n-1$, we have

$$\frac{d^k I_n}{dx^k} = (n-1)! I_1. \tag{14.7.5}$$

Differentiating (14.7.5) w.r.t. 'x' and using (14.7.3), we obtain

$$\frac{d^n I_n}{dx^n} = (n-1)! y(x) \tag{14.7.6}$$

From (14.7.1), (14.7.4), and (14.7.5), it follows that $I_n(x)$ and its first $n-1$ derivatives all vanish when $x = a$. Hence using (14.7.3) and (14.7.6), we obtain

$$I_1(x) = \int_a^x y(t_1)\ dt_1, \quad I_2(x) = \int_a^x I_1(t_2)\ dt_2 = \int_a^x \int_a^{t_2} y(t_1)\ dt_1\ dt_2$$

Proceeding likewise, we obtain

$$I_n(x) = (n-1)! \int_a^x \int_a^{t_n} \cdots \int_a^{t_3} \int_a^{t_2} y(t_1)\ dt_1\ dt_2\ \cdots\ dt_{n-1}\ dt_n \tag{14.7.7}$$

Combining (14.7.1) and (14.7.7), we obtain

$$\int_a^x \int_a^{t_n} \cdots \int_a^{t_3} \int_a^{t_2} y(t_1)\ dt_1\ dt_2\ \cdots\ dt_{n-1}\ dt_n = \frac{1}{(n-1)!} \int_a^x (x-t)^{n-1} y(t)\ dt \tag{14.7.8}$$

From (14.7.8), we obtain

$$\int_a^x y(t)\ dt^n = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} y(t)\ dt.$$

Example 14.7.1. Show that the function $y(x) = (1+x^2)^{-3/2}$ is a solution of the Volterra integral equation

$$y(x) = \frac{1}{1+x^2} - \int_0^x \frac{t}{1+x^2} y(t) dt$$

Solution. Given integral equation is

$$y(x) = \frac{1}{1+x^2} - \int_0^x \frac{t}{1+x^2} y(t) dt \tag{14.7.9}$$

Also given

$$y(x) = (1+x^2)^{-3/2} \tag{14.7.10}$$

From (14.7.10),

$$y(t) = (1+t^2)^{-3/2} \tag{14.7.11}$$

The RHS of (14.7.9)

$$\begin{aligned}
&= \frac{1}{1+x^2} - \int_0^x \frac{t}{1+x^2}(1+t^2)^{-3/2}, \text{ using (14.7.11)} \\
&= \frac{1}{1+x^2} - \frac{1}{1+x^2}\int_0^x (1+u)^{-3/2}\cdot\frac{1}{2}du \quad \left[\text{on putting } t^2 = u \text{ and } 2tdt = du\right] \\
&= \frac{1}{1+x^2} - \frac{1}{1+x^2}\cdot\frac{1}{2}\left[\frac{(1+u)^{-1/2}}{-1/2}\right]_0^{x^2} \\
&= \frac{1}{1+x^2} + \frac{1}{1+x^2}\left[\frac{1}{(1+u)^{1/2}}\right]_0^{x^2} \\
&= (1+x^2)^{-3/2} = y(x) = \text{LHS of (14.7.9)}.
\end{aligned}$$

Hence, $y(x) = (1+x^2)^{-3/2}$ is a solution of the given integral equation. ■

14.8 Method of converting an initial value problem into a volterra integral equation

While searching for the representation formula for the solution of an ordinary differential equation in such a manner so as to include the boundary conditions or initial conditions explicitly, we always arrive at integral equations. Thus, a boundary value or an initial value problem is converted to an integral equation. After converting an initial value problem into an integral equation, it can be solved by shorter methods of solving integral equations.

The method is illustrated with the help of the following examples.

Example 14.8.1. Convert the following differential equation into integral equation

$$y'' + y = 0 \quad \text{when} \quad y(0) = y'(0) = 0.$$

Solution. Given $y''(x) + y(x) = 0$, with initial conditions $y(0) = 0$ and $y'(0) = 0$. From the given differential equation we have

$$y''(x) = -y(x) \tag{14.8.1}$$

Integrating both sides of (14.8.1) w.r.t. 'x' from 0 to x, we have

$$\int_0^x y''(x)\,dx = -\int_0^x y(x)\,dx \quad \text{or} \quad [y'(x)]_0^x = -\int_0^x y(x)\,dx$$

$$\text{or} \quad y'(x) - y'(0) = -\int_0^x y(x)\,dx \quad \text{or} \quad y'(x) = -\int_0^x y(x)\,dx$$

Integrating both sides of w.r.t. 'x' from 0 to x, we have

$$\int_0^x y'(x)\,dx = -\int_0^x y(x)dx^2 \quad \text{or} \quad [y(x)]_0^x = -\int_0^x y(x)dx^2$$
$$\text{or} \quad y(x) - y(0) = -\int_0^x y(x)\,dx^2 \quad \text{or} \quad y(x) = -\int_0^x y(t)dt^2$$
$$\text{or} \quad y(x) = -\int_0^x (x-t)y(t)\,dt$$

which is the desired integral equation.

Example 14.8.2. Convert $y'' - \sin xy' + e^x y = x$ with initial conditions $y(0) = 1, y'(0) = -1$ to a Volterra integral equation of the second kind. Conversely, derive the original differential equation with the initial conditions from the integral equation obtained.

Solution. Given $y''(x) - \sin xy'(x) + e^x y(x) = x$ with initial conditions $y(0) = 1$ and $y'(0) = -1$. From the given differential equation, one can write

$$y''(x) = x - e^x y(x) + \sin xy'(x). \tag{14.8.2}$$

Integrating both sides of (14.8.2) w.r.t. 'x' from 0 to x, we have

$$\int_0^x y''(x)\,dx = \int_0^x x\,dx - \int_0^x e^x y(x)\,dx + \int_0^x \sin xy'(x)\,dx$$
$$[y'(x)]_0^x = \frac{x^2}{2} - \int_0^x e^x y(x)\,dx + [\sin x\,y(x)]_0^x - \int_0^x \cos x\,y(x)\,dx$$
$$y'(x) - y'(0) = \frac{x^2}{2} - \int_0^x e^x y(x)\,dx + \sin xy(x) - \int_0^x \cos x\,y(x)\,dx$$
$$y'(x) + 1 = \frac{x^2}{2} + \sin xy(x) - \int_0^x (e^x + \cos x)\,y(x)\,dx$$
$$y'(x) = \frac{x^2}{2} - 1 + \sin x\,y(x) - \int_0^x (e^x + \cos x)\,y(x)\,dx \tag{14.8.3}$$

Integrating both sides of (14.8.3) w.r.t. 'x' from 0 to x, we have

$$\int_0^x y'(x)\,dx = \int_0^x \left(\frac{x^2}{2}-1\right)\,dx + \int_0^x \sin x\, y(x)\,dx - \int_0^x \left(e^x+\cos x\right)\,y(x)\,dx^2$$

$$[y(x)]_0^x = \left[\frac{x^3}{6}-x\right]_0^x + \int_0^x \sin t\, y(t)\,dt - \int_0^x \left(e^t+\cos t\right)\,y(t)\,dt^2$$

$$y(x)-y(0) = \frac{x^3}{6}-x+\int_0^x \sin t\, y(t)\,dt - \int_0^x (x-t)\left(e^t+\cos t\right)y(t)\,dt$$

$$y(x)-1 = \frac{x^3}{6}-x+\int_0^x \left\{\sin t-(x-t)\left(e^t+\cos t\right)\right\}y(t)\,dt,$$

$$y(x) = \frac{x^3}{6}-x+1+\int_0^x \left[\sin t-(x-t)\left(e^t+\cos t\right)\right]y(t)\,dt \tag{14.8.4}$$

which is the required Volterra integral equation of the second kind.

Second Part: Derivation of the given differential equation together with given initial conditions from integral equation (14.8.4).

Differentiating both sides of (14.8.4) w.r.t. 'x', we get

$$y'(x) = \frac{x^2}{2}-1+\frac{d}{dx}\int_0^x \left[\sin t-(x-t)\left(e^t+\cos t\right)\right]y(t)\,dt$$

$$y'(x) = \frac{x^2}{2}-1+\int_0^x \frac{\partial}{\partial x}\left[\left\{\sin t-(x-t)\left(e^t+\cos t\right)\right\}y(t)\right]\,dt$$

$$+\left[\sin x-(x-x)\left(e^x+\cos x\right)\right]y(x)\frac{dx}{dx}-\left[\sin 0-(x-0)\left(e^0+\cos 0\right)\right]y(0)\frac{d0}{dx}$$

[using Leibnitz's rule of differentiation under integral sign

$$y'(x) = \frac{x^2}{2}-1-\int_0^x \left(e^t+\cos t\right)y(t)\,dt+\sin x\,y(x). \tag{14.8.5}$$

Differentiating both sides of (14.8.5) with respect to 'x' we get

$$y''(x) = x+\cos x y(x)+\sin x y'(x)-\frac{d}{dx}\int_0^x \left(e^t+\cos t\right)y(t)\,dt$$

$$y''(x) = x+\cos x\,y(x)+\sin x y'(x)-\left[\int_0^x \frac{\partial}{\partial x}\left\{\left(e^t+\cos t\right)y(t)\right\}\,dt\right.$$

$$\left.+\left(e^x+\cos x\right)y(x)\frac{dx}{dx}-\left(e^0+\cos 0\right)y(0)\frac{d0}{dx}\right],$$

$$y''(x) = x+\cos x y(x)+\sin x y'(x)-\left[0+\left(e^x+\cos x\right)y(x)+0\right]$$

$$y''(x)-\sin x \quad y'(x)+e^x \quad y(x) = x,$$

which is the same as the given differential equation.

Example 14.8.3. Convert $y''(x) - 3y'(x) + 2y(x) = 4\sin x$ with initial conditions $y(0) = 1,\ y'(0) = -2$ into a Volterra integral equation of the second kind. Conversely, derive the original differential equation with initial conditions from the integral equation obtained.

Solution. Given

$$y''(x) - 3y'(x) + 2y(x) = 4\sin x, \quad y(0) = 1,\ y'(0) = -2 \tag{14.8.6}$$

From (14.8.6),

$$y''(x) = 4\sin x - 2y(x) + 3y'(x). \tag{14.8.7}$$

Integrating both sides of (**??**) w.r.t. 'x' from 0 to x, we have

$$\int_0^x y''(x)\,dx = 4\int_0^x \sin x\,dx - 2\int_0^x y(x)\,dx + 3\int_0^x y'(x)\,dx$$

$$\left[y'(x)\right]_0^x = 4\left[-\cos x\right]_0^x - 2\int_0^x y(x)\,dx + 3\left[y(x)\right]_0^x$$

$$y'(x) - y'(0) = 4(-\cos x + 1) - 2\int_0^x y(x)\,dx + 3\left[y(x) - y(0)\right]$$

$$y'(x) + 2 = 4 - 4\cos x - 2\int_0^x y(x)\,dx + 3y(x) - 3$$

$$y'(x) = -1 - 4\cos x - 3y(x) - 2\int_0^x y(x)\,dx \tag{14.8.8}$$

Integrating both sides of (14.8.8) w.r.t. 'x' from 0 to x, we have

$$\int_0^x y'(x)\,dx = -\int_0^x dx - 4\int_0^x \cos x\,dx + 3\int_0^x y(x)\,dx - 2\int_0^x y(x)\,dx^2$$

$$\left[y(x)\right]_0^x = -x - 4\left[\sin x\right]_0^x + 3\int_0^x y(x)\,dx - 2\int_0^x y(t)dt^2$$

$$y(x) - y(0) = -x - 4\sin x + 3\int_0^x y(t)\,dt - 2\int_0^x (x-t)y(t)\,dt,$$

$$y(x) = 1 - x - 4\sin x + \int_0^x \left[3 - 2(x-t)\right]y(t)\,dt, \tag{14.8.9}$$

which is the required Volterra integral equation of the second kind.

Second Part: Derivation of the given differential equation together with given initial conditions from integral equation (14.8.9).

Differentiating both sides of (14.8.9) w.r.t. 'x', we get

$$
\begin{aligned}
& y'(x) = -1 - 4\cos x + \frac{d}{dx}\left[\int_0^x [3-2(x-t)]\right] y(t)\,dt \\
\Rightarrow\ & y'(x) = -1 - 4\cos x + \int_0^x \frac{\partial}{\partial x}\left[\{3-2(x-t)\}y(t)\,dt\right]dt + [3-2(x-x)]y(x)\frac{dx}{dx} - \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [3-2(x-0)]y(0)\frac{d0}{dx} \\
\Rightarrow\ & y'(x) = -1 - 4\cos x + \int_0^x (-2)y(t)\,dt + 3y(x) \\
\Rightarrow\ & y'(x) = -1 - 4\cos x + 3y(x) - 2\int_0^x y(t)\,dt \qquad (14.8.10)
\end{aligned}
$$

Differentiating both sides of (14.8.10) w.r.t. 'x', we get

$$
\begin{aligned}
& y''(x) = 4\sin x + 3y'(x) - 2\frac{d}{dx}\int_0^x y(t)\,dt \\
\Rightarrow\ & y''(x) = 4\sin x + 3y'(x) - 2\left[\int_0^x \frac{\partial}{\partial x}y(t)dt + y(x)\frac{dx}{dx} - y(0)\frac{d0}{dx}\right] \\
\Rightarrow\ & y''(x) = 4\sin x + 3y'(x) - 2[0 + y(x) - 0] \\
\Rightarrow\ & y''(x) - 3y'(x) + 2y(x) = 4\sin x, \qquad (14.8.11)
\end{aligned}
$$

which is the same as the given differential equation. Putting $x = 0$ in (14.8.9), we get $y(0) = 1$. Further putting $x = 0$ in (14.8.10), we get $y'(0) = -1 - 4 + 3y(0) = -1 - 4 + 3 = -2 = -2$. Thus,

$$
y(0) = 1 \quad \text{and} \quad y'(0) = -2 \qquad (14.8.12)
$$

(14.8.11) and (14.8.12) together give us the given differential equation and initial conditions.

14.8.1 Initial Value Problem

When an ordinary differential equation is to be solved under conditions involving dependent variable and its derivative at the same value of the independent variable, then the problem under consideration is said to be an initial value problem. For example

$$
\frac{d^2y}{dx^2} + y = x,\ \ y(0) = 2,\ \ y'(0) = 3
$$

and

$$
\frac{d^2y}{dx^2} + y = x,\ \ y(1) = 2,\ \ y'(1) = 2
$$

are both initial value problems.

Method of converting an initial value problem into a Volterra integral equation

This method is illustrated with the help of the following example.

Example 14.8.4. Convert the following differential equation into integral equation:

$$y'' + y = 0 \text{ when } y(0) = y'(0) = 0.$$

Solution. Given

$$y'' + y = 0 \tag{14.8.13}$$

with initial conditions

$$y(0) = 0 \tag{14.8.14}$$
$$y'(0) = 0 \tag{14.8.15}$$

From (14.8.13),

$$y''(x) = -y(x) \tag{14.8.16}$$

Integrating both sides of (14.8.16) with respect to 'x' from 0 to x, we have

$$\begin{aligned} &\int_0^x y''(x)dx = -\int_0^x y(x)dx \\ \text{or,}\quad &[y'(x)]_0^x = -\int_0^x y(x)dx \\ \text{or,}\quad &y'(x) - y'(0) = -\int_0^x y(x)dx \\ \text{or,}\quad &y'(x) = -\int_0^x y(x)dx \;\; \text{[using (14.8.15)]} \end{aligned} \tag{14.8.17}$$

Integrating both sides of (14.8.17) with respect to 'x' from 0 to x, we have

$$\begin{aligned} &\int_0^x y'(x)dx = -\int_0^x y(x)dx^2 \\ \text{or,}\quad &[y(x)]_0^x = -\int_0^x y(x)dx^2 \\ \text{or,}\quad &y(x) - y(0) = -\int_0^x y(x)dx^2 \\ \text{or,}\quad &y(x) = -\int_0^x y(t)dt^2 \;\; \text{[using (14.8.14)]} \qquad (14.8.18) \\ \text{or,}\quad &y(x) = -\int_0^x (x-t)y(t)dt, \end{aligned}$$

using the result $\int_a^x y(t)dt^n = \int_a^x \frac{(x-t)^{n-1}}{(n-1)!}y(t)dt$, which is the desired result. ■

14.8.2 Boundary Value Problem

When an ordinary differential equation is to be solved under conditions involving dependent variable and its derivatives at two different values of independent variable, then the problem under consideration is said to be a boundary value problem. For example,

$$\frac{d^2y}{dx^2} + y = 0, \;\; y(a) = y_1, \;\; y(b) = y_2$$

is a boundary value problem.

Method of converting a boundary value problem into a Fredholm integral equation:

We explain this method with the help of the following solved examples.

Example 14.8.5. Reduce the following boundary value problem into an integral equation:

$$\frac{d^2y}{dx^2} + \lambda y = 0, \;\; y(0) = 0, \;\; y(l) = 0.$$

Solution. Given

$$\begin{aligned} y''(x) + \lambda y(x) &= 0 && (14.8.19)\\ \text{with} \quad y(0) &= 0 && (14.8.20)\\ y(l) &= 0. && (14.8.21) \end{aligned}$$

From (14.8.19),

$$y''(x) = -\lambda y(x). \tag{14.8.22}$$

Integrating both sides of (14.8.22) with respect to 'x' from 0 to x, we have

$$\begin{aligned} \int_0^x y''(x)dx &= -\lambda \int_0^x y(x)dx \\ \text{or,} \quad [y'(x)]_0^x &= -\lambda \int_0^x y(x)dx \\ \text{or,} \quad y'(x) - y'(0) &= -\lambda \int_0^x y(x)dx. && (14.8.23) \end{aligned}$$

Let $y'(0) = c$, a constant. Then (14.8.23) becomes

$$y'(x) = c - \lambda \int_0^x y(x)dx. \tag{14.8.24}$$

Integrating both sides of (14.8.24) with respect to 'x' from 0 to x, we have

$$\begin{aligned} \int_0^x y'(x)dx &= c\int_0^x dx - \lambda \int_0^x y(x)dx^2 \\ \text{or,} \quad [y(x)]_0^x &= cx - \lambda \int_0^x y(t)dt^2 \\ \text{or,} \quad y(x) - y(0) &= cx - \lambda \int_0^x (x-t)y(t)dt \\ \text{or,} \quad y(x) - 0 &= cx - \lambda \int_0^x (x-t)y(t)dt \;\; [\text{using } (14.8.20)] && (14.8.25)\\ \text{or,} \quad y(x) &= cx - \lambda \int_0^x (x-t)y(t)dt && (14.8.26) \end{aligned}$$

Putting $x = l$ in (14.8.26) we get

$$\begin{aligned} y(l) &= cl - \lambda \int_0^l (l-t)y(t)dt \\ \text{or,} \quad 0 &= cl - \lambda \int_0^l (l-t)y(t)dt \;\; [\text{using } (14.8.21)] \\ \text{or,} \quad c &= \frac{\lambda}{l}\int_0^l (l-t)y(t)dt && (14.8.27) \end{aligned}$$

Using (14.8.27), (14.8.26) reduces to

$$y(x) = \frac{\lambda}{l}\int_0^l (l-t)y(t)dt - \lambda\int_0^x (x-t)y(t)dt \tag{14.8.28}$$

$$\text{or,}\quad y(x) = \int_0^l \frac{\lambda x(l-t)}{l}y(t)dt - \lambda\int_0^x (x-t)y(t)dt$$

$$\text{or,}\quad y(x) = \int_0^x \frac{\lambda x(l-t)}{l}y(t)dt + \int_x^l \frac{\lambda x(l-t)}{l}y(t)dt - \int_0^x \lambda(x-t)y(t)dt$$

$$\text{or,}\quad y(x) = \lambda\int_0^x \left[\frac{x(l-t)}{l} - (x-t)\right]y(t)dt + \lambda\int_x^l \frac{x(l-t)}{l}y(t)dt$$

$$\text{or,}\quad y(x) = \lambda\int_0^x \frac{x(l-t)-l(x-t)}{l}y(t)dt + \lambda\int_x^l \frac{x(l-t)}{l}y(t)dt$$

$$\text{or,}\quad y(x) = \lambda\left[\int_0^x \frac{t(l-x)}{l}y(t)dt + \int_x^l \frac{x(l-t)}{l}y(t)dt\right]$$

$$\text{or,}\quad y(x) = \lambda\int_0^l K(x,t)y(t)dt, \tag{14.8.29}$$

where

$$\begin{aligned} K(x,t) &= \frac{t}{l}(l-x), \ \text{ if } 0 < t < x \\ &= \frac{x}{l}(l-t), \ \text{ if } x < t < l \end{aligned} \tag{14.8.30}$$

where (14.8.29) is the required Fredholm integral equation where $K(x,t)$ is given by (14.8.30). ■

14.9 Volterra Integral Equations

Definition 14.9.1. Volterra Integral Equation : A linear integral equation of the form

$$g(x)y(x) = f(x) + \lambda\int_a^x K(x,t)y(t)dt, \tag{14.9.1}$$

where a, b are both constants, $f(x)$, $g(x)$ and $K(x,t)$ are known functions while $y(x)$ is unknown function; λ is a non-zero real or complex parameter is called *Volterra integral equation of third kind*. The function $K(x,t)$ is known as the kernel of the integral equation.

- Setting $g(x) = 0$ in Eq.(14.9.1), we have the *Volterra integral equation of the first kind*.

- Setting $g(x) = 1$ in Eq.(14.9.1), we have the *Volterra integral equation of the second kind*.

- Setting $g(x) = 1$ and $f(x) = 0$ in Eq.(14.9.1), we have the *Homogeneous Volterra integral equation of the second kind*.

14.9.1 Solution of Volterra integral equations

14.9.2 Determination of Resolvent kernel for Volterra integral equations

Example 14.9.2. Find the resolvent kernel of the Volterra integral equation with the kernel $K(x,t) = 1$.

Solution : Iterated kernels $K_n(x,t)$ are given by

$$K_1(x,t) = K(x,t) \tag{14.9.2}$$

$$\text{and} \quad K_n(x,t) = \int_t^x K(x,z)K_{n-1}(z,t)\,dz, \quad n = 1,2,3,\dots \tag{14.9.3}$$

Given $K(x,t) = 1$. Thus we have

$$K_1(x,t) = K(x,t) = 1$$

Putting $n = 2$ in Eq.(14.9.3) we have

$$K_2(x,t) = \int_t^x K(x,z)K_1(z,t)\,dz = \int_t^x dz = [z]_t^x = x - t$$

Putting $n = 3$ in Eq.(14.9.3) we have

$$K_3(x,t) = \int_t^x K(x,z)K_2(z,t)\,dz = \int_t^x 1\cdot(z-t)\,dz = \left[\frac{(z-t)^2)}{2}\right]_t^x = \frac{(x-t)^2}{2!}$$

Putting $n = 4$ in Eq.(14.9.3) we have

$$K_4(x,t) = \int_t^x K(x,z)K_3(z,t)\,dz = \int_t^x 1\cdot\frac{(z-t)^2}{2!}\,dz = \frac{1}{2!}\left[\frac{(z-t)^3)}{3}\right]_t^x = \frac{(x-t)^3}{3!}$$

Observing above, we find by mathematical induction, that

$$K_n(x,t) = \frac{(x-t)^{n-1}}{(n-1)!}, \quad n = 1,2,3,\dots$$

Now by the definition of the resolvent kernel, we have

$$\begin{aligned} R(x,t;\lambda) &= \sum_{m=1}^{\infty} K_m(x,t) = K_1(x,t) + \lambda K_2(x,t) + \lambda^2 K_3(x,t) + \cdots \\ &= 1 + \frac{\lambda(x-t)}{1!} + \frac{[\lambda(x-t)]^2}{2!} + \frac{[\lambda(x-t)]^3}{3!} + \cdots \\ &= e^{\lambda(x-t)} \end{aligned}$$

Exercise 14.9.3. Find the resolvent kernel of the Volterra integral equation with the kernel

$$i) \quad K(x,t) = e^{x-t} \qquad ii) \quad K(x,t) = (2+\cos x)/(2+\cos t)$$

Answers :

$$i) \quad R(x,t;\lambda) = e^{(x-t)(1+\lambda)}, \qquad (ii) \quad R(x,t;\lambda) = \frac{2+\cos x}{2+\cos t}e^{\lambda(x-t)}$$

Unit 15

Course Structure

- Homogeneous Fredholm integral equation of the second kind with separable or degenerate kernel; classical Fredholm theory- Fredholm alternative, Fredholm theorem.

15.1 Introduction

In Unit 11, we obtained the solution of the Fredholm integral equation of the second kind

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{15.1.1}$$

as a uniformly convergent power series in the parameter λ for $|\lambda|$ suitably small. Fredholm derived the solution of (15.1.1) in general form which is valid for all values of the parameter λ. He gave three important results which are known as Fredholm's first, second and third fundamental theorems. In the present unit we propose to discuss these theorems.

Objective

The objective of this course is to learn the students all of the above topics and by the end of it students should be able to

- know different fundamental theorems of Fredholm integral equation.
- know the method of solution of Fredholm integral equation using fundamental theorems.
- know various aspects of Hilbert-Schmidt theory
- know fundamental properties of eigenvalues and eigenfunctions for symmetric kernels

15.2 Classical Fredholm Theory

15.2.1 Fredholm's First Fundamental Theorem

The non-homogeneous Fredholm integral equation of second kind

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{15.2.1}$$

where the functions $f(x)$ and $y(t)$ are integrable, has a unique solution

$$y(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, dt \tag{15.2.2}$$

where the resolvent kernel $R(x,t;\lambda)$ is given by

$$R(x,t;\lambda) = \frac{D(x,t;\lambda)}{D(\lambda)} \tag{15.2.3}$$

with $D(\lambda) \neq 0$, is a meromorphic function of the complex variable λ, being the ratio of two entire functions defined by the series

$$\begin{aligned} D(x,t;\lambda) &= K(x,t) + \sum_{p=1}^{\infty} \frac{(-\lambda)^p}{p!} \int \cdots \int K\begin{pmatrix} x, & z_1, & \cdots, & z_p \\ t, & z_1, & \cdots, & z_p \end{pmatrix} dz_1 \cdots dz_p \\ \text{and } D(\lambda) &= 1 + \sum_{p=1}^{\infty} \frac{(-\lambda)^p}{p!} \int \cdots \int K\begin{pmatrix} z_1, & \cdots, & z_p \\ z_1, & \cdots, & z_p \end{pmatrix} dz_1 \cdots dz_p, \end{aligned} \tag{15.2.4}$$

both of which converge for all values of λ. Also, note the following symbol for the determinant formed by the values of the values of the kernel at all points (x_i, t_i)

$$\begin{vmatrix} K(x_1,t_1) & K(x_1,t_2) & \cdots & K(x_1,t_n) \\ K(x_2,t_1) & K(x_2,t_2) & \cdots & K(x_2,t_n) \\ \vdots & \vdots & \ldots & \vdots \\ K(x_n,t_1) & K(x_n,t_2) & \cdots & K(x_n,t_n) \end{vmatrix} = K\begin{pmatrix} x_1, & x_2, & \cdots, & x_n \\ t_1, & t_2, & \cdots, & t_n \end{pmatrix} \tag{15.2.5}$$

which is known as the *Fredholm determinant.* In particular, the solution of the Fredholm homogeneous equation

$$y(x) = \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{15.2.6}$$

is identically zero.

Result 15.2.1. For Fredholm integral equation

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{15.2.7}$$

the resolvent kernel is given by

$$R(x,t;\lambda) = \frac{D(x,t;\lambda)}{D(\lambda)} \tag{15.2.8}$$

where

$$D(x,t;\lambda) = K(x,t) + \sum_{m=1}^{\infty} \frac{(-\lambda)^m}{m!} B_m(x,t) \tag{15.2.9}$$

$$\text{and } D(\lambda) = 1 + \sum_{m=1}^{\infty} \frac{(-\lambda)^m}{m!} C_m \tag{15.2.10}$$

where

$$B_n(x,t) = \underbrace{\int_a^b \cdots \int_a^b}_{n} \begin{vmatrix} K(x,t) & K(x,z_1) & \cdots & K(x,z_n) \\ K(z_1,t) & K(z_1,z_1) & \cdots & K(z_1,z_n) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ K(z_n,t) & K(z_n,z_1) & \cdots & K(z_n,z_n) \end{vmatrix} dz_1\, dz_2 \cdots dz_n, \tag{15.2.11}$$

$$\text{and } C_n = \underbrace{\int_a^b \cdots \int_a^b}_{n} \begin{vmatrix} K(z_1,z_1) & K(z_1,z_2) & \cdots & K(z_1,z_n) \\ K(z_2,z_1) & K(z_2,z_2) & \cdots & K(z_2,z_n) \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ K(z_n,z_1) & K(z_n,z_2) & \cdots & K(z_n,z_n) \end{vmatrix} dz_1\, dz_2 \cdots dz_n, \tag{15.2.12}$$

The function $D(x,t;\lambda)$ is called the *Fredholm minor* and $D(\lambda)$ is called the *Fredhom determinant*.

15.2.2 Alternative Procedure of calculating $\mathbf{B_m(x,t)}$ and $\mathbf{C_m}$

The following results will be used

$$\begin{aligned} C_0 &= 1 \\ C_p &= \int_a^b B_{p-1}(s,s)\, ds, \;\; p \geq 1 \\ B_0(x,t) &= K(x,t) \\ B_p(x,t) &= C_p\, K(x,t) - p \int_a^b K(x,z)\, B_{p-1}(z,t)\, dz, \;\; p \geq 1. \end{aligned} \tag{15.2.13}$$

After getting $R(x,t;\lambda)$, the required solution is given by

$$y(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, dt \tag{15.2.14}$$

Example 15.2.2. Using Fredholm determinants, find the resolvent kernel and hence solve the following integral equation

$$y(x) = f(x) + \lambda \int_0^1 x e^t\, y(t)\, dt, \; (\lambda \neq 1)$$

Solution : Here

$$K(x,t) = x e^t$$

From Eq.(15.2.11)

$$B_1(x,t) = \int_0^1 \begin{vmatrix} K(x,t) & K(x,z_1) \\ K(z_1,t) & K(z_1,z_1) \end{vmatrix} dz_1 = \int_0^1 \begin{vmatrix} xe^t & xe^{z_1} \\ z_1e^t & z_1e^{z_1} \end{vmatrix} dz_1 = 0$$

$$B_2(x,t) = \int_0^1 \int_0^1 \begin{vmatrix} K(x,t) & K(x,z_1) & K(x,z_2) \\ K(z_1,t) & K(z_1,z_1) & K(z_1,z_2) \\ K(z_2,t) & K(z_2,z_1) & K(z_2,z_2) \end{vmatrix} dz_1\, dz_2 = \int_0^1 \int_0^1 \begin{vmatrix} xe^t & xe^{z_1} & xe^{z_2} \\ z_1e^t & z_1e^{z_1} & z_1e^{z_2} \\ z_2e^t & z_2e^{z_1} & z_2e^{z_2} \end{vmatrix} dz_1\, dz_2 = 0$$

Since $B_1(x,t) = B_2(x,t) = 0$, it follows that $B_n(x,t) = 0$, for $n \geq 1$. Now from Eq.(15.2.12), we have

$$C_1 = \int_0^1 K(z_1,z_1)\, dz_1 = \int_0^1 z_1e^{z_1}\, dz_1 = \left[z_1e^{z_1}\right]_0^1 - \int_0^1 e^{z_1}dz_1 = e - \left[e^{z_1}\right]_0^1 = e - (e-1) = 1.$$

$$C_2 = \int_0^1 \int_0^1 \begin{vmatrix} K(z_1,z_1) & K(z_1,z_2) \\ K(z_2,z_1) & K(z_2,z_2) \end{vmatrix} dz_1\, dz_2 = \int_0^1 \int_0^1 \begin{vmatrix} z_1e^{z_1} & z_1e^{z_2} \\ z_2e^{z_1} & z_2e^{z_2} \end{vmatrix} dz_1\, dz_2 = 0$$

It follows that $C_m = 0$ for all $m \geq 2$. Now Eq.(15.2.9) and Eq.(15.2.10) respectively gives

$$\begin{aligned} D(x,t;\lambda) &= K(x,t) - \lambda B_1(x,t) + \frac{\lambda^2}{2!}B_2(x,t) - \cdots = xe^t \\ D(\lambda) &= 1 - \lambda C_1 + \frac{\lambda^2}{2!}C_2 - \ldots = 1 - \lambda \end{aligned}$$

Hence, Eq.(15.2.8) yields

$$R(x,t;\lambda) = \frac{D(x,t;\lambda)}{D(\lambda)} = \frac{xe^t}{1-\lambda}$$

Hence the required solution is

$$\begin{aligned} y(x) &= f(x) + \lambda \int_0^1 R(x,t;\lambda)\, f(t)\, dt \\ \Rightarrow y(x) &= f(x) + \lambda \int_0^1 \frac{xe^t}{1-\lambda} f(t)\, dt \\ \Rightarrow y(x) &= f(x) + \frac{\lambda x}{1-\lambda} \int_0^1 e^t\, f(t)\, dt. \end{aligned}$$

Alternative Method : We shall use the results of Eqs.(15.2.13) to compute $R(x,t;\lambda)$ as follows. First write down these results for complete solution. Here

$$C_0 = 1$$

$$B_0(x,t) = K(x,t) = xe^t$$

$$C_1 = \int_0^1 B_0(s,s)\, ds = \int_0^1 se^s\, ds = \left[se^s\right]_0^1 - \int_0^1 e^s\, ds = e - \left[e^s\right]_0^1 = e - (e-1) = 1$$

$$B_1 = C_1K(x,t) - \int_0^1 K(x,z)B_0(z,t)\, dz = xe^t - \int_0^1 xe^zze^t\, dz == xe^t - xe^t \int_0^1 ze^z\, dz = 0$$

$$C_2 = \int_0^1 B_1(s,s)\, ds = 0$$

$$B_2(x,t) = C_2K(x,t) - 2\int_0^1 K(x,z)\, B_1(z,t)\, dz = 0$$

$$\therefore B_m(x,t) = 0 \quad \text{for all } m \geq 1 \quad \text{and} \quad C_m = 0 \quad \text{for all } m \geq 2.$$

Now we proceed as before to determine $R(x,t;\lambda)$ and can solve the given Fredholm integral equation.

Important Observation : The reader will find that the above alternative method is a short cut. However, he should find the required quantities strictly in the following order :

$$C_0, \quad B_0(x,t), \quad C_1, \quad B_1(x,t), \quad C_2, \quad B_2(x,t) \quad \text{and so on.}$$

Exercise 15.2.3. Using Fredholm determinants, find the resolvent kernel and hence solve the following integral equation

$$i)\ y(x) = e^{-x} + \lambda \int_0^1 xe^t\, y(t)\, dt, \qquad ii)\ y(x) = 1 + \int_0^1 (1-3xt)\, y(t)\, dt,$$

$$iii)\ y(x) = \sin x + \lambda \int_4^{10} x\, y(t)\, dt, \qquad iv)\ y(x) = 1 + \int_0^{\pi} \sin(x+t)\, y(t)\, dt$$

Answers :

$$i)\ y(x) = e^{-x} + \frac{\lambda x}{1-\lambda},\ \text{if } \lambda \neq 1 \qquad (ii)\ y(x) = \frac{8-6x}{3},$$

$$(iii)\ y(x) = \sin x + \frac{2\lambda x \sin 7 \sin 3}{1-42\lambda}, \qquad (iv)\ y(x) = 1 + \frac{4}{4-\pi^2}(2\cos x + \pi \sin x)$$

15.2.3 Fredholm Second Fundamental Theorem

If λ_0 is a zero of multiplicity m of the function $D(\lambda)$, then the homogeneous integral equation

$$y(x) = \lambda_0 \int_a^b K(x,t)\, y(t)\, dt \tag{15.2.15}$$

possesses at least one, and the most m, linearly independent solutions

$$y_i(x) = D_r\left(\begin{matrix} x_1, & \cdots, & x_{i-1}, & x, & x_{i+1}, & \cdots, & x_r \\ t_1, & \cdots, & t_{i-1}, & t, & t_{i+1}, & \cdots, & t_r \end{matrix} \,\middle|\, \lambda_0\right), \quad i = 1,2,\ldots,r; \quad 1 \le r \le m \tag{15.2.16}$$

not identically zero. Any other solution of this equation is a linear combination of these solutions. Here, we have to remember the following definition of the Fredholm minor

$$D_n\begin{pmatrix} x_1, & x_2, & \cdots, & x_n \\ t_1, & t_2, & \cdots, & t_n \end{pmatrix} = K\begin{pmatrix} x_1, & x_2, & \cdots, & x_n \\ t_1, & t_2, & \cdots, & t_n \end{pmatrix}$$
$$+ \sum_{p=1}^{\infty} \frac{(-\lambda)^p}{p!} \int_a^b \cdots \int_a^b K\begin{pmatrix} x_1, & \cdots, & x_n, & z_1 & \cdots & z_p \\ t_1, & \cdots, & t_n, & z_1 & \cdots & z_p \end{pmatrix} dz_1 \cdots dz_p, \tag{15.2.17}$$

where $\{x_i\}$ and $\{t_i\},\ i = 1,2,\ldots,n$, are two sequences of arbitrary variables. Series (15.2.17) converges for all values of λ and hence it is an entire function of λ.

15.2.4 Fredholm Third Fundamental Theorem

For an inhomogeneous integral equation

$$y(x) = f(x) + \lambda_0 \int_a^b K(x,t)\, y(t)\, dt \tag{15.2.18}$$

to possesses a solution in the case $D(\lambda_0) = 0$, it is necessary and sufficient that the given function $f(x)$ be orthogonal to all the eigenfunctions $z_i(x), i = 1, 2, \ldots, \nu$, of the transposed homogeneous equation corresponding to the eigenvalue λ_0. The general solution has the form

$$y(x) = f(x) + \lambda \int_a^b \frac{D_{r+1}\left(\begin{array}{ccccc|c} x, & x_1, & x_2, & \cdots, & x_r & \\ & & & & & \lambda_0 \\ t, & t_1, & t_2, & \cdots, & t_r & \end{array}\right)}{D_r\left(\begin{array}{cccc|c} x_1, & x_2, & \cdots, & x_r & \\ & & & & \lambda_0 \\ t_1, & t_2, & \cdots, & t_r & \end{array}\right)} f(t)\, dt + \sum_{h=1}^{r} C_h \Phi_h(x), \tag{15.2.19}$$

where $\Phi_i(x)$ are given by

$$\Phi_i(x) = \frac{D_r\left(\begin{array}{ccccccc|c} x_1, & \cdots, & x_{i-1}, & x, & x_{i+1}, & \cdots, & x_r & \\ & & & & & & & \lambda_0 \\ t_1, & \cdots & \cdots & & \cdots & \cdots & t_r & \end{array}\right)}{D_r\left(\begin{array}{ccccccc|c} x_1, & \cdots, & x_{i-1}, & x_i, & x_{i+1}, & \cdots, & x_r & \\ & & & & & & & \lambda_0 \\ t_1, & \cdots & \cdots & & \cdots & \cdots & t_r & \end{array}\right)}, \qquad i = 1, 2, \ldots, r. \tag{15.2.20}$$

Unit 16

Course Structure

- Method of successive approximations: Solution of Fredholm and Voltera integral equation of the second kind by successive substitutions & Iterative method (Fredholm integral equation only), reciprocal function, determination of resolvent kernel and solution of Fredholm integral equation.

16.1 Solution of Fredholm integral equations

16.1.1 Method of Successive approximations :

Consider the Fredholm integral equation of the second kind given by Eq. (16.1.4). As a zero-order approximation to the required solution $y(x)$, let us take $y_0(x) = f(x)$. Further, if $y_n(x)$ and $y_{n-1}(x)$ are the n-th and $(n-1)$-th order approximations respectively, then these are connected by

$$y_n(x) = f(x) + \lambda \int_a^b K(x,t) y_{n-1}(t) dt. \tag{16.1.1}$$

We know that the iterated kernels (or iterated functions) $K_n(x,t), (n = 1, 2, 3, \ldots)$ are defined by

$$\begin{aligned} K_1(x,t) &= K(x,t) \\ \text{and} \quad K_n(x,t) &= \int_a^b K(x,z) K_{n-1}(z,t) dz. \end{aligned}$$

Putting $n = 1$ in Eq.(16.1.1), the first-order approximation $y_1(x)$ is given by

$$\begin{aligned} y_1(x) &= f(x) + \lambda \int_a^b K(x,t) y_0(t) dt. \\ \Rightarrow y_1(x) &= f(x) + \lambda \int_a^b K(x,t) f(t) dt. \end{aligned} \tag{16.1.2}$$

Putting $n = 2$ in Eq.(16.1.1), the second-order approximation $y_2(x)$ is given by

$$\begin{aligned}
y_2(x) &= f(x) + \lambda \int_a^b K(x,t) y_1(t)\, dt. \\
\Rightarrow y_2(x) &= f(x) + \lambda \int_a^b K(x,z) y_1(z)\, dz \\
\Rightarrow y_2(x) &= f(x) + \lambda \int_a^b K(x,z) \left[f(z) + \lambda \int_a^b K(z,t)\, f(t)\, dt \right] dz \\
\Rightarrow y_2(x) &= f(x) + \lambda \int_a^b K(x,t)\, f(t)\, dt + \lambda^2 \int_a^b f(t) \left[\int_a^b K(x,z) K(z,t)\, dz \right]\, dt \\
\Rightarrow y_2(x) &= f(x) + \lambda \int_a^b K_1(x,t)\, f(t)\, dt + \lambda^2 \int_a^b K_2(x,t) f(t)\, dt
\end{aligned}$$

$$\Rightarrow y_2(x) = f(x) + \sum_{m=1}^{2} \lambda^m \int_a^b K_m(x,t) f(t)\, dt.$$

Proceeding likewise, we easily obtain by Mathematical induction the n-th approximate solution $y_n(x)$ as

$$y_n(x) = f(x) + \sum_{m=1}^{n} \lambda^m \int_a^b K_m(x,t)\, f(t)\, dt. \tag{16.1.3}$$

16.1.2 Resolvent kernel :

Suppose solution of Fredholm integral equation of the second kind

$$y(x) = f(x) + \lambda \int_a^b K(x,t) y(t) dt \tag{16.1.4}$$

takes the form

$$y(x) = f(x) + \lambda \int_a^b R(x,t;\lambda) f(t) dt, \tag{16.1.5}$$

then $R(x,t;\lambda)$ is known as the resolvent kernal of (16.1.4). If $K_n(x,t)$ be iterated kernals then

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t)$$

Example 16.1.1. Find the iterated kernels for the following kernels

$$K(x,t) = \sin(x-t),\ 0 \le x \le 2\pi,\ 0 \le t \le 2\pi$$

Solution : Iterated kernel $K_n(x,t)$ are given by

$$K_1(x,t) = K(x,t) \tag{16.1.6}$$

$$\text{and} \quad K_n(x,t) = \int_0^{2\pi} K(x,z) K_{n-1}(z,t)\, dz, \ \ (n = 2, 3, \ldots) \tag{16.1.7}$$

From Eq.(16.1.6) $K_1(x,t) = K(x,t) = \sin(x-2t)$. Putting $n=2$ in Eq.(16.1.7), we have

$$\begin{aligned} K_2(x,t) &= \int_0^{2\pi} K(x,z)K_1(z,t)\,dz = \int_0^{2\pi} \sin(x-2z)\sin(z-2t)\,dz \\ &= \frac{1}{2}\int_0^{2\pi} [\cos(x+2t-3z)-\cos(x-2t-z)]\,dz = \frac{1}{2}\left[-\frac{1}{3}\sin(x+2t-3z)+\sin(x-2t-z\right] \\ &= 0, \quad \text{on simplification.} \end{aligned}$$

Putting $n=3$ in Eq.(16.1.7), we have

$$K_3(x,t) = \int_0^{2\pi} K(x,z)K_2(z,t)\,dz = 0 \quad [\because K_2(z,t)=0]$$

Thus, $K_1(x,t) = \sin(x-2t)$ and $K_n(x,t) = 0$ for $n = 2,3,4,\ldots$.

Example 16.1.2. Determine the resolvent kernels for the Fredholm integral equation having kernels

$$K(x,t) = e^{x+t};\ a=0,\ b=1.$$

Solution : Iterated kernels $K_m(x,t)$ are given by

$$K_1(x,t) = K(x,t) \tag{16.1.8}$$

$$K_m(x,t) = \int_0^1 K(x,z)K_{m-1}(z,t)\,dz \tag{16.1.9}$$

From Eq.(16.1.8) $K_1(x,t) = K(x,t) = e^{x+t}$.

Putting $n=2$ in Eq.(16.1.9), we have

$$\begin{aligned} K_2(x,t) &= \int_0^1 K(x,z)K_1(z,t)\,dz = \int_0^1 e^{x+z}e^{z+t}\,dz \\ &= e^{x+t}\int_0^1 e^{2z}\,dz = e^{x+t}\left[\frac{1}{2}e^{2z}\right] = e^{x+t}\left(\frac{e^2-1}{2}\right), \end{aligned}$$

Putting $n=3$ in Eq.(16.1.9), we have

$$\begin{aligned} K_3(x,t) &= \int_0^1 K(x,z)K_2(z,t)\,dz = \int_0^1 e^{x+z}e^{z+t}\left(\frac{e^2-1}{2}\right)\,dz \\ &= e^{x+t}\left(\frac{e^2-1}{2}\right)\int_0^1 e^{2z}\,dz = e^{x+t}\left[\frac{1}{2}e^{2z}\right] = e^{x+t}\left(\frac{e^2-1}{2}\right)^2 \quad \text{and so on,} \end{aligned}$$

Observing above, we may write

$$K_m(x,t) = e^{x+t}\left(\frac{e^2-1}{2}\right)^{m-1}, \quad m = 1,2,3,\ldots$$

Now, the required resolvent kernel is given by

$$R(x,t;\lambda) = \sum_{m=1}^{\infty}\lambda^{m-1}K_m(x,t) = \sum_{m=1}^{\infty}\lambda^{m-1}e^{x+t}\left(\frac{e^2-1}{2}\right)^{m-1} = e^{x+t}\sum_{m=1}^{\infty}\left\{\frac{\lambda(e^2-1)}{2}\right\}^{m-1}$$

$$\text{But} \quad \sum_{m=1}^{\infty}\left\{\frac{\lambda(e^2-1)}{2}\right\}^{m-1} = 1 + \frac{\lambda(e^2-1)}{2} + \left\{\frac{\lambda(e^2-1)}{2}\right\}^2 + \cdots$$

which is an infinite geometric series with common ratio $\{\lambda(e^2-1)\}/2$.

$$\therefore \sum_{m=1}^{\infty}\left\{\frac{\lambda(e^2-1)}{2}\right\}^{m-1} = \frac{1}{1-\{\lambda(e^2-1)\}/2} = \frac{2}{2-\lambda(e^2-1)},$$

$$\text{provided} \quad \left|\frac{\lambda(e^2-1)}{2}\right| < 1 \quad \text{or} \quad |\lambda| < \frac{2}{e^2-1}$$

$$\text{Therefore} \quad R(x,t;\lambda) = \frac{2e^{x+t}}{2-\lambda(e^2-1)}, \quad \text{provided} \quad |\lambda| < \frac{2}{e^2-1}$$

16.1.3 Solution in terms of resolvent kernel :

Let the Fredholm integral is given by Eq.(16.1.4). Let $K_m(x,t)$ be the m-th iterated kernel and let $R(x,t;\lambda)$ be the resolvent kernel of Eq.(16.1.4). Then we have

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) \tag{16.1.10}$$

Suppose the sum of the infinite series (16.1.10) exists and so $R(x,t;\lambda)$ can be obtained in the closed form. Then, the required solution of Eq.(16.1.4) is given by

$$y(x) = f(x) + \lambda \int_a^b R(x,t;\lambda)\, f(t)\, dt \tag{16.1.11}$$

Example 16.1.3. Solve

$$y(x) = x + \int_0^{1/2} y(t)\, dt$$

Solution : Comparing the given equation with

$$y(x) = f(x) + \lambda \int_0^{1/2} K(x,t)\, y(t)\, dt,$$

$$\text{we have} \quad f(x) = x, \quad \lambda = 1, \quad K(x,t) = 1 \tag{16.1.12}$$

Let $K_m(x,t)$ be the m-th iterated kernel. Then, we have

$$K_1(x,t) = K(x,t) \tag{16.1.13}$$

$$K_m(x,t) = \int_0^1 K(x,z) K_{m-1}(z,t)\, dz \tag{16.1.14}$$

From (16.1.12), $K_1(x,t) = K(x,t) = 1$.

Putting $m=2$ in (16.1.14), we have

$$K_2(x,t) = \int_0^{1/2} K(x,z)K_1(z,t)\, dz = \int_0^{1/2} dz = [z]_0^{1/2} = \frac{1}{2}.$$

Putting $m=3$ in (16.1.14), we have

$$K_3(x,t) = \int_0^{1/2} K(x,z)K_2(z,t)\, dz = \int_0^{1/2} \frac{1}{2}\, dz = \left(\frac{1}{2}\right)^2.$$

Observing above we find

$$K_m(x,t) = \left(\frac{1}{2}\right)^{m-1}$$

Now, the resolvent kernel $R(x,t;\lambda)$ is given by

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t) = \sum_{m=1}^{\infty} \left(\frac{1}{2}\right)^{m-1}$$

$$\text{But} \quad \sum_{m=1}^{\infty} \left(\frac{1}{2}\right)^{m-1} = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \cdots$$

which is an infinite geometric series with common ratio $1/2$.

$$\therefore \quad \sum_{m=1}^{\infty} \left(\frac{1}{2}\right)^{m-1} = \frac{1}{1-(1/2)} = 2 \quad \text{and hence} \quad R(x,t;\lambda) = 2$$

Finally, the required solution of the given equation is given by

$$\begin{aligned} y(x) &= f(x) + \lambda \int_0^{1/2} R(x,t;\lambda)\, f(t)\, dt \\ \Rightarrow y(x) &= x + \int_0^{1/2} (2t)\, dt \\ \Rightarrow y(x) &= x + 2\left[\frac{t^2}{2}\right]_0^{1/2} = x + \frac{1}{4} \end{aligned}$$

Exercise 16.1.4. Solve

$$i) \quad y(x) = e^x - \frac{1}{2}e + \frac{1}{2} + \frac{1}{2}\int_0^1 y(t)\, dt \qquad ii) \quad y(x) = \frac{5x}{6} + \frac{1}{2}\int_0^1 xt\, y(t)\, dt$$

$$iii) \quad y(x) = \sin x - \frac{x}{4} + \frac{1}{4}\int_0^{\pi/2} xt\, y(t)\, dt \qquad iv) \quad y(x) = \frac{3}{2}e^x - \frac{1}{2}xe^x - \frac{1}{2} + \frac{1}{2}\int_0^1 t\, y(t)\, dt$$

Answers :

$$i) \quad y(x) = e^x, \quad (ii) \quad y(x) = x, \quad (iii) \quad y(x) = \sin x, \quad (iv) \quad y(x) = \frac{3e^x}{2} - \frac{xe^x}{2} - \frac{e}{3} + 1$$

16.1.4 Iterative scheme for Fredholm integral equations

When the resolvent kernel cannot be obtained in closed form i.e., the sum of infinite series occurring in the formula of the resolvent kernel can not be determined, we use the method of successive approximations to find solutions upto third order.

Let the given Fredholm integral equation of the second kind be

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{16.1.15}$$

As zero-order approximation, we take $y_0(x) = f(x)$. If n-th order approximation be $y_n(x)$, then

$$y_n(x) = f(x) + \lambda \int_a^b K(x,t) y_{n-1}(x) \tag{16.1.16}$$

Sometimes the zero-order approximation is mentioned in the problem. In that case, we will modify the scheme accordingly.

Example 16.1.5. Solve the following integral equation

$$y(x) = 1 + \lambda \int_0^1 (x+t)\, y(t)\, dt,$$

by the method of successive approximation to third order.

Solution : Given

$$y(x) = 1 + \lambda \int_0^1 (x+t)\, y(t)\, dt \tag{16.1.17}$$

Let $y_0(x)$ denote the zero-order approximation. Then we may take

$$y_0 = 1.$$

If $y_n(x)$ denotes the n-th order approximation, then we know that

$$y_n(x) = 1 + \lambda \int_0^1 (x+t)\, y_{n-1}(t)\, dt, \tag{16.1.18}$$

Putting $n = 1$ in (16.1.18),

$$\begin{aligned} y_1(x) &= 1 + \lambda \int_0^1 (x+t)\, y_0(t)\, dt = 1 + \lambda \int_0^1 (x+t)\, dt, \\ \Rightarrow y_1(x) &= 1 + \lambda \left[xt + \frac{1}{2} t^2 \right]_0^1 = 1 + \lambda \left(x + \frac{1}{2} \right). \end{aligned} \tag{16.1.19}$$

Next, putting $n = 2$ in (16.1.18), we have

$$\begin{aligned} y_2(x) &= 1 + \lambda \int_0^1 (x+t)\, y_1(t)\, dt = 1 + \lambda \int_0^1 (x+t) \left\{ 1 + \lambda \left(t + \frac{1}{2} \right) \right\}, \\ &= 1 + \lambda \int_0^1 (x+t) \left\{ \left(1 + \frac{\lambda}{2} \right) + \lambda t \right\} dt = 1 + \lambda \int_0^1 \left[x \left(1 + \frac{\lambda}{2} \right) + t \left(1 + \frac{\lambda}{2} + \lambda x \right) + \lambda t^2 \right] dt \\ &= 1 + \lambda \left[x \left(1 + \frac{\lambda}{2} \right) t + \frac{t^2}{2} \left(1 + \frac{\lambda}{2} + \lambda x \right) + \frac{\lambda t^3}{3} \right]_0^1 \\ &= 1 + \lambda \left[x \left(1 + \frac{\lambda}{2} \right) t + \frac{1}{2} \left(1 + \frac{\lambda}{2} + \lambda x \right) + \frac{\lambda}{3} \right] = 1 + \lambda \left(x + \frac{1}{2} \right) + \lambda^2 \left(x + \frac{7}{12} \right) \end{aligned}$$

Finally, putting $n = 3$ in (16.1.18), we have

$$\begin{aligned}
y_3(x) &= 1+\lambda\int_0^1 (x+t)\,y_2(t)\,dt = 1+\lambda\int_0^1 (x+t)\left\{1+\lambda\left(t+\frac{1}{2}\right)+\lambda^2\left(t+\frac{7}{2}\right)\right\},\\
&= 1+\lambda\int_0^1 (x+t)\left\{\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}\right)+\lambda t(1+\lambda)\right\}dt\\
&= 1+\lambda\int_0^1 \left[x\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}\right)+t\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}+\lambda x+\lambda^2 x\right)+\lambda t^2(1+\lambda)\right]dt\\
&= 1+\lambda\left[x\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}\right)t+\frac{t^2}{2}\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}+\lambda x+\lambda^2 x\right)+\frac{\lambda 1}{3}\lambda t^3(1+\lambda)\right]_0^1\\
&= 1+\lambda x\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}\right)t+\frac{\lambda}{2}\left(1+\frac{\lambda}{2}+\frac{7\lambda^2}{12}+\lambda x+\lambda^2 x\right)+\frac{1}{3}\lambda^2(1+\lambda)
\end{aligned}$$

$$\text{Therefore,}\quad y_3(x) = 1+\lambda\left(x+\frac{1}{2}\right)+\lambda^2\left(x+\frac{7}{12}\right)+\lambda^3\left(\frac{13}{12}x+\frac{5}{8}\right)$$

Exercise 16.1.6. Exercise : Solve the inhomogeneous Fredholm integral equation of the second kind

$$y(x) = 2x+\lambda\int_0^1 (x+t)\,y(t)\,dt,$$

by the method of successive approximations to the third order by taking $y_0(x) = 1$.

Answers :

$$y_3(x) = 2x+\lambda\left(x+\frac{2}{3}\right)+\lambda^2\left(\frac{7}{6}x+\frac{2}{3}\right)+\lambda^3\left(\frac{13}{12}x+\frac{5}{8}\right)$$

16.1.5 Solution of Homogeneous Fredholm Integral Equation of the second kind with separable (or Degenerate) kernel

Consider a homogeneous Fredholm integral equation of the second kind

$$y(x) = \lambda\int_a^b K(x,t)y(t)dt. \tag{16.1.20}$$

Since kernel $K(x,t)$ is separable, we take

$$K(x,t) = \sum_{i=1}^{n} f_i(x)g_i(t). \tag{16.1.21}$$

Using (16.1.21), (16.1.20) reduces to

$$y(x) = \lambda\int_a^b\left[\sum_{i=1}^{n} f_i(x)g_i(t)\right]y(t)dt$$

$$\text{or,}\quad y(x) = \lambda\sum_{i=1}^{n} f_i(x)\int_a^b g_i(t)y(t)dt\ \ \text{[Interchanging the order of summation and integration]} \tag{16.1.22}$$

Let

$$\int_a^b g_i(t)y(t)dt = C_i,\quad \text{where } i = 1,2,\ldots,n. \tag{16.1.23}$$

Using (16.1.23), (16.1.22) reduces to

$$y(x) = \lambda \sum_{i=1}^{n} C_i f_i(x) \tag{16.1.24}$$

where constants C_i $(i = 1, \dots, n)$ are to be determined in order to find the solution of (16.1.20) in the form given by (16.1.24).

We now proceed to evaluate C_i's as follows:

Multiplying both sides of (16.1.24) successively by $g_1(x), g_2(x), \dots, g_n(x)$ and integrating over the interval (a, b), we have the following system

$$\begin{aligned}
\int_a^b g_1(x)y(x)dx &= \lambda \sum_{i=1}^{n} C_i \int_a^b g_1(x)f_i(x)dx \\
\int_a^b g_2(x)y(x)dx &= \lambda \sum_{i=1}^{n} C_i \int_a^b g_2(x)f_i(x)dx \\
&\dots \\
&\dots \\
\int_a^b g_n(x)y(x)dx &= \lambda \sum_{i=1}^{n} C_i \int_a^b g_n(x)f_i(x)dx.
\end{aligned}$$

Let

$$\alpha_{ji} = \int_a^b g_j(x)f_i(x)dx, \quad \text{where } i, j = 1, \dots, n. \tag{16.1.25}$$

Using (16.1.23) and (16.1.25), the first equation of the above system reduces to

$$\begin{aligned}
& C_1 = \lambda \sum_{i=1}^{n} C_i \alpha_{1i} \\
\text{or,} \quad & C_1 = \lambda[C_1\alpha_{11} + C_2\alpha_{12} + \dots + C_n\alpha_{1n}] \\
\text{or,} \quad & (1 - \lambda\alpha_{11})C_1 - \lambda\alpha_{12}C_2 - \dots - \lambda\alpha_{1n}C_n = 0.
\end{aligned}$$

Similarly, simplifying the other equations, we finally obtain the following system of homogeneous linear equations to determine $C_1, C_2, \dots, C_n$.

$$\begin{aligned}
(1 - \lambda\alpha_{11})C_1 - \lambda\alpha_{12}C_2 - \dots - \lambda\alpha_{1n}C_n &= 0 \\
-\lambda\alpha_{21}C_1 + (1 - \lambda\alpha_{22})C_2 - \dots - \lambda\alpha_{2n}C_n &= 0 \\
&\dots \\
&\dots \\
-\lambda\alpha_{n1}C_1 - \lambda\alpha_{n2}C_2 - \dots + (1 - \lambda\alpha_{nn})C_n &= 0.
\end{aligned}$$

The determinant $D(\lambda)$ of this system is

$$D(\lambda) = \begin{vmatrix} 1 - \lambda\alpha_{11} & -\lambda\alpha_{12} & \dots & -\lambda\alpha_{1n} \\ -\lambda\alpha_{21} & 1 - \lambda\alpha_{22} & \dots & -\lambda\alpha_{2n} \\ \dots & & & \\ \dots & & & \\ -\lambda\alpha_{n1} & -\lambda\alpha_{n2} & \dots & 1 - \lambda\alpha_{nn} \end{vmatrix}. \tag{16.1.26}$$

If $D(\lambda) \neq 0$, the system of equations for C_i's have only trivial solution $C_1 = C_2 = \dots = C_n = 0$ and hence from (16.1.24), we notice that equation (16.1.20) has only zero or trivial solution $y(x) = 0$.

However, if $D(\lambda) = 0$, at least one of the C_i's can be assigned arbitrarily , and the remaining C_i's can be determined accordingly. Hence, when $D(\lambda) = 0$, infinitely many solutions of the equation (16.1.20) exist.

Those values of λ for which $D(\lambda) = 0$ are called the eigen values, and any non-trivial solution of (16.1.20) is called a corresponding eigen function of (16.1.20).

The eigen values of (16.1.20) are gievn by $D(\lambda) = 0$, that is,

$$\begin{vmatrix} 1-\lambda\alpha_{11} & -\lambda\alpha_{12} & \dots & -\lambda\alpha_{1n} \\ -\lambda\alpha_{21} & 1-\lambda\alpha_{22} & \dots & -\lambda\alpha_{2n} \\ \dots & & & \\ \dots & & & \\ -\lambda\alpha_{n1} & -\lambda\alpha_{n2} & \dots & 1-\lambda\alpha_{nn} \end{vmatrix} = 0 \tag{16.1.27}$$

So the degree of (16.1.27) in λ is $m \leq n$. It follows that if integral equation (16.1.20) has separable kernel given by (16.1.21) , then (16.1.20) has at most n eigen values.

Example 16.1.7. Solve the homogeneous Fredholm equation

$$y(x) = \lambda \int_0^1 \mathrm{e}^x \, \mathrm{e}^t \, y(t)dt.$$

OR

Find the eigen values and eigen functions of the homogeneous integral equation

$$y(x) = \lambda \int_0^1 \mathrm{e}^x \, \mathrm{e}^t \, y(t)dt.$$

Solution. Given

$$y(x) = \lambda \int_0^1 \mathrm{e}^x \, \mathrm{e}^t \, y(t)dt$$

$$\text{or,} \quad y(x) = \lambda \, \mathrm{e}^x \int_0^1 \mathrm{e}^t \, y(t)dt. \tag{16.1.28}$$

Let

$$C = \int_0^1 \mathrm{e}^t \, y(t)dt. \tag{16.1.29}$$

Then (16.1.28) reduces to

$$y(x) = \lambda C \, \mathrm{e}^x \, . \tag{16.1.30}$$

From (16.1.30),

$$y(t) = \lambda C \, \mathrm{e}^t \, . \tag{16.1.31}$$

Using (16.1.31), (16.1.29) becomes

$$C = \int_0^1 \mathrm{e}^t \left(\lambda C \, \mathrm{e}^t\right) dt$$

$$\text{or,} \quad C = \lambda C \left[\frac{\mathrm{e}^{2t}}{2}\right]_0^1$$

$$\text{or,} \quad C = \frac{\lambda C}{2}(\mathrm{e}^2 - 1)$$

$$\text{or,} \quad C \left[1 - \frac{\lambda}{2}(\mathrm{e}^2 - 1)\right] = 0. \tag{16.1.32}$$

If $C = 0$, then (16.1.31) gives $y(x) = 0$. We therefore assume that for non-zero solution of (16.1.28), $C \neq 0$. Then (16.1.32) gives

$$1 - \frac{\lambda}{2}(e^2 - 1) = 0$$

$$\text{or,} \quad \lambda = \frac{2}{\mathrm{e}^2 - 1}, \tag{16.1.33}$$

which is an eigen value of (16.1.28).

Putting the values of λ given by (16.1.33) in (16.1.30), the corresponding eigen function is given by

$$y(x) = \frac{2C}{e^2 - 1}\,\mathrm{e}^x\,.$$

Hence, corresponding to the eigen value $\frac{2}{\mathrm{e}^2-1}$, the eigen function is e^x. ■

Example 16.1.8. Find the eigen values and the corresponding eigen functions of the homogeneous integral equation

$$y(x) = \lambda \int_0^1 \sin \pi x \cos \pi t y(t) dt.$$

Solution. Given

$$y(x) = \lambda \int_0^1 \sin \pi x \cos \pi t y(t) dt$$

$$\text{or,} \quad y(x) = \lambda \sin \pi x \int_0^1 \cos \pi t y(t) dt. \tag{16.1.34}$$

Let

$$C = \int_0^1 \cos \pi t y(t) dt. \tag{16.1.35}$$

Then (16.1.34) reduces to

$$y(x) = C\lambda \sin \pi x. \tag{16.1.36}$$

From (16.1.36)

$$y(t) = C\lambda \sin \pi t. \tag{16.1.37}$$

Using (16.1.37), (16.1.35) becomes

$$C = \int_0^1 \cos \pi t(\lambda C \sin \pi t) dt$$

$$\text{or,} \quad C = \frac{\lambda C}{2} \int_0^1 \sin 2\pi t dt$$

$$\text{or,} \quad C = \frac{\lambda C}{2}\left[-\frac{1}{2\pi} + \frac{1}{2\pi}\right] = 0.$$

So from (16.1.36), $y(x) = 0$. Thus for any λ, (16.1.34) has only zero solution $y(x) = 0$. Therefore, (16.1.34) does not possess any characteristic number or eigen function. ■

Exercise 16.1.9. 1. Find the eigen values and eigen functions of the integral equation

$$y(x) = \lambda \int_0^1 (2xt - 4x^2) y(t) dt.$$

ANS: $\lambda_1 = -3$, $\lambda_2 = -3$, $y(x) = (x - 2x^2)$.

2. Find the eigen values and eigen functions of the homogeneous integral equation

$$y(x) = \int_{-1}^{1} (5xt^3 + 4x^2t + 3xt)y(t)dt.$$

ANS: $\lambda = \frac{1}{4}$, $y(x) = x^2 + \frac{3}{2}x$.

3. Solve:

$$y(x) = \frac{1}{50} \int_{0}^{10} ty(t)dt.$$

ANS: $y(x) = 0$.

16.1.6 Some Problems on Fredholm Integral Equation of the Second kind with separable kernels

Example 16.1.10. Find the solution of the integral equation

$$g(s) = s + \int_{0}^{1} su^2 g(u)du.$$

Solution. Given

$$g(s) = s + s\int_{0}^{1} u^2 g(u)du. \tag{16.1.38}$$

Let

$$C = \int_{0}^{1} u^2 g(u)du. \tag{16.1.39}$$

Using (16.1.39), (16.1.38) yields

$$g(s) = s + Cs = s(1 + C). \tag{16.1.40}$$

From (16.1.40),

$$g(u) = u(1 + C). \tag{16.1.41}$$

Using (16.1.41), (16.1.39) yields

$$C = \int_{0}^{1} u^3(1 + C)du$$

$$\text{or,} \quad C = \frac{1}{3}.$$

Hence,

$$g(s) = s\left(1 + \frac{1}{3}\right) = \frac{4s}{3}$$

and so $g(t) = \frac{4t}{3}$. ■

Example 16.1.11. Solve

$$y(x) = \cos x + \lambda \int_{0}^{\pi} \sin xy(t)dt.$$

Solution. Given

$$
\begin{aligned}
& y(x)=\cos x+\lambda \int_{)}^{\pi} \sin x y(t) d t \\
\text{or,} \quad & y(x)=\cos x+\lambda \sin x \int_{)}^{\pi} y(t) d t. \qquad (16.1.42)
\end{aligned}
$$

Let

$$C=\int_{0}^{\pi} y(t) d t. \qquad (16.1.43)$$

Using (16.1.43), (16.1.42) becomes

$$y(x)=\cos x+\lambda C \sin x. \qquad (16.1.44)$$

From (16.1.44),

$$y(t)=\cos t+\lambda C \sin t. \qquad (16.1.45)$$

Using (16.1.45), (16.1.43) reduces to

$$
\begin{aligned}
& C=\int_{0}^{\pi}(\cos t+\lambda C \sin t) d t \\
\text{or,} \quad & \quad =[\sin t]_{0}^{\pi}+\lambda C[-\cos t]_{0}^{\pi} \\
\text{or,} \quad & C(1-2 \lambda)=0.
\end{aligned}
$$

So that $C=0$ if $\lambda \neq 1/2$. Hence, by (16.1.44), the required solution is

$$y(x)=\cos x,$$

provided $\lambda \neq 1/2$. ■

Example 16.1.12. Solve

$$y(x)=f(x)+\lambda \int_{0}^{1} x t y(t) d t.$$

Solution. Given,

$$
\begin{aligned}
& y(x)=f(x)+\lambda \int_{0}^{1} x t y(t) d t \\
\text{or,} \quad & y(x)=f(x)+\lambda x \int_{0}^{1} t y(t) d t. \qquad (16.1.46)
\end{aligned}
$$

Let

$$C=\int_{0}^{1} t y(t) d t. \qquad (16.1.47)$$

Then (16.1.46) reduces to

$$y(x)=f(x)+\lambda C x. \qquad (16.1.48)$$

From (16.1.48),

$$y(t)=f(t)+\lambda C t. \qquad (16.1.49)$$

Using (16.1.49), (16.1.46) reduces to

$$C = \int_0^1 t[f(t) + \lambda C t]dt$$

$$\text{or,} \quad C = \int_0^1 tf(t)dt + \lambda C \left[\frac{t^3}{3}\right]_0^1$$

$$\text{or,} \quad C = \int_0^1 tf(t)dt + \frac{\lambda C}{3}$$

$$\text{or,} \quad C\left(1 - \frac{\lambda}{3}\right) = \int_0^1 tf(t)dt$$

$$\text{or,} \quad C = \frac{3}{3-\lambda}\int_0^1 tf(t)dt$$

where $\lambda \neq 3$. Putting this value of C in (16.1.48), the required solution is

$$y(x) = f(x) + \frac{3\lambda x}{3-\lambda}\int_0^1 tf(t)dt$$

where $\lambda \neq 3$. ■

Exercise 16.1.13. Solve the following integral equations:

1. $y(x) = \tan x + \int_{-1}^1 \mathrm{e}^{\sin^{-1} x}\, y(t)dt;$
2. $y(x) = \sec^2 x + \lambda \int_0^1 y(t)dt;$
3. $y(x) = \dfrac{1}{\sqrt{1-x^2}} + \lambda \int_0^1 \cos^{-1} t y(t)dt;$
4. $y(x) = \mathrm{e}^x + \lambda \int_0^1 2\,\mathrm{e}^x\,\mathrm{e}^t\, y(t)dt.$

16.2 Solution of Volterra integral equations

16.2.1 Solution of Volterra integral equation in terms of resolvent kernel

Working Rule : Let

$$y(x) = f(x) + \lambda \int_a^x K(x,t)\, y(t)\, dt \tag{16.2.1}$$

be given Volterra integral equation. Let $K_m(x,t)$ be the m-th iterated kernel and let $R(x,t;\lambda)$ be the resolvent kernel of (16.2.1). Then we have

$$R(x,t;\lambda) = \sum_{m=1}^{\infty} \lambda^{m-1} K_m(x,t). \tag{16.2.2}$$

Suppose the sum of infinite series (16.2.2) exists and so $R(x,t;\lambda)$ can be obtained in the closed form. Then the required solution of (16.2.1) is given by

$$y(x) = f(x) + \lambda \int_a^x R(x,t;\lambda)\, f(t)\, dt. \tag{16.2.3}$$

Example 16.2.1. With the aid of the resolvent kernel, find the solution of the integral equation

$$y(x) = e^{x^2} + \int_0^x e^{x^2-t^2} y(t)\,dt.$$

Solution : Comparing the given equation with

$$y(x) = f(x) + \lambda \int_0^x K(x,t)\,y(t)\,dt$$

we have

$$f(x) = e^{x^2}, \qquad \lambda = 1, \qquad K(x,t) = e^{x^2-t^2}$$

Let $K_m(x,t)$ be the m-th iterated kernel. Then we have

$$K_1(x,t) = K(x,t)$$

$$\text{and} \quad K_m(x,t) = \int_t^x K(x,z) K_{m-1}(z,t)\,dz \tag{16.2.4}$$

Thus we have

$$K_1(x,t) = K(x,t) = e^{x^2-t^2} \tag{16.2.5}$$

Putting $m = 2$ in (16.2.4), we have

$$K_2(x,t) = \int_t^x K(x,z)K_1(z,t)\,dz = \int_t^x e^{x^2-z^2} e^{z^2-t^2} dz = e^{x^2-t^2} \int_t^x dz = e^{x^2-t^2}(x-t)$$

Putting $m = 3$ in (16.2.4), we have

$$\begin{aligned} K_3(x,t) &= \int_t^x K(x,z)K_2(z,t)\,dz = \int_t^x e^{x^2-z^2} e^{z^2-t^2}(z-t)\,dz = e^{x^2-t^2} \int_t^x (z-t)\,dz \\ &= e^{x^2-t^2} \left[\frac{(z-t)^2}{2}\right]_t^x = e^{x^2-t^2} \frac{(x-t)^2}{2!} \end{aligned}$$

Putting $m = 4$ in (16.2.4), we have

$$\begin{aligned} K_4(x,t) &= \int_t^x K(x,z)K_3(z,t)\,dz = \int_t^x e^{x^2-z^2} e^{z^2-t^2} \frac{(z-t)^2}{2!}\,dz \\ &= \frac{e^{x^2-t^2}}{2!} \left[\frac{(z-t)^3}{3}\right]_t^x = e^{x^2-t^2} \frac{(x-t)^3}{3!} \end{aligned}$$

Observing above by mathematical induction we may write

$$K_m(x,t) = e^{x^2-t^2} \frac{(x-t)^{m-1}}{(m-1)!}, \quad m = 1, 2, 3, \ldots \tag{16.2.6}$$

Now, by the definition of the resolvent kernel, we have

$$\begin{aligned} R(x,t;\lambda) &= \sum_{m=1}^{\infty} K_m(x,t) = K_1(x,t) + \lambda K_2(x,t) + \lambda^2 K_3(x,t) + \cdots \\ &= e^{x^2-t^2} + e^{x^2-t^2} \frac{(x-t)}{1!} + e^{x^2-t^2} \frac{(x-t)^2}{2!} + \cdots \\ &= e^{x^2-t^2} \left[1 + \frac{(x-t)}{1!} + \frac{(x-t)^2}{2!} + \cdots\right] \\ &= e^{x^2-t^2} e^{x-t} \end{aligned}$$

Finally, the required solution of the given equation is given by

$$\begin{aligned}
y(x) &= f(x) + \lambda \int_0^x R(x,t;\lambda)\, f(t)\, dt = e^{x^2} + \int_0^x e^{x^2-t^2} e^{x-t} e^{t^2}\, dt \\
&= e^{x^2} + e^{x^2+x} \int_0^x e^{-t}\, dt = e^{x^2} + e^{x^2+x}\big[-e^{-t}\big]_0^x \\
&= e^{x^2} + e^{x^2+x}[-e^{-x}+1] = e^{x^2} - e^{x^2} + e^{x^2+x} = e^{x^2+x}
\end{aligned}$$

Exercise 16.2.2. Solve the following integral equation by means of resolvent kernel

$$i) \quad y(x) = e^x \sin x + \int_0^x \frac{2+\cos x}{2+\cos t} y(t)\, dt \qquad ii) \quad y(x) = \cos x - x - 2 + \int_0^x (t-x)\, y(t)\, dt$$

Answers :

$$i) \quad y(x) = e^x \sin x - e^x(2+\cos x)\log\left(\frac{2+\cos x}{3}\right), \qquad (ii) \quad y(x) = -\cos x - \sin x - \frac{x}{2}\sin x$$

16.2.2 Method of Successive approximations for solving Volterra integral equation

Working Rule : Let $f(x)$ be continuous in $[0,a]$ and $K(x,t)$ be continuous for $0 \le x \le a,\ \ 0 \le t \le x$. We start with some function $y_0(x)$ continuous in $[0,a]$. Replacing $y(t)$ on R.H.S of (14.9.1) by $y_0(x)$, we obtain

$$y_1(x) = f(x) + \lambda \int_0^x K(x,t)\, y_0(t)\, dt. \tag{16.2.7}$$

$y_1(x)$ given by (16.2.7) is itself continuous in $[0,a]$. Proceeding likewise we arrive at a sequence of functions $y_0(x),\ y_1(x), \ldots, y_n(x), \ldots$, where

$$y_n(x) = f(x) + \lambda \int_0^x K(x,t)\, y_{n-1}(t)\, dt. \tag{16.2.8}$$

In view of continuity of $f(x)$ and $K(x,t)$, the sequence $\{y_n(x)\}$ converges, as $n \to \infty$ to obtain the solution of $y(x)$ of given integral equation (14.9.1). It should be note that when $y_0(x) = f(x)$, we obtain the so called *Neumann series*.

Example 16.2.3. Using the method of successive approximations, solve the integral equation

$$y(x) = 1 + \int_0^x y(t)\, dt, \quad \text{taking } \ y_0(x) = 0.$$

Solution : Comparing the given equation with

$$y(x) = f(x) + \lambda \int_0^x K(x,t)\, y(t)\, dt,$$

we find

$$f(x) = 1, \qquad \lambda = 1, \qquad K(x,t) = 1$$

The n-th order approximation is given by

$$y_n(x) = 1 + \int_0^x y_{n-1}(t)\, dt, \tag{16.2.9}$$

Putting $n = 1$ in (16.2.9), we have

$$y_1(x) = 1 + \int_0^x y_0(t)\,dt = 1 + \int_0^x (0)\,dt = 1.$$

Putting $n = 2$ in (16.2.9), we have

$$y_2(x) = 1 + \int_0^x y_1(t)\,dt = 1 + \int_0^x dt = 1 + [t]_0^x = 1 + x.$$

Putting $n = 3$ in (16.2.9), we have

$$y_3(x) = 1 + \int_0^x y_2(t)\,dt = 1 + \int_0^x (1+t)\,dt = 1 + \left[t + \frac{t^2}{2}\right]_0^x = 1 + x + \frac{x^2}{2!}.$$

Putting $n = 4$ in (16.2.9), we have

$$y_4(x) = 1 + \int_0^x y_3(t)\,dt = 1 + \int_0^x \left(1 + t + \frac{t^2}{2!}\right)\,dt = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}.$$

Observing the above trend, we find

$$y_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n-1}}{(n-1)!}$$

Making $n \to \infty$, we find the required solution is given by

$$\begin{aligned} y(x) &= \lim_{n\to\infty} y_n(x) \\ &= 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\cdots = e^x \end{aligned}$$

Exercise 16.2.4. Using the method of successive approximations, solve the integral following integral equations.

$i)$ $y(x) = 1 + x - \int_0^x y(t)\,dt,$ taking $y_0(x) = 1$ $ii)$ $y(x) = x - \int_0^x (x-t)\,y(t)\,dt,$ taking $y_0(x) = 0$

$iii)$ $y(x) = 1 + \int_0^x (x-t)\,y(t)\,dt,$ taking $y_0(x) = 1$ $iv)$ $y(x) = \frac{1}{2}x^3 - 2x - \int_0^x y(t)\,dt,$ taking $y_0(x) = x^2$

Answers :

$i)$ $y(x) = 1,$ (ii) $y(x) = \sin x,$ (ii) $y(x) = \cosh x,$ (iv) $y(x) = x^2 - 2x$

Unit 17

Course Structure

- Hilbert-Schmidt theory: Orthonormal system of function, fundamental properties of eigen value and function for symmetric kernel, Hilbert theorem, Hilbert-Schmidt theorem.

17.1 Hilbert-Schmidt Theory

17.1.1 Symmetric Kernels

A kernel is called *symmetric* if it coincides with its own complex conjugate. Such a kernel is characterized by the identity

$$K(x,t) = \overline{K}(t,x),$$

where the bar denotes the complex conjugate. If the kernel is real, then its *symmetry* is defined by the identity $K(x,t) = K(t,x)$. An integral equation with a symmetric kernel is called a symmetric equation.

Remark 17.1.1. For a symmetric kernel that is not identically zero, at least one eigenvalue will always exist. This is an important characteristic of symmetric kernel. An eigenvalue is *simple* if there is only one corresponding eigenfunction, otherwise the eigenvalues are degenerate. The *spectrum* of the kernel $K(x,t)$ is the set of all its eigenvalues. Thus the *spectrum of a symmetric kernel is never empty.*

17.1.2 Orthogonal system of functions

A finite or an infinite set $\{\phi_k(x)\}$ defined on an interval $a \le n \le b$ is said to be an orthogonal set if

$$(\phi_i, \phi_j) = 0 \quad \text{or} \quad \int_a^b \phi_i(x)\phi_j(x) = 0, \quad i \neq j. \tag{17.1.1}$$

If none of the elements of this set is a zero vector, then it is called a *proper orthogonal set.* The set $\{\phi_i(x)\}$ is *orthonormal* if

$$(\phi_i, \phi_j) = \int_a^b \phi_i(x)\phi_j(x)\,dx = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \tag{17.1.2}$$

Any function $\phi(x)$ for which $||\phi(x)|| = 1$ is said to be *normalized.*

Some examples of the complete orthogonal and orthonormal systems.

(i) The system $\phi_n(x) = (2\pi)^{-1/2}e^{inx}$, where n takes every integer value from $-\infty$ to ∞, is orthonormal in the interval $(-\pi, \pi)$.

(ii) The functions $1,\ \cos x,\ \cos 2x,\ \cos 3x, \ldots$ form an orthogonal system in the interval $(0, \pi)$. Again the functions $\sin x,\ \sin 2x,\ \sin 3x, \ldots$ also form an orthogonal system in $(0, \pi)$.

(iii) The Legendre polynomials given by

$$P_0(x) = 1, \quad P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2-1)^n)}{dx^n}, \quad n = 1, 2, 3, \ldots$$

are orthogonal in the interval $(-1, 1)$. It can be shown that

$$\int_{-1}^{1} P_m(x)P_n(x)\, dx = \begin{cases} 0, & \text{if } m \neq n, \\ 2/(2n+1), & \text{if } m = n. \end{cases}$$

(iv) The Chebychev polynomials $T_n(x) = 2^{1-n}\cos(n\ \cos^{-1} x),\ n = 0, 1, 2, 3, \ldots$ are orthogonal with weight $r(x) = 1/(1-x^2)^{1/2}$ in the interval $(-1, 1)$. They can be normalized by multiplying $T_n(x)$ by the quality $(2^{2n-1}/\pi)^{1/2}$.

17.1.3 Fundamental properties of eigenvalues and eigenfunctions of symmetric kernels

Theorem 17.1.2. *If a kernel is symmetric then all its iterated kernels are also symmetric.*

Proof. Let kernel $K(x, t)$ be symmetric. Then by definition

$$K(x, t) = \overline{K}(t, x). \tag{17.1.3}$$

By definition, the iterated kernels $K_n(x, t),\ n = 1, 2, 3, \ldots$ are defined as follows:

$$K_1(x, t) = K(x, t) \tag{17.1.4}$$

$$K_n(x, t) = \int_a^b K(x, z)K_{n-1}(z, t)\, dz,\ n = 2, 3, \ldots \tag{17.1.5}$$

$$\text{and } K_n(x, t) = \int_a^b K_{n-1}(x, z)K(z, t)\, dz,\ n = 2, 3, \ldots \tag{17.1.6}$$

We shall use mathematical induction to prove the required result. Now

$$\begin{aligned} K_2(x, t) &= \int_a^b K(x, z)K_1(z, t)\, dz = \int_a^b K(x, z)K(z, t)\, dz \\ &= \int_a^b \overline{K}(z, x)\overline{K}_1(t, z)\, dz = \int_a^b \overline{K}(t, z)\overline{K}_1(z, x)\, dz = \overline{K}_2(t, x) \end{aligned}$$

Thus,

$$K_2(x, t) = \overline{K}_2(t, x) \tag{17.1.7}$$

showing that $K_2(x, t)$ is symmetric. Hence the required result is true for $n = 1$ and $n = 2$.

Let $K_n(x,t)$ be symmetric for $n = m$. Then by definition, we have

$$K_m(x,t) = \overline{K}_m(t,x). \tag{17.1.8}$$

We shall now prove that $K_n(x,t)$ is also symmetric for $n = m+1$, i.e.,

$$K_{m+1}(x,t) = \overline{K}_{m+1}(t,x). \tag{17.1.9}$$

$$\begin{aligned}\text{L.H.S of (17.1.9)} = K_{m+1}(x,t) &= \int_a^b K(x,z)K_m(z,t)\,dz = \int_a^b \overline{K}(z,x)\overline{K}_m(t,z)\,dz \\ &= \int_a^b \overline{K}_m(t,z)\overline{K}(z,x)\,dz = \overline{K}_{m+1}(t,x) = \text{R.H.S of (17.1.9)}\end{aligned}$$

Thus iterated Kernel $K_n(x,t)$ is symmetric for $n = 1$ and $n = 2$. Moreover, $K_n(x,t)$ is symmetric for $n = m+1$ whenever it is symmetric for $n = m$. Hence, by the mathematical induction, $K_n(x,t)$ is symmetric for $n = 1, 2, 3, \ldots$.

□

Theorem 17.1.3. Hilbert Theorem

Every symmetric kernel with a norm not equal to zero has at least one eigenvalue.
OR, *If the kernel $K(x,t)$ is symmetrical and not identically equal to zero, then it has at least one eigenvalue.*

Theorem 17.1.4. *The eigenvalues of a symmetric kernel are real.* OR, *If K(x,t) is real, symmetric, continuous and identically not equal to zero, then all the characteristic constants (eigenvalues) are real.*

Proof. Let λ an $\phi(x)$ be an eigenvalue and a corresponding eigenfunction of the kernel $K(x,t)$. Then by definition

$$\phi(x) = \lambda \int_a^b K(x,t)\phi(t)\,dt \tag{17.1.10}$$

Multiplying (17.1.10) by $\overline{\phi}(x)$ and integrating with respect x from to x from a to b.

$$\int_a^b \phi(x)\overline{\phi}(x)\,dx = \lambda \int_a^b \left\{ \int_a^b K(x,t)\phi(t)\,dt \right\} \overline{\phi}(x)\,dx \tag{17.1.11}$$

By definition of Fredholm operator K, we have

$$K\phi = \int_a^b K(x,t)\phi(t)\,dt \tag{17.1.12}$$

$$\text{Also,} \quad ||\phi(x)|| = \int_a^b \{\phi(x)\overline{\phi}(x)\,dx\}^{1/2} \tag{17.1.13}$$

Using (17.1.12) and (17.1.13) and the definition of inner product, (17.1.11) reduces to

$$||\phi(x)||^2 = \lambda(K\phi,\phi) \quad \text{so that} \quad \lambda = ||\phi(x)||^2/(K\phi,\phi)$$

Since bothe the numerator and denominator are real, it follows that λ is also real and thus the required result is proved. □

Theorem 17.1.5. *The eigenfunctions of a symmetric kernel, corresponding to different eigenvalues are orthogonal.* OR *The fundamental functions (i.e. eigenfunctions) $\phi_m(x)$ and $\phi_n(x)$ of the symmetric kernel $K(x,t)$ for corresponding eigenvalues λ_m and λ_n ($\lambda_m \neq \lambda_n$) are orthogonal in the domain (a,b).*

Proof. Since $\phi_m(x)$ and $\phi_n(x)$ are eigenfunctions corresponding to eigenvalues λ_m and λ_n respectively, where $\lambda_m \neq \lambda_n$. Then, by definition, we have

$$\phi_m(x) = \lambda_m \int_a^b K(x,t)\phi_m(t)\, dt \tag{17.1.14}$$

$$\text{and} \quad \phi_n(x) = \lambda_n \int_a^b K(x,t)\phi_n(t)\, dt \tag{17.1.15}$$

Since λ_n is real, (17.1.15) may be re-written as $$\overline{\phi}(x) = \lambda_n \int_a^b \overline{K}(x,t)\overline{\phi}_n(t)\, dt \tag{17.1.16}$$

Since $K(x,t)$ is symmetric, we have $$\overline{K}(x,t) = K(x,t) \tag{17.1.17}$$

Using (17.1.17), (17.1.16) may be re-written as $$\overline{\phi}(x) = \lambda_n \int_a^b K(t,x)\overline{\phi}_n(t)\, dt \tag{17.1.18}$$

Interchanging x and t in (17.1.18), $we have$ $$\overline{\phi}_n(t) = \lambda_n \int_a^b K(x,t)\overline{\phi}_n(x)\, dx \tag{17.1.19}$$

Multiplying both sides of (17.1.14) by $\overline{\phi}_n(x)$ and then integrating the both sides w.r.t. 'x' from a to b, we have

$$\begin{aligned}
\int_a^b \phi_m(x)\overline{\phi}_n(x)\, dx &= \lambda_m \int_a^b \left\{ \int_a^b K(x,t)\phi_m(t)\, dt \right\} \overline{\phi}_n(x)\, dx \\
&= \lambda_m \int_a^b \left\{ \int_a^b K(x,t)\phi_n(x)\, dx \right\} \overline{\phi}_m(t)\, dt \quad \text{[on changing the order of integration]} \\
&= (\lambda_m/\lambda_n) \int_a^b \phi_m(t)\overline{\phi}_n(t)\, dt, \quad \text{by Eq.(17.1.19)}
\end{aligned}$$

$$\begin{aligned}
&\therefore \quad \int_a^b \phi_m(x)\overline{\phi}_n(x)\, dx = \lambda_m \int_a^b \phi_m(x)\overline{\phi}_n(x)\, dx \\
&\Rightarrow (\lambda_n - \lambda_m) \int_a^b \phi_m(x)\overline{\phi}_n(x)\, dx = 0, \\
&\Rightarrow (\lambda_n - \lambda_m)(\phi_m, \phi_n) = 0
\end{aligned}$$

Since $\lambda_n \neq \lambda_m$, $(\lambda_n - \lambda_m) \neq 0$ and so we have $(\phi_m, \phi_n) = 0$, showing that the eigenfunctions ϕ_m and ϕ_n are orthogonal. □

17.1.4 Hilbert-Schmidt Theorem

Theorem 17.1.6. Let $F(x)$ be generated from a continuous function $y(x)$ y the operator

$$\lambda \int_a^b K(x,t)\, y(t)\, dt$$

where $K(x,t)$ is continuous, real and symmetric, so that

$$F(x) = \lambda \int_a^b K(x,t)\, y(t)\, dt.$$

Then $F(x)$ can be represented over interval (a,b) by a linear combination of the normalized eigenfunctions of homogeneous integral equation

$$y(x) = \lambda \int_a^b K(x,t)\, y(t)\, dt,$$

having $K(x,t)$ as its kernel.

Result 17.1.7. Schmidt's Solution of non-homogeneous fredholm integral equation of second kind

Let

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, dt \tag{17.1.20}$$

be a non-homogeneous Fredholm integral equation of the second kind in which $K(x,t)$ is continuous, real and symmetric and λ is not an eigenvalue. Then the solution of Eq.(17.1.20) may be expressed as

$$y(x) = f(x) + \lambda \sum_m \frac{f_m}{\lambda_m - \lambda} \phi_m(x) \tag{17.1.21}$$

where $f_m = \int_a^b f(t)\phi_m(t)\, dt$. Solution exists *uniquely* if and only if λ does not take on an eigenvalue. If $\lambda = \lambda_k$, where λ_k is the k-th eigenvalue and eigenfunction $\phi_k(x)$ is not orthogonal to $f(x)$, then *no solution* exists. Finally, if $\lambda = \lambda_k$ and eigenfunction $\phi_k(x)$ is orthogonal to $f(x)$, then we have *infinitely many solutions of Eq.*(17.1.20).

Example 17.1.8. Solve the symmetric integral equation

$$y(x) = (x+1)^2 + \int_{-1}^{1} (xt + x^2t^2) y(t)\, dt,$$

by using Hilbert-Schmidt theorem.

Solution:

$$\text{Given} \quad y(x) = (x+1)^2 + \int_{-1}^{1} (xt + x^2t^2) y(t)\, dt, \tag{17.1.22}$$

$$\text{Comparing (17.1.22) with} \quad y(x) = f(x) + \lambda \int_{-1}^{1} (xt + x^2t^2) y(t)\, dt, \tag{17.1.23}$$

$$\text{here} \quad f(x) = (x+1)^2 \quad \text{and} \quad \lambda = 1 \tag{17.1.24}$$

We begin with determining eigenvalues and the corresponding normalized eigenfunctions of

$$y(x) = \lambda \int_{-1}^{1} (xt + x^2t^2) y(t)\, dt \tag{17.1.25}$$

$$\Rightarrow y(x) = \lambda x \int_{-1}^{1} t y(t)\, dt + \lambda x^2 \int_{-1}^{1} t^2 y(t)\, dt.$$

$$\text{Let} \quad C_1 = \int_{-1}^{1} t y(t)\, dt \quad \text{and} \quad C_2 = \int_{-1}^{1} t^2 y(t)\, dt$$

Thus we have from Eq.(17.1.25)

$$y(x) = \lambda C_1 x + \lambda C_2 x^2 \quad \Rightarrow \quad y(t) = \lambda C_1 t + \lambda C_2 t^2 \tag{17.1.26}$$

$$\begin{aligned} \text{Hence} \quad C_1 &= \int_{-1}^{1} t(\lambda C_1 + \lambda C_2 t^2)\, dt \quad \Rightarrow \quad C_1 = C_1\lambda \left[\frac{t^3}{3}\right]_{-1}^{1} + C_2\lambda \left[\frac{t^4}{4}\right]_{-1}^{1} \\ \Rightarrow C_1 &= \frac{2C_1\lambda}{3} + 0 \quad \Rightarrow \quad C_1\left(1 - \frac{2\lambda}{3}\right) + 0.C_2 = 0. \end{aligned} \tag{17.1.27}$$

$$\begin{aligned} \text{Again} \quad C_2 &= \int_{-1}^{1} t^2(\lambda C_1 + \lambda C_2 t^2)\, dt \quad \Rightarrow \quad C_2 = C_1\lambda \left[\frac{t^4}{4}\right]_{-1}^{1} + C_2\lambda \left[\frac{t^5}{5}\right]_{-1}^{1} \\ \Rightarrow C_2 &= 0 + \frac{2C_2\lambda}{5} \quad \Rightarrow \quad 0.C_1 + \left(1 - \frac{2\lambda}{5}\right) C_2 = 0. \end{aligned} \tag{17.1.28}$$

Equations (17.1.27) and (17.1.28) have a nontrivial solution only if

$$D(\lambda) = \begin{vmatrix} 1-(2\lambda/3) & 0 \\ 0 & 1-(2\lambda/5) \end{vmatrix} = 0$$

$$\Rightarrow \{1-(2\lambda/3)\}\{1-(2\lambda/5)\} = 0 \quad \text{giving} \quad \lambda = 3/2 \quad \text{and} \quad \lambda = 5/2$$

Hence the required eigenvalues are $\lambda_1 = 3/2$ and $\lambda_2 = 5/2$.

Determination of eigenfunction corresponding to $\lambda_1 = 3/2$

Putting $\lambda = \lambda_1 = 3/2$ in (17.1.27) and (17.1.28), we obtain

$$C_1 \cdot 0 + 0 \cdot C_2 = 0 \quad \text{and} \quad 0 \cdot C_1 + \left[1 - \left(\frac{2}{5} \times \frac{3}{2}\right)\right] C_2 = 0$$

Hence $C_2 = 0$ and C_1 is arbitrary. Putting these values in (17.1.26) and noting that $\lambda = 3/2$, we have the required eigenfunction $y_1(x)$ is given by

$$y(x) = (3/2) \times C_1 x$$

Setting $(3/2) \times C_1 = 1$, we may take $y_1(x) = x$. Now, the corresponding normalized eigenfunction $\phi_1(x)$ is given by

$$\phi_1(x) = \frac{y_1(x)}{\left[\int_{-1}^{1} \{y_1(x)\}^2\right]^{1/2}} = \frac{x}{\left[\int_{-1}^{1} x^2 dx\right]^{1/2}} = \frac{x}{\left\{[x^3/3]_{-1}^{1}\right\}^2} = \frac{x}{\sqrt{(2/3)}} = x \times \left(\frac{3}{2}\right)^{1/2} = \frac{x\sqrt{6}}{2}.$$

Determination of eigenfunction corresponding to $\lambda_1 = 5/2$

Putting $\lambda = \lambda_1 = 5/2$ in (17.1.27) and (17.1.28), we obtain

$$\left[1 - \left(\frac{2}{3} \times \frac{5}{2}\right)\right] C_1 + 0 \cdot C_2 = 0 \quad \text{and} \quad 0 \cdot C_1 + 0 \cdot C_2 = 0$$

Hence $C_1 = 0$ and C_2 is arbitrary. Putting these values in (17.1.26) and noting that $\lambda = 5/2$, we have the required eigenfunction $y_2(x)$ is given by

$$y(x) = (5/2) \times C_2 x^2$$

Setting $(5/2) \times C_2 = 1$, we may take $y_2(x) = x^2$. Now, the corresponding normalized eigenfunction $\phi_2(x)$ is given by

$$\phi_2(x) = \frac{y_2(x)}{\left[\int_{-1}^{1} \{y_2(x)\}^2\right]^{1/2}} = \frac{x^2}{\left[\int_{-1}^{1} x^4 dx\right]^{1/2}} = \frac{\sqrt{10}}{2} x^2.$$

Also,

$$\begin{aligned} f_1 &= \int_{-1}^{1} f(x)\phi_1(x)\, dx = \int_{-1}^{1} (x+1)^2 \left(\frac{\sqrt{6}}{2} x\right) dx \\ &= \frac{\sqrt{6}}{2} \int_{-1}^{1} (x^2 + 2x + 1)x\, dx = \frac{2\sqrt{6}}{3} \\ f_2 &= \int_{-1}^{1} f(x)\phi_2(x)\, dx = \int_{-1}^{1} (x+1)^2 \left(\frac{\sqrt{10}}{2} x^2\right) dx = \frac{8}{15}\sqrt{10} \end{aligned} \tag{17.1.29}$$

Now we have $\lambda = 1$. Also $\lambda_1 = 3/2$ and $\lambda_2 = 5/2$. Hence $\lambda \neq \lambda_1$ and $\lambda \neq \lambda_2$. Therefore the unique solution given by

$$\begin{aligned} y(x) &= f(x) + \lambda \sum_{m=1}^{2} \frac{f_m}{\lambda_m - \lambda} \phi_m(x) \\ \Rightarrow y(x) &= (x+1)^2 + \frac{f_1\phi_1(x)}{\lambda_1 - 1} + \frac{f_2\phi_2(x)}{\lambda_2 - 1} \\ \Rightarrow y(x) &= (x+1)^2 + \frac{(2\sqrt{6}/3) \times (x\sqrt{6}/2)}{(3/2) - 1} + + \frac{(8\sqrt{10}/15) \times (x^2\sqrt{10}/2)}{(5/2) - 1} \\ \Rightarrow y(x) &= (x+1)^2 + 4x + (16/9) \times x^2 = x^2 + 2x + 1 + 4x + (16/9) \times x^2 \\ \Rightarrow y(x) &= \frac{25}{9} x^2 + 6x + 1. \end{aligned}$$

Exercise 17.1.9. Using Hilbert-Schmidt theorem, find the solution of the symmetric integral equation

$i)\ y(x) = x^2 + 1 + \frac{3}{2} \int_{-1}^{1} (xt + x^2 t^2)\, y(t)\, dt,$ $\qquad ii)\ y(x) = 1 + \int_{0}^{\pi} \cos(x+t)\, y(t)\, dt,$

Answers :

$i)\ y(x) = 5x^2 + Cx + 1,$ where C is constant $\qquad (ii)\ y(x) = 1 + C\ \cos x - (2/\pi) \sin x,$

Unit 18

Course Structure

- Integral Transform: Laplace transforms of elementary functions & their derivatives and Dirac-delta function, Laplace integral, Lerch's theorem (statement only), property of differentiation, integration and convolution, inverse transform, application to the solution of ordinary differential equation, integral equation and BVP.

18.1 Introduction

The integral transform methods are very convenient in solving integral equations of some special forms. Suppose that a relationship of the form

$$y(x) = \int_a^b \int_a^b \Gamma[x, z] K(z, t) y(t)\, dt\, dz \tag{18.1.1}$$

be known to be valid and that this double integral can be evaluated as an iterated integral. Then from (18.1.1), it follows that if

$$F(x) = \int_a^b K(x, t)\; y(t)\; dt \tag{18.1.2}$$

we also have

$$y(x) = \int_a^b \Gamma(x, t)\; F(t)\; dt. \tag{18.1.3}$$

Thus, if Eq.(18.1.2) is an integral equation in y, a solution is given by Eq., whereas if Eq.(18.1) is regarded as an integral equation in F a solution is given by Eq.(18.1.2). It is conventional to refer to one of the function as the *transform* of the second function, and to the second function as an *inverse transform* of the first.

Definition 18.1.1. Function of exponential order: A function $f(x)$ is said to be of exponential order a as $x \to \infty$ if

$$\lim_{x\to\infty} e^{-ax} f(x) = \text{finite quantity}$$

i.e., if given a positive integer n_0, there exists a real number $M > 0$ s.t.

$$|e^{-ax} f(x)| < M \quad \forall x \geq n_0 \qquad \text{or} \qquad |f(x)| < Me^{ax} \quad \forall x \geq n_0$$

Example 18.1.2. Show that x^n is of exponential order as $x \to \infty$, n being any positive integer.

Solution:

$$\lim_{x\to\infty} e^{-ax}x^n = \lim_{x\to\infty} \frac{x^n}{e^{ax}} = \lim_{x\to\infty} \frac{n!}{a^n e^{ax}} = \frac{n!}{\infty} = 0$$

Example 18.1.3. Show that $F(t) = e^{t^2}$ is not of exponential order as $t \to \infty$.

Solution:

$$\lim_{t\to\infty} e^{-at}F(t) = \lim_{t\to\infty} e^{t(t-a)} = \infty$$

Hence $F(t)$ is not of exponential order.

Definition 18.1.4. Laplace transform: Suppose $F(t)$ is a real valued function defined over the interval $(-\infty, \infty)$ such that $F(t) = 0$. The *Laplace transform* of $F(t)$, denoted by $L\{F(t)\}$, is defined as

$$L\{F(t)\} = f(s) = \int_0^\infty e^{-st}F(t)\,dt \tag{18.1.4}$$

Sometimes we use symbol p for the parameter s. The Laplace transform is said to exist if the integral (18.1.4) is convergent for some value of s.

Table of Laplace transform of some elementary functions

Serial Number	F(t)	L{F(t)} or $\overline{F}(p)$ or $f(p)$
1	1	$1/p,\ p > 0$
2	$t^n,\ n > -1$	$\Gamma(n+1)/p^{n+1},\ p > 0$
3	t^n (n is positive integer)	$n!/p^{n+1},\ p > 0$
4	e^{at}	$1/(p-a),\ p > a$
5	$\sin at$	$a/(p^2+a^2),\ p > 0$
6	$\cos at$	$p/(p^2+a^2),\ p > 0$
7	$\sinh at$	$a/(p^2-a^2),\ p > \lvert a\rvert$
8	$\cosh at$	$p/(p^2-a^2),\ p > \lvert a\rvert$

Theorem 18.1.5. Linear Property: Suppose $f_1(s)$ and $f_2(s)$ are Laplace forms of $F_1(t)$ and $F_2(t)$ respectively. Then

$$L\{c_1F_1(t) + c_2F_2(t)\} = c_1L\{F_1(t)\} + c_2L\{F_2(t)\}$$

where c_1 and c_2 are constants.

Proof. Let

$$L\{F_1(t)\} = f_1(s) = \int_0^\infty e^{-st}F_1(t)\,dt \quad \text{and} \quad L\{F_2(t)\} = f_2(s) = \int_0^\infty e^{-st}F_2(t)\,dt$$

Also let c_1 and c_2 be arbitrary constants. Now

$$\begin{aligned} L\{c_1F_1(t) + c_2F_2(t)\} &= \int_0^\infty e^{-st}\Big[c_1F_1(t) + c_2F_2(t)\Big]\,dt \\ &= c_1\int_0^\infty e^{-st}F_1(t)\,dt \ + \ c_2\int_0^\infty e^{-st}F_2(t)\,dt \\ &= c_1L\{f_1(t)\} + c_2L\{f_2(t)\} \end{aligned}$$

□

Theorem 18.1.6. First Shifting Theorem (First Translation): If $L\{F(t)\} = f(s)$, then $L\{e^{at}F(t)\} = f(s-a)$.

Proof. We know that

$$\begin{aligned} L\{e^{at}F(t)\} &= \int_0^\infty e^{-st}e^{at}F(t)\,dt = \int_0^\infty e^{-(s-a)t}F(t)\,dt = \int_0^\infty e^{-ut}F(t)\,dt, \text{ where } u = s-a > 0 \\ &= f(u) = f(s-a) \end{aligned}$$

□

Example 18.1.7. Find $L\{e^{-t}(3\sinh 2t - 5\cosh 2t)\}$

Solution: We know that

$$L\{\sinh 2t\} = \frac{2}{s^2-2^2}, \qquad L\{\cosh 2t\} = \frac{s}{s^2-2^2}$$

Therefore

$$L\{e^{-t}(3\sinh 2t - 5\cosh 2t)\} \;=\; 3\frac{2}{(s+1)^2-2^2} - 5\frac{s+1}{(s+1)^2-2^2} = \frac{1-5s}{s^2+2s-3}$$

Theorem 18.1.8. Second Shifting Theorem (Second Translation)

$$\text{If } L\{F(t)\} = f(s) \text{ and } G(t) = \begin{cases} F(t-a) & t > a \\ 0 & t < a \end{cases} \quad \text{Then } L\{G(t)\} = e^{-as}f(s)$$

Proof.

$$\text{Let } L\{F(t)\} = f(s) \text{ and } G(t) = \begin{cases} F(t-a) & t > a \\ 0 & t < a \end{cases}$$

Now

$$\begin{aligned} L\{G(t)\} &= \int_0^\infty e^{-st}G(t)\,dt = \int_0^a e^{-st}G(t)\,dt + \int_a^\infty e^{-st}G(t)\,dt \\ &= \int_0^a e^{-st}\cdot 0\,dt + \int_0^\infty e^{-st}F(t-a)\,dt = 0 + \int_a^\infty e^{-st}F(t-a)\,dt \end{aligned}$$

Now putting $t-a=p$ so that $dt = dp$. If $t=a$, then $p = t-a = a-a = 0$ and if $t=\infty$, then $p = \infty - a = \infty$.

$$\therefore L\{G(t)\} = \int_0^\infty e^{-s(p+a)}F(p)\,dp = e^{-sa}\int_0^\infty e^{-sp}F(p)\,dp = e^{-sa}f(s)$$

□

Example 18.1.9. Find the Laplace transform of $F(t)$, where $F(t) = \begin{cases} \cos\left(t-\frac{2\pi}{3}\right) & \text{if } t > \frac{2\pi}{3} \\ 0 & \text{if } t > \frac{2\pi}{3} \end{cases}$

Solution: Let $a = \frac{2\pi}{3}$, and $G(t) = \cos t$, then

$$L\{G(t)\} = \frac{p}{p^2+1} = g(p), \quad \text{as} \quad L\{\cos at\} = \frac{p}{p^2+a^2}. \quad \text{Also } F(t) = \begin{cases} G(t-a) & t > a \\ 0 & t < a \end{cases}$$

By second shifting theorem,

$$L\{F(t)\} = e^{-ap}g(p) = e^{-\frac{2\pi p}{3}}\frac{p}{p^2+1}$$

Theorem 18.1.10. Change of Scale Property

$$\text{If } \ L\{F(t)\} = f(s), \quad \text{then} \quad L\{F(at)\} = \frac{1}{a} f\left(\frac{s}{a}\right)$$

Proof.

$$\begin{aligned}
\text{Let } \ L\{F(t)\} = f(s), \ \text{ then } \ L\{F(at)\} &= \int_0^\infty e^{-st} F(at)\, dt = \int_0^\infty e^{-\frac{sx}{a}} F(x) \frac{dx}{a}, \ \text{ Putting } \ x = at \\
&= \frac{1}{a} \int_0^\infty e^{-\frac{sx}{a}} F(x)\, dx = \frac{1}{a} \int_0^\infty e^{-\frac{st}{a}} F(t)\, dt \\
&= \frac{1}{a} \int_0^\infty e^{-pt} F(t)\, dt, \ \text{ where } \ p = \frac{s}{a} \ \ = \frac{1}{a} f(p) = \frac{1}{a} f\left(\frac{s}{a}\right)
\end{aligned}$$

□

Example 18.1.11. If $L\{\cos^2 t\} = \frac{s^2+2}{s(s^2+4)}$, then find $L\{cos^2(at)\}$. **Answer:** $\frac{(s^2+2a^2)}{s(s^2+4a^2)}$

18.1.1 Laplace Transform of Derivatives

Theorem 18.1.12. If $L\{F(t)\} = f(s)$, then $L\{F^{(n)}(t)\} = s^n f(s) - s^{n-1}F(0) - s^{n-2}F'(0) - \ldots - sF^{(n-2)}(0) - F^{(n-1)}(0)$ where $F^{(n)}(t)$ stands for $\frac{d^n F(t)}{dt^n}$.

18.1.2 The Dirac Delta Function

The Dirac delta or Dirac's delta is a mathematical construct introduced by theoretical physicist Paul Dirac. Informally, it is a generalised function representing an infinitely sharp peak bounding unit area: a 'function' $\delta(x)$ that has the value zero everywhere except at $x = 0$ where its value is infinitely large in such a way that its total integral is 1.

The Dirac delta is not strictly a function, while for many purposes it can be manipulated as such, formally it can be determined as a distribution. In many applications, the Dirac delta function is regarded as limit of a sequence of functions having a tall spike at the origin. The approximating functions of the sequence are thus "approximate" or "nascent" delta functions.

$$\begin{aligned}
\delta(x) &= +\infty, \ x = 0 \\
&= 0, \ x \neq 0
\end{aligned}$$

and

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1.$$

18.1.3 Laplace Transform of Integral

Theorem 18.1.13. If $L\{F(t)\} = f(s)$, then $\frac{1}{s} f(s) = L\left\{\int_0^t F(u)\, du\right\}$.

18.1.4 Multiplication by Powers of t

Theorem 18.1.14. If $L\{F(t)\} = f(s)$, then $L\{t^n F(t)\} = (-1)^n \frac{d^n}{ds^n} f(s)$ for $n = 1, 2, 3, \ldots$.

18.1.5 Division by t

Theorem 18.1.15. If $L\{F(t)\} = f(s)$, then $L\{\frac{F(t)}{t}\} = \int_s^\infty f(x)\, dx$.

Note: Proofs of the above theorems are left for the readers.

Example 18.1.16. Find $L\{\cos^2 t\}$ and $L\{\sin^2 t\}$.

Solution:

$$L\{\cos^2 t\} = \frac{1}{2}L\{1+\cos 2t\} = \frac{1}{2}\left[\frac{1}{s} + \frac{s}{s^2+2^2}\right] = \frac{(s^2+2)}{s(s^2+4)}$$
$$L\{\sin^2 t\} = \frac{1}{2}L\{1-\cos 2t\} = \frac{1}{2}\left[\frac{1}{s} - \frac{s}{s^2+2^2}\right] = \frac{2}{s(s^2+4)}$$

Example 18.1.17. Using Laplace transform prove that (i) $\int_0^\infty \left(\frac{\sin t}{t}\right) dt = \frac{\pi}{2}$ and (ii) $\int_0^\infty te^{-3t}\sin t\, dt = \frac{3}{50}$.

Solution: (i) We know, $L\{\sin t\} = \frac{1}{p^2+1}$, then

$$L\left\{\frac{\sin t}{t}\right\} = \int_p^\infty \frac{dp}{1+p^2} = \Big(\tan^{-1} p\Big)_p^\infty = \frac{\pi}{2} - \tan^{-1}(p) = \cot^{-1}(p) = \tan^{-1}\left(\frac{1}{p}\right)$$

Putting $p = 0$, we get $\int_0^\infty \left(\frac{\sin t}{t}\right)\ dt = \tan^{-1}(\infty) = \frac{\pi}{2}$.

(ii) Since $L\{\sin t\} = \frac{1}{p^2+1}$ and so $L\{t\sin t\} = (-1)^1\frac{d}{dp}\left(\frac{1}{p^2+1}\right) = \frac{2p}{(p^2+1)^2}$

$$\Rightarrow \int_0^\infty e^{-pt}(t\ \sin t)\, dt = \frac{2p}{(p^2+1)^2} \quad \text{Putting } p=3, \text{ we have } \int_0^\infty e^{-3t}(t\ \sin t)\, dt = \frac{3}{50}$$

Example 18.1.18. Find $L\{F_\varepsilon(t)\}$ where $F_\varepsilon(t)$ is dirac delta function.

Solution:

$$\begin{aligned}
F_\varepsilon(t) &= \begin{cases} 1/\varepsilon & \text{if } 0 \le t \le \varepsilon \\ 0 & \text{if } t > \varepsilon \end{cases} \\
L\{F_\varepsilon(t)\} &= \int_0^\infty e^{-st}F_\varepsilon(t)\, dt = \int_0^\varepsilon e^{-st}F_\varepsilon(t)\, dt + \int_\varepsilon^\infty e^{-st}F_\varepsilon(t)\, dt \\
&= \int_0^\varepsilon \frac{e^{-st}}{\varepsilon}dt + \int_\varepsilon^\infty e^{-st}\cdot 0\ \cdot dt = \frac{1}{\varepsilon}\left[\frac{e^{-st}}{-s}\right]_{t=0}^{\varepsilon} + 0 \\
&= \frac{1}{\varepsilon s}(1 - e^{-\varepsilon s}).
\end{aligned} \tag{18.1.5}$$

18.1.6 Inverse Laplace transform

Let $L\{F(t)\} = \overline{F}(p)$. Then $F(t)$ is called an *inverse Laplace transform* of $\overline{F}(p)$, and we write $F(t) = L^{-1}\{\overline{F}(p)\}$, in which L^{-1} is known as the *inverse Laplace transformation operator.*

Theorem 18.1.19. Lerch's Theorem: Let $L\{F(t)\} = f(s)$. Let $F(t)$ be piecewise continuous in every finite interval $0 \le t \le a$ and of exponential order for $t > a$, then the inverse Laplace transform of $F(t)$ is unique $f(s)$.

Table of inverse Laplace transform of some elementary functions

Serial Number	$\overline{F}(p)$	$L^{-1}\{\overline{F}(p)\}$
1	$1/p$	1
2	$1/p^{n+1}, n > -1$	$t^n/\Gamma(n+1)$
3	$1/p^{n+1}$ (n is positive integer)	$t^n/n!$
4	$1/(p-a)$	e^{at}
5	$1/(p^2+a^2)$	$(\sin at)/a$
6	$p/(p^2+a^2)$	$\cos at$
7	$1/(p^2-a^2)$	$(\sin at)/a$
8	$p/(p^2-a^2)$	$\cos at$

Theorem 18.1.20. Inverse Laplace transform of derivatives: If $L^{-1}\{f(s)\} = F(t)$, then $L^{-1}\{f^{(n)}(s)\} = (-1)^n t^n F(t)$

Theorem 18.1.21. First Shifting theorem: If $L^{-1}\{f(s)\} = F(t)$, then $L^{-1}\{f(s-a)\} = e^{at}F(t)$.

Theorem 18.1.22. Second Shifting theorem: If $L^{-1}\{f(s)\} = F(t)$, then $L^{-1}\{e^{-as}f(s)\} = G(t)$, where $G(t) = \begin{cases} F(t-a) & \text{if } t > a \\ 0 & \text{if } t < a \end{cases}$

Example 18.1.23.

$$\text{Find} \quad L^{-1}\left\{\frac{s-2}{(s-2)^2+5^2} + \frac{s+4}{(s+4)^2+9^2} + \frac{1}{(s+2)^2+3^2}\right\}$$

Solution:

$$\begin{aligned}
&L^{-1}\left\{\frac{s-2}{(s-2)^2+5^2} + \frac{s+4}{(s+4)^2+9^2} + \frac{1}{(s+2)^2+3^2}\right\} \\
&= L^{-1}\left\{\frac{s-2}{(s-2)^2+5^2}\right\} + L^{-1}\left\{\frac{s+4}{(s+4)^2+9^2}\right\} + L^{-1}\left\{\frac{1}{(s+2)^2+3^2}\right\} \\
&= e^{2t}L^{-1}\left\{\frac{s}{s^2+5^2}\right\} + e^{-4t}L^{-1}\left\{\frac{s}{s^2+9^2}\right\} + e^{-2t}L^{-1}\left\{\frac{1}{s^2+3^2}\right\} \\
&= e^{2t}\cos 5t + e^{-4t}\cos 9t + \frac{e^{-2t}}{3}\sin 3t
\end{aligned}$$

Example 18.1.24. *Find* $L^{-1}\left\{\frac{e^{4-3p}}{(p+4)^{5/2}}\right\}$

Solution:

$$L^{-1}\left[\frac{1}{(p+4)^{5/2}}\right] = e^{-4t}L^{-1}\{\frac{1}{(p-4+4)^{5/2}}\} = e^{-4t}L^{-1}\left\{\frac{1}{p^{(3/2)+1}}\right\} = e^{-4t}\frac{t^{3/2}}{\Gamma(5/2)}$$

But $\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi} = \frac{3}{4}\sqrt{\pi}$,

$$\text{Therefore,} \quad L^{-1}\left[\frac{1}{(p+4)^{5/2}}\right] = \frac{4}{3\sqrt{\pi}}e^{-4t}t^{3/2} \;\Rightarrow\; L^{-1}\left[\frac{e^{4-3p}}{(p+4)^{5/2}}\right] = e^4L^{-1}\left[\frac{e^{-3p}}{(p+5)^{5/2}}\right]$$

Using second shifting theorem

$$L^{-1}\left[\frac{e^{4-3p}}{(p+4)^{5/2}}\right] = \begin{cases} \frac{4e^4}{3\sqrt{\pi}}e^{-4(t-3)}(t-3)^{3/2} & \text{if } t > 3 \\ 0 & \text{if } t < 4 \end{cases} = \frac{4e^4}{3\sqrt{\pi}}e^{-4(t-3)}(t-3)^{3/2}H(t-3).$$

Note 18.1.25. In the above example $H(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases}$ is *Heaviside unit step function.*

Definition 18.1.26. Convolution (or Faltung): The convolution of $F(t)$ and $G(t)$ is denoted and defined as

$$F * G = \int_0^t F(x)G(t-x)\,dx \quad \text{or} \quad F * G = \int_0^t F(t-x)G(x)\,dx$$

Theorem 18.1.27. Convolution theorem or Convolution property: If $L^{-1}\{\overline{F}(p)\} = F(t)$ and $L^{-1}\{\overline{G}(p)\} = G(t)$, then $L^{-1}\{\overline{F}(p)\overline{G}(p)\} = \int_0^t F(x)G(t-x)\,dx = F * G$ or $L^{-1}\{\overline{F}(p)\overline{G}(p)\} = \int_0^t F(t-x)G(x)\,dx = F * G$.

Example 18.1.28. Find $L^{-1}\left\{\frac{1}{(s+a)(s+b)}\right\}$.

Solution: Let $f(s) = \frac{1}{s-a}$, $g(s) = \frac{1}{s+b}$. Then $F(t) = e^{-at}$, $G(t) = e^{-bt}$. Hence

$$\begin{aligned}
L^{-1}\{f(s) \cdot g(s)\} &= \int_0^t F(u)G(t-u)\,du \\
&= \int_0^t e^{-au}e^{-b(t-u)}\,du = e^{-bt}\int_0^t e^{u(b-a)}du \\
&= e^{-bt}\left[\frac{e^{u(b-a)}}{b-a}\right]_{u=0}^t = \frac{e^{-bt}}{b-a}[e^{t(b-a)} - 1] = \frac{e^{-at} - e^{=bt}}{b-a}
\end{aligned}$$

Example 18.1.29. Apply the convolution theorem to show that

$$\int_0^t \sin u \cos(t-u)\,du = \frac{1}{2}(t \sin t).$$

Solution: Let $F(t) = \int_0^t \sin u \cos(t-u)\,du$. Then, by convolution theorem,

$$\begin{aligned}
&L\{F(t)\} = L\{\sin t\}L\{\cos t\} = \frac{1}{s^2+1} \cdot \frac{s}{s^2+1} = \frac{s}{(s^2+1)^2} \\
&\therefore F(t) = L^{-1}\left\{\frac{s}{(s^2+1)^2}\right\}
\end{aligned}$$

In order to calculate the above inverse Laplace transform, let $F(t) = \sin t$. Therefore,

$$\begin{aligned}
f'(s) = -\frac{2s}{(s^2+1)^2} \quad &\Rightarrow \quad L^{-1}\{f'(s)\} = L^{-1}\left\{-\frac{2s}{(s^2+1)^2}\right\} \\
&\Rightarrow \quad (-1)^1 t^1 L^{-1}\{f(s)\} = L^{-1}\left\{-\frac{2s}{(s^2+1)^2}\right\} \\
&\Rightarrow \quad \frac{t \sin t}{2} = L^{-1}\left\{\frac{s}{(s^2+1)^2}\right\}
\end{aligned}$$

Hence $F(t) = \frac{t \sin t}{2}$.

Exercise 18.1.30. (i) Evaluate $L^{-1}\left\{\frac{1}{(s+1)(s+2)}\right\}$ **Answer:** $\frac{e^{2t}-e^{-t}}{3}$.

(ii) Evaluate $L^{-1}\left\{\frac{1}{(s-3)(s+4)}\right\}$ **Answer:** $\frac{1}{7}(e^{3t} - e^{-4t})$.

18.1.7 Some special types of integral equations

Definition 18.1.31. (i) Integro-differential equation: An integral equation in which various derivatives of the unknown function $y(t)$ can also be present is said to be an *integro-differential equation.* For example, the following integral equation is an integro-differential equation.

$$y'(t) = y(t) + f(t) + \int_0^t \sin(t-x)y(x)\,dx.$$

Definition 18.1.32. (i) Integral equation of convolution type: The integral equation

$$y(t) = f(t) + \int_0^t K(t-x)y(x)\,dx,$$

in which the kernel $K(t-x)$ is a function of the difference $(t-x)$ only, is known as integral equation of the convolution type. Using the definition of convolution, we may re-write it as

$$y(t) = f(t) + K(t) * y(t).$$

Example 18.1.33. Solve the integro-differential equation

$$y'(t) = \sin t + \int_0^t y(t-x)\cos x\,dx, \quad \text{where} \quad y(0) = 0$$

Solution: Rewriting the given equation, we have

$$y'(t) = \sin t + y(t) * \cos t, \qquad y(0) = 0$$

Applying the Laplace transform on both sides, we obtain

$$\begin{aligned}
&L\{y't)\} = L\{\sin t\} + L\{y(t) * \cos t\}\\
&\Rightarrow pL\{y(t)\} - y(0) = \frac{1}{p^2+1} + L\{y(t)\}L\{\cos t\}\\
&\Rightarrow \left(1 - \frac{1}{p^2+1}\right) pL\{y(t)\} = \frac{1}{p^2+1}\\
&\Rightarrow L\{y(t)\} = \frac{1}{p^3}, \quad \text{Inverting we have} \quad y(t) = L^{-1}\{\frac{1}{p^3}\} = \frac{t^2}{2!} = \frac{t^2}{2}.
\end{aligned}$$

Exercise 18.1.34. (i) Solve $y'(t) = t + \int_0^t y(t-x)\cos x\,dx, \quad y(0) = 4$ **Answer:** $y(t) = 4 + \frac{5}{2}t^2 + \frac{1}{24}t^4$.

Example 18.1.35. Solve the integral equation $y(t) = 1 + \int_0^t y(x)\sin(t-x)\,dx, \quad .$

Solution: The given integral equation can be re-written as $y(t) = 1 + y(t) * \sin t$. Applying Laplace transform, we obtain

$$\begin{aligned}
&L\{y(t)\} = L\{1\} + L\{y(t)\} \cdot L\{\sin t\}\\
&\Rightarrow L\{y(t)\} = \frac{1}{p} + L\{y(t)\} \times \frac{1}{p^2+1} \Rightarrow \left(1 - \frac{1}{p^2+1}\right) L\{y(t)\} = \frac{1}{p}\\
&\Rightarrow L\{y(t)\} = \frac{p^2+1}{p^3} = \frac{1}{p} + \frac{1}{p^3}
\end{aligned}$$

Inverting the above equation, we have

$$y(t) = L^{-1}\{y(t)\} = L^{-1}\{\frac{1}{p}\} + L^{-1}\{\frac{1}{p^3}\} = 1 + \frac{t^2}{2!} = 1 + \frac{t^2}{2}.$$

Exercise 18.1.36. (i) Solve $y(t) = a\ \sin t - \int_0^t y(x)\cos(t-x)\ dx$, **Answer:** $y(t) = a\ t\ e^{-t}$.
(ii) Solve $y(t) = e^{-t} - 2\int_0^t \cos(t-x)\ y(x)\ dx$ **Answer:** $y(t) = e^{-t}(1-t)^2$.
(iii) Solve $y(t) = t + 2\int_0^t \cos(t-x)y(x)\ dx$ **Answer:** $y(t) = 2e^t(t-1)+2+t$.
(iv) Solve $t = \int_0^t e^{t-x}y(x)\ dx$ **Answer:** $y(t) = 1-t$.

18.1.8 Application to solve ordinary differential equations

Consider the differential equation

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + x = F(t) \tag{18.1.6}$$

Case I: When a, b are constants. Taking Laplace transform on the both sides of Eq. (18.1.6), we have

$$aL\{x''\} + bL\{x'\} + L(x) = L\{F(t)\}. \tag{18.1.7}$$

Letting $L\{x(t)\} = \overline{x}(s)$, we have from Eq. (18.1.7)

$$a\{s^2\overline{x} - s\ x(0) - x'(0)\} + b\{s\overline{x} - x(0)\} + \overline{x} = f(s)$$

Required solution is obtained by taking the inverse Laplace transform of $\overline{x}(s)$.

Example 18.1.37. Solve by using Laplace transformation:

$$(D^2+9)y = \cos(2t) \quad \text{if} \quad y(0) = 1,\ y(\pi/2) = -1$$

Solution: Taking Laplace transform, we get

$$\begin{aligned}
&p^2\overline{y} - py(0) - y'(0) + 9\overline{y} = \frac{p}{p^2+2^2}\\
&\Rightarrow (p^2+9)\overline{y} = p + a + \frac{p}{p^2+2^2} \quad \text{where } \ y'(0) = a\\
&\Rightarrow \overline{y} = \frac{p}{p^2+3^2} + \frac{a}{p^2+3^2} + \frac{p}{(p^2+2^2)(p^2+3^2)}\\
&\Rightarrow \overline{y} = \frac{p}{p^2+3^2} + \frac{a}{p^2+3^2} + \frac{1}{5}\left[\frac{p}{p^2+2^2} - \frac{p}{p^2+3^2}\right]\\
&\Rightarrow \overline{y} = \frac{4}{5}\frac{p}{p^2+3^2} + \frac{a}{p^2+3^2} + \frac{1}{5}\frac{p}{p^2+2^2}
\end{aligned}$$

Taking inverse Laplace transform, we get

$$\begin{aligned}
y &= \frac{4}{5}\cos 3t + \frac{a}{3}\sin 3t + \frac{1}{5}\cos 2t \qquad \Rightarrow -1 = y(\pi/2) = 0 - \frac{a}{3} - \frac{1}{5} \Rightarrow \frac{a}{3} = \frac{4}{5}\\
\therefore y &= \frac{4}{5}\cos 3t + \frac{4}{5}\sin 3t + \frac{1}{5}\cos 2t.
\end{aligned}$$

Example 18.1.38.

$$\text{Solve } \ 2\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 2y = e^{-2t},\ \ y(0) = 1,\ y'(0) = 1$$

Solution: Taking Laplace transform of the equation

$$(2D^2+5D+2)y=e^{-2t},$$

we get $2[s^2\overline{y}-sy(0)-y'(0)]+5[s\overline{y}-y(0)]+2\overline{y}=\dfrac{1}{s-2}$

Putting $y(0)=1=y'(0)$, we get

$$2[s^2\overline{y}-s-1]+5[s\overline{y}-1]+2\overline{y}=\frac{1}{s-2}$$

$$\Rightarrow (2s^2+5s+2)\overline{y}-2s-7=\frac{1}{s-2}$$

$$\Rightarrow \overline{y}=\frac{1}{(s+2)(2s^2+5s+2)}+\frac{2s+7}{2s^2+5s+2}$$

$$\Rightarrow \overline{y}=\frac{1}{(s+2)^2(2s+1)}+\frac{2s+7}{(s+2)(2s+1)}$$

Now $L^{-1}\left\{\dfrac{1}{(s+2)^2}(2s+1)\right\}=e^{-2t}L^{-1}\left\{\dfrac{1}{s^2\{2(s-2)+1\}}\right\}=e^{-2t}L^{-1}\left\{\dfrac{1}{s^2(2s-3)}\right\}$

But $\dfrac{1}{s(2s-3)}=\dfrac{1}{3}\left[\dfrac{2}{2s-3}-\dfrac{1}{s}\right]=\dfrac{1}{3}\left[\dfrac{1}{s-\frac{3}{2}}-\dfrac{1}{s}\right]$

$\Rightarrow L^{-1}\left\{\dfrac{1}{s(2s-3)}\right\}=\dfrac{1}{3}\left[e^{3t/2}-1\right]\Rightarrow L^{-1}\left\{\dfrac{1}{s^2(2s-3)}\right\}=\dfrac{1}{3}\displaystyle\int_0^t\left[e^{3t/2}-1\right]dx$

$\Rightarrow L^{-1}\left\{\dfrac{1}{s^2(2s-3)}\right\}=\dfrac{2}{3}\left[e^{3t/2}-1\right]-\dfrac{1}{3}t$

and $\dfrac{2s+7}{(s+2)(2s+1)}=\dfrac{4}{2s+1}-\dfrac{1}{s+2}=\dfrac{2}{s+\frac{1}{2}}-\dfrac{1}{s+2}$

$\Rightarrow L^{-1}\left\{\dfrac{2s+7}{(s+2)(2s+1)}\right\}=2e^{-t/2}-e^{-2t}.$

Using these inverse Laplace transforms we get

$$y(t)=(2e^{-t/2}-e6-2t)+e^{-2t}\left[\frac{2}{9}(e^{3t/2}-1)-\frac{t}{3}\right]=\frac{20}{9}e^{-t/2}-e^{-2t}\left(\frac{11}{9}+\frac{t}{3}\right)$$

Exercise 18.1.39. (i) Solve $y''+25y=10\cos(5t)$, $y(0)=2$, $y'(0)=0$, **Answer:** $y(t)=t\sin 5t+2\cos(5t)$.
(ii) Solve $(D+D^2)x=2$ when $x(0)=3$, $x'(0)=1$ **Answer:** $y(t)=e^{-t}+2t+2$.
(iii) Solve $(D^2+D)y=t^2+2t$ when $y(0)=4$, $y'(0)=-2$ **Answer:** $2+2e^{-t}+\frac{t^3}{3}$.

Case II: When a,b are functions of t, i.e., of the form

$$t^2\frac{d^2x}{dt^2}+t\frac{dx}{dt}+x=F(t) \tag{18.1.8}$$

In this case, we use the theorem

$$L\left\{t^m \frac{d^n x}{dt^n}\right\} = L\{t^m x^{(n)}(t)\} = (-1)^m \frac{d^m}{ds^m} L\{x^{(n)}\} \tag{18.1.9}$$

Taking the Laplace transform of Eq. (18.1.8),

$$(-1)^2 \frac{d^2}{ds^2}\Big[s^2\overline{x} - s\,x(0) - x'(0)\Big] - \frac{d}{ds}\Big\{s\overline{x} - x(0)\Big\} + \overline{x} = f(s)$$

The required solution is obtained by taking inverse Laplace transform of $\overline{x}(s)$.

Example 18.1.40. Using Laplace transform solve the following differential equation.

$$y'' + ty' - y = 0 \quad \text{if} \quad y(0) = 0,\ y'(0) = 1$$

Solution: Taking Laplace transform of $y'' + ty' - y = 0$, we get

$$p^2\overline{y} - py(0) - y'(0) + (-1)^1 \frac{d}{dp}\Big[p\overline{y} - y(0)\Big] - \overline{y} = 0$$

Putting $y(0) = 0, y'(0) = 1$, we get

$$\begin{aligned} p^2\overline{y} - 1 - \frac{d}{dp}\Big[p\,\overline{y}\Big] - \overline{y} &= 0 \\ \Rightarrow (p^2 - 1)\overline{y} - \left(\overline{y} + p\frac{d\overline{y}}{dp}\right) &= 1 \\ \Rightarrow (p^2 - 2)\overline{y} - p\frac{d\overline{y}}{dp} &= 1 \\ \Rightarrow \frac{d\overline{y}}{dp} + \left(\frac{2}{p} - p\right)\overline{y} &= -\frac{1}{p} \end{aligned}$$

Therefore the integrating factor

$$I.F = e^{\int\left(\frac{2}{p} - p\right)dp} = e^{2\log p - p^2/2} = e^{\log p^2} \cdot e^{-p^2/2} = p^2 e^{-p^2/2}$$

Therefore

$$\overline{y}p^2 e^{-p^2/2} = c + \int (p^2 e^{-p^2/2})\left(-\frac{1}{p}dp\right) = c - \int p e^{p^2/2} dp.$$

Put $p^2/2 = z$, then $p\,dp = dz$. Therefore,

$$\begin{aligned} \overline{y}p^2 e^{-p^2/2} &= c - \int e^{-z}dz = c + e^{-z} = c + e^{-p^2/2} \\ \Rightarrow \overline{y} &= \frac{c}{p^2}e^{p^2/2} + \frac{1}{p^2} \end{aligned}$$

Taking inverse transform,

$$y(t) = t + cL^{-1}\left\{\frac{1}{p^2}e^{p^2/2}\right\}$$

Subjecting this to the condition $y = 0$ when $t = 0$, we get $c = 0$. Therefore, the solution is

$$y(t) = t$$

Exercise 18.1.41. (i) Solve $ty'' + (1 - 2t)y' - 2y = 0,\ y(0) = 1,\ y'(0) = -2$ **Answer:** $y(t) = e^{2t}$.

Unit 19

Course Structure

- Fourier Transform: Fourier transform of some elementary functions and their derivatives, inverse Fourier transform, convolution theorem & Parseval's relation and their application, Fourier sine and cosine transform.

19.1 Introduction

The Fourier transform is a generalization of the Fourier series representation of functions. The Fourier series is limited to periodic functions, while the Fourier transform can be used for a larger class of functions which are not necessarily periodic. Since the transform is essential to the understanding of several exercises, we briefly explain some basic Fourier transform concepts in this unit.

Objective

After reading this unit readers will able to know fundamental mathematical properties of the Fourier transform including linearity, shift, symmetry, scaling, modulation and convolution. Further, the reader will be able to calculate the Fourier transform or inverse transform of common functions.

19.1.1 The Infinite Fourier Transform

Definition 19.1.1. Infinite Fourier sine transform: The Fourier sine transform of $F(x)$ on $0 < x < \infty$ is denoted by $f_s(s)$ or $F_s\{F(x)\}$ and is defined as

$$f_s(s) = F_s\{F(x)\} = \int_0^\infty F(x) \sin sx \, dx.$$

The inverse formula for infinite Fourier sine transform is given by

$$F(x) = F_s^{-1}\{f_s(s)\} = \frac{2}{\pi} \int_0^\infty f_s(s) \, \sin sx \, ds.$$

Definition 19.1.2. Infinite Fourier cosine transform: The Fourier sine transform of $F(x)$ on $0 < x < \infty$ is denoted by $f_c(s)$ or $F_c\{F(x)\}$ and is defined as

$$f_c(s) = F_c\{F(x)\} = \int_0^\infty F(x) \cos sx \, dx.$$

The inverse formula for infinite Fourier cosine transform is given by

$$F(x) = F_c^{-1}\{f_c(s)\} = \frac{2}{\pi} \int_0^\infty f_c(s) \; \cos sx \, ds.$$

Definition 19.1.3. Infinite Fourier transform: The infinite Fourier transform of $F(x)$ on $0 < x < \infty$ is denoted by $f(s)$ or $F\{F(x)\}$ and is defined as

$$f(s) = F\{F(x)\} = \int_{-\infty}^\infty F(x) e^{-isx} \, dx.$$

The inverse formula for infinite Fourier sine transform is given by

$$F(x) = F^{-1}\{f(s)\} = \frac{1}{2\pi} \int_{-\infty}^\infty f(s) \; e^{isx} \, ds.$$

19.1.2 Relationship between Fourier transform and Laplace transform

We define a function as follows:

$$F(t) = \begin{cases} e^{-xt}\phi(t), & , \quad t > 0 \\ 0 & , \quad t < 0 \end{cases}$$

Taking infinite Fourier transform of this function we obtain

$$\begin{aligned} F\{F(t)\} &= \int_{-\infty}^\infty e^{-iyt} F(t) \, dt = \int_{-\infty}^0 e^{-iyt} F(t) \, dt + \int_0^\infty e^{-iyt} F(t) \, dt \\ &= \int_{-\infty}^0 e^{-iyt} \cdot 0 \, dt + \int_0^\infty e^{-iyt} e^{-xt} \phi(t) \, dt = \int_0^\infty e^{-(x+iy)t} \phi(t) \, dt \\ &= e^{-st} \phi(t) \, dt = L\{\phi(t)\}, \quad \text{where } \; s = x + iy \end{aligned}$$

Therefore $F\{F(t)\} = L\{\phi(t)\}$. This is the required relation between Fourier transform and Laplace transform.

19.1.3 Some theorems

Theorem 19.1.4. Linear Property: If c_1 and c_2 are arbitrary constants, then

$$F\{c_1 F(x) + c_2 G(x)\} = c_1 F\{F(x)\} + c_2 F\{G(x)\}$$

Theorem 19.1.5. Change of Scale Property: If $f(s)$ is the Fourier transform of $F(x)$, then $\frac{1}{a} f\left(\frac{s}{a}\right)$ is the Fourier transform of $F(ax)$.

Theorem 19.1.6. Shifting Property: If $f(s)$ is the Fourier transform of $F(x)$, then $e^{-ias} f(s)$ is the Fourier transform of $F(x - a)$.

Theorem 19.1.7. Modulation Theorem: If $F(x)$ has the Fourier transform $f(s)$, then $F(x)\cos ax$ has the Fourier transform
$$\frac{1}{2}f(s-a)+\frac{1}{2}f(s+a).$$

Theorem 19.1.8. Derivative Theorem: The Fourier transform of $F'(x)$, the derivative of $F(x)$, is $is\ f(s)$, where $f(s)$ is the Fourier transform of $F(x)$. Moreover,
$$F\left\{\frac{d^nF}{dx^n}\right\}=(is)^nf(s),\ \text{where } F\{F(x)\}=f(s)$$
if the first $(n-1)$ derivative of $F(x)$ vanish identically as $x\to\pm\infty$.

- Proofs are left as exercise.

Theorem 19.1.9. Convolution Theorem: The convolution for the Fourier transform is defined as
$$F*G=\int_{-\infty}^{\infty}F(u)\,G(x-u)\,du.$$
If $F\{f(x)\}$ and $F\{g(x)\}$ are the Fourier transforms of functions $f(x)$ and $g(x)$ respectively, then Fourier transform of the convolution of $f(x)$ and $g(x)$ is the product of the their Fourier transforms, i.e.,
$$F\{f(x)*g(x)\}=F\{f(x)\}\cdot F\{g(x)\}$$

Proof.

$$\begin{aligned}\text{We have } F\{f(x)*g(x)\} &= \int_{-\infty}^{\infty}\{f(x)*g(x)\}e^{-isx}dx\\ &= \int_{x=-\infty}^{\infty}\left[\int_{u=-\infty}^{\infty}f(u)\,g(x-u)\,du\right]e^{-isx}dx\\ &= \int_{u=-\infty}^{\infty}f(u)\left[\int_{x=-\infty}^{\infty}g(x-u)e^{-isx}dx\right]du \quad \text{(Changing order of integration)}\\ &= \int_{u=-\infty}^{\infty}f(u)e^{-isu}F\{g(x)\}du \quad \text{(Using Shifting Property)}\\ &= F\{g(x)\}\int_{u=-\infty}^{\infty}f(u)e^{-isu}du\\ &= F\{g(x)\}\cdot F\{f(x)\}=F\{f(x)\}\cdot F\{g(x)\}\end{aligned}$$

□

Example 19.1.10. Find the Fourier transform of $f(x)=\begin{cases} x & |x|\le a\\ 0 & |x|>a.\end{cases}$

Solution: Given that $f(t) = \begin{cases} t & -a \le t \le a \\ 0 & |t| > a. \end{cases}$ Now

$$\begin{aligned}
F\{f(t)\} &= \int_{-\infty}^{\infty} e^{-ist} f(t)\, dt \\
&= \int_{-\infty}^{-a} e^{-ist} f(t)\, dt + \int_{-a}^{a} e^{-ist} f(t)\, dt + \int_{a}^{\infty} e^{-ist} f(t)\, dt \\
&= \int_{\infty}^{a} e^{isy} f(-y)\,(-dy) + \int_{-a}^{a} e^{-ist}\, t\, dt + \int_{a}^{\infty} e^{-ist} \cdot 0 \cdot dt \\
&= \int_{a}^{\infty} e^{isy} \cdot 0 \cdot dy + \int_{-a}^{a} e^{-ist}\, t\, dt \\
&= \left(-\frac{a}{is}\right)\{e^{-isa} + e^{isa}\} + \frac{1}{s^2}\{e^{-isa} - e^{isa}\} \\
&= \left(\frac{2ai}{s}\right)\cos(sa) - \frac{2}{s^2}\sin(sa)
\end{aligned}$$

Example 19.1.11. Find the Fourier transform of $f(x) = \begin{cases} 1 & |x| < a \\ 0 & |x| > a. \end{cases}$ and hence evaluate

$$(i) \int_{-\infty}^{\infty} \frac{\sin sa \cdot \cos sx}{s} dx, \qquad (ii) \int_{0}^{\infty} \frac{\sin s}{s} ds.$$

Solution: For the first part, proceed as the above example and find the answer is $\frac{2}{s}\sin sa$. For the second part, let $F\{f(x)\} = \overline{f}(s)$. We know that if

$$\overline{f}(s) = F\{f(x)\} = \int_{-\infty}^{\infty} f(x) e^{-isx} dx \quad \text{then} \quad f(x) = F^{-1}\{\overline{f}(s)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{f}(s) e^{isx} ds.$$

$$\therefore \int_{-\infty}^{\infty} \overline{f}(s) e^{isx} ds = 2\pi f(x) = \begin{cases} 2\pi & \text{if} |x| < a \\ 0 & \text{if} |x| > a. \end{cases}$$

But $\overline{f}(s) = \frac{2\sin sa}{s}$, by first part.

$$\therefore \int_{-\infty}^{\infty} \frac{2\sin sa}{s}(\cos sx + i \sin sx)\, ds = \begin{cases} 2\pi & \text{if} & |x| < a \\ 0 & \text{if} & |x| > a. \end{cases}$$

$$\Rightarrow \int_{-\infty}^{\infty} \frac{\sin sa \cos sx}{s}\, ds + i \int_{-\infty}^{\infty} \frac{\sin sa \sin sx}{s}\, ds = \begin{cases} \pi & \text{if} & |x| < a \\ 0 & \text{if} & |x| > a. \end{cases}$$

Equating real parts on the both sides, we obtain

$$\int_{-\infty}^{\infty} \frac{\sin sa \sin sx}{s}\, ds = \begin{cases} \pi & \text{if} & |x| < a \\ 0 & \text{if} & |x| > a. \end{cases}$$

Now if $x = 0$ and $a = 1$, then the second part gives

$$\int_{-\infty}^{\infty} \frac{\sin s}{s}\, ds = \pi \Rightarrow 2\int_{0}^{\infty} \frac{\sin s}{s}\, ds = \pi \Rightarrow \int_{0}^{\infty} \frac{\sin s}{s}\, ds = \frac{\pi}{2}.$$

Exercise 19.1.12. (i) Find the Fourier transform of $F(x) = \begin{cases} (1-x^2), & |x| < 1 \\ 0 & |x| > 1. \end{cases}$
and hence evaluate $\int_0^\infty \left(\frac{x\cos x - \sin x}{x^3}\right)\cos\left(\frac{x}{2}\right)dx$. **Answer:** $f(s) = \frac{4}{s^3}(\sin s - s\cos s); -\frac{3\pi}{16}$.
(ii) Show that the Fourier transform of $f(x) = e^{-x^2/2}$ is $\sqrt{2\pi}e^{-s^2/2}$
(iii) Find the complex Fourier transform of $f(x) = e^{-a|x|}$ **Answer:** $f(s) = \frac{2a}{s^2+a^2}$.
(iv) Find the inverse Fourier transform of $\overline{f}(s) = e^{-|s|y}$ **Answer:** $F^{-1}\{\overline{f}(s)\} = \frac{y}{\pi(y^2+x^2)}$.

19.1.4 Problems related to Integral Equations

Example 19.1.13. Solve the integral equation $\int_0^\infty f(x)\cos sx\, dx = e^{-\lambda}$.

Solution: We have $\int_0^\infty f(x)\cos sx\, dx = e^{-s}$. By definition,

$$F_c\{f(x)\} = \int_0^\infty f(x)\cos sx\, dx = \overline{f}_c(s) \quad \text{and} \quad F_c^{-1}\{\overline{f}_c(s)\} = f(x) = \frac{2}{\pi}\int_0^\infty \overline{f}_c(s)\cos sx\, ds$$

Comparing this with the given equation, we have $\overline{f}_c(s) = e^{-s}$. Using this, we obtain

$$f(x) = \frac{2}{\pi}\int_0^\infty e^{-s}\cos xs\, ds = \frac{2}{\pi}\frac{1}{1+x^2}.$$

Example 19.1.14. Solve the integral equation $\int_0^\infty F(x)\sin(xt)\, dx = F(x) = \begin{cases} 1 & , & 0 \le t < 1 \\ 2 & , & 1 \le t < 2 \\ 0 & , & t \ge 2 \end{cases}$

Solution: By the definition, we know

$$F_s\{F(x)\} = \int_0^\infty F(x)\sin(sx)\, dx = f_c(s). \tag{19.1.1}$$

Then $f_s(s) = \begin{cases} 1 & , & 0 \le t < 1 \\ 2 & , & 1 \le t < 2 \\ 0 & , & t \ge 2 \end{cases}$
The sine inversion formula relative to (19.1.2) is

$$F_s^{-1}\{f_s(s)\} = F(x) = \frac{2}{\pi}\int_0^\infty f_s(s)\sin sx\, ds$$

From which we have

$$\begin{aligned}
\frac{\pi}{2}F(x) &= \int_0^\infty f_s(s)\ \sin sx\, ds \\
\Rightarrow \frac{\pi}{2}F(x) &= \int_0^1 f_s(s)\sin sx\, dx + \int_1^2 f_s(s)\ \sin sx\, ds + \int_2^\infty f_s(s)\ \sin sx\, ds \\
\Rightarrow \frac{\pi}{2}F(x) &= \int_0^1 1\cdot\sin sx\, dx + \int_1^2 2\cdot\ \sin sx\, ds + \int_2^\infty 0\cdot\ \sin sx\, ds \\
\Rightarrow \frac{\pi}{2}F(x) &= \frac{1}{x}\Big[(1-\cos x) + 2(\cos x - \cos 2x)\Big] = \frac{1+\cos x - 2\cos 2x}{x} \\
\Rightarrow F(x) &= \frac{2}{\pi x}(1+\cos x - 2\cos 2x)
\end{aligned}$$

Exercise 19.1.15. (i) Show that $\int_0^\infty \frac{\cos\lambda x}{\lambda^2+1} d\lambda = \frac{\pi}{2} e^{-x}$

(ii) Solve the integral equation $\int_0^\infty F(x)\cos\lambda x\, dx = \begin{cases} 1-\lambda, & 0 \le \lambda \le 1 \\ 0, & \lambda > 1. \end{cases}$

Answer: $F(x) = \frac{2}{\pi x^2}(1-\cos x)$

19.1.5 The finite Fourier Transform

Definition 19.1.16. The finite Fourier sine transform of $F(x)$: The finite Fourier sine transform of $F(x)$ on $0 < x < l$, is defined by

$$F_s\{F(x)\} = f_s(s) = \int_0^l F(x)\sin\frac{s\pi x}{l} dx$$

where s is a positive integer. The function $F(x)$ is then called the *inverse finite Fourier sine transform* of $f_s(s)$ and is given by

$$F_s^{-1}\{f_s(s)\} = F(x) = \frac{2}{l}\sum_{s=1}^{\infty} f_s(s)\ \sin\frac{s\pi x}{l}$$

This formula is obtained from Fourier sine series $f(x) = \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l}$.

Definition 19.1.17. The finite Fourier cosine transform of $F(x)$: The finite Fourier cosine transform of $F(x)$ on $0 < x < l$, is defined by

$$F_c\{F(x)\} = f_c(s) = \int_0^l F(x)\cos\frac{s\pi x}{l} dx$$

where s is a positive integer or zero. The function $F(x)$ is then called the *inverse finite Fourier cosine transform* of $f_c(s)$ and is given by

$$F_c^{-1}\{f_c(s)\} = F(x) = \frac{1}{l} f_c(0) + \frac{2}{l}\sum_{s=1}^{\infty} f_c(s)\ \cos\frac{s\pi x}{l}$$

This formula is obtained from Fourier cosine series $F(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} b_n \cos\frac{n\pi x}{l}$.

Theorem 19.1.18. Fourier Integral Theorem: It sates that if $f(x)$ satisfies the following conditions:

- $f(x)$ satisfies the Dirichlet conditions in every interval $-l \le x \le l$.
- $\int_{-\infty}^{\infty} |f(x)|\, dx$ converges, i.e., $f(x)$ is absolutely integrable in the interval $-\infty < x < \infty$, then

$$f(x) = \frac{1}{2\pi}\int_{s=-\infty}^{\infty}\int_{t=-\infty}^{\infty} f(t)\cos s(x-t)\, ds\, dt.$$

The integral on R.H.S is called *Fourier integral* or *Fourier integral expansion of* $f(x)$.

Theorem 19.1.19. Different forms of Fourier integral formula:

$$(i) f(x) = \frac{1}{\pi}\int_{s=0}^{\infty}\int_{t=-\infty}^{\infty} f(t)\cos s(x-t)\; ds\; dt.$$

$$(ii) f(x) = \frac{2}{\pi}\int_{0}^{\infty}\int_{0}^{\infty} f(t)\;\cos st\cos sx\; ds\; dt \qquad \text{(Cosine Form)}$$

$$(iii) f(x) = \frac{2}{\pi}\int_{0}^{\infty}\int_{0}^{\infty} f(t)\;\sin st\sin sx\; ds\; dt \qquad \text{(Sine Form)}$$

$$(iv) f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\; e^{-ist}e^{isx}\; ds\; dt \qquad \text{(Exponential Form)}$$

Theorem 19.1.20. Parseval's identity for Fourier series: Suppose the Fourier series corresponding to $f(x)$ converges uniformly to $f(x)$ in the interval $-l < x < l$, then

$$\frac{1}{l}\int_{-l}^{l}\Big[f(x)\Big]^2 dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2),$$

where the integral on L.H.S is supposed to exist.

Proof. Let the Fourier series of $f(x)$ converges uniformly to $f(x)$ at every point of the interval $-l < x < l$, so that

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\Big(a_n\cos\frac{n\pi x}{l} + b_n\sin\frac{n\pi x}{l}\Big) \tag{19.1.2}$$

and that term by term integration of this series is possible. Here

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\; dx\;\; (n = 0, 1, 2, 3, \ldots) \quad \text{and} \quad b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}\; dx\;\; (n = 1, 2, 3, \ldots)$$

Multiplying (19.1.2) by $f(x)$ and integrating term by term from $-l$ to l, we get

$$\begin{aligned}
\int_{-l}^{l}\Big[f(x)\Big]^2 dx &= \frac{a_0}{2}\int_{-l}^{l} f(x)\; dx + \sum_{n=1}^{\infty} a_n\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}dx + \sum_{n=1}^{\infty} b_n\int_{-l}^{l} f(x)\sin\frac{n\pi x}{l}dx\\
\Rightarrow \int_{-l}^{l}\Big[f(x)\Big]^2 dx &= \frac{a_0}{2}\cdot la_0 + \sum_{n=1}^{\infty} l(a_n^2 + b_n^2)\\
\Rightarrow \frac{1}{l}\int_{-l}^{l}\Big[f(x)\Big]^2 dx &= \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)
\end{aligned}$$

□

Theorem 19.1.21. Parseval's identity for Fourier transform. Rayleigh's Theorem: If $f(p)$ and $g(p)$ are complex Fourier transforms of $F(x)$ and $G(x)$ respectively, then

$$(i)\; \frac{1}{2\pi}\int_{-\infty}^{\infty} f(p)\;\overline{g(p)}\; dp = \int_{-\infty}^{\infty} F(x)\;\overline{G(x)}\; dx \quad \text{and} \quad (ii)\; \frac{1}{2\pi}\int_{-\infty}^{\infty} |f(p)|^2\; dp = \int_{-\infty}^{\infty} |F(x)|^2\; dx$$

where bar represents the complex conjugate.

Proof. Using the inversion formula for Fourier transform, we get

$$G(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(p)\, e^{ipx}\, dp \tag{19.1.3}$$

Taking conjugate complex of the both sides in (19.1.3), we obtain

$$\overline{G(x)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{g(p)} e^{-ipx}\, dp \tag{19.1.4}$$

Therefore,

$$\begin{aligned}
\int_{-\infty}^{\infty} F(x)\,\overline{G(x)}\,dx &= \int_{-\infty}^{\infty} F(x)\,dx \cdot \left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{g(p)}e^{-ipx}\,dp\right\}, && \text{(Using (19.1.4))}\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)\,\overline{g(p)}e^{-ipx}\,dx\,dp && \text{(On changing order of integration)}\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x)\,\overline{g(p)}e^{-ipx}\,dp\,dx\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{g(p)}\,dp\left[\int_{-\infty}^{\infty} F(x)\,e^{-ipx}\,dx\right]\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{g(p)}\,dp\,\{f(p)\} && \text{(By the def. of Fourier transform)}\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(p)\,\overline{g(p)}\,dp
\end{aligned}$$

Thus, we have proved that

$$\int_{-\infty}^{\infty} F(x)\,\overline{G(x)}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(p)\,\overline{g(p)}\,dp. \tag{19.1.5}$$

This proves the first part. Now putting $G(x) = F(x)$ in (19.1.5), we get

$$\int_{-\infty}^{\infty} F(x)\,\overline{F(x)}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(p)\,\overline{f(p)}\,dp \quad\Rightarrow\quad \int_{-\infty}^{\infty} |F(x)|^2\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |f(p)|^2\,dp.$$

□

Example 19.1.22. Use Parseval's identity to prove that

$$(i)\ \int_0^{\infty}\frac{dt}{(a^2+t^2)(b^2+t^2)} = \frac{\pi}{2ab(a+b)},\quad (ii)\ \int_0^{\infty}\frac{\sin(at)\,dt}{t(a^2+t^2)} = \frac{\pi}{2}\left(\frac{1-e^{-a^2}}{a^2}\right).$$

Solution: (i) Let $F(x) = e^{-ax}, G(x) = e^{-bx}$. Now,

$$f_c(p) = \int_0^{\infty} F(x)\,\cos(px)\,dx = \int_0^{\infty} e^{-ax}\cos(px)\,dx = \frac{a}{a^2+p^2}$$

Similarly, we can find $g_c(p) = \frac{b}{b^2+p^2}$. By Parseval's identity for Fourier transform, we get

$$\frac{2}{\pi}\int_0^{\infty} f_c(p)\,g_c(p)\,dp = \int_0^{\infty} F(x)\,G(x)\,dx.$$

Putting values, we get

$$\frac{2}{\pi}\int_0^\infty \frac{a}{(a^2+p^2)}\cdot\frac{b}{(b^2+p^2)}dp = \int_0^\infty e^{-ax}\cdot e^{-bx}\,dx$$
$$\Rightarrow \int_0^\infty \frac{dp}{(b^2+p^2)(a^2+p^2)} = \frac{\pi}{2ab}\left[\frac{e^{-(a+b)x}}{-(a+b)}\right]_{x=0}^\infty = \frac{\pi}{2ab(a+b)}\{1-0\}$$
$$\Rightarrow \int_0^\infty \frac{dt}{(b^2+t^2)(a^2+t^2)} = \frac{\pi}{2ab(a+b)}$$

(ii) Let $F(x)=e^{-ax}$, then $f_c(p)=\frac{a}{a^2+p^2}$. Also let $G(x)=\begin{cases} 1 &, \quad 0<x<a \\ 0 &, \quad x>a. \end{cases}$
Then

$$\begin{aligned} g_c(p) &= \int_0^\infty G(x)\ \cos(px)\,dx \\ &= \int_0^a G(x)\ \cos(px)\,dx + \int_a^\infty G(x)\ \cos(px)\,dx \\ &= \int_0^a \cos(px)\,dx + \int_a^\infty 0\cdot\ \cos(px)\,dx \\ &= \left[\frac{\sin(px)}{p}\right]_{x=0}^a = \frac{\sin(pa)}{p} \end{aligned}$$

$$\text{Since}\quad \frac{2}{\pi}\int_0^\infty f_c(p)\,g_c(p)\,dp = \int_0^\infty F(x)\,G(x)\,dx$$
$$\Rightarrow \frac{2}{\pi}\int_0^\infty \frac{a}{a^2+p^2}\frac{\sin(pa)}{p}\,dp = \int_0^\infty e^{-ax}\,G(x)\,dx$$
$$\Rightarrow \frac{2a}{\pi}\int_0^\infty \frac{\sin(pa)\,dp}{p(a^2+p^2)}\,dp = \int_0^a e^{-ax}\,G(x)\,dx + \int_a^\infty e^{-ax}\,G(x)\,dx$$
$$\Rightarrow \frac{2a}{\pi}\int_0^\infty \frac{\sin(pa)\,dp}{p(a^2+p^2)}\,dp = \int_0^a e^{-ax}\cdot 1\,dx + \int_a^\infty e^{-ax}\cdot 0\,dx$$
$$\Rightarrow \frac{2a}{\pi}\int_0^\infty \frac{\sin(pa)\,dp}{p(a^2+p^2)}\,dp = \frac{1}{a}(1-e^{-a^2})$$
$$\Rightarrow \int_0^\infty \frac{\sin(pa)\,dp}{p(a^2+p^2)}\,dp = \frac{\pi}{2a^2}(1-e^{-a^2}).$$

Example 19.1.23. Find the Fourier transform of $f(x)$ defined by

$$f(x)=\begin{cases} 1 &, \quad |x|<a \\ 0 &, \quad |x|>a. \end{cases}$$

and hence prove that

$$\int_0^\infty \frac{\sin^2(ax)}{x^2}dx = \frac{\pi a}{2}.$$

Solution: First Part:

$$\begin{aligned}
F\{f(x)\} &= \int_{-\infty}^{\infty} e^{-ipx} f(x)\, dx \\
&= \int_{-\infty}^{-a} e^{-ipx}\ f(x)\ dx + \int_{-a}^{a} e^{-ipx}\ f(x)\ dx + \int_{a}^{\infty} e^{-ipx}\ f(x)\ dx \\
&= \int_{\infty}^{a} e^{ipy}\ f(-y)\ (-dy) + \int_{-a}^{a} e^{-ipx}\ dx + \int_{a}^{\infty} e^{-ipx}\ \cdot 0\ dx \\
&= \int_{a}^{\infty} e^{ipy} \cdot 0 \cdot dy + \frac{1}{-ip}\left(e^{-ipx}\right)_{-a}^{a} + 0 \\
&= \frac{e^{ipa} - e^{-ipa}}{ip} = \frac{2}{p}\sin pa = \overline{f}(p)
\end{aligned}$$

Second Part: Using Parseval's identity for Fourier integral, we get

$$\begin{aligned}
\textstyle\int_{-\infty}^{\infty} |f(x)|^2 dx &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\overline{f}(p)|^2 dp \\
\Rightarrow \textstyle\int_{-a}^{a} 1^2 dx &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{4}{p^2}\sin^2 pa\ dp \\
\Rightarrow 2a &= \frac{2}{2\pi}\int_{0}^{\infty} \frac{4}{p^2}\sin^2 pa\ dp \\
\Rightarrow \textstyle\int_0^{\infty} \frac{\sin^2(ax)}{x^2} dx &= \frac{\pi a}{2}.
\end{aligned}$$

19.1.6 Problems related to finite Fourier Sine and Cosine transform:

Example 19.1.24. Find finite Fourier $sine$ and $cosine$ transform of

$$f(x) = x^2, \quad 0 < x < 4.$$

Solution: (i)

$$\begin{aligned}
F_s\{f(x)\} &= \int_0^c f(x)\ \sin\frac{n\pi x}{c} dx = \int_0^4 x^2 \sin\frac{n\pi x}{4} dx \\
&= \left[-\frac{4}{n\pi}x^2\cos\frac{n\pi x}{4}\right]_{x=0}^{4} + \int_0^4 2x\frac{4}{n\pi}\cos\frac{n\pi x}{4} dx \\
&= -\frac{4^3}{n\pi}\cos n\pi + \frac{8}{n\pi}\left[\frac{4x}{n\pi}\sin\frac{n\pi x}{4} + \frac{4^2}{n^2\pi^2}\cos\frac{n\pi x}{4}\right]_{x=0}^{4} \\
&= -\frac{4^3}{n\pi}\cos n\pi + \frac{8\cdot 4^2}{n\pi\cdot n^2\pi^2}(\cos n\pi - 1) \\
\therefore \overline{f}_s(n) &= -\frac{64}{n\pi}\cos n\pi + \frac{128}{n^3}
\end{aligned}$$

(ii)

$$\begin{aligned}
F_c\{f(x)\} &= \overline{f}_c(n) = \int_0^4 f(x)\cos\frac{n\pi x}{4} = \int_0^4 x^2\cos\frac{n\pi x}{4}dx \\
&= \left[\frac{4x^2}{n\pi}\sin\frac{n\pi x}{4}\right]_{x=0}^4 - \int_0^4 \frac{4}{n\pi}2x\sin\frac{n\pi x}{4}dx \\
&= 0 - \frac{8}{n\pi}\left[-\frac{4x}{n\pi}\cos\frac{npx}{4} + \frac{4^2}{n^2\pi^2}\sin\frac{n\pi x}{4}\right]_{x=0}^4 \\
&= \frac{128}{n^2\pi^2}\cos n\pi
\end{aligned}$$

Example 19.1.25. Find $f(x)$ if its finite $sine$ transform is given by

$$\overline{f}_s(s) = \frac{1-\cos s\pi}{s^2\pi^2}, \text{ where } 0 < x < \pi,\ s = 1, 2, 3, \ldots$$

Solution: We know that

$$f(x) = \frac{2}{l}\sum_{n=1}^{\infty}\overline{f}_s(n)\sin\frac{n\pi x}{l}.$$

In our case this becomes

$$f(x) = \frac{2}{\pi}\sum_{s=1}^{\infty}\overline{f}_s(s)\sin\left(\frac{s\pi x}{\pi}\right) = \frac{2}{\pi}\sum_{s=1}^{\infty}\frac{1-\cos s\pi}{s^2\pi^2}\sin sx$$

$$\Rightarrow f(x) = \frac{2}{\pi^3}\sum_{s=1}^{\infty}\left(\frac{1-\cos\pi s}{s^2}\right)\sin xs.$$

Exercise 19.1.26. (i) Find the finite cosine transform of $\left(1-\frac{x}{\pi}\right)^2$. **Answer:** $f_c(s) = \begin{cases} \frac{\pi}{3}, & s = 0 \\ \frac{2}{\pi s^2}, & s = 1, 2, 3, \ldots \end{cases}$

(ii) Show that the finite sine transform of $\frac{x}{\pi}$ is $(-1)^{s+1}\frac{1}{s}$

(iii) When $f(x) = \sin mx$, where, m is a positive integer, show that $f_s(p) = \begin{cases} 0, & p \neq m \\ \frac{\pi}{2}, & p = m \end{cases}$

(iv) If $f_s(n) = 2\pi\frac{(-1)^{n-1}}{n^2}, n = 1, 2, 3, \ldots$ where $0 < x < \pi$, then find $f(x)$. **Answer:** $\frac{2}{s}\sin\frac{s\pi}{2},\ s > 0$

(v) Find the finite cosine transform of $f(x)$ if $f(x) = \begin{cases} 1, & 0 < x < \frac{\pi}{2} \\ -1, & \frac{\pi}{2} < x < \pi \end{cases}$

(vi) Show that $f_c\left\{\frac{x^2}{2\pi} - \frac{\pi}{6}\right\} = \begin{cases} 0, & n = 0 \\ (-1)^n/n^2, & n = 1, 2, 3, \ldots \end{cases}$

Unit 20

Course Structure

- Hankel Transform, inversion formula and Finite Hankel transform, solution of two-dimensional Laplace and one-dimensional diffusion & wave equation by integral transform.

20.1 Introduction

Hankel transforms are integral transformations whose kernels are Bessel functions. They are sometimes referred to as Bessel transforms. When we are dealing with problems that show circular symmetry, Hankel transforms may be very useful. Laplace's partial differential equation in cylindrical coordinates can be transformed into an ordinary differential equation by using the Hankel transform. Because the Hankel transform is the two-dimensional Fourier transform of a circularly symmetric function, it plays an important role in optical data processing.

20.1.1 Definition: Infinite Hankel Transform

The *infinite Hankel transform* of a function $f(x),\ 0 < x < \infty$, is defined as

$$H\{f(x)\} = \overline{f}(s) = \int_0^\infty f_n(x) \cdot xJ_n(sx)\, dx \tag{20.1.1}$$

where $J_n(sx)$ is the Bessel function of the first kind of order n. Also here $\overline{f}(s)$ is defined as *Hankel transform of order n* of the function $f(x)$.

Remark: In the integral (20.1.1), $xJ_n(sx)$ is called the Kernel of the transformation.

20.1.2 Definition: Inverse Hankel Transform

If $\overline{f}(s)$ is the infinite Hankel transform of order n of the function $f(x)$, then we write

$$H\{f(x)\} = \overline{f}(s) = \int_0^\infty f(x) \cdot xJ_n(sx)\, dx. \tag{20.1.2}$$

Here $f(x)$ is called the *inverse transform* of the function $\overline{f}(s)$ and we write $f(x) = H^{-1}\{\overline{f}(s)\}$. The inverse formula for inverse Hankel Transform is

$$f(x) = H^{-1}\{\overline{f}(s)\} = \int_0^\infty \overline{f}(s) \cdot sJ_n(sx)\, ds. \tag{20.1.3}$$

20.1.3 Some Important Results on Bessel functions

I. Bessel function of first kind: $J_n(x) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!\Gamma(n+r+1)} \left(\frac{x}{2}\right)^{n+2r}$

II. Recurrence formula for $J_n(x)$:

$(i)\ xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x)$ $\qquad (ii)\ xJ_n'(x) = -nJ_n(x) + xJ_{n-1}(x)$

$(iii)\ 2J_n(x) = J_{n-1}(x) - J_{n+1}(x)$ $\qquad (iv)\ 2nJ_n'(x) = x[J_{n-1}(x) + J_{n+1}(x)]$

$(v)\ \frac{d}{dx}\left[x^{-n}J_n(x)\right] = -x^{-n}J_{n+1}(x)$ $\qquad (vi)\ \frac{d}{dx}\left[x^{n}J_n(x)\right] = x^{n}J_{n-1}(x)$

III. Infinite Integrals Involving Bessel Functions

$(i)\ \int_0^\infty e^{-ax}J_0(sx)\, dx = (a^2+s^2)^{-1/2}$ $\qquad (ii)\ \int_0^\infty e^{-ax}J_1(sx)\, dx = \frac{1}{s} - \frac{a}{s(a^2+s^2)^{1/2}}$

$(iii)\ \int_0^\infty xe^{-ax}J_0(sx)\, dx = a(a^2+s^2)^{-3/2}$ $\qquad (iv)\ \int_0^\infty e^{-ax}J_1(sx)\, dx = s(a^2+s^2)^{-3/2}$

$(v)\ \int_0^\infty \frac{e^{-ax}}{x}J_1(sx)\, dx = \frac{(a^2+s^2)^{1/2}-a}{s}$

Theorem: If $f(x)$ and $g(x)$ are two functions and a, b two constants, then

$$H\{af(x)+bg(x)\} = aH\{f(x)\} + bH\{g(x)\}$$

Proof:

$$\begin{aligned} H\{af(x)+bg(x)\} &= \int_0^\infty x[af(x)+bg(x)]J_n(sx)\, dx \\ &= a\int_0^\infty xf(x)J_n(sx)\, dx \ + \ b\int_0^\infty xg(x)J_n(sx)\, dx \\ &= aH\{f(x)\} + bH\{g(x)\} \end{aligned}$$

Theorem: If $H\{f(x)\} = \overline{f}(s)$, then $H\{f(ax)\} = \frac{1}{a^2}\overline{f}\left(\frac{s}{a}\right)$, a being a constant.

Proof: Let $H\{f(x)\} = \overline{f}(s)$. Then

$$\begin{aligned} H\{f(ax)\} &= \int_0^\infty xf(ax)J_n(sx)\, dx \\ &= \int_0^\infty \frac{y}{a}f(y)J_n\left(\frac{sy}{a}\right)\frac{dy}{a}, \quad \text{where } ax = y \\ &= \frac{1}{a^2}\int_0^\infty yf(y)J_n\left(\frac{s}{a}y\right)\, dy \\ &= \frac{1}{a^2}\overline{f}\left(\frac{s}{a}\right) \end{aligned}$$

Theorem: *Hankel transform of the derivatives of a function*

Let $\overline{f}(s)$ be the Hankel transform of order n of the function $f(x)$ and $\overline{f'}_n(s)$ is the transform of $f'(x)$. Then

$$\overline{f'}_n(s) = -\frac{s}{2n}\left[(n+1)\overline{f}_{n-1}(s) - (n-1)\overline{f}_{n+1}(s)\right].$$

Proof:

$$\begin{aligned} \overline{f}_n(s) &= \int_0^\infty x f(x) J_n(sx)\, dx \\ \text{and} \quad \overline{f'}_n(s) &= \int_0^\infty x \frac{df}{dx} J_n(sx)\, dx \quad \Rightarrow \quad \overline{f'}_n(s) = \int_0^\infty \frac{df}{dx}\left[xJ_n(sx)\right]\, dx \end{aligned}$$

Integrating by parts and assuming that $x\, f(x) \to 0$ as $x \to 0$ and $x\, f(x) \to 0$ as $x \to \infty$, we obtain

$$\begin{aligned} \overline{f'}_n(s) &= \Big[f(x)\cdot xJ_n(sx)\Big]_0^\infty - \int_0^\infty f(x)\cdot \frac{d}{dx}[xJ_n(sx)]\, dx \\ &= 0 - \int_0^\infty f(x)\,[J_n(sx) + sx\, J_n'(sx)]\, dx \end{aligned} \tag{20.1.4}$$

But $xJ_n'(x) = -nJ_n(x) + xJ_{n-1}(x)$. Replacing x by sx, we get

$$\begin{aligned} & sxJ_n'(sx) = -nJ_n(sx) + sxJ_{n-1}(sx) \\ & \Rightarrow sxJ_n'(sx) + J_n(sx) = (1-n)J_n(sx) + sxJ_{n-1}(sx) \end{aligned}$$

Using this in Eq.(20.1.4), we have

$$\begin{aligned} \overline{f'}_n(sx) &= -\int_0^\infty f(x)\Big[(1-n)J_n(sx) + sxJ_{n-1}(sx)\Big]\, dx \\ \Rightarrow \overline{f'}_n(sx) &= -(1-n)\int_0^\infty f(x)J_n(sx)\, dx - s\int_0^\infty xf(x)J_{n-1}(sx)\, dx \\ \Rightarrow \overline{f'}_n(sx) &= -(1-n)\int_0^\infty f(x)J_n(sx)\, dx - s\overline{f}_{n-1}(s) \end{aligned} \tag{20.1.5}$$

We know from Recurrence formula for $J_n(x)$, that

$$2nJ_n(x) = x\Big[J_{n-1}(x) + J_{n+1}(x)\Big]$$

Replacing x by sx, we have

$$2nJ_n(sx) = sx\Big[J_{n-1}(sx) + J_{n+1}(sx)\Big]$$

Multiplying this by $f(x)$ and then integrating, we obtain

$$\begin{aligned} 2n\int_0^\infty f(x)J_{sx}\, dx &= s\left[\int_0^\infty xf(x)J_{n-1}(sx)\, dx + \int_0^\infty xf(x)J_{n+1}(sx)\, dx\right] = s\Big[\overline{f}_{n-1}(s) + \overline{f}_{n+1}(s)\Big] \\ \Rightarrow \int_0^\infty f(x)J_n(sx)\, dx &= \frac{s}{2n}\Big[\overline{f}_{n-1}(s) + \overline{f}_{n+1}(s)\Big] \end{aligned}$$

In this event Eq.(20.1.5) becomes

$$
\begin{aligned}
\overline{f'}_n(s) &= -\frac{(1-n)s}{2n}\Big[\overline{f}_{n-1}(s) + f_{n+1}(s)\Big] - s\overline{f}_{n-1}(s) \\
\Rightarrow \overline{f'}_n(s) &= \frac{s}{2n}\Big[(n-1)\overline{f}_{n-1}(s) + (n-1)\overline{f}_{n+1} - 2n\overline{f}_{n+1}(s)\Big] \\
\Rightarrow \overline{f'}_n(s) &= \frac{s}{2n}\Big[-(n+1)\overline{f}_{n-1}(s) + (n-1)\overline{f}_{n+1}\Big] \\
\Rightarrow \overline{f'}_n(s) &= -\frac{s}{2n}\Big[(n+1)\overline{f}_{n-1}(s) - (n-1)\overline{f}_{n+1}(s)\Big] \qquad (20.1.6)
\end{aligned}
$$

Remark 1. When $n=1$, from Eq.(20.1.6) we have

$$
\begin{aligned}
\overline{f'}_1(s) &= -s\overline{f}_0(s) \\
\Rightarrow H\{f'(x), n=1\} &= -sH\{f(x), n=0\}
\end{aligned}
$$

Remark 2. When $n=2$, from Eq.(20.1.6) we have

$$\overline{f'}_2(s) = -\frac{s}{4}\Big[3\overline{f}_1(s) - \overline{f}_3(s)\Big]$$

Remark 3. When $n=3$, from Eq.(20.1.6) we have

$$\overline{f'}_3(s) = -\frac{s}{6}\Big[4\overline{f}_2(s) - 2\overline{f}_4(s)\Big]$$

Result 1. Prove that

$$\overline{f''}_n(s) = \frac{s^2}{4}\left[\left(\frac{n+1}{n-1}\right)\overline{f}_{n-2}(s) - 2\left(\frac{n^2-3}{n^2-1}\right)\overline{f}_n(s) + \left(\frac{n-1}{n+1}\right)\overline{f}_{n+2}(s)\right].$$

Proof: From Eq.(20.1.6) we have

$$\overline{f'}_n(s) = -s\left[\left(\frac{n+1}{2n}\right)\overline{f}_{n-1}(s) - \left(\frac{n-1}{2n}\right)\overline{f}_{n+1}(s)\right] \qquad (20.1.7)$$

Replacing f by f', we get

$$\overline{f''}_n(s) = -s\left[\left(\frac{n+1}{2n}\right)\overline{f'}_{n-1}(s) - \left(\frac{n-1}{2n}\right)\overline{f'}_{n+1}(s)\right] \qquad (20.1.8)$$

Replacing n by $(n-1)$, and $(n+1)$ respectively in Eq.(20.1.7), we get

$$
\begin{aligned}
\overline{f'}_{n-1}(s) &= -s\left[\left(\frac{n}{2(n-1)}\right)\overline{f}_{n-2}(s) - \left(\frac{n-2}{2(n-1)}\right)\overline{f}_n(s)\right] \\
\overline{f'}_{n+1}(s) &= -s\left[\left(\frac{n+2}{2(n+1)}\right)\overline{f}_n(s) - \left(\frac{n}{2(n+1)}\right)\overline{f}_{n+2}(s)\right]
\end{aligned}
$$

Writing Eq.(20.1.8) with the help of these two equations, we obtain

$$
\begin{aligned}
\overline{f''}_n(s) &= \frac{s^2}{4}\left[\left(\frac{n+1}{n}\right)\left\{\frac{n}{n-1}\overline{f}_{n-2}(s) - \frac{n-2}{n-1}\overline{f}_n(s)\right\} - \left(\frac{n-1}{n}\right)\left\{\left(\frac{n+2}{n+1}\right)\overline{f}_n(s) - \left(\frac{n}{n+1}\right)\overline{f}_{n+2}(s)\right\}\right] \\
&= \frac{s^2}{4}\left[\left(\frac{n+1}{n-1}\right)\overline{f}_{n-2}(s) \quad 2\left(\frac{n^2-3}{n^2-1}\right)\overline{f}_n(s) + \left(\frac{n-1}{n+1}\right)\overline{f}_{n+2}(s)\right].
\end{aligned}
$$

Example: Find the Hankel transform of $\frac{df}{dx}$, when $f = \frac{e^{-ax}}{x}$ and $n = 1$.

Solution: Let $f(x) = \frac{e^{-ax}}{x}$. To determine $H\left\{\frac{df}{dx}, n = 1\right\} = \overline{f}'_1(s)$, we know that

$$\begin{aligned}
\overline{f}'_1(s) = -s f_0(s) &= -s \int_0^\infty x f(x) J_0(sx)\, dx \\
&= -s \int_0^\infty x \frac{e^{-ax}}{x} J_0(sx)\, dx \\
&= -s \int_0^\infty e^{-ax} J_0(sx)\, dx \\
&= -s(a^2 + s^2)^{-1/2}.
\end{aligned}$$

Example: Find the Hankel transform of $x^{-2}e^{-x}$ of order one.

Solution:

$$\begin{aligned}
H\{x^{-2}e^{-x}, n = 1\} &= \int_0^\infty x^{-2}e^{-x} x J_1(sx)\, dx \\
&= \int_0^\infty \frac{e^{-x}}{x} J_1(sx)\, dx = \frac{(1+s^2)^{1/2} - 1}{s}.
\end{aligned}$$

Example: Evaluate $H^{-1}\{s^{-2}e^{as}\}$ when $n = 1$, that is, find out inverse Hankel transform of $s^{-2}e^{-as}$ of order one.

Solution:

$$\begin{aligned}
H^{-1}\{s^{-2}e^{as}, n = 1\} &= \int_0^\infty s^{-2}e^{-as} s J_1(sx)\, ds \\
&= \int_0^\infty \frac{e^{-as}}{s} J_1(sx)\, ds \\
&= \frac{(a^2 + x^2)^{1/2} - a}{x}
\end{aligned}$$

Example: Find the Hankel transformation of

$$f(x) = \begin{cases} 1 & 0 < x < a, \quad n = 0, \\ 0 & x > a, \qquad\ \ n = 0. \end{cases}$$

Solution:

$$\begin{aligned}
H\{f(x), n = 0\} &= \int_0^\infty f(x) \cdot x J_0(sx)\, dx \\
&= \int_0^a f(x) \cdot x J_0(sx)\, dx + \int_0^\infty f(x) \cdot x J_0(sx)\, dx \\
&= \int_0^a 1 \cdot x J_0(sx)\, dx + \int_0^\infty 0 \cdot x J_0(sx)\, dx \\
&= \int_0^a x J_0(sx)\, dx \qquad (20.1.9)
\end{aligned}$$

By Recurrence formula for Bessel's function, we have

$$\frac{d}{dx}\{x^n J_n(x)\} = x^n J_{n-1}(x).$$

Replacing n and x by 1 and sx respectively,

$$\begin{aligned} \frac{d}{s\,dx}\{sxJ_1(sx)\} &= sxJ_0(sx) \\ \Rightarrow \frac{d}{dx}\{xJ_1(sx)\} &= sxJ_0(sx). \end{aligned}$$

Integrating this form $x = 0$ to $x = a$,

$$\Big[xJ_1(sx)\Big]_0^a = s\int_0^a xJ_0(sx)\,dx \quad \Rightarrow \int_0^a xJ_0(sx)\,dx = \frac{a}{s}J_1(sa). \tag{20.1.10}$$

Now using Eq.(20.1.10) in Eq.(20.1.9) we obtain

$$H\{f(x), n = 0\} = \frac{a}{s}J_1(sa)$$

Exercise 20.1.1. (i) Find the Fourier transform of $F(x) = \begin{cases} (1-x^2), & |x| < 1 \\ 0 & |x| > 1. \end{cases}$
and hence evaluate $\int_0^\infty \left(\frac{x\cos x - \sin x}{x^3}\right)\cos\left(\frac{x}{2}\right)dx$. **Answer:** $f(s) = \frac{4}{s^3}(\sin s - s\cos s); -\frac{3\pi}{16}$.

20.1.4 Finite Hankel Transform

Definition 20.1.2. Finite Hankel Transform: If $f(x)$ satisfies Dirichlet's conditions in the closed interval $[0, a]$, then its finite Hankel transform $\overline{f}(s_i)$ of order n is given by

$$\overline{f}(s_i) = \int_0^a f(x)\cdot xJ_n(xs_i)\,dx, \tag{20.1.11}$$

where a is positive root of the transcendental equation

$$J_n(as_i) = 0 \tag{20.1.12}$$

If the function $f(x)$ is continuous at any point of the interval $[0, a]$, then the inversion formula for $\overline{f}(s_i)$ is

$$f(x) = \frac{2}{a^2}\sum_i \overline{f}(s_i)\frac{J_n(xs_i}{[J_n'(ac_i)]^2} \tag{20.1.13}$$

where the sum is taken over all the positive roots of the Eq.(20.1.12). If $f(x)$ is represented by generalised Fourier Bessel series

$$f(x) = \sum_i c_iJ_n(xs_i), \quad 0 \le x \le a, \tag{20.1.14}$$

then the coefficient c_i is given by

$$\begin{aligned} c_i &= \frac{2}{a^2J_{n+1}^2(as_i)}\int_0^a f(x)\cdot xJ_n(xs_i)\,dx \\ &= \frac{2\overline{f}}{a^2[J_{n+1}(as_i)]^2} = \frac{2\overline{f}(s_i)}{a^2[J_n'(as_i)]^2} \end{aligned}$$

The recurrence formula,

$$\begin{aligned}
& xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x). \\
& \therefore as_iJ_n'(as_i) = nJ_n(as_i) - as_iJ_{n+1}(as_i), \quad \text{[Replacing } x \text{ by } as_i] \\
& \Rightarrow as_iJ_n'(as_i) = -as_iJ_{n+1}(as_i) \quad \text{[Using (20.1.12)]} \\
& \Rightarrow J_n'(as_i) = -J_{n+1}(as_i).
\end{aligned}$$

Consequently,

$$f(x) = \frac{2}{a^2}\sum_i \overline{f}(s_i)\frac{J_n(xs_i)}{[J_{n+1}(as_i)]^2}$$

Remark 20.1.3. It has been found in practice that the choice of unity as the upper limit of the integral defining the transform is more convenient. Therefore the definition of finite Hankel transform becomes

$$\overline{f}(s_i) = \int_0^1 f(x)\cdot xJ_n(xs_i)\,dx$$

Theorem 20.1.4. Finite Hankel transform of $\frac{df}{dx}$, i.e.,

$$H_n\left(\frac{df}{dx}\right) = \int_0^a \frac{df}{dx}xJ_n(sx)\,dx,$$

where s is any root of $J_n(sa) = 0$. To show that

$$H_n\left\{\frac{df}{dx}\right\} = \frac{s}{2n}[(n-1)H_{n+1}\{f(x)\} - (n+1)H_{n-1}\{f(x)\}]$$

Proof. The finite Hankel transform of $\frac{df}{dx}$ of order n is denoted by $H_n\left\{\frac{df}{dx}\right\}$.

$$\begin{aligned}
H_n\left\{\frac{df}{dx}\right\} &= \int_0^a \frac{df}{dx}\cdot xJ_n(sx)\,dx \\
&= \Big[f(x)\cdot xJ_n(sx)\Big]_{x=0}^{x=a} - \int_0^a f(x)\cdot\frac{d}{dx}\{xJ_n(sx)\}\,dx \\
&= -\int_0^a f(x)\frac{d}{dx}\{xJ_n(sx)\}\,dx \qquad (20.1.15)
\end{aligned}$$

By Recurrence formula,

$$2J_n'(x) = J_{n-1}(x) - J_{n+1}(x) \quad \text{and} \quad 2nJ_n(x) = x[J_{n-1}(x) + J_{n+1}(x)]$$

Replacing x by sx in both equations,

$$2J_n'(sx) = J_{n-1}(sx) - J_{n+1}(sx) \quad \text{and} \quad 2nJ_n(sx) = sx[J_{n-1}(sx) + J_{n+1}(sx)]$$

Now

$$\begin{aligned}
\frac{d}{dx}\{xJ_n(sx)\} &= J_n(sx) + sxJ_n'(sx) \\
&= \frac{sx}{2n}[J_{n-1}(sx) + J_{n+1}(sx)] + \frac{sx}{2}[J_{n-1}(sx) - J_{n+1}(sx)] \\
&= \frac{sx}{2n}[J_{n-1}(sx)\cdot(1+n) + (1-n)J_{n+1}(sx)]
\end{aligned}$$

Now from Eq.(20.1.15) we have

$$\begin{aligned}
H_n\left\{\frac{df}{dx}\right\} &= -\int_0^a f(x)\cdot\frac{sx}{2n}[J_{n-1}(sx)\cdot(n-1)+(1-n)J_{n+1}(sx)] \\
&= -\frac{s}{2n}\int_0^a [f(x)\cdot xJ_{n-1}(sx)\cdot(1+n)-(n-1)f(x)xJ_{n+1}(sx)]dx \\
&= \frac{s}{2n}[(n-1)H_{n+1}\{f(x)\}-(1+n)H_{n-1}\{f(x)\}]
\end{aligned}$$

□

Corollary 20.1.5. If $n=1$, the the last gives

$$H_1\left\{\frac{df}{dx}\right\} = \frac{s}{2}[0-2H_0\{f(x)\}] = -sH_0\{f(x)\}$$

Theorem 20.1.6.

$$H\left\{\frac{d^2f}{dx^2}+\frac{1}{x}\frac{df}{dx}\right\} = \frac{s}{2n}\left[-H_{n-1}\left\{\frac{df}{dx}\right\}+H_{n+1}\left\{\frac{df}{dx}\right\}\right]$$
$$H\left\{\frac{d^2f}{dx^2}+\frac{1}{x}\frac{df}{dx}-\frac{n^2}{x^2}f\right\} = -sa\ f(a)J_n'(sa)-s^2H_n\{f(x)\}$$

Proof. Proof of the above theorems are left as exercise. □

Example 20.1.7. Show that

$$H_0(c) = \frac{ca}{s}J_1(as)$$

Solution:

$$H_0\{c\} = \int_0^a cxJ_0(sx)\,dx = c\int_0^a xJ_0(sx)\,dx \tag{20.1.16}$$

By recurrence formula No. (vi), we know that

$$\frac{d}{dx}[x^nJ_n(x)] = x^nJ_{n-1}(x).$$

Putting $n=1$, $\quad \frac{d}{dx}[xJ_1(x)] = xJ_0(x).$

Replacing x by sx, we have

$$\frac{d}{s\,dx}[sxJ_1(sx)] = sxJ_0(sx)$$
$$\Rightarrow \frac{d}{dx}\{xJ_1(sx)\} = sxJ_0(sx).$$

Using this in (20.1.16), we get

$$H_0\{c\} = \frac{c}{s}\int_0^a \frac{d}{dx}\{xJ_1(sx)\} = \frac{c}{s}\Big[xJ_1(sx)\Big]_0^\infty = \frac{ca}{s}J_1(sa).$$

Example 20.1.8. Find finite Hankel transform of x^2 if $xJ_0(sx)$ is the Kernel of the transform.

Solution: By recurrence formula No. (iv) and (vi) we have

$$2nJ_n(x) = x[J_{n-1}(x) + J_{n+1}(x)] \quad \text{and} \quad \frac{d}{dx}[x^n J_n(x)] = x^n J_{n-1}(x)$$

Replacing x by sx, we have

$$2nJ_n(sx) = sx[J_{n-1}(sx) + J_{n+1}(sx)] \tag{20.1.17}$$

$$\frac{d}{s\,dx}[x^n J_n(sx)] = x^n J_{n-1}(sx). \tag{20.1.18}$$

Now

$$H_0\{x^2\} = \int_0^a x^2 \cdot xJ_0(sx)\,dx = \int_0^a x^2 \cdot \frac{d}{s\,dx}\{xJ_1(sx)\}dx, \quad \text{according to (20.1.18)}$$

Integrating by parts, we obtain

$$\begin{aligned}
H_0\{x^2\} &= \frac{1}{s}\Big[x^2 \cdot xJ_1(sx)\Big]_0^a - \frac{1}{s}\int_0^a 2x \cdot xJ_1(sx)\,dx \\
&= \frac{a^3}{s}J_1(sa) - \frac{2}{s}\int_0^a x^2 J_1(sx)\,dx \\
&= \frac{a^3}{s}J_1(sa) - \frac{2}{s}\int_0^a \frac{d}{s\,dx}[x^2 J_2(sx)]\,dx, \quad \text{[according to (20.1.18)]} \\
&= \frac{a^3}{s}J_1(sa) - \frac{2}{s^2}\Big[x^2 J_2(sx)\Big]_0^a \\
&= \frac{a^3}{s}J_1(sa) - \frac{2a^2}{s^2}J_2(sa).
\end{aligned} \tag{20.1.19}$$

Putting $n = 1,\ x = a$ in (20.1.17), we have

$$2J_1(sa) = sa[J_0(sa) + J_2(sa)] \quad \Rightarrow \quad \frac{2}{sa}J_1(sa) - J_0(sa) = J_2(sa).$$

Putting this in Eq.(20.1.19), we obtain

$$\begin{aligned}
H_0\{x^2\} &= \frac{a^2}{s}J_1(sa) - \frac{2a^2}{s^2}\left[\frac{2}{sa}J_1(sa) - J_0(sa)\right] \\
&= \frac{a^3}{s}J_1(sa) - \frac{4a^2}{s^3 a}J_1(sa) + \frac{2a^2}{s^2}J_0(sa) \\
&= \frac{a^2}{s^2}\left[\left(as - \frac{4}{as}\right)J_1(sa) + 2J_0(sa)\right]
\end{aligned}$$

Example 20.1.9. Find the finite Hankel transform of

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) - \frac{n^2 V}{r^2}, \quad \text{where} \quad V = \begin{cases} 0 & \textit{when } r = 0 \\ V_1 & \textit{when } r = 1. \end{cases}$$

Solution: From the problem it is clear that we should take the limits $x = 0$ and $x = 1$ of the transform.

$$\text{Let } f(r) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) - \frac{n^2 V}{r^2}$$

Then $f(r)$ is expressible as

$$\begin{aligned} f(r) &= \frac{1}{r}\left(\frac{\partial V}{\partial r} + r\frac{\partial^2 V}{\partial r^2}\right) - \frac{n^2 V}{r^2} \\ &= \frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} - \frac{n^2 V}{r^2}. \end{aligned} \tag{20.1.20}$$

By Theorem 15.6, we can write

$$H_n\{f(r)\} = -sf(1)J_n'(s) - s^2 H_n\{f\} = -sV_1 J_n'(s) - s^2 H_n\{f\}$$

Exercise 20.1.10. (i) Find the finite Hankel transform of x^n, $(n > -1)$ if $xJ_n(sx)$ is the Kernel of the transform. **Answer:** $H_n\{x^n\} = \frac{a^{n+1}}{s} J_{n+1}(sa)$.
(ii) Find the finite Hankel transform of $(1 - x^2)$, taking $xJ_0(sx)$ as the kernel.
Answer: $H_0\{1 - x^2\} = \frac{a}{s} J_1(as) - \frac{a^2}{s^2}\left[\left(as - \frac{4}{as}\right) J_1(sa) + 2J_0(sa)\right]$
(iii) Find the Hankel transform of $(a^2 - x^2)$ if $xJ_0(sx)$ is the kernel of the transform.
Answer: $H_0\{a^2 - x^2\} = \frac{4a}{s^3} J_1(sa) - \frac{2a^2}{s^2} J_0(sa)$
(iv) Show that

$$\int_0^a r^3 J_0\{pr\}\, dr = \frac{a^2}{p^2}\left[2J_0(pa) + \left(ap - \frac{4}{ap} J_1(pa)\right)\right]$$

20.2 Laplace transforms and Fourier transforms to solve some partial differential equations

The given partial differential equations are given along with certain prescribed conditions on the functions which arise from the physical situation. The conditions which are given at $t = 0$ are known as *initial conditions* whereas the conditions given at the boundary of the region or interval are called *boundary conditions.* Most of the well known partial differential equations like Laplace equation, Heat equation and Wave equation can be solved by using the method of integral transform. Readers are suggested to familiar with the following important partial differential equations.

1. One dimensional heat conduction or diffusion equation:

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}, \quad 0 < x < \infty, t > 0$$

Here $u(x, t)$ is the temperature in a solid at position x at time t. The constant k is called the diffusivity of the material of the solid. Again $k = K/\sigma\rho$, where the thermal conductivity K, the specific heat σ and the density ρ are assumed constant. The amount of heat per unit area per unit time conducted across a plane is given by $-K\ u_x(x, t)$.

2. One dimensional wave equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad x > 0, t > 0$$

This equation is applicable to the small transverse vibrations of a taut flexible string initially located on the $x-$axis and set into motion. Here $u(x,t)$ is the transverse displacement of the string at any time t. Again, $c^2 = T/\rho$, where T is constant tension in the string and ρ is constant mass per unit length of the string.

3. Two dimensional Laplace's equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

We can solve one dimensional heat and wave equation by the method of Laplace transform as well as Fourier transform while only Fourier transform method is used to solved boundary value problem governed by Laplace equation.

20.2.1 Solution of two dimensional Laplace Equation using Finite Fourier Transform (FFT)

In this subsection, we will solve the two dimensional Laplace equation over a finite region using the finite Fourier transform. Let us begin with the following example. For infinite or semi infinite range, one may use infinite cosine or sine transform.

Example 20.2.1. Determine a function $V(x,y)$ which is harmonic in the open square $0 < x < \pi,\ 0 < y < \pi$, takes a constant value V_0 on the edge $y = \pi$ and vanishes on the other edges of the square.

or

Find the steady temperature $V(x,y)$ in a long square bar of side π when one face is kept at constant temperature V_0 and the other faces at zero temperature. Also $V(x,y)$ is bounded.

Solution: The steady temperature $V(x,y)$ is governed by the Laplace equation (since V is harmonic)

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0 \tag{20.2.1}$$

with the conditions (i) $V(x,\pi) = V_0$, (ii) $V(0,y) = 0 = V(\pi,y)$ for every y and (iii) $V(x,y)$ is bounded. Taking finite Fourier sine transform of Eq. (20.2.1), we have

$$\int_0^\pi \frac{\partial^2 V}{\partial x^2} \sin sx\, dx + \int_0^\pi \frac{\partial^2 V}{\partial y^2} \sin sx\, dx = 0$$

$$\Rightarrow \left[\frac{\partial V}{\partial x} \sin sx\right]_0^\pi - s\int \frac{\partial V}{\partial x} \cos sx\, dx + \frac{\partial^2}{\partial y^2} \int_0^\pi V \ \sin sx\, dx = 0$$

$$\Rightarrow \frac{d^2 \overline{V_s}}{dy^2} + 0 - s\left[\Big(V \cos sx\Big)_0^\pi + s\int_0^\pi V \sin sx\, dx\right] = 0$$

$$\Rightarrow \frac{d^2 \overline{V_s}}{dy^2} - s^2 \overline{V_s} - s\Big[V(\pi,y)\cos s\pi - V(0,y)\cos 0\Big] = 0$$

$$\Rightarrow \frac{d^2 \overline{V_s}}{dy^2} - s^2 \overline{V_s} - s\Big[V(\pi,y)\cos s\pi - V(0,y)\cos 0\Big] = 0$$

$$\Rightarrow \frac{d^2 \overline{V_s}}{dy^2} - s^2 \overline{V_s} = 0 \qquad \text{[using boundary condition (ii)]}$$

The solution of this equation is

$$V_s = A\cosh sy + B\sinh sy. \tag{20.2.2}$$

$$\begin{aligned}
\text{Now,}\quad V(x,\pi) = V_0 \quad &\Rightarrow \quad F_s\{V(x,\pi)\} = F_s\{V_0\} \\
&\Rightarrow \quad \int_0^\pi V(x,\pi)\sin sx\,dx = \int_0^\pi V_0 \sin sx\,dx \\
&\Rightarrow \quad \overline{V_s}(s,\pi) = \frac{V_0}{s}\Big(-\cos sx\Big)_0^\pi = V_0\left(\frac{1-\cos s\pi}{s}\right) \\
&\therefore \quad \overline{V_s}(s,\pi) = V_0\frac{1-\cos s\pi}{s}
\end{aligned}$$

$$\begin{aligned}
\text{Again,}\quad V(x,0) = 0 \quad &\Rightarrow \quad \overline{V_s}(s,0) = 0 \\
&\Rightarrow \quad A\cdot 1 + B\cdot 0 = 0 \quad [\text{Using Eq.(20.2.2)}] \\
&\Rightarrow \quad A = 0
\end{aligned}$$

$$\therefore \text{ From Eq. (20.2.2), we have } \overline{V_s} = B\sinh sy, \tag{20.2.3}$$

$$\begin{aligned}
&\Rightarrow \quad \overline{V_s}(s,\pi) = B\sinh(s\pi) \\
&\Rightarrow \quad V_0\left(\frac{1-\cos s\pi}{s}\right) = B\sinh s\pi \\
&\Rightarrow \quad B = V_0\frac{1-\cos s\pi}{s\ \sinh s\pi}
\end{aligned}$$

Now Eq.(20.2.3) takes the form

$$\overline{V_s} = \frac{V_0}{s}\frac{(1-\cos s\pi)}{\sinh s\pi}\sinh(sy)$$

Taking inverse finite sine transform, we obtain

$$\begin{aligned}
V(x,y) &= \frac{2}{\pi}\sum_{s=1}^{\infty}\frac{V_0(1-\cos s\pi)\sinh sy \sin sx}{s\ \sinh s\pi} \\
&= \frac{2V_0}{\pi}\sum_{s=1}^{\infty}\frac{[1-(-1)^s]\sinh sy\sin sx}{s\sinh s\pi} \\
&= \frac{4V_0}{\pi}\sum_{n=0}^{\infty}\frac{\sinh(2n+1)y\cdot\sin(2n+1)x}{(2n+1)\sinh(2n+1)\pi}.
\end{aligned}$$

Example 20.2.2. Use a cosine transform to show that the steady temperature in the semi-infinite solid $y > 0$ when the temperature on the surface $y = 0$ is kept at unity over the strip $|x| < a$ and at zero outside the strip, is

$$\frac{1}{\pi}\left[\tan^{-1}\left(\frac{a+x}{y}\right) + \tan^{-1}\left(\frac{a-x}{y}\right)\right]$$

The result $\displaystyle\int_0^\infty e^{-sx}x^{-1}\sin rx\,dx = \tan^{-1}\frac{r}{s}$, $r > 0, s > 0$ may be assumed.

Solution: We know that the steady temperature in the semi-infinite solid is represented by two-dimensional Laplace equation

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0 \quad 0 < y < \infty, \quad -\infty < x < \infty \tag{20.2.4}$$

subject to the boundary conditions:

$$U(x, 0) = 1, \quad |x| < a \quad i.e. \; -a < x < a \tag{20.2.5}$$

$$U(x, 0) = 1, \quad x < -a \quad or \quad x > a \tag{20.2.6}$$

Taking the Fourier cosine transform of (20.2.4), we get

$$\int_0^\infty \frac{\partial^2 U}{\partial x^2} \cos sx \, dx + \int_0^\infty \frac{\partial^2 U}{\partial y^2} \cos sx \, dx = 0$$

$$\Rightarrow \left[\frac{\partial U}{\partial x} \cos sx\right]_0^\infty - \int_0^\infty \frac{\partial U}{\partial x}(-s \, \sin sx) \, dx + \frac{d^2}{dy^2} \int_0^\infty U(x, y) \cos sx \, dx = 0$$

$$\Rightarrow s \int_0^\infty \frac{\partial U}{\partial x} \sin sx \, dx + \frac{d^2 \overline{U_c}}{dy^2} = 0, \quad \text{where } \overline{U_c}(s, y) = \int_0^\infty U(x, y) \, \cos sx \, dx$$

$$\left[\because \text{ due to symmetry, } \frac{\partial U}{\partial x} \to 0 \text{ as } x \to \infty \text{ and } \frac{\partial U}{\partial x} \to 0 \text{ as } x \to 0.\right]$$

$$\Rightarrow s \left\{ [U(x, y) \sin sx]_0^\infty - \int_0^\infty U(x, y) \; s \; \cos sx \, dx \right\} + \frac{d^2 \overline{U_c}}{dy^2} = 0$$

$$\Rightarrow -s^2 \overline{U_c} + \frac{d^2 \overline{U_c}}{dy^2} = 0 \qquad \text{if } U(x, y) \to 0 \text{ as } x \to \infty$$

$$\Rightarrow (D^2 - s^2)\overline{U_c} = 0 \qquad \text{where } D \equiv \frac{d}{dy}$$

whose general solution is

$$\overline{U_c}(s, y) = C_1 e^{sy} + C_2 e^{-sy}, \quad C_1 \text{ and } C_2 \text{ being arbitrary constants.} \tag{20.2.7}$$

Since $\overline{U_c}(s, y)$ is finite, we must take $C_1 = 0$ in (20.2.7), otherwise $\overline{U_c}(s, y)$ would become infinite as $y \to \infty$. Hence (20.2.7) reduces to

$$\overline{U_c}(s, y) = C_2 e^{-sy} \tag{20.2.8}$$

Again

$$\int_0^\infty U(x, 0) \, \cos sx \, dx = \int_0^a U(x, 0) \, \cos sx \, dx + \int_a^\infty U(x, 0) \, \cos sx \, dx$$

$$\Rightarrow \overline{U_c}(s, 0) = \int_0^a \cos sx \, dx = \frac{\sin sa}{s} = C_2 \tag{20.2.9}$$

Hence from (20.2.8), we finally find

$$\overline{U_c}(s, y) = \frac{\sin sa}{s} e^{-sy} \tag{20.2.10}$$

Now taking the inverse Fourier cosine transform, we get

$$\begin{aligned} U(x,y) &= \frac{2}{\pi}\int_0^\infty \frac{\sin sa}{s}e^{-sy}\cos sx\,ds = \frac{1}{\pi}\int_0^\infty \frac{e^{-sy}}{s}\Big[\sin(a+x)s+\sin(a-x)s\Big]\,ds \\ &= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{a+x}{y}\right)+\tan^{-1}\left(\frac{a-x}{y}\right)\right] \end{aligned}$$

Exercise 20.2.3. (i) Using the finite Fourier transform, solve $\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0,\ 0<x<\pi,\ 0<y<y_0$ subject to $u(0,y)=0, u(\pi,y)=1, u_y(x,0)=0, u(x,y_0)=0$ **Answer:** $u(x,y)=\dfrac{2}{\pi}+\dfrac{2}{\pi}\sum_{n=1}^{\infty}\dfrac{(-1)^n}{n}\dfrac{\cosh ny}{\cosh ny_0}\sin nx$

(ii) Solve the boundary value problem in the half-plane $y>0$, described by $\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0,\ -\infty<x<\infty,\ y>0$ subject to $u(x,0)=f(x),\ -\infty<x<\infty$, u is bounded as $y\to\infty$, u and $\dfrac{\partial u}{\partial x}$ both vanish as $|x|\to\infty$. **Answer:** $u(x,y)=\dfrac{y}{\pi}\displaystyle\int_{-\pi}^{\pi}\dfrac{f(\xi)}{(\xi-x)^2+y^2}d\xi$

(iii) Solve the boundary value problem in the half-plane $x>0$, described by $\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0,\ -\infty<y<\infty,\ x>0$ subject to $u(0,y)=f(y),\ -\infty<y<\infty$, u is bounded as $x\to\infty$, u and $\dfrac{\partial u}{\partial y}$ both vanish as $|y|\to\infty$. **Answer:** $u(x,y)=\dfrac{x}{\pi}\displaystyle\int_{-\pi}^{\pi}\dfrac{f(\xi)}{x^2+(y-\xi)^2}d\xi$

20.2.2 Application to Heat Conduction and Wave Equations

Formulae for Laplace transform method

In order to solve heat equation using the method of Laplace transform, the following results will be used frequently

$$\begin{aligned} &L\{u(x,t)\}=\overline{u}(x,s) \\ &L\left\{\frac{\partial u}{\partial x}\right\}=\frac{d\overline{u}}{dx},\quad L\left\{\frac{\partial^2 u}{\partial x^2}\right\}=\frac{d^2\overline{u}}{dx^2}, \\ &L\left\{\frac{\partial u}{\partial t}\right\}=s\overline{u}-u(x,0) \\ &L\left\{\frac{\partial^2 u}{\partial t^2}\right\}=s^2\overline{u}-s\,u(x,0)-u_t(x,0). \end{aligned}$$

Example 20.2.4. Find the temperature $u(x,t)$ in a slab whose ends $x=0$ and $x=a$ are kept at temperature zero and whose initial temperature is $\sin(\pi x/a)$.

Solution: We have to solve one-dimensional heat conduction equation

$$\frac{\partial u}{\partial t}=k\frac{\partial^2 u}{\partial x^2},\ 0<x<a,\ t>0, \tag{20.2.11}$$

$u(x,t)$ being the temperature in the slab at any point x at any time t and k being the diffusivity of the material of the bar, subject to the boundary conditions $u(0,t) = 0$, $u(a,t) = 0$ and initial condition $u(x,0) = \sin(\pi x/a)$. Let $L\{u(x,t)\} = \overline{u}(x,s)$. Taking the Laplace transform of both sides of (20.2.11), we have

$$L\left\{\frac{\partial u}{\partial t}\right\} = kL\left\{\frac{\partial^2 u}{\partial x^2}\right\} \quad \Rightarrow \quad s\overline{u}(x,s) - u(x,0) = k\frac{d^2\overline{u}}{dx^2}$$
$$\Rightarrow s\overline{u} - \sin\left(\frac{\pi x}{a}\right) = k\frac{d^2\overline{u}}{dx^2} \quad \Rightarrow \quad \left(D^2 - \frac{s}{k}\right)\overline{u} = -\frac{1}{k}\sin\frac{\pi x}{a}$$

Calculating the complementary function corresponding to the homogeneous part and the particular solution using classical methods of ordinary differential equations, we may write the general solution of the aforesaid ODE as

$$\overline{u}(x,s) = c_1 e^{x\sqrt{s/k}} + c_1 e^{-x\sqrt{s/k}} + \frac{1}{s + (\pi^2 k/a^2)}\sin\frac{\pi x}{a}. \tag{20.2.12}$$

Taking the Laplace transform of boundary conditions, we have

$$\overline{u}(0,s) = 0 \quad \text{and} \quad \overline{u}(a,s) = 0 \tag{20.2.13}$$

Now using the conditions (20.2.13) in (20.2.2), we obtain $c_1 = c_2 = 0$ and therefore (20.2.2) reduces to

$$\overline{u}(x,s) = \frac{\sin(\pi x/a)}{s + (\pi^2 k/a^2)} \quad \text{so that} \quad u(x,t) = sin\frac{\pi x}{a}L^{-1}\left\{\frac{1}{s + (\pi^2 k/a^2)}\right\} = \sin\frac{\pi x}{a}e^{-(\pi^2 kt/a^2)}$$

Example 20.2.5. The faces $x = 0$ and $x = 1$ of a slab of material for which $k = 1$ are kept at temperature 0 and 1 respectively until the temperature distribution becomes $u = x$. After time $t = 0$ both faces are held at temperature 0. Determine the temperature formula. It is given that

$$L^{-1}\left\{\frac{\sinh x\sqrt{s}}{s\,\sinh a\sqrt{s}}\right\} = \frac{x}{a} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}e^{-n^2\pi^2 t/a^2}\sin\left(\frac{n\pi x}{a}\right).$$

Solution: The temperature $u(x,t)$ in the slab is governed by the partial differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{20.2.14}$$

with the boundary conditions (i) $u(0,t) = 0$, (ii) $u(1,t) = 0$, (iii) $u(x,0) = x$. From Eq. (20.2.14) we have

$$L\left\{\frac{\partial u}{\partial t}\right\} = L\left\{\frac{\partial^2 u}{\partial x^2}\right\}$$
$$\Rightarrow s\,\overline{u} - u(x,0) = \frac{d^2\overline{u}}{dx^2}$$
$$\Rightarrow \frac{d^2\overline{u}}{dx^2} - s\,\overline{u} = -x \qquad [\because u(x,0) = x]$$
$$\Rightarrow (D^2 - s)\overline{u} = -x$$

The solution of it is

$$\begin{aligned}
\overline{u} &= a\,e^{-x\sqrt{s}} + b\,e^{x\sqrt{s}} + \frac{1}{D^2 - s}(-x) \\
&= a\,e^{-x\sqrt{s}} + b\,e^{x\sqrt{s}} + \frac{1}{s}\left(1 - \frac{D^2}{s}\right)^{-1} x \\
&= a\,e^{-x\sqrt{s}} + b\,e^{x\sqrt{s}} + \frac{x}{s}.
\end{aligned}$$

It is also expressed as

$$\overline{u} = a\ \cosh x\sqrt{s} + b\ \sinh x\sqrt{s} + \frac{x}{s}.$$

Now

$$(i) \Rightarrow L\{u(0,t)\} = 0 \Rightarrow \overline{u}(0,s) = 0 \Rightarrow a = 0$$

Hence,

$$\overline{u} = b\ \sinh x\sqrt{s} + \frac{x}{s}.$$

$$(ii) \Rightarrow L\{u(1,t)\} = 0 \Rightarrow \overline{u}(1,s) = 0 \Rightarrow 0 = b\ \sinh\sqrt{s} + \frac{1}{s} \Rightarrow b = -\frac{1}{s\ \sinh\sqrt{s}}$$

Using this we obtain

$$\begin{aligned}
\overline{u} &= \frac{x}{s} - \frac{\sinh x\sqrt{s}}{s\ \sinh\sqrt{s}} \\
\Rightarrow u &= L^{-1}\left\{\frac{x}{s}\right\} - L^{-1}\left\{\frac{\sinh x\sqrt{s}}{s\ \sinh\sqrt{s}}\right\} \\
\Rightarrow u &= x - \left[x + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}e^{-n^2\pi^2 t}\sin(n\pi x)\right] \\
\Rightarrow u &= -\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}e^{-n^2\pi^2 t}\sin(n\pi x)
\end{aligned}$$

Example 20.2.6. Solve the wave equation $\frac{\partial^2 u}{\partial t^2} + c^2\frac{\partial^2 u}{\partial x^2}, x > 0, t > 0$, where $u(x,0) = 0,\ u_t(x,0) = 0,\ x > 0$ and $u(0,t) = F(t)$. $\lim_{x\to\infty} u(x,t) = 0,\ t > 0$.

Solution: We have to solve one-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2} \tag{20.2.15}$$

subject to boundary condition: $u(0,t) = F(t), \quad \lim_{x\to\infty} u(x,t) = 0$
and initial conditions: $u(x,0) = 0, \quad u_t(x,0) = 0.$
Let $L\{u(x,t)\} = \overline{u}(x,s)$. Applying Laplace transform to Eq.(20.2.15), we have

$$s^2\overline{u}(x,s) - su(x,0) - u_t(x,0) = c^2\frac{d^2\overline{u}}{dx^2} \quad \Rightarrow \quad \frac{d^2\overline{u}}{dx^2} - \frac{s^2}{c^2}\overline{u} = 0$$

Its solution is

$$\overline{u}(x,s) = c_1\ e^{sx/c} + c_2\ e^{-sx/c}, \quad c_1, c_2 \text{ being the arbitrary constants}$$

Now using the above boundary conditions we have $\overline{u}(0,s) = f(s)$ where $f(s) = L\{F(t)\}$ and $\overline{u}(x,s) = 0$ as $x \to \infty$. Since $\overline{u}(x,s) = 0$ as $x \to \infty$, we must choose $c_1 = 0$. Hence the solution reduces to

$$\overline{u}(x,s) = c_2\ e^{-sx/c}$$

Putting $x = 0$ in the above equation and using $\overline{u}(0,s) = f(s)$, we get $c_2 = f(s)$. Then the solution reduces to

$$\overline{u}(x,s) = f(s)\ e^{-sx/c}$$

Taking inverse Laplace transform, we obtain

$$u(x,t) = L^{-1}\{f(s)\, e^{-sx/c}\} = F(t - x/c)H(t - x/c)$$

where $H(t - x/c)$ is the Heaviside unit step function.

Exercise 20.2.7. (i) A string is stretched between two fixed points $(0,0)$ and $(a,0)$. If it is displaced into the curve $u = b\ \sin(\pi x/a)$ and released from rest in that position at time $t = 0$, find its displacement at any time $t < 0$ and at any point $0 < x < a$. **Answer:** $u(x,t) = b\ \sin\dfrac{\pi x}{a}\cos\dfrac{\pi ct}{a}$

(ii) Solve the boundary value problem $\dfrac{\partial^2 u}{\partial t^2} = a^2\dfrac{\partial^2 u}{\partial x^2} - g,\ x > 0,\ t > 0$ with the boundary conditions $u(x,0) = 0 = u_t(x,0),\ x > 0;\ u(0,t) = 0,\ \lim\limits_{x\to\infty} u_x(x,t) = 0,\ t \geq 0.$

Answer: $u(x,t) = \dfrac{1}{2}g(t - x/a)^2 H(t - x/a) - \dfrac{1}{2}gt^2$

References

1. Topics in Algebra; *I. N. Herstein*.
2. Linear Algebra; *K. Hoffman and R. Kunze*.
3. Linear Algebra; *J. H. Kwak and S. Hong*.
4. Linear Algebra and Matrix Theory; *E. D. Nering*.
5. Module Theory; *T. S. Blyth*.
6. Modules; *I. S. Luthar and I. B. S. Passi*.
7. Special Functions of Mathematical Physics and Chemistry; *I. N. Sneddon.*
8. Special Functions & Their Applications; *Lebedev & Silverman*
9. Special Functions; *E. D. Rainville.*
10. Integral Equations and Boundary Value Problems; *M. D. Raisinghania*
11. A First Course in Integral Equations; *Wazwaz and Abdul*
12. Integral Equations; *D. C. Sharma and M. C. Goyal*
13. Advanced Differential Equations; *M.D. Raisinghania*