Lecture Notes on Mathematical Method of Physics I

Dr. A. N. Njah,
Department of Physics,
University of Agriculture,
Abeokuta.

PHS 471: Linear Algebra: Transformation in linear vector spaces and matrix theory.
Functional analysis; Hilbert space, complete sets of orthogonal functions; Linear operations.
Special functions: Gamma, hypergometric, Legendre, Bessel, Hermite and Laguerre functions. The Dirac delta function
Integral transform and Fourier series: Fourier series and Fourier transform; Application of transform methods to the solution of elementary differential equations in Physics and Engineering.

Suggested reading.

1. Advanced Engineering Mathematics by E. Kreyszig.
2. Further Engineering Mathematics by H.A. Stroud
3. Mathematical Physics by B.D. Gupta
4. Advanced Mathematics for Engineers and Scientist (Schaum series) by Spiegel

Contents

1 LINEAR ALGEBRA **5**

1.1 Vector Space or Linear Space 5

1.1.1 Algebraic Operations on Vectors 6

1.1.2 Linearly Dependent and Independent sets of vectors . . 6

1.2 Matrix Theory 7

1.2.1 Determinant of a Matrix 9

1.2.2 General properties of determinants 10

1.2.3 Adjugate Matrix or Adjoint of a Matrix 10

1.2.4 Reciprocal Matrix or Inverse of a Matrix 11

1.2.5 The Transpose of a Matrix 11

1.2.6 Symmetric, Skew-symmetric and Orthogonal Matrices . 11

1.3 Complex Matrices 12

1.3.1 The Conjugate of a Matrix 12

1.3.2 The Conjugate transpose or Hermitian Conjugate of a Matrix 12

1.3.3 Hermitian, Skew-Hermitian and Unitary Matrices . . . 13

1.4 Matrix Algebra 13

1.4.1 Rank of a Matrix 13

1.5 Consistency of equations 15

1.5.1 Homogeneous and Non-Homogeneous Linear Equations 15

1.5.2 Uniqueness of Solutions 16

1.6 Solution of sets of Equations 17

1.6.1 Inverse method 17

1.6.2 Row Transformation method 19

1.6.3 Gaussian elimination method 20

1.6.4 Triangular Decomposition method: LU-decomposition . 20

1.6.5 Cramer's Rule 20

1.7 Eigenvalues and Eigenvectors of a Matrix 20

1.7.1 Nature of the eigenvalues and eigenvectors of special types of matrices . . . 22
1.7.2 Diagonalisation of a matrix . . . 24
1.8 Transformation . . . 25
1.8.1 Transformation . . . 25
1.8.2 Resultant of two linear transformation . . . 26
1.8.3 Similarity transformation . . . 26
1.8.4 Unitary transformation . . . 26
1.8.5 Orthogonal transformation . . . 26
1.8.6 Orthogonal set . . . 27
1.9 Bases and dimension . . . 27
1.9.1 Linear Maps . . . 27

2 FUNCTIONAL ANALYSIS 28
2.1 Normed Spaces . . . 28
2.1.1 Cauchy Sequences . . . 28
2.1.2 Completeness . . . 28
2.1.3 Pre-Hilbert spaces . . . 29
2.1.4 Hilbert Spaces . . . 29
2.1.5 Geometry of Hilbert space . . . 29

3 SPECIAL FUNCTIONS 30
3.1 The gamma and beta functions . . . 30
3.1.1 The gamma function Γ . . . 30
3.1.2 The beta function, β . . . 32
3.1.3 Application of gamma and beta functions . . . 32
3.2 Bessel's Functions . . . 32
3.3 Legendre's Polynomials . . . 35
3.4 Hermite Polynomials . . . 36
3.5 Laguerre Polynomials . . . 38
3.5.1 Hypergeometric Function . . . 39

4 INTEGRAL TRANSFORM AND FOURIER SERIES 41
4.1 Laplace Transform . . . 41
4.1.1 Inverse Transform . . . 44
4.1.2 Solution of Differential Equations by Laplace Transform 46
4.2 The Dirac Delta Function − the impulse function . . . 48

4.2.1 Integration involving the impulse function 49
4.2.2 Differential equations involving the impulse function . . 49
4.3 Fourier Series 50
4.3.1 Fourier series of functions of period 2π 51
4.3.2 Half-range series 54
4.3.3 Functions with arbitrary period T 54
4.3.4 Sum of a Fourier series at a point of finite discontinuity 55
4.4 Fourier Integrals 55
4.4.1 The Fourier integral 55

Chapter 1

LINEAR ALGEBRA

1.1 Vector Space or Linear Space

A vector space V over a field F is a set of elements called vectors which may be combined by two operations - addition and scalar multiplication; such that

a) if the vectors $\vec{a}$ and $\vec{b}$ belong to V, then

(i) $\vec{a} + \vec{b}$ also belongs to V. (This is known as *closure property*)

(ii) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ (*commutative law of addition*)

(iii) $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}) = \vec{b} + (\vec{c} + \vec{a})$ (*associative law of addition*)

(iv) there exist an *additive identity* vector $\vec{0}$ known as the null vector such that $\vec{a} + \vec{0} = \vec{a}$

(v) to every vector $\vec{a}$ in V, there corresponds a vector $-\vec{a}$ known as the *additive inverse* vector, such that $\vec{a} + (-\vec{a}) = \vec{0}$

b) if $m,\ n$ (elements of F) are any two scalars and $\vec{a}$ is a vector in V, then

(i) $(m + n)\vec{a} = m\vec{a} + n\vec{a}$ (*distributive law*)

(ii) $m(\vec{a} + \vec{b}) = m\vec{a} + m\vec{b}$ (*distributive law*)

(iii) $m(n\vec{a}) = (mn)\vec{a} = n(m\vec{a})$ (*associative law of multiplication*)

(iv) to every vector $\vec{a}$ in V, there corresponds a multiplicative identity scalar 1, such that $1\vec{a} = \vec{a}$

Conditions for a physical quantity to be representable by a vector:

- It must obey the parallelogram law of addition ($\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|$)
- It must have magnitude as well as direction independent of any choice of coordinate axes.

1.1.1 Algebraic Operations on Vectors

If $\vec{A}$ and $\vec{B}$ are two vectors with components $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ respectively then

(i) $\vec{A} \pm \vec{B} = (a_1 \pm b_1, a_2 \pm b_2, \ldots, a_n \pm b_n)$

(ii) $k\vec{A} = (ka_1, ka_2, \ldots, ka_n), \quad k$ a scalar

(iii) $\vec{A} \cdot \vec{B} = a_1b_1 + a_2b_2 + \ldots + a_nb_n$

(iv) A vector will be a unit vector if the magnitude $|\vec{A}| = 1$

(v) The vectors $\vec{A}$ and $\vec{B}$ will be orthogonal if $\vec{A} \cdot \vec{B} = 0$

1.1.2 Linearly Dependent and Independent sets of vectors

A set a vectors $\vec{A}_1, \vec{A}_2, \ldots \vec{A}_n$ is said to be *linearly dependent* if there exist a set of n scalars $k_1, k_2, \ldots, k_n$ (of which at least one is non-zero) such that $k_1\vec{A}_1 + k_2\vec{A}_2 + \ldots + k_n\vec{A}_n = \vec{0}$
In the case when $k_1 = k_2 = \ldots = k_n = 0$, in order that $k_1\vec{A}_1 + k_2\vec{A}_2 + \ldots + k_n\vec{A}_n = \vec{0}$
the set of n vectors $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_n$ is said to be *linearly independent*
A vector $\vec{A}$ is known as a *linear combination* of the set $\vec{A}_1, \vec{A}_2, \ldots \vec{A}_n$ if it is expressable as $\vec{A} = k_1\vec{A}_1 + k_2\vec{A}_2 + \ldots + k_n\vec{A}_n$
Examples:

1. The set of vectors (1,2,3), (2,-2,0) is linearly independent since $k_1(1, 2, 3) + k_2(2, -2, 0) = (0, 0, 0)$ is equivalent to the set of equations $k_1 + 2k_2 = 0,\ 2k_1 - 2k_2 = 0$ and $3k_1 = 0$ which gives $k_1 = k_2 = 0$.

2. The set of vectors (2,4,10), (3,6,15) is linearly dependent since $k_1(2, 4, 10) + k_2(3, 6, 15) = (0, 0, 0)$ gives the system $2k_1 + 3k_2 = 0,\ 4k_1 + 6k_2 = 0,\ 10k_1 + 15k_2 = 0 \quad \Rightarrow k_1 = 3,\ k_2 = -2$

1.2 Matrix Theory

A set of numbers arranged in a rectangular array of m rows and n columns such as

$$\begin{pmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots\dots\dots \\ a_{m1} & a_{m2} \dots a_{mn} \end{pmatrix}$$

is called a matrix of order $m \times n$ or an $m \times n$ matrix. If $m = n$ (i.e. number of rows = number of columns) it is called a *square matrix* of order n. $a_{ij}, (i = 1, 2, \dots, m;\ j = 1, 2, \dots, n)$ are called its *elements* or *constituents* or *entries*. a_{ij} represents the element in the i^{th} row and j^{th} column of the matrix. The element a_{ij} $(i = j)$ of a square matrix $\mathbf{A}$ lie on the *main diagonal* or *principal diagonal* and are called its diagonal elements. The sum of the diagonal elements is called the *trace* of $\mathbf{A}$ and is denoted by $tr\mathbf{A} = \sum_{i=1}^{n} a_{ii}$

A matrix can be represented by $\mathbf{A}$ or $[a_{ij}]$

Null matrix: A matrix having all of its elements zero

Row matrix and column matrix: A row matrix is a matrix having only a single row i.e. a $1 \times n$ matrix. A column matrix is one having a single column, i.e. $m \times 1$ matrix

Equality of matrices: Two matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$ are equal if both are of the same order (size) $m \times n$ and each element a_{ij} of $\mathbf{A}$ is equal to the corresponding element b_{ij} of $\mathbf{B}$.

Equivalent relation on matrices:

1. Reflexivity $\mathbf{A} = \mathbf{A}$
2. Symmetry $\mathbf{A} = \mathbf{B} \quad \Rightarrow \mathbf{B} = \mathbf{A}$
3. Transitivity $\mathbf{A} = \mathbf{B}$ and $\mathbf{B} = \mathbf{C} \Rightarrow \mathbf{A} = \mathbf{C}$

Addition and Subtraction of Matrices: Addition and subtraction are defined only for matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order or size and are done by adding and subtracting corresponding entries. Addition of matrices is commutative, associative, distributive by a scalar, has an identity and an inverse, obeys the cancellation law $(\mathbf{A} + \mathbf{B} = \mathbf{A} + \mathbf{C} \;\Rightarrow\; \mathbf{B} = \mathbf{C})$

Multiplication of a matrix by a scalar: The product of a matrix $\mathbf{A}$ by a scalar c is obtained by multiplying each entry of $\mathbf{A}$ by c

Multiplication of a matrix by a matrix: For two matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$ to be multiplied (i.e. for $\mathbf{C}$=$\mathbf{AB}$ to be defined) the number of columns of $\mathbf{A}$ must be equal the number of rows of $\mathbf{B}$; i.e. if $\mathbf{A}$ is a $p \times n$ matrix $\mathbf{B}$ must be an $n \times q$ matrix; $\mathbf{C}$=$\mathbf{AB}$ is a $p \times q$ matrix whose elements c_{ij} in the i^{th} row and j^{th} column is the algebraic sum of the products of the elements in the i^{th} row of $\mathbf{A}$ by the corresponding elements in the j^{th} column of $\mathbf{B}$
$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$
Matrix multiplication is:

- not commutative i.e. $\mathbf{AB} \neq \mathbf{BA}$
- distributive; $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- associative; $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

Matrix multiplication differ from multiplication of numbers in that:

1. it is not commutative i.e. $\mathbf{AB} \neq \mathbf{BA}$
2. $\mathbf{AB} = 0$ does not necessarily imply $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = 0$ or $\mathbf{BA} = 0$
3. $\mathbf{AC} = \mathbf{AD}$ does not necessarily imply $\mathbf{C} = \mathbf{D}$

Upper and lower triangular Matrices: A square matrix $\mathbf{U}$ of order n is said to be an upper triangular matrix if its elements $u_{ij} = 0$ for $i > j$. On the other hand a square matrix $\mathbf{L}$ is said to be lower triangular matrix if its elements $l_{ij} = 0$ for $i < j$
Examples

$$\mathbf{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \qquad \mathbf{L} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{23} & l_{33} \end{pmatrix}$$

Diagonal Matrix: It is a square matrix $\mathbf{D}$ which is both upper and lower triangular. If the diagonal elements are equal to a scalar quantity λ i.e. $d_{11} = d_{22} = \ldots = d_{nn} = \lambda$ then the matrix is called a *scalar matrix* $\mathbf{S}$ (because multiplication of any square matrix $\mathbf{A}$ of the same size by $\mathbf{S}$ has the same effect as the multiplication by a scalar λ, that is $\mathbf{AS} = \mathbf{SA} = \lambda\mathbf{A}$). In particular, if $\lambda = 1$ we have a *unit matrix* or *identity matrix* $\mathbf{I}$

Examples

$$\mathbf{D} = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

1.2.1 Determinant of a Matrix

Determinants were originally introduced for solving linear systems (by Cramer's rule). They have important engineering application in eigenvalue problems, differential equations, vector algebra, etc.
The determinant of a 2×2 matrix $\mathbf{A}$ is called a *second order determinant* and is denoted and defined by
$|\mathbf{A}|$ or $det\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$
Similarly the determinant of a 3×3 matrix $\mathbf{A}$ is referred to as a *third order determinant* and is defined by

$$det\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$
$$= a_{11}M_{11} - a_{21}M_{21} + a_{31}M_{31}$$

Here we have expanded the determinant by the first column. Determinant can be expanded by any row or column

Minor and Cofactors of a third order determinant

$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$, $M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$ and $M_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$
are called the *minors* of a_{11}, a_{21} and a_{31} respectively while their *cofactors* are $C_{11} = +M_{11}$, $C_{21} = -M_{21}$ and $C_{31} = +M_{31}$ respectively. $\Rightarrow$ the cofactor of any one element is its minor together with its "place sign". The "place signs" in C_{ij} form a checkerboard pattern as follows:

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

In general a *determinant of n-order* is a scalar associated with an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$, which is written

$$Det\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots\dots\dots \\ a_{m1} & a_{m2} \dots a_{mn} \end{vmatrix}$$

1.2.2 General properties of determinants

Theorem 1: (Behaviour of n^{th} order determinant under elementary row operation)

(a) Interchange of two rows multiplies the value of the determinant by -1.

(b) Addition of a multiple of a row to another row does not alter the value of the determinant.

(c) Multiplication of a row by c multiplies the value of the determinant by c

Theorem 2: (further properties of the n^{th}-order determinant)

(a)-(c) in theorem 1 hold also for columns

(d) Transposition leaves the value of a determinent unaltered

(e) A zero row or column renders the value of a determinant zero.

(f) Proportional rows or columns render the value of a determinant zero. In particular, a determinant with two identical rows or columns has the value zero.

Singular and Non-singular matrices
A square matrix $\mathbf{A}$ is known as singular matrix if its determinant $|\mathbf{A}| = 0$. In case $|\mathbf{A}| \neq 0$ then $\mathbf{A}$ is known as non-singular matrix

1.2.3 Adjugate Matrix or Adjoint of a Matrix

Let $\mathbf{A} = [a_{ij}]$ be a square matrix of order n and c_{ij} represents the cofactor of the element a_{ij} in the determinant $|\mathbf{A}|$ then the transpose of the matrix of cofactors ($\mathbf{C} = [c_{ij}]$) denoted by $\mathbf{C}' = [c_{ji}]$ is called the *adjugate* or *adjoint* of $\mathbf{A}$ and is denoted $adj\mathbf{A}$.

Properties:

- $\mathbf{A}(adj\mathbf{A}) = (adj\mathbf{A})\mathbf{A}$............................$*$
  i.e. multiplication of $\mathbf{A}$ by $adj\mathbf{A}$ is commutative and their product is a scalar matrix having every diagonal element as $|\mathbf{A}|$.

- $|adj\mathbf{A}| = |\mathbf{A}|^{n-1}$, where n is the order of the matrix.

- If $|\mathbf{A}| = 0$ then $\mathbf{A}(adj\mathbf{A}) = (adj\mathbf{A})\mathbf{A} = 0$
- $adj(\mathbf{AB}) = adj\mathbf{B}adj\mathbf{A}$
  or $adj(\mathbf{ABC}) = adj\mathbf{C}adj\mathbf{B}adj\mathbf{A}$ (prove!!)

1.2.4 Reciprocal Matrix or Inverse of a Matrix

If $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ (the identity matrix) then $\mathbf{B}$ is said to be the inverse of $\mathbf{A}$ and vice versa and is denoted by $\mathbf{A}^{-1}$. That is $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$
When the inverse of $\mathbf{A}$ exists, then $\mathbf{A}$ is said to be invertible. The necessary and sufficient condition for a square matrix to be invertible is that it is non-singular. From eq.(*) $\mathbf{A}^{-1}$ is given by

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} adj\mathbf{A}$$

1.2.5 The Transpose of a Matrix

If $\mathbf{A}$ is a matrix of order $m \times n$ the transpose of $\mathbf{A}$ denoted by $\mathbf{A}'$ or $\mathbf{A}^T$, is a matrix of order $n \times m$ obtained by interchanging its rows and columns
Examples

If $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$ then $\mathbf{A}' = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}$

Properties

- $(\mathbf{A}')' = \mathbf{A}$
- $|\mathbf{A}'| = |\mathbf{A}|$
- $(k\mathbf{A})' = k\mathbf{A}'$ for k a scalar
- $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

1.2.6 Symmetric, Skew-symmetric and Orthogonal Matrices

A square matrix $\mathbf{A}$ is said to be:

(i) *symmetric* if $\mathbf{A}' = \mathbf{A}$ i.e. if $a_{ij} = a_{ji}$ for all i and j, i.e. if transposition leaves it unchanged

(ii) *skew-symmetric* if $\mathbf{A}' = -\mathbf{A}$ i.e. if transposition gives the negative of $\mathbf{A}$

(iii) *orthogonal* if $\mathbf{A}' = \mathbf{A}^{-1}$ i.e. if transposition gives the inverse of $\mathbf{A}$.

Every square matrix $\mathbf{A} = \mathbf{P} + \mathbf{Q}$
where $\mathbf{P} = \frac{1}{2}(\mathbf{A} + \mathbf{A}')$ is a symmetric matrix
and $\mathbf{Q} = \frac{1}{2}(\mathbf{A} - \mathbf{A}')$ is a skew-symmetric matrix

1.3 Complex Matrices

1.3.1 The Conjugate of a Matrix

If the elements of a matrix $\mathbf{A}$ are complex quantities, then the matrix obtained from $\mathbf{A}$, on replacing its elements by the corresponding conjugate complex numbers, is said to be the *conjugate matrix* of $\mathbf{A}$ and is denoted by $\bar{\mathbf{A}}$ or $\mathbf{A}^*$, with the following properties:

- $(\mathbf{A}^*)^* = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$
- If α is a complex number and $\mathbf{A}$ a matrix of any order say $m \times n$, then $(\alpha\mathbf{A})^* = \alpha^*\mathbf{A}^*$
- $(\mathbf{AB})^* = \mathbf{A}^*\mathbf{B}^*$

1.3.2 The Conjugate transpose or Hermitian Conjugate of a Matrix

The matrix, which is the conjugate of the transpose of a matrix $\mathbf{A}$ is said to be the *conjugate transpose* of $\mathbf{A}$ and denoted by $\mathbf{A}^\dagger$ (called $\mathbf{A}$ dagger)
<u>*Example:*</u>

$$\mathbf{A} = \begin{pmatrix} -2+3i & 3-4i & i \\ -5i & -5-3i & 4+i \end{pmatrix}$$

$$\mathbf{A}^* = \begin{pmatrix} -2-3i & 3+4i & -i \\ 5i & -5+3i & 4-i \end{pmatrix}$$

$$\mathbf{A}^\dagger = \begin{pmatrix} -2-3i & 5i \\ 3+4i & -5+3i \\ -i & 4-i \end{pmatrix}$$

Properties:

- $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$
- $(\mathbf{A}+\mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger$
- If α is a complex number and $\mathbf{A}$ a matrix, then $(\alpha\mathbf{A})^\dagger = \alpha^*\mathbf{A}^\dagger$
- $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$ (prove!!)

1.3.3 Hermitian, Skew-Hermitian and Unitary Matrices

A matrix $\mathbf{A}$ is said to be

(i) Hermitian if $\mathbf{A}^\dagger = \mathbf{A}$

(ii) skew-Hermitian if $\mathbf{A}^\dagger = -\mathbf{A}$

(iii) unitary if $\mathbf{A}^\dagger = \mathbf{A}^{-1}$

Examples:

$$\mathbf{A} = \begin{pmatrix} 5 & 2+3i & -i \\ 2-3i & 3 & -3-4i \\ i & -3+4i & 0 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 3i & 2+i \\ -2+i & -i \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} \frac{1}{2}i & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2}i \end{pmatrix}$$

are Hermitian, skew-Hermitian, and unitary respectively.
Every square complex matrix $\mathbf{A} = \mathbf{P} + \mathbf{Q}$
where $\mathbf{P} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\dagger)$ is a Hermitian matrix
and $\mathbf{Q} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^\dagger)$ is a skew-Hermitian matrix

1.4 Matrix Algebra

1.4.1 Rank of a Matrix

It is the order of the largest determinant that can be formed from the elements of the matrix. A matrix $\mathbf{A}$ is said to have rank r if it contains at least one square submatrix of r rows with a non-zero determinant, while all square submatrices of $(r+1)$ rows, or more, have zero determinants.
Examples:

1. $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 1 & 5 \end{pmatrix}$ is of rank 2 since $|\mathbf{A}| = \begin{vmatrix} 4 & 2 \\ 1 & 5 \end{vmatrix} = 18$ i.e. not zero.

2. $\mathbf{B} = \begin{pmatrix} 6 & 3 \\ 8 & 4 \end{pmatrix}$ gives $|\mathbf{B}| = \begin{vmatrix} 6 & 3 \\ 8 & 4 \end{vmatrix} = 0$.
   Therefore $\mathbf{B}$ is not of rank 2. It is, however, of rank 1, since the submatrices (6),(3),(8),(4) are not zero. This implies that only a null matrix is of rank zero.

3. Determine the ranks of (a) $\mathbf{A} = \begin{pmatrix} 1 & 2 & 8 \\ 4 & 7 & 6 \\ 9 & 5 & 3 \end{pmatrix}$ and (b) $\mathbf{B} = \begin{pmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.
   Since $|\mathbf{A}| = \begin{vmatrix} 1 & 2 & 8 \\ 4 & 7 & 6 \\ 9 & 5 & 3 \end{vmatrix} = -269$ is not zero the rank of $\mathbf{A}$ is 3.
   Since $|\mathbf{B}| = \begin{vmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} = 0$ the rank of $\mathbf{B}$ is not 3. We now try submatrices of order 2:
   $\begin{vmatrix} 3 & 4 \\ 1 & 2 \end{vmatrix} = 2 \neq 0$, therefore, rank of $\mathbf{B}$ is 2. We could equally well have tested
   $\begin{vmatrix} 4 & 5 \\ 2 & 3 \end{vmatrix}, \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}, \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}, \begin{vmatrix} 3 & 5 \\ 1 & 3 \end{vmatrix}, \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}, \begin{vmatrix} 4 & 5 \\ 5 & 6 \end{vmatrix}, \begin{vmatrix} 3 & 5 \\ 4 & 6 \end{vmatrix}, \begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix}$. i.e. we test all possible second order minors to find one that is not zero.

For a rectangular matrix of order $m \times n$ the rank is given by the order of the largest square sub-matrix formed by the elements.
<u>*Example:*</u>
For a 3×4 matrix $\begin{pmatrix} 2 & 2 & 3 & 1 \\ 0 & 8 & 2 & 4 \\ 1 & 7 & 3 & 2 \end{pmatrix}$ the largest square sub-matrix cannot be greater than order 3. We try $\begin{vmatrix} 2 & 2 & 3 \\ 0 & 8 & 2 \\ 1 & 7 & 3 \end{vmatrix} = 0$
But we must also try other 3×3 sub-matrices, e.g,
$\begin{vmatrix} 2 & 3 & 1 \\ 8 & 2 & 4 \\ 7 & 3 & 2 \end{vmatrix} = 30 \neq 0$, therefore, $\mathbf{B}$ is of rank 3

1.5 Consistency of equations

1.5.1 Homogeneous and Non-Homogeneous Linear Equations

A set of m simultaneous linear equations in n unknowns
$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$
$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$
. . . .
$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$
can be written in the matrix form as follows

$$\begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots & \ldots\ldots\ldots \\ a_{m1} & a_{m2} \ldots a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$$

i.e. $\mathbf{Ax} = \mathbf{b}$
This set of equations is *homogeneous* if $\mathbf{b} = \mathbf{0}$, i.e. $(b_1, b_2, \ldots, b_m) = (0, 0, \ldots, 0)$ otherwise it is said to be *non-homogeneous*. This set of equations is said to be *consistent* if solutions for $x_1, x_2, \ldots, x_n$ exist and *inconsistent* if no such solutions can be found.
The *Augmented coefficient matrix* $\mathbf{A}_b$ of $\mathbf{A}$ is

$$\mathbf{A}_b = \begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} & b_1 \\ a_{21} & a_{22} \ldots a_{2n} & b_2 \\ \ldots & \ldots\ldots\ldots & \ldots \\ a_{m1} & a_{m2} \ldots a_{mn} & b_m \end{pmatrix}$$

which is an $m \times (n+1)$ matrix formed by writing the constant terms as an $(n+1)^{th}$ column of the coefficient matrix $\mathbf{A}$. Note:

- If the rank of the coefficient matrix $\mathbf{A}$ equal to the rank of the augmented matrix $\mathbf{A}_b$ then the equations are consistent.
- If the rank of $\mathbf{A}$ is less than the rank of $\mathbf{A}_b$ then the equations are inconsistent and have no solution.
- If the rank of the $m \times n$ matrix is r, then it has r linearly independent column vectors and the remaining $n - r$ column vectors is a linear combination of these r column vectors.

Example:

If $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ then $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ and $\mathbf{A}_b = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \end{pmatrix}$

Rank of **A:** $\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0 \quad \Rightarrow$ rank of $\mathbf{A} = 1$

Rank of $\mathbf{A}_b$: $\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0$ as before

but $\begin{vmatrix} 3 & 4 \\ 6 & 5 \end{vmatrix} = -9 \quad \Rightarrow$ rank of $\mathbf{A}_b = 2$ In this case rank of **A** is less than rank of $\mathbf{A}_b \quad \Rightarrow$ no solution exists.

1.5.2 Uniqueness of Solutions

1. With a set of n equations in n unknowns, the equations are consistent if the coefficient matrix **A** and the augmented matrix $\mathbf{A}_b$ are each of rank n. There is then a unique solution for the n equations.

2. If the rank of **A** and that of $\mathbf{A}_b$ is m, where $m < n$, then the matrix **A** is singular, i.e. $|\mathbf{A}| = 0$, and there will be an infinite number of solutions for the equations.

3. If the rank of **A** is less than the rank of $\mathbf{A}_b$, then no solution exists. That is, with a set of n equations in n unknowns

(i) a unique solution exists if rank $\mathbf{A}$ = rank $\mathbf{A}_b = n$

(ii) an infinite number of solution exists if rank $\mathbf{A}$ = rank $\mathbf{A}_b = m < n$

(iii) no solution exists if rank **A**<rank A_b

Example:

Show that $\begin{pmatrix} -4 & 5 \\ -8 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \end{pmatrix}$ has an infinite number of solutions.

$\mathbf{A} = \begin{pmatrix} -4 & 5 \\ -8 & 10 \end{pmatrix}$ and $\mathbf{A}_b = \begin{pmatrix} -4 & 5 & -3 \\ -8 & 10 & -6 \end{pmatrix}$ Rank **A**: $= \begin{vmatrix} -4 & 5 \\ -8 & 10 \end{vmatrix} = 0 \quad \Rightarrow$ rank $\mathbf{A} = 1$ Rank $\mathbf{A}_b$: $= \begin{vmatrix} -4 & 5 \\ -8 & 10 \end{vmatrix} = 0$, $\begin{vmatrix} 5 & -3 \\ 10 & -6 \end{vmatrix} = 0$, $\begin{vmatrix} -4 & -3 \\ -8 & -6 \end{vmatrix} = 0 \quad \Rightarrow$ rank of $\mathbf{A}_b = 1$. Therefore rank of **A**=rank of $\mathbf{A}_b = 1 < n = 2$.

Therefore an infinite number of solutions exist.
For the homogeneous linear equations

$$\begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots & \ldots\ldots\ldots \\ a_{m1} & a_{m2} \ldots a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$$

i.e. $\mathbf{Ax} = 0$ (∗)

Let r be the rank of the matrix $\mathbf{A}$ of order $m \times n$. We have the following results:

1. A system of homogeneous linear equations always have one or more solutions. The two cases are: $r = n$ or $r < n$. For $r = n$ the eq(*) will have no linearly independent solutions, for in that case trivial (zero) solution is the only solution, while in the case $r < n$ there will be $(n-r)$ independent solutions and therefore the eq(*) will have more than one solution.

2. The number of linearly independent solutions of $\mathbf{Ax} = \mathbf{0}$ is $(n-r)$ i.e. if we assign arbitrary values to $(n-r)$ of the variables, then the values of the others can be uniquely determined.
   Since the rank of $\mathbf{A}$ is r, it has r linearly independent columns.

3. If the number of equations is less than the number of variables, then the solution is always other than $x_1 = x_2 = \ldots = x_n = 0$ (i.e. the solution is always non-trivial solution)

4. If the number of equations is equal to the number of variables a necessary and sufficient condition for solutions other than $x_1 = x_2 = \ldots = x_n = 0$ is that the determinant of the coefficients must be zero.

1.6 Solution of sets of Equations

1.6.1 Inverse method

To solve $\mathbf{Ax} = \mathbf{b}$ we use $\mathbf{x} = \mathbf{A^{-1}b}$. To find $\mathbf{A}^{-1}$ proceed as follows:

(i) evaluate $|\mathbf{A}|$. If $|\mathbf{A}| = 0$ then stop (no solution) else proceed.

(ii) Form $\mathbf{C}$, the matrix of cofactors of $\mathbf{A}$ (the cofactor of any element is its minor together with its 'place sign')

(iii) Write $\mathbf{C}'$, the transpose of $\mathbf{C}$

(iv) Then $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \times \mathbf{C}'$

Example: To solve the system
$3x_1 + 2x_2 - x_3 = 4$
$2x_1 - x_2 + 2x_3 = 10$
$x_1 - 3x_2 - 4x_3 = 5$
we rewrite it in matrix form as

$$\begin{pmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 1 & -3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 5 \end{pmatrix}$$

i.e. $\mathbf{Ax} = \mathbf{b}$

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 1 & -3 & -4 \end{pmatrix}$$

(i) $|\mathbf{A}| = \begin{vmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 1 & -3 & -4 \end{vmatrix} = 55$

(ii) $\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$

where $c_{11} = \begin{vmatrix} -1 & 2 \\ -3 & -4 \end{vmatrix} = 10, \quad c_{12} = -\begin{vmatrix} 2 & 2 \\ 1 & -4 \end{vmatrix} = 10, \quad c_{13} = \begin{vmatrix} 2 & -1 \\ 1 & -3 \end{vmatrix} = -5,$

$c_{21} = -\begin{vmatrix} 2 & -1 \\ -3 & -4 \end{vmatrix} = 11, \quad c_{22} = \begin{vmatrix} 3 & -1 \\ 1 & -4 \end{vmatrix} = -11, \quad c_{23} = -\begin{vmatrix} 3 & 2 \\ 1 & -3 \end{vmatrix} = 11,$

$c_{31} = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3, \quad c_{32} = -\begin{vmatrix} 3 & -1 \\ 2 & 2 \end{vmatrix} = -8, \quad c_{33} = \begin{vmatrix} 3 & 2 \\ 2 & -1 \end{vmatrix} = -7$

So $\quad \mathbf{C} = \begin{pmatrix} 10 & 10 & -5 \\ 11 & -11 & 11 \\ 3 & -8 & -7 \end{pmatrix}$

(iii) $\mathbf{C}' = \begin{pmatrix} 10 & 11 & 3 \\ 10 & -11 & -8 \\ -5 & -11 & -7 \end{pmatrix}$ (i.e. the adjoint of $\mathbf{A}$)

(iv) $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{C}' = \frac{1}{55}\begin{pmatrix} 10 & 11 & 3 \\ 10 & -11 & -8 \\ -5 & -11 & -7 \end{pmatrix}$

So $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{A}^{-1}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \frac{1}{55}\begin{pmatrix} 10 & 11 & 3 \\ 10 & -11 & -8 \\ -5 & -11 & -7 \end{pmatrix}\begin{pmatrix} 4 \\ 10 \\ 5 \end{pmatrix} = \frac{1}{55}\begin{pmatrix} 40+110+15 \\ 40-110-40 \\ -20+10-35 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$

Therefore $x_1 = 3, \quad x_2 = -2, \quad x_3 = 1$

1.6.2 Row Transformation method

Elementary row transformation are:

(a) interchange any two rows

(b) multiply (or divide) every element in a row by a non-zero scalar (constant) k

(c) add to (or subtract from) all the elements of any row k times the corresponding elements of any other row.

Equivalent Matrices: Two matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be equivalent if $\mathbf{B}$ can be obtained from $\mathbf{A}$ by a sequence of elementary transformations.
Theorem: Elementary operations do not change the rank of a matrix.
Corollary 1: Equivalent matrices have the same rank
Corollary 2: The rank of $\mathbf{A}'$ is equal the rank of $\mathbf{A}$
Solution of equations:
Example:

The system $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 3 & -2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix}$

is written as $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 3 & -2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix}$

and $\begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 3 & -2 & 4 & 0 & 0 & 1 \end{pmatrix}$ is transformed to $\begin{pmatrix} 1 & 0 & 0 & \frac{8}{17} & \frac{-2}{17} & \frac{1}{17} \\ 0 & 1 & 0 & \frac{-10}{17} & \frac{11}{17} & \frac{3}{17} \\ 0 & 0 & 1 & \frac{11}{17} & \frac{-7}{17} & \frac{-5}{17} \end{pmatrix}$.

We now have $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{17} \begin{pmatrix} 8 & -2 & 1 \\ -10 & 11 & 3 \\ 11 & -7 & 5 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ -4 \end{pmatrix}$

Which gives $x_1 = 2, \quad x_2 = -3, \quad x_3 = 4$

1.6.3 Gaussian elimination method

Example:

For the system $\begin{pmatrix} 2 & -3 & 2 \\ 3 & 2 & -1 \\ 1 & -4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \\ 6 \end{pmatrix}$

We write the augmented matrix as $\begin{pmatrix} 2 & -3 & 2 & : 9 \\ 3 & 2 & -1 & : 4 \\ 1 & 4 & 2 & : 6 \end{pmatrix}$

which is reduced to upper triangular matrix by elementary transformations to give

$$\begin{pmatrix} 1 & -4 & 2 & : & 6 \\ 0 & 1 & -\frac{2}{5} & : & -\frac{3}{5} \\ 0 & 0 & 1 & : & 4 \end{pmatrix}$$

We now rewrite the system as $\begin{pmatrix} 1 & -4 & 2 \\ 0 & 1 & -\frac{2}{5} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ -\frac{3}{5} \\ 4 \end{pmatrix}$

By backward substitution $x_3 = 4, \ x_2 = 1, \ x_1 = 2$

1.6.4 Triangular Decomposition method: LU-decomposition

1.6.5 Cramer's Rule

1.7 Eigenvalues and Eigenvectors of a Matrix

A non-zero vector $\mathbf{X}$ is called an *eigenvector* or *characteristic vector* of a matrix $\mathbf{A}$, if there is a number λ called the *eigenvalue* or *characteristic value* or *characteristic root* or *latent root* such that $\mathbf{AX} = \lambda\mathbf{X}$
i.e. $\mathbf{AX} = \lambda\mathbf{IX}$, where $\mathbf{I}$ is a unit matrix.

or $\quad (\mathbf{A}-\lambda\mathbf{I})\mathbf{X}=\mathbf{0}$
Since $\mathbf{X}\neq\mathbf{0}$, the matrix $(\mathbf{A}-\lambda\mathbf{I})$ is singular so that its determinant $|\mathbf{A}-\lambda\mathbf{I}|$, which is known as the *characteristic determinant* of $\mathbf{A}$ is zero. This leads to the *characteristics equation* of $\mathbf{A}$ as
$$|\mathbf{A}-\lambda\mathbf{I})|=0 \qquad (*)$$
which follows that every characteristic root λ of a matrix $\mathbf{A}$ is a root of its characteristic equation, eq(*)

Example:

If $\mathbf{A}=\begin{pmatrix}5 & 4\\ 1 & 2\end{pmatrix}$ *find its characteristic roots and vectors.*

The characteristic equation of $\mathbf{A}$ is given by $|A-\lambda I|=0$

i.e. $\left|\begin{pmatrix}5 & 4\\ 1 & 2\end{pmatrix}-\lambda\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right|=\begin{vmatrix}5-\lambda & 4\\ 1 & 2-\lambda\end{vmatrix}=0$

i.e. $\lambda^2-7\lambda+6=0$

or $(\lambda-1)(\lambda-6)=0$

Therefore $\lambda_1=1,\ \lambda_2=6$ are the eigenvalues of $\mathbf{A}$.

Now the eigenvector $\vec{X}_1=\begin{pmatrix}x_1\\ x_2\end{pmatrix}$ of $\mathbf{A}$ corresponding to $\lambda_1=1$ is given by

$(\mathbf{A}-1\mathbf{I})\vec{X}_1=\vec{0}$

i.e. $\left(\begin{bmatrix}5 & 4\\ 1 & 2\end{bmatrix}-1\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$

or $\begin{pmatrix}4 & 4\\ 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$

i.e. $4x_1+4x_2=0$

and $x_1+x_2=0$

which yield $x_1=-x_2$.

If we take $x_1=1$, then $x_2=-1$

Therefore $\vec{X}_1=\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}1\\ -1\end{pmatrix}$

Again the eigenvector $\vec{X}_2=\begin{pmatrix}x_1\\ x_2\end{pmatrix}$ of $\mathbf{A}$ corresponding to $\lambda_2=6$ is given by

$(\mathbf{A}-6\mathbf{I})\vec{X}_2=\vec{0}$

i.e. $\left(\begin{bmatrix}5 & 4\\ 1 & 2\end{bmatrix}-6\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$

or $\begin{pmatrix}-1 & 4\\ 1 & -4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$

i.e. $-x_1 + 4x_2 = 0$
and $x_1 - 4x_2 = 0$
which yield $x_1 = 4x_2$.
If we take $x_1 = 4$, then $x_2 = 1$
Therefore $\vec{X}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$
In general we write $\vec{X}_1 = C_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\vec{X}_2 = C_2 \begin{pmatrix} 4 \\ 1 \end{pmatrix}$, where C_1 and C_2 are non-zero scalars called the *normalization constants*.
Normalization: The constants C_1 and C_2 can be fixed by normalization as follows:
$\vec{X}_1^\dagger \vec{X}_1 = 1$ i.e. $C_1 \begin{pmatrix} 1 & -1 \end{pmatrix} C_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1$ or $2C_1^2 = 1 \quad \Rightarrow C_1 = \frac{1}{\sqrt{2}}$,
hence $\vec{X}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
$\vec{X}_2^\dagger \vec{X}_2 = 1$ i.e. $C_2 \begin{pmatrix} 4 & 1 \end{pmatrix} C_2 \begin{pmatrix} 4 \\ 1 \end{pmatrix} = 1$ or $17C_2^2 = 1 \quad \Rightarrow C_2 = \frac{1}{\sqrt{17}}$,
hence $\vec{X}_2 = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 \\ 1 \end{pmatrix}$

1.7.1 Nature of the eigenvalues and eigenvectors of special types of matrices

Theorem 1: The eigenvalues of a Hermitian matrix are all real.
Proof: Let λ be an eigenvalue of a Hermitian matrix $\mathbf{A}$. By definition there exists a vector $\vec{X} \neq 0$, such that
$\mathbf{A}\vec{X} = \lambda \vec{X}$
$\Rightarrow \quad \vec{X}^\dagger \mathbf{A} \vec{X} = \vec{X}^\dagger \lambda \vec{X} = \lambda \vec{X}^\dagger \vec{X} = \lambda \vec{X}^\dagger \mathbf{I} \vec{X}$
so that $\lambda = \frac{\vec{X}^\dagger \mathbf{A} \vec{X}}{\vec{X}^\dagger \mathbf{I} \vec{X}}$
But $\vec{X}^\dagger \mathbf{A} \vec{X} = \overline{\vec{X}^\dagger \mathbf{A} \vec{X}}$ which implies $\vec{X}^\dagger \mathbf{A} \vec{X}$ is a real number and therefore $\vec{X}^\dagger \mathbf{I} \vec{X}$ is also a real number. Hence λ is real.
Corollary: The eigenvalues of a real symmetric matrix are all real.
Theorem 2: The eigenvalues of a skew-Hermitian matrix are purely imaginary or zero.
Corollary: The eigenvalues of a real skew-symmetric matrix are either zero or purely imaginary.

Theorem 3: The modulus of each eigenvalue of a unitary matrix is unity, i.e. the eigenvalues of a unitary form have absolute value 1

Corollary: The eigenvalues of an orthogonal matrix have the absolute value unity and are real, or complex conjugate in pairs.

Theorem 4: Any two eigenvectors corresponding to two distinct eigenvalues of a Hermitian matrix are orthogonal.

Proof: Let $\vec{X}_1$ and $\vec{X}_2$ be two eigenvectors corresponding to two distinct eigenvalues λ_1 and λ_2 of a Hermitian matrix $\mathbf{A}$; then

$\mathbf{A}\vec{X}_1 = \lambda_1\vec{X}_1$..............................(1)

$\mathbf{A}\vec{X}_2 = \lambda_2\vec{X}_2$..............................(2)

From theorem 1 λ_1 and λ_2 are real. Premultiplying (1) and (2) by $\vec{X}_2^\dagger$ and $\vec{X}_1^\dagger$ respectively

$\vec{X}_2^\dagger\mathbf{A}\vec{X}_1 = \lambda_1\vec{X}_2^\dagger\vec{X}_1$..............................(3)

$\vec{X}_1^\dagger\mathbf{A}\vec{X}_2 = \lambda_2\vec{X}_1^\dagger\vec{X}_2$..............................(4)

But $(\vec{X}_2^\dagger\mathbf{A}\vec{X}_1)^\dagger = \vec{X}_1^\dagger\mathbf{A}\vec{X}_2$

Therefore for a Hermitian matrix $\mathbf{A}^\dagger = \mathbf{A}$ and also $(\vec{X}_2^\dagger)^\dagger = \vec{X}_2$; therefore we have, from (3) and (4),

$(\lambda_1\vec{X}_2^\dagger\vec{X}_1)^\dagger = \lambda_2\vec{X}_1^\dagger\vec{X}_2$

or $\lambda_1\vec{X}_1^\dagger\vec{X}_2 = \lambda_2\vec{X}_1^\dagger\vec{X}_2$

or $(\lambda_1 - \lambda_2)\vec{X}_1^\dagger\vec{X}_2 = 0$

Since $\lambda_1 - \lambda_2 \neq 0$ for distinct roots we have $\vec{X}_1^\dagger\vec{X}_2 = 0 \quad \Rightarrow \quad \vec{X}_1$ and $\vec{X}_2$ are orthogonal

Corollary: Any two eigenvectors corresponding to two distinct eigenvalues of a real symmetric matrix are orthogonal.

Theorem 5: Any two eigenvectors corresponding to two distinct eigenvalues of a unitary matrix are orthogonal.

Theorem 6: The eigenvectors corresponding to distinct eigenvalues of a matrix are linearly independent.

Theorem 7: The characteristic polynomial and hence the eigenvalues of similar matrices are the same. Also if $\vec{X}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ, then $\mathbf{P}^{-1}\vec{X}$ is an eigenvector of $\mathbf{B}$ corresponding to the eigenvalue λ where $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$

Proof: Let $\mathbf{A}$ and $\mathbf{B}$ be two similar matrices. Then there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Consider

$$\begin{aligned}\mathbf{B} - \lambda\mathbf{I} &= \mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}\\ &= \mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}\end{aligned}$$

since $\mathbf{P}^{-1}(\lambda\mathbf{I})\mathbf{P} = \lambda\mathbf{P}^{-1}\mathbf{P} = \lambda\mathbf{I}$
Therefore

$$\begin{aligned}
|\mathbf{B} - \lambda\mathbf{I}| &= |\mathbf{P}^{-1}|.|\mathbf{A} - \lambda\mathbf{I}|.|\mathbf{P}| \\
&= |\mathbf{P}^{-1}||\mathbf{P}|.|\mathbf{A} - \lambda\mathbf{I}| \;\; \text{since scalar quantities commute} \\
&= |\mathbf{P}^{-1}\mathbf{P}|.|\mathbf{A} - \lambda\mathbf{I}| \;\; \text{where} \;\; |CD| = |C||D| \\
&= |\mathbf{A} - \lambda\mathbf{I}| \;\; \text{where} \;\; |\mathbf{P}^{-1}\mathbf{P}| = |\mathbf{I}| = 1
\end{aligned}$$

which follows that $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic polynomial and so they have the same eigenvalues.
Corollary 1: The eigenvalues of a matrix are invariant under similarity transformation.
Corollary 2: If $\mathbf{A}$ is similar to a diagonal matrix $\mathbf{D}$ then the diagonal elements of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$.

1.7.2 Diagonalisation of a matrix

Modal matrix: If the eigenvectors of a matrix $\mathbf{A}$ are arranged as columns of a square matrix, the modal matrix of $\mathbf{A}$ denoted by $\mathbf{M}$, is formed.
Example:
For $\mathbf{A} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{pmatrix}$, $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_3 = -5$ and the corresponding eigenvectors are $\vec{X}_1 = \begin{pmatrix} 4 \\ -7 \\ 1 \end{pmatrix}$, $\vec{X}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $\vec{X}_3 = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$
Then the modal matrix $M = \begin{pmatrix} 4 & 2 & 2 \\ -7 & 0 & 0 \\ 1 & 1 & -3 \end{pmatrix}$
Spectral matrix: Also we define the spectral matrix of $\mathbf{A}$, i.e. $\mathbf{S}$, as a diagonal matrix with the eigenvalues only on the main diagonal
$S = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{pmatrix}$
Note that the eigenvalues of $\mathbf{S}$ and $\mathbf{A}$ are the same as indicated in corollary 2 above.

Show that $\mathbf{AM} = \mathbf{MS}$

$\Rightarrow \mathbf{M}^{-1}\mathbf{AM} = \mathbf{S}$ (i.e. similarity transformation of $\mathbf{A}$ to $\mathbf{S}$ which implies that $\mathbf{S}$ and $\mathbf{A}$ are similar matrices)

Note

1. $\mathbf{M}^{-1}\mathbf{AM}$ transforms the square matrix $\mathbf{A}$ into a diagonal matrix $\mathbf{S}$
2. A square matrix $\mathbf{A}$ of order n can be so transformed if the matrix has n independent eigenvectors.
3. A matrix $\mathbf{A}$ always has n linearly independent eigenvectors if it has n distinct eigenvalues or if it is a symmetric matrix.
4. If the matrix has repeated eigenvalues, it may or may not have linearly independent eigenvectors

1.8 Transformation

Linear form: An expression of the form $\sum_{j=1}^{n} a_{ij}x_j$ is said to be linear form of the variable x_j.

1.8.1 Transformation

If a_{ij} are the given constants and x_j the variables then the set of equation

$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad (\text{for } i = 1, 2, 3 \ldots, n) \qquad (1)$

is called a linear transformation connecting the variables x_j and the variables y_i. The square matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \ldots & \ldots\ldots\ldots \\ a_{n1} & a_{n2} \ldots a_{nn} \end{pmatrix}$$

is said to be the matrix of transformations. The determinant of the matrix $|\mathbf{A}|$ is said to be the determinant or modulus of transformation. For short eq.(1) is written as $\vec{y} = \mathbf{A}\vec{x}$

When $|\mathbf{A}| = 0$ the transformation is called singular and when $|\mathbf{A}| \neq 0$ the transformation is said to be non-singular. For non-singular matrices the transformation may be expressed as $\vec{x} = \mathbf{A}^{-1}\vec{y}$

If $\mathbf{A}$ is an identity matrix $\mathbf{I}$ then we have the identical transformation and its determinant is unity. In this case $y_1 = x_1,\ y_2 = x_2,\ y_3 = x_3,\ \ldots, y_n = x_n$

1.8.2 Resultant of two linear transformation

Let two successive transformations be $\vec{y} = \mathbf{P}\vec{x}$ and $\vec{z} = \mathbf{Q}\vec{y}$, then the resultant transformation is $z = \mathbf{QP}x$

1.8.3 Similarity transformation

If $\mathbf{A}$, $\mathbf{B}$ are two non-singular matrices and there exist two non-singular matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{QAP}$ with $\mathbf{Q} = \mathbf{P}^{-1}$ so that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP} \tag{2}$$

then the transformation of $\mathbf{A}$ into $\mathbf{B}$ is termed similarity transformation and the matrices $\mathbf{A}$ and $\mathbf{B}$ are known as similar matrices.

A matrix equation $\mathbf{A}\vec{x} = \mathbf{B}$ preserves its structure (or form) under similarity transformation, $\mathbf{P}^{-1}(\mathbf{A}\vec{x})\mathbf{P} = \mathbf{P}^{-1}\mathbf{BP}$

$\Rightarrow\ \ \mathbf{P}^{-1}\mathbf{APP}^{-1}\vec{x}\mathbf{P} = \mathbf{P}^{-1}\mathbf{BP} \qquad$ as $\mathbf{PP}^{-1} = \mathbf{I}$

$\Rightarrow\ \ (\mathbf{P}^{-1}\mathbf{AP})(\mathbf{P}^{-1}\vec{x}\mathbf{P}) = \mathbf{P}^{-1}\mathbf{BP}$

$\Rightarrow \mathbf{C}\vec{y} = \mathbf{D}$, (where $\mathbf{C} = \mathbf{P}^{-1}\mathbf{AP},\ \vec{y} = \mathbf{P}^{-1}\vec{x}\mathbf{P}$ and $\mathbf{D} = \mathbf{P}^{-1}\mathbf{BP}$). By eq.(2) this is of the form $\mathbf{A}\vec{x} = \mathbf{B}$

The trace of a matrix is invariant under similarity transformation.

1.8.4 Unitary transformation

Let $\mathbf{A}$ be a unitary matrix of order $n \times n$ and $\vec{y},\ \vec{x}$ are column vectors of order $n \times 1$, then the linear transformation $\vec{y} = \mathbf{A}\vec{x}$ is known as unitary transformation. Since $\mathbf{A}^\dagger\mathbf{A} = \mathbf{AA}^\dagger = \mathbf{I}$

Therefore $\vec{y}^\dagger\vec{y} = (\mathbf{A}\vec{x})^\dagger(\mathbf{A}\vec{x}) = \vec{x}^\dagger\mathbf{A}^\dagger\mathbf{A}\vec{x} = \vec{x}^\dagger\vec{x}$

$\Rightarrow$ the norm of vectors is invariant under similarity transformation. In eq.(2), if $\mathbf{P}$ is unitary, i.e., $\mathbf{P}^\dagger\mathbf{P} = \mathbf{PP}^\dagger = \mathbf{I}$, i.e. $\mathbf{P}^\dagger = \mathbf{P}^{-1}$ then the transformation $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ is also unitary.

1.8.5 Orthogonal transformation

Any transformation $\vec{y} = \mathbf{A}\vec{x}$ that transforms $\sum y^2$ into $\sum x^2$ is said to be an orthogonal transformation and the matrix $\mathbf{A}$ is known as an orthogonal matrix. The necessary and sufficient condition for a square matrix to be orthogonal is $\mathbf{AA}' = \mathbf{I}$. In eq.(2) if $\mathbf{P}$ is orthogonal, then $\mathbf{P}^{-1} = \mathbf{P}'$ and eq.(2) is an orthogonal transformation. The product of two orthogonal transformations is an orthogonal transformation. Two n-vectors $\vec{x}$ and $\vec{y}$ are orthogonal to each other if $\vec{x}.\vec{y} = \langle \vec{x}, \vec{y} \rangle = 0$ i.e. if $\vec{x}^\dagger \vec{y} = 0$.

1.8.6 Orthogonal set

A set of vectors is said to be orthonormal if:

- each vector of the set is a normal vector
- any two vectors of the set are orthogonal

1.9 Bases and dimension

Let X be a linear space over K. Every subset of a linearly independent subset of X is automatically linearly independent. X possesses a maximal linearly independent subset called a *basis* of X. The *cardinality* of a basis is an invariant of X, since any two bases possess the same cardinality. This invariant is called the dimension of X. If the dimension of X is finite, then X is called a finite dimensional linear space, otherwise X is an infinite dimensional linear space.

1.9.1 Linear Maps

Let X, Y be two linear spaces over K and D a subspace of X. A transformation $T : D \rightarrow Y$ is called a *linear map* or a *linear operator* if
$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2)$
where $x_1, x_2 \in D$ and $\alpha, \beta \in K$. The set D is called the *domain* of T and is often written as $D(T)$. The set $R(T) = \{Tx : x \in D(T)\}$ is the *range* or *codomain* of T. $R(T)$ is automatically a linear space over K

Chapter 2

FUNCTIONAL ANALYSIS

2.1 Normed Spaces

A normed space over a field K is a pair $(X, \| \cdot \|)$ consisting of a linear space X over K and a map $\| \cdot \|: X \to \Re$ with the following properties:

(i) $\| x \| \geq 0 \forall x \in X$ i.e. $\| \cdot \|$ is non-negative

(ii) $\| x \| = 0$ implies $x = 0$ (i.e. $\ker(\| \cdot \|) = \{0\}$)

(iii) $\| \lambda x \| = |\lambda| \, \| x \|$ for all $x \in X$, $\lambda \in K$ (i.e. $\| \cdot \|$ is positively homogeneous)

(iv) $\| x+y \| \leq \| x \| + \| y \|$ for all $x, \ y \in X$ (i.e. $\| \cdot \|$ satisfies the triangular inequality).

For $x \in X$, the number $\| x \|$ is called the norm of x

2.1.1 Cauchy Sequences

Let $(X, \| \cdot \|)$ be normed space and x_n a sequence of number of X. Then x_n is a Cauchy sequence if given any $\varepsilon > 0$, there is natural number $N(\varepsilon)$, such that $\| x_n - x_m \| < \varepsilon$ whenever $n, m > N(\varepsilon)$

2.1.2 Completeness

A normed space in which every Cauchy sequence has a limit is said to be complete.

2.1.3 Pre-Hilbert spaces

An inner product space or a pre-Hilbert space is a pair $(H, \langle\cdot,\cdot\rangle)$ consisting of a linear space H over K and a functional $\langle\cdot,\cdot\rangle : H \times H \to K$, called the inner product of H, with the following properties:

(i) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle, \quad \forall x, y, z \in H$

(ii) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle, \quad \forall x \in H,\ \alpha \in K$

(iii) $\langle x, y\rangle = \langle y, x\rangle,\ \forall x, y \in H$

(iv) $\langle x, x\rangle \geq 0,\ \forall x \in H$ and $\langle x, x\rangle = 0$ iff $x = 0$.

<u>*Remark*</u>

1. For $x, y \in H$ the number $\langle x, y\rangle$ is called the inner product of x and y.
2. For $x \in H$, define $\| x \|$ by $\| x \| = \sqrt{\langle x, x\rangle}, \qquad x \in H$ Then, $\| \cdot \|$ is a norm on H, whence $(H, \| \cdot \|)$ is a normed space. $\| \cdot \|$ is called the norm induced by the inner product $\langle\cdot,\cdot\rangle$
3. With $\| \cdot \|$ as in 2., one can show that
   $\| x+y \|^2 + \| x-y \|^2 = 2 \| x \|^2 + 2 \| y \|^2$ for all $x, y \in H$.
   This result is called the *parallelogram law* and is a characterizing property of prer-Hilbert spaces, i.e. if a norm does not satisfy the parallelogram law, then it is not induced by an inner product

2.1.4 Hilbert Spaces

A Hilbert space is a pre-Hilbert space $(H, \langle\cdot,\cdot\rangle)$ such that the pair $(H, \| \cdot \|)$, where $\| \cdot \|$ is the norm induced by $\langle\cdot,\cdot\rangle$, is a complete normed space
Example of Hilbert space: Define $\langle\cdot,\cdot\rangle : K^n \times K^n \to K$ by $\langle x, y\rangle = \sum_{j=1}^{n} \bar{x}_j y_j$ where $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ Then $(K^n, \langle\cdot,\cdot\rangle)$ is a Hilbert space finite dimension.

2.1.5 Geometry of Hilbert space

Chapter 3

SPECIAL FUNCTIONS

3.1 The gamma and beta functions

3.1.1 The gamma function Γ

The gamma function $\Gamma(x)$ is defined by the integral
$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ (1)
and is convergent for $x > 0$. It follows from eq.(1) that
$\Gamma(x+1) = \int_0^\infty t^{x}e^{-t}dt$
Integrating by parts
$\Gamma(x+1) = \left[t^x(\frac{e^{-t}}{-1})\right]_0^\infty + x\int_0^\infty t^{x-1}e^{-t}dt$
$\Gamma(x+1) = x\Gamma(x)$ (2)
This is a fundamental recurrence relation for gamma functions. It can also be written as $\Gamma(x) = (x-1)\Gamma(x-1)$.
A number of other results can be derived from this as follows:
If $x = n$, a positive integer, i.e. if $n \geq 1$, then

$$\begin{aligned}
\Gamma(n+1) &= n\Gamma(n). \\
&= n(n-1)\Gamma(n-1) \qquad \text{since } \Gamma(n) = (n-1)\Gamma(n-1) \\
&= n(n-1)(n-2)\Gamma(n-2) \qquad \text{since } \Gamma(n-1) = (n-2)\Gamma(n-2) \\
&= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&= n(n-1)(n-2)(n-3)\ldots 1\Gamma(1) \\
&= n!\Gamma(1)
\end{aligned}$$

But $\Gamma(1) = \int_0^\infty t^0 e^{-t}dt = [-e^{-t}]_0^\infty = 1$
$\Rightarrow \quad \Gamma(n+1) = n!$ (3)
Examples:

$\Gamma(7) = 6! = 720, \quad \Gamma(8) = 7! = 5040, \quad \Gamma(9) = 40320$

We can also use the recurrence relation in reverse

$\Gamma(x+1) = x\Gamma(x) \quad \Rightarrow \quad \Gamma(x) = \frac{\Gamma(x+1)}{x}$

Example:

If $\Gamma(7) = 720$ then $\Gamma(6) = \frac{\Gamma(6+1)}{6} = \frac{\Gamma(7)}{6} = \frac{720}{6} = 120$

If $x = \frac{1}{2}$ it can be shown that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (F.E.M. 150)

Using the recurrence relation $\Gamma(x+1) = x\Gamma(x)$ we can obtain the following:

$\Gamma(\frac{3}{2}) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{1}{2}(\sqrt{\pi}) \quad \Rightarrow \quad \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$

$\Gamma(\frac{5}{2}) = \frac{3}{2}\Gamma(\frac{3}{2}) = \frac{3}{2}(\frac{\sqrt{\pi}}{2}) \quad \Rightarrow \quad \Gamma(\frac{5}{2}) = \frac{3\sqrt{\pi}}{4}$

Negative values of x

Since $\Gamma(x) = \frac{\Gamma(x+1)}{x}$, then as $x \to 0,\ \Gamma(x) \to \infty \quad \Rightarrow\ \Gamma(0) = \infty$

The same result occurs for all negative integral values of x

Examples:

At $x = -1, \quad \Gamma(-1) = \frac{\Gamma(0)}{-1} = \infty$

At $x = -2, \quad \Gamma(-2) = \frac{\Gamma(-1)}{-2} = \infty$ etc.

Also at $x = -\frac{1}{2}, \quad \Gamma(-\frac{1}{2}) = \frac{\Gamma(\frac{1}{2})}{-\frac{1}{2}} = -2\sqrt{\pi}$

and at $x = -\frac{3}{2}, \quad \Gamma(-\frac{3}{2}) = \frac{\Gamma(-\frac{1}{2})}{-\frac{3}{2}} = \frac{4}{3}\sqrt{\pi}$

Gragh of $y = \Gamma(x)$

Examples:

1. Evaluate $\int_0^\infty x^7 e^{-x} dx$

$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$

Let $\Gamma(v) = \int_0^\infty x^{v-1} e^{-x} dx \quad \Rightarrow v = 8$

i.e. $\int_0^\infty x^7 e^{-x} dx = \Gamma(8) = 7! = 5040$

2. Evaluate $\int_0^\infty x^3 e^{-4x} dx$

Since $\Gamma(v) = \int_0^\infty x^{v-1} e^{-x} dx$ we use the substitution $y = 4x \quad \Rightarrow dy = 4dx$

$\Rightarrow \ I = \frac{1}{4^4} \int_0^\infty y^3 e^{-y} dy = \frac{1}{4^4}\Gamma(v)$ where $v = 4 \quad \Rightarrow \ I = \frac{3}{128}$

3. Evaluate $\int_0^\infty x^{\frac{1}{2}} e^{-x^2} dx$

Use $y = x^2$ therefore $dy = 2xdx$. Limits $x = 0, y = 0 \quad x = \infty, y = \infty$

$x = y^{\frac{1}{2}} \quad \Rightarrow x^{\frac{1}{2}} = y^{\frac{1}{4}}$
$I = \int_0^\infty \frac{y^{\frac{1}{4}}e^{-y}}{2x}dy = \int_0^\infty \frac{y^{\frac{1}{4}}e^{-y}}{2y^{\frac{1}{2}}}dy = \frac{1}{2}\int_0^\infty y^{-\frac{1}{4}}e^{-y}dy = \frac{1}{2}\int_0^\infty y^{v-1}e^{-y}dy$ where $v = \frac{3}{4} \quad \Rightarrow \ I = \frac{1}{2}\Gamma(\frac{3}{4})$
From tables, $\Gamma(0.75) = 1.2254 \quad \Rightarrow \ I = 0.613$

3.1.2 The beta function, β

The beta function $\beta(m,n)$ is defined by $\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$
It can be shown that the beta function and the gamma function are related as $\beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \frac{(m-1)!(n-1)!}{(m+n-1)!}$

3.1.3 Application of gamma and beta functions

Examples:
1. Evaluate $I = \int_0^1 x^5(1-x)^4dx$
Comparing this with $\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$
then $m - 1 = 5 \quad \Rightarrow m = 6$ and $n - 1 = 4 \quad \Rightarrow n = 5$
$I = \beta(6,5) = \frac{5!4!}{10!} = \frac{1}{1260}$
2.Evaluate $I = \int_0^1 x^4\sqrt{1-x^2}dx$
Comparing this with $\beta(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx$
we see that we have x^2 in the root, instead of a single x. Therefore, put $x^2 = y \quad \Rightarrow x = y^{\frac{1}{2}}$ and $dx = \frac{1}{2}y^{-\frac{1}{2}}dy$
The limits remain unchanged.
$I = \int_0^1 y^2(1-y)^{\frac{1}{2}}\frac{1}{2}y^{-\frac{1}{2}}dy = \frac{1}{2}\int_0^1 y^{\frac{3}{2}}(1-y)^{\frac{1}{2}}dy$
$m - 1 = \frac{3}{2} \quad \Rightarrow m = \frac{5}{2}$ and $n - 1 = \frac{1}{2} \quad \Rightarrow n = \frac{3}{2}$
Therefore, $I = \frac{1}{2}\beta(\frac{5}{2},\frac{3}{2}) = \frac{1}{2}\frac{\Gamma(\frac{5}{2})\Gamma(\frac{3}{2})}{\Gamma(\frac{5}{2}+\frac{3}{2})} = \frac{1}{2}\frac{(\frac{3}{4}\sqrt{\pi})(\frac{1}{2}\sqrt{\pi})}{3!} = \frac{\pi}{32}$
3. Evaluate $I = \int_0^3 \frac{x^3dx}{\sqrt{3-x}}$ (F.E.M. 170)

3.2 Bessel's Functions

Bessel's functions are solutions of the Bessel's differential equation

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - v^2)y = 0 \qquad (1)$$

where v is a real constant.
By the Frobenius method we assume a series solution of the form

$$y = x^c(a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_rx^r + \ldots) \text{ or } y = x^c\sum_{r=0}^{\infty} a_rx^r$$

$$\text{i.e. } y = a_0x^c + a_1x^{c+1} + a_2x^{c+2} + \ldots + a_rx^{c+r} + \ldots \quad \text{or } y = \sum_{r=0}^{\infty} a_rx^{c+r} \quad (2)$$

where $c, a_0, a_1, a_2, \ldots, a_r$ are constants. a_0 is the first non-zero coefficient. c is called the indicial constant.

$$\frac{dy}{dx} = a_0cx^{c-1} + a_1(c+1)x^c + a_2(c+2)x^{c+1} + \ldots + a_r(c+r)x^{c+r-1} + \ldots \quad (3)$$

$$\frac{d^2y}{dx^2} = a_0c(c-1)x^{c-2} + a_1c(c+1)x^{c-1} + a_2(c+1)(c+2)x^c + \ldots + a_r(c+r-1)(c+r)x^{c+r-2} + \ldots \quad (4)$$

Substituting eqs.(2),(3) and (4) into (1) and equating coefficients of equal powers of x, we have $c = \pm v$ and $a_1 = 0$.
The recurrence relation is $a_r = \frac{a_{r-2}}{v^2-(c+r)^2}$ for $r \geq 2$.
It follows that $a_1 = a_3 = a_5 = a_7 = \ldots = 0$ so that when $c = +v$
$a_2 = \frac{-a_0}{2^2(v+1)}$
$a_4 = \frac{a_0}{2^4\times 2!(v+1)(v+2)}$
$a_6 = \frac{-a_0}{2^6\times 3!(v+1)(v+2)(v+3)}$
$a_r = \frac{(-1)^{\frac{r}{2}}a_0}{2^r\times \frac{r}{2}!(v+1)(v+2)...(v+\frac{r}{2})}$
for r even. The resulting solution is

$$y_1 = a_0x^v\left\{1 - \frac{x^2}{2^2(v+1)} + \frac{x^4}{2^4\times 2!(v+1)(v+2)} - \frac{x^6}{2^6\times 3!(v+1)(v+2)(v+3)} + \ldots\right\}$$

This is valid provided v is not a negative integer.
Similarly, when $c = -v$

$$y_2 = a_0x^{-v}\left\{1 + \frac{x^2}{2^2(v-1)} + \frac{x^4}{2^4\times 2!(v-1)(v-2)} + \frac{x^6}{2^6\times 3!(v-1)(v-2)(v-3)} + \ldots\right\}$$

This is valid provided v is not a positive integer. The complete solution is

$$y = Ay_1 + By_2$$

with the two arbitrary constants A and B.
Besssel's functions:

Let $a_0 = \frac{1}{2^v\Gamma(v+1)}$ then the solution y_1 gives for $c = v = n$ (where n is a positive integer) Bessel's functions of the first kind of order n denoted by $J_n(x)$ where

$$\begin{aligned} J_n(x) &= (\frac{x}{2})^n \left\{ \frac{1}{\Gamma(n+1)} - \frac{x^2}{2^2(1!)\Gamma(n+2)} + \frac{x^4}{2^4(2!)\Gamma(n+3)} - \dots \right\} \\ &= (\frac{x}{2})^n \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}(k!)\Gamma(n+k+1)} \\ &= (\frac{x}{2})^n \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}(k!)(n+k)!} \end{aligned}$$

Similarly for $c = -v = -n$ (a negative integer)

$$\begin{aligned} J_{-n}(x) &= (\frac{x}{2})^{-n} \left\{ \frac{1}{\Gamma(1-n)} - \frac{x^2}{2^2(1!)\Gamma(2-n)} + \frac{x^4}{2^4(2!)\Gamma(3-n)} - \dots \right\} \\ &= (\frac{x}{2})^{-n} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}(k!)\Gamma(k-n+1)} \\ &= (-1)^n(\frac{x}{2})^n \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}(k!)(n+k)!} \quad \text{(for details see F.E.M. 247)} \\ &= (-1)^n J_n(x) \end{aligned}$$

$\Rightarrow$ The two solutions $J_n(x)$ and $J_{-n}(x)$ dependent on each other. Further more the series for $J_n(x)$ is

$$J_n(x) = (\frac{x}{2})^n \left\{ \frac{1}{n!} - \frac{1}{(n+1)!}(\frac{x}{2})^2 + \frac{1}{(2!)(n+2)!}(\frac{x}{2})^4 - \dots \right\}$$

From this we obtain two commonly used functions

$$J_0(x) = 1 - \frac{1}{(1!)^2}(\frac{x}{2})^2 + \frac{1}{(2!)^2}(\frac{x}{2})^4 - \frac{1}{(3!)^2}(\frac{x}{2})^6 + \dots$$

$$J_1(x) = \frac{x}{2}\left\{ 1 - \frac{1}{(1!)(2!)}(\frac{x}{2})^2 + \frac{1}{(2!)(3!)}(\frac{x}{2})^4 + \dots \right\}$$

Graphs of Bessel's functions $J_0(x)$ and $J_1(x)$

Remark: Note that $J_0(x)$ and $J_1(x)$ are similar to $\cos x$ and $\sin x$ respectively.

Generating function: If we want to study a certain sequence $\{f(x)\}$ and can find a function $G(t,x) = \sum_{n=0}^{\infty} f_n(x)t^n$ we may obtain the properties of $\{f_n(x)\}$ from those of G which "generates" this sequence and is called a *generating function* of it.

The generating function for $J_n(x)$ is $e^{\frac{x}{2}(t-\frac{1}{t})} = \sum_{-\infty}^{\infty} J_n(x)t^n$

Recurrence formula: $J_n(x)$ can also be obtained from the recurrence formula $\rightarrow J_{n+1}(x) = \frac{2n}{x}\left[J_n(x) - J_{n-1}(x)\right]$

For $(0 < x < 1)$ $J_n(x)$ are orthogonal

3.3 Legendre's Polynomials

These are solutions of the Legendre's differential equation

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + k(k+1)y = 0$$

where k is a real constant. Solving it by the Frobenius method as before we obtain $c = 0$ and $c = 1$ and the corresponding solutions are

a) $c = 1 : y = a_0\left\{1 - \frac{k(k+1)}{2!}x^2 + \frac{k(k-2)(k+1)(k+3)}{4!}x^4 - \ldots\right\}$

b) $c = 0 : y = a_1\left\{x - \frac{(k-1)(k-2)}{3!}x^3 + \frac{(k-1)(k-3)(k+2)(k+4)}{5!}x^5 - \ldots\right\}$

where a_0 and a_1 are the usual arbitrary constants. When k is an integer n, one of the solution series terminates after a finite number of terms. The resulting polynomial in x denoted by $P_n(x)$ is called *Legendre polynomial* with a_0 and a_1 being chosen so that the polynomial has unit value when $x = 1$. $(-1 < x < 1)$ orthogonality

e.g. $P_0(x) = a_0\{1 - 0 + 0 - \ldots\} = a_0$. We choose $a_0 = 1$ so that $P_0(x) = 1$

$P_1(x) = a_1\{x - 0 + 0 - \ldots\} = a_1 x$

a_1 is then chosen to make $P_1(x) = 1$ when $x = 1$ $\Rightarrow$ $a_1 = 1 \Rightarrow P_1(x) = x$

$p_2(x) = a_0\left\{1 - \frac{2x^3}{2!}x^2 + 0 + 0 + \ldots\right\}$ $= a_0\{1 - 3x^2\}$

If $P_2(x) = 1$ when $x = 1$ then $a_0 = -\frac{1}{2} \Rightarrow P_2(x) = \frac{1}{2}(3x^2 - 1)$

Using the same procedure obtain:

$P_3(x) = \frac{1}{2}(5x^3 - 3x)$
$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$
$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$ etc.
Legendre polynomials can also be expressed by Rodrigue's formula given by

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n$$

(Use this formula to obtain $P_0(x), P_1(x), P_2(x), P_3(x)$, etc)
The generating function is

$$\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n$$

To show this, start from the binomial expansion of $\frac{1}{\sqrt{1-v}}$ where $v = 2xt - t^2$, multiply the powers of $2xt - t^2$ out, collect all the terms involving t^n and verify that the sum of these terms is $P_n(x)t^n$.
The recurrence formula for Legendre polynomials is

$$P_{n+1}(x) = \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x)$$

This means that if we know $P_{n-1}(x)$ and $P_n(x)$ we can calculate $P_{n+1}(x)$, e.g. given that $P_0(x) = 1$ and $P_1(x) = x$ we can calculate $P_2(x)$ using the recurrence formula by taking $P_{n-1} = P_0,\ P_n = P_1$ and $P_{n+1} = P_2 \quad \Rightarrow n = 1$.
Substituting these in the formula,
$P_2(x) = \frac{2\times1+1}{1+1}xP_1(x) - \frac{1}{1+1}P_0(x) = \frac{1}{2}(3x^2 - 1)$
Similarly to find $P_3(x)$ we set $P_{n-1} = P_1,\ P_n = P_2$ and $P_{n+1} = P_3$ where $n = 2$. Substituting these in the formula we have

$$\begin{aligned} P_3(x) &= \frac{2\times2+1}{2+1}xP_2(x) - \frac{2}{2+1}P_1(x) \\ &= \frac{5}{3}x \times \frac{1}{2}(3x^2 - 1) - \frac{2}{3}x \\ &= \frac{1}{2}(5x^3 - 3x) \end{aligned}$$

(Using the recurrence formula obtain $P_4(x)$ and $P_5(x)$)

3.4 Hermite Polynomials

They are solutions of the Hermite differential equation

$$\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2vy = 0 \qquad (*)$$

where v is a parameter. Using the Frobenius method the solution is

$$y = \sum_{r=0}^{\infty} a_r x^{c+r}, \text{ where } a_0 \neq 0$$
$$\frac{dy}{dx} = \sum_{r=0}^{\infty} a_r (c+r) x^{c+r-1} \text{ and}$$
$$\frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} a_r (c+r)(c+r-1) x^{c+r-2}$$

Substituting these in eq.(*) and equating coefficients of like terms we have

$a_0 c(c-1) = 0 \quad \Rightarrow c = 0, \text{ or } c = 1$

and $a_{r+2} = \frac{2(c+r-v)a_r}{(c+r+2)(c+r+1)} \quad \Rightarrow a_1 = a_3 = a_5 = \ldots = 0$

When $c = 0$ (M.P. by Gupta 8.94),

$y_1 = a_0 \left\{ 1 - \frac{2v}{2!}x^2 + \frac{2^2 v(v-2)}{4!}x^4 - \ldots + \frac{(-1)^r 2^r v(v-2)\ldots(v-2r+2)}{2r!}x^{2r} + \ldots \right\}$

(where $a_1 = 0$). When c=1,

$y_2 = a_0 x \left\{ 1 - \frac{2(v-1)}{2!}x^2 + \frac{2^2(v-1)(v-3)}{4!}x^4 - \ldots + \frac{(-1)^r 2^r (v-1)(v-3)\ldots(v-2r+1)}{(2r+1)!} + \ldots \right\}$

The complete solution of eq.(*) is then given by $y = Ay_1 + By_2$ i.e.

$y = A\left\{1 - \frac{2v}{2!}x^2 + \frac{2^2 v(v-2)}{4!}x^4 - \ldots\right\} + Bx\left\{1 - \frac{2(v-1)}{3!}x^2 + \frac{2^2(v-1)(v-3)}{5!}x^4 - \ldots\right\}$

where A and B are arbitrary constants. When $v = n$, an integer, the series terminates after a few terms. The resulting polynomials $H_n(x)$ are called *Hermite polynomials*. The first 5 of them are:

$H_0(x) = 1$

$H_1(x) = 2x$

$H_2(x) = 4x^2 - 2$

$H_3(x) = 8x^3 - 12x$

$H_4(x) = 16x^4 - 48x^2 + 12$

$H_5(x) = 32x^5 - 160x^3 + 120x$

They can also be given by a corresponding Rodrigue's formula

$$H_n(x) = e^{x^2}(-1)^n \frac{d^n}{dx^n}(e^{-x^2})$$

The generating function is given by

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n$$

This can be proved using the formula for the coefficients of a Maclaurin series and noting that $tx - \frac{1}{2}t^2 = \frac{1}{2}x^2 - \frac{1}{2}(x-t)^2$

Hermite polynomials satisfy the recursion formula

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$$

(Given that $H_0 = 1$ and $H_1 = 2x$ use this formula to obtain H_2, H_3, H_4 and H_5).

3.5 Laguerre Polynomials

They are solutions of the Laguerre differential equation
$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + vy = 0$ (*1)
Using the Frobenius method again we have

$$\begin{aligned} y &= \sum_{r=0}^{\infty} a_r x^{c+r}, \text{ where } a_0 \neq 0 \\ \frac{dy}{dx} &= \sum_{r=0}^{\infty} a_r (c+r) x^{c+r-1} \text{ and} \\ \frac{d^2y}{dx^2} &= \sum_{r=0}^{\infty} a_r (c+r)(c+r-1) x^{c+r-2} \end{aligned}$$

Substituting these in eq.(*1) and equating coefficients of like terms we have
$c^2 = 0 \quad \Rightarrow \quad c = 0$
and $a_{r+1} = \frac{c+r-v}{(c+r+1)^2} a_r = \frac{r-v}{(r+1)^2} a_r$
$y = a_0 \left\{ 1 - vx + \frac{v(v-1)}{(2!)^2} x^2 - \ldots + (-1)^r \frac{v(v-1)\ldots(v-r+1)}{(r!)^2} x^r + \ldots \right\}$ (*2)
In case $v = n$ (a positive integer) and $a_0 = n!$ the solution eq.(*2) is said to be the Laguerre polynomial of degree n and is denoted by $L_n(x)$ i.e.

$$L_n(x) = (-1)^n \left\{ x^n - \frac{n^2}{1!} x^{n-1} + \frac{n^2(n-1)^2}{2!} x^{n-2} + \ldots + (-1)^n n! \right\} \quad (*3)$$

Then the solution of Laguerre equation for v to be a positive integer is

$$y = AL_n(x)$$

From eq.(*3) it is easy to show that
$L_0(x) = 1$
$L_1(x) = 1 - x$
$L_2(x) = x^2 - 4x + 2$
$L_3(x) = -x^3 + 9x^2 - 18x + 6$
$L_4(x) = x^4 - 16x^3 + 72x^2 - 96x + 48$
They can also be given by the Rodrigue's formula

$$L_n(x) = e^x \frac{d^n}{dx^n}(x^n e^{-x})$$

Their generating function is

$$e^{-\frac{xt}{(1-t)}} = \sum_{n=0}^{\infty} \frac{L_n(x)}{n!} t^n$$

They satisfy the recursion formula

$$L_{n+1} = (2n + 1 - x)L_n(x) - n^2 L_{n-1}(x)$$

They are orthogonal for $0 < x < \infty$

3.5.1 Hypergeometric Function

The solutions of Gauss's hypergeometric differential equation

$$x(1-x)y'' + [\gamma - (\alpha+\beta+1)x]y' - aby = 0 \tag{1}$$

where α, β and γ are parametric constants, and by choosing specific values we can obtain a large number of elementary and higher functions as solutions of (1). This accounts for the practical importance of (1)

a) Hypergeometric series

Using the Frobenius method the indicial equation of (1) has roots $c = 0$ and $c = 1-\gamma$; $a_{r+1} = \frac{(c+r+\beta)(c+r+\alpha)}{(c+r+1)(c+r+\gamma)} a_r$

For $c_1 = 0, \quad a_{r+1} = \frac{(\alpha+r)(\beta+r)}{(r+1)(r+\gamma)} a_r$ and the Frobenius method gives

$$y_1(x) = a_0 \left\{1 + \frac{\alpha\beta}{1!\gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{2!\gamma(\gamma+1)}x^2 + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{3!\gamma(\gamma+1)(\gamma+2)}x^3 + \dots\right\} \tag{2}$$

where $\gamma \neq 0, -1, -2, \dots$. This series is called the *hypergeometric series.* Its sum $y_1(x)$ is denoted by $F(\alpha, \beta, \gamma; x)$ and is called the *hypergeometric function.*

b) Special cases: It can be shown that

(i) $\frac{1}{1-x} = F(1,1,1;x) = F(1,\beta,\beta;x) = F(\alpha,1,\alpha;x)$

(ii) $(1+x)^n = F(-n,\beta,\beta;-x)$ or $F(-n,1,1;-x)$

(iii) $(1-x)^n = 1 - nxF(1-n,1,2;x)$

(iv) $\arctan x = xF(\frac{1}{2},1,\frac{3}{2};-x^2)$

(v) $\arcsin x = xF(\frac{1}{2},\frac{1}{2},\frac{3}{2};x^2)$

(vi) $e^x = \lim_{n\to\infty} F(n,1,1;\frac{x}{n})$

(vii) $\ln(1+x) = xF(1,1,2;-x)$

(viii) $\ln\frac{1+x}{1-x} = 2xF(\frac{1}{2},1,\frac{3}{2};x^2)$ etc

c) Second solution

For $c_2 = 1-\gamma, \quad a_{r+1} = \frac{(\alpha+r+1-\gamma)(\beta+r+1-\gamma)}{(2-\gamma+r)(r+1)}$ and the Frobenius method yields the following solution (where $\gamma \neq 2, 3, 4, \dots$)

$$y_2(x) = a_0 x^{1-\gamma}\left\{1 + \frac{(\alpha-\gamma+1)(\beta-\gamma+1)}{1!(-\gamma+2)}x + \frac{(\alpha-\gamma+1)(\alpha-\gamma+2)(\beta-\gamma+1)(\beta-\gamma+2)}{2!(-\gamma+2)(-\gamma+3)}x^2 + \dots\right\}$$

which gives $\quad y_2(x) = x^{1-\gamma} F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma; x)$. The complete solution is

$$y = Ay_1(x) + By_2(x) = AF(\alpha, \beta, \gamma; x) + Bx^{1-\gamma} F(\alpha - \gamma + 1, \beta - \gamma + 1, 2 - \gamma; x)$$

Chapter 4

INTEGRAL TRANSFORM AND FOURIER SERIES

4.1 Laplace Transform

LT is used in solving ordinary differential equations (ode). It has the following advantages:

- Solution of the ode is obtained by algebraic processes.
- the initial conditions are involved from the early stages so that the determination of the particular solution is considerably shortened.
- the method enables us to deal with situations where the function is discontinuous.

The LT of a function $f(t)$ is denoted by $\mathcal{L}\{f(t)\}$ or $F(s)$ and is defined by the integral $\int_0^\infty f(t)e^{-st}dt$
i.e. $\mathcal{L}\{f(t)\}$ or $F(s) = \int_0^\infty f(t)e^{-st}dt$
where s is a positive constant such that $f(t)e^{-st}$ converges as $t \to \infty$
<u>*Examples*</u>

1. To find the LT of a constant function $f(t) = a$
   $\mathcal{L}\{a\} = \int_0^\infty ae^{-st}dt = a\left[\frac{e^{-st}}{-s}\right]_0^\infty = -\frac{a}{s}\left[e^{-st}\right]_0^\infty = -\frac{a}{s}[0-1] = \frac{a}{s}$

$$\Rightarrow \boxed{\mathcal{L}\{a\} = \frac{a}{s}} \qquad (1)$$

   e.g. for $a = 1$, $\mathcal{L}\{1\} = \frac{1}{s}$

2. If $f(t) = e^{at}$

$\mathcal{L}\{e^{at}\} = \int_0^\infty e^{at}e^{-st}dt = \int_0^\infty e^{-(s-a)t}dt = \left[\frac{e^{-(s-a)t}}{-(s-a)}\right]_0^\infty = -\frac{1}{s-a}[0-1] = \frac{1}{s-a}$

$$\Rightarrow \boxed{\mathcal{L}\{e^{at}\} = \frac{1}{s-a}} \tag{2}$$

$$\text{Similarly } \boxed{\mathcal{L}\{e^{-at}\} = \frac{1}{s+a}} \tag{3}$$

3. If $f(t) = \sin at$

$$\begin{aligned}
\mathcal{L}\{\sin at\} = \int_0^\infty \sin(at)e^{-st}dt &= \int_0^\infty \left(\frac{e^{iat} - e^{-iat}}{2i}\right) e^{-st}dt \\
&= \frac{1}{2i}\left(\int_0^\infty e^{-(s-ia)t}dt - \int_0^\infty e^{-(s+ia)t}dt\right) \\
&= \frac{1}{2i}\left(\frac{1}{s-ia} - \frac{1}{s+ia}\right) \\
&= \frac{1}{2}\left(\frac{1}{a+is} + \frac{1}{a-is}\right) \\
&= \frac{1}{2}\frac{a-is+a+is}{(a+is)(a-is)} \\
&= \frac{a}{s^2+a^2}
\end{aligned}$$

$$\Rightarrow \boxed{\mathcal{L}\{\sin at\} = \frac{a}{s^2+a^2}} \tag{4}$$

e.g. $\mathcal{L}\{\sin 2t\} = \frac{2}{s^2+4}$

Similarly (*show that*):

4. if $f(t) = \cos at$

$$\Rightarrow \boxed{\mathcal{L}\{\cos at\} = \frac{s}{s^2+a^2}} \tag{5}$$

e.g. $\mathcal{L}\{\cos 4t\} = \frac{s}{s^2+16}$

5. if $f(t) = t^n$

$$\Rightarrow \boxed{\mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}}} \tag{6}$$

e.g. $\mathcal{L}\{t^3\} = \frac{3!}{s^{3+1}} = \frac{6}{s^4}$.

6. If $f(t) = \sinh at$

$$
\begin{aligned}
\mathcal{L}\{\sinh at\} = \int_0^\infty \sinh(at)e^{-st}dt &= \int_0^\infty \left(\frac{e^{at}-e^{-at}}{2}\right)e^{-st}dt \\
&= \frac{1}{2}\left(\int_0^\infty e^{-(s-a)t}dt - \int_0^\infty e^{-(s+a)t}dt\right) \\
&= \frac{1}{2}\left(\frac{1}{s-a} - \frac{1}{s+a}\right) \\
&= \frac{a}{s^2-a^2}
\end{aligned}
$$

$$
\Rightarrow \boxed{\mathcal{L}\{\sinh at\} = \frac{a}{s^2-a^2}} \qquad (7)
$$

e.g. $\mathcal{L}\{\sinh 2t\} = \frac{2}{s^2-4}$
Similarly (show that)

7. if $f(t) = \cosh at = \frac{1}{2}(e^{at} + e^{-at})$

$$
\Rightarrow \boxed{\mathcal{L}\{\cosh at\} = \frac{s}{s^2-a^2}} \qquad (8)
$$

e.g. $\mathcal{L}\{4\cosh 3t\} = 4\frac{s}{s^2-3^2} = \frac{4s}{s^2-9}$

Existence theorem for Laplace Transforms

Let $f(t)$ be a function that is *piecewise continuous* on every finite interval in the range $t \geq 0$ and satisfies $|f(t)| \leq Me^{-kt}$ for all $t \geq 0$ and for some constants k and M. Then the LT of $f(t)$ exists for all $s > k$

A function $f(t)$ is said to be *piecewise continuous* in an interval (a, b) if

(i) the interval can be divided into a finite number of subintervals in each of which $f(t)$ is continuous.

(ii) the limits of $f(t)$ as t approaches the endpoints of each subinterval are finite.

In other words a piecewise continuous function is one that has a finite number of finite discontinuities. e.g see fig.

4.1.1 Inverse Transform

Given a LT, $F(s)$ one can find the function $f(t)$ by inverse transform $f(t) = \mathcal{L}^{-1}\{F(s)\}$ where $\mathcal{L}^{-1}$ indicates inverse transform.
e.g. $\mathcal{L}^{-1}\left\{\frac{1}{s-2}\right\} = e^{2t}$
$\mathcal{L}^{-1}\left\{\frac{4}{s}\right\} = 4$
$\mathcal{L}^{-1}\left\{\frac{s}{s^2+25}\right\} = \cos 5t$

$$\begin{aligned}
\mathcal{L}^{-1}\left\{\frac{3s+1}{s^2-s-6}\right\} &= \mathcal{L}^{-1}\left\{\frac{1}{s+2} + \frac{2}{s-3}\right\} \text{ (by partial fractions)} \\
&= \mathcal{L}^{-1}\left\{\frac{1}{s+2}\right\} + \mathcal{L}^{-1}\left\{\frac{2}{s-3}\right\} \\
&\quad \text{(Note: } \mathcal{L},\ \mathcal{L}^{-1} \text{ are linear operators. } \textit{Prove it}\text{)} \\
&= e^{-2t} + 2e^{3t}
\end{aligned}$$

Rules of Partial Fractions

1. The numerator must be of lower degree than the denominator. If it is not then we first divide out
2. Factorize the denominator into its prime factors. These determine the shapes of the partial fractions.
3. A linear factor $(s+a)$ gives a partial fraction $\frac{A}{s+a}$ where A is a constant to be determined.
4. A repeated factor $(s+a)^2$ gives $\frac{A}{s+a} + \frac{B}{(s+a)^2}$
5. Similarly $(s+a)^3$ gives $\frac{A}{s+a} + \frac{B}{(s+a)^2} + \frac{C}{(s+a)^3}$
6. Quadratic factor (s^2+ps+q) gives $\frac{As+B}{s^2+ps+q}$
7. repeated quadratic factor $(s^2+ps+q)^2$ gives $\frac{As+B}{s^2+ps+q} + \frac{Cs+D}{(s^2+ps+q)^2}$

Examples
1. $\frac{s^2-15s+41}{(s+2)(s-3)^2} = \frac{3}{s+2} - \frac{2}{s-3} + \frac{1}{(s-3)^2}$
2. $\mathcal{L}^{-1}\left\{\frac{4s^2-5s+6}{(s+1)(s^2+4)}\right\}$

but $\frac{4s^2-5s+6}{(s+1)(s^2+4)} \equiv \frac{A}{s+1} + \frac{Bs+C}{s^2+4} = \frac{3}{s+1} + \frac{s-6}{s^2+4} = \frac{3}{s+1} + \frac{s}{s^2+4} - \frac{6}{s^2+4}$

$\Rightarrow\ f(t) = \mathcal{L}^{-1}\left\{\frac{4s^2-5s+6}{(s+1)(s^2+4)}\right\} = \mathcal{L}^{-1}\left\{\frac{3}{s+1} + \frac{s}{s^2+4} - \frac{6}{s^2+4}\right\}$
$\Rightarrow\ f(t) = 3e^{-t} + \cos 2t - 3\sin 2t$

PROPERTIES OF LAPLACE TRANSFORM

Linearity: $\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}$ (*Prove!*)
(1) The first shift theorem (or s-shifting):
It states that if $\mathcal{L}\{f(t)\} = F(s)$ then

$$\Rightarrow \boxed{\mathcal{L}\{e^{-at}f(t)\} = F(s+a)} \tag{9}$$

i.e. $\mathcal{L}\{e^{-at}f(t)\}$ is the same as $\mathcal{L}\{f(t)\}$ with s replaced by $(s+a)$
Examples
1. If $\mathcal{L}\{\sin 2t\} = \frac{2}{s^2+4}$ then $\mathcal{L}\{e^{-3t}\sin 2t\} = \frac{2}{(s+3)^2+4} = \frac{2}{s^2+6s+13}$
2. If $\mathcal{L}\{t^2\} = \frac{2}{s^3}$ then $\mathcal{L}\{t^2e^{4t}\} = \frac{2}{(s-4)^3}$

(2) Theorem 2: Multiplying by t (or derivative of LT):
If $\mathcal{L}\{f(t)\} = F(s)$ then

$$\Rightarrow \boxed{\mathcal{L}\{tf(t)\} = -\frac{d}{ds}F(s)} \tag{10}$$

e.g. if $\mathcal{L}\{\sin 2t\} = \frac{2}{s^2+4} \quad \Rightarrow \mathcal{L}\{t\sin 2t\} = -\frac{d}{ds}\left(\frac{2}{s^2+4}\right) = \frac{4s}{(s^2+4)^2}$
(3) Theorem 3: Dividing by t: If $\mathcal{L}\{f(t)\} = F(s)$ then

$$\Rightarrow \boxed{\mathcal{L}\left\{\frac{f(t)}{t}\right\} = \int_s^\infty F(s)ds} \tag{11}$$

If limit of $\frac{f(t)}{t}$ as $t \to 0$ exists, we use l'Hopital's rule to find out if it does
e.g. $\mathcal{L}\left\{\frac{\sin at}{t}\right\}$; here $\lim_{t\to 0}\left\{\frac{\sin at}{t}\right\} = \frac{0}{0}$ (undefined).
By l'Hopital's rule, we differentiate top and bottom separately and substitute $t = 0$ in the result to ascertain the limit of the new function.
$\lim_{t\to 0}\left\{\frac{\sin at}{t}\right\} = \lim_{t\to 0}\left\{\frac{a\cos at}{1}\right\} = a,$
i.e. the limit exists. The theorem can, therefore, be applied.
Since $\mathcal{L}\{\sin at\} = \frac{a}{s^2+a^2}$
then $\mathcal{L}\left\{\frac{\sin at}{t}\right\} = \int_s^\infty \frac{a}{s^2+a^2}ds = \left[\arctan\left(\frac{s}{a}\right)\right]_s^\infty = \frac{\pi}{2} - \arctan\left(\frac{s}{a}\right) = \arctan\left(\frac{a}{s}\right)$
(4) Transform of derivative
Let $\frac{df(t)}{dt} = f'(t)$ and $\frac{d^2f(t)}{dt^2} = f''(t)$
Then $\mathcal{L}\{f'(t)\} = \int_0^\infty e^{-st}f'(t)dt$ by definition.
Integrating by parts,

$\mathcal{L}\{f'(t)\} = [e^{-st}f(t)]_0^\infty - \int_0^\infty f(t)\{-se^{-st}\}dt$
i.e. $\mathcal{L}\{f'(t)\} = -f(0) + s\mathcal{L}\{f(t)\}$

$$\Rightarrow \quad \boxed{\mathcal{L}\{f'(t)\} = sF(s) - f(0)} \tag{12}$$

Similarly $\mathcal{L}\{f''(t)\} = -f'(0) + s\mathcal{L}\{f'(t)\} = -f'(0) + s[-f(0) + s\mathcal{L}\{f(t)\}]$

$$\Rightarrow \quad \boxed{\mathcal{L}\{f''(t)\} = s^2F(s) - sf(0) - f'(0)} \tag{13}$$

$$\text{Similarly} \Rightarrow \boxed{\mathcal{L}\{f'''(t)\} = s^3F(s) - s^2f(0) - sf'(0) - f''(0)} \tag{14}$$

Alternative notation:
Let $x = f(t),\ f(0) = x_0,\ f'(0) = x_1,\ f''(0) = x_2, \ldots, f^{(n)}(0) = x_n$ and
$\overline{x} = \mathcal{L}\{x\} = \mathcal{L}\{f(t)\} = F(s)$ we now have
$\mathcal{L}\{x\} = \bar{x}$
$\mathcal{L}\{\dot{x}\} = s\bar{x} - x_0$
$\mathcal{L}\{\ddot{x}\} = s^2\bar{x} - sx_0 - x_1$
$\mathcal{L}\{\dddot{x}\} = s^3\bar{x} - s^2x_0 - sx_1 - x_2$

4.1.2 Solution of Differential Equations by Laplace Transform

Procedure:

a) Rewrite the equation in terms of LT.

b) Insert the given initial conditions.

c) Rearrange the equation algebraically to give the transform of the solution.

d) Determine the inverse transform to obtain the particular solution

Solution of first order differential equations
Example
Solve the equation $\frac{dx}{dt} - 2x = 4$, given that at $t = 0,\ x = 1$.
We go through the four stages as follows:

a) $\mathcal{L}\{x\} = \bar{x},\quad \mathcal{L}\{\dot{x}\} = s\bar{x} - x_0,\quad \mathcal{L}\{4\} = \frac{4}{s}$
Then the equation becomes $(s\bar{x} - x_0) - 2\bar{x} = \frac{4}{s}$

b) Insert the initial condition that at $t = 0,\ x = 1$, i.e., $x_0 = 1$
$\Rightarrow \quad s\bar{x} - 1 - 2\bar{x} = \frac{4}{s}$

c) Now we rearrange this to give an expression for $\bar{x}$:
i.e. $\bar{x} = \frac{s+4}{s(s-2)}$

d) finally, we take inverse transform to obtain x:
$\frac{s+4}{s(s-2)} = \frac{A}{s} + \frac{B}{s-2} \;\Rightarrow\; s+4 = A(s-2) + Bs$
(i) Put $(s-2) = 0,$ i.e., $s = 2 \;\Rightarrow\; 6 = 2B$ or $B = 3$
(ii) Put $s = 0 \;\Rightarrow\; s = -2A$ or $A = -2$
$\bar{x} = \frac{s+4}{s(s-2)} = \frac{3}{s-2} - \frac{2}{s}$
$\Rightarrow\; x = 3e^{2t} - 2$

Solve the following equations:
1. $\frac{dx}{dt} + 2x = 10e^{3t}$, given that at $t = 0,\; x = 6$
2. $\frac{dx}{dt} - 4x = 2e^{2t} + e^{4t}$, given that at $t = 0,\; x = 0$

Solution of second order differential equation

Example

Solve the equation $\frac{d^2x}{dt^2} - 3\frac{dx}{dt} + 2x = 2e^{3t}$, given that at $t = 0,\; x = 5$ and $\frac{dx}{dt} = 7$

a) $\mathcal{L}\{x\} = \bar{x}$
$\mathcal{L}\{\dot{x}\} = s\bar{x} - x_0$
$\mathcal{L}\{\ddot{x}\} = s^2\bar{x} - sx_0 - x_1$
The equation becomes $(s^2\bar{x} - sx_0 - x_1) - 3(s\bar{x} - x_0) + 2\bar{x} = \frac{2}{s-3}$

b) Insert the initial conditions. In this case $x_0 = 5$ and $x_1 = 7$
$(s^2\bar{x} - 5s - 7) - 3(s\bar{x} - 5) + 2\bar{x} = \frac{2}{s-3}$

c) Rearrange to obtain $\bar{x}$ as $\bar{x} = \frac{5s^2-23s+26}{(s-1)(s-2)(s-3)}$

d) Now for partial fractions
$\frac{5s^2-23s+26}{(s-1)(s-2)(s-3)} = \frac{A}{s-1} + \frac{B}{s-2} + \frac{C}{s-3}$
$\Rightarrow 5s^2 - 23s + 26 = A(s-2)(s-3) + B(s-1)(s-3) + C(s-1)(s-2)$
$\Rightarrow\; A = 4,\; B = 0,\; C = 1$
$\Rightarrow\; \bar{x} = \frac{4}{s-1} + \frac{1}{s-3}$
$\Rightarrow\; x = 4e^{t} + e^{3t}$

Solve $\frac{d^2x}{dt^2} - 4x = 24\cos 2t$, given that at $t = 0,\; x = 3$ and $\frac{dx}{dt} = 4$

Solution of simultaneous differential equations

Example

Solve the pair of simultaneous equations
$\dot{y} - x = e^t$
$\dot{x} + y = e^{-t}$
given that at $t = 0,\ x = 0$ and $y = 0$

a) $(s\bar{y} - y_0) - \bar{x} = \frac{1}{s-1}$
$(s\bar{x} - x_0) + \bar{y} = \frac{1}{s+1}$

b) Insert the initial conditions $x_0 = 0$ and $y_0 = 0$
$s\bar{y} - \bar{x} = \frac{1}{s-1}$
$s\bar{x} + \bar{y} = \frac{1}{s+1}$

c) Eliminating $= \bar{y}$ we have
$s\bar{y} - \bar{x} = \frac{1}{s-1}$
$s\bar{y} + s^2\bar{x} = \frac{s}{s+1}$
$\Rightarrow\ \bar{x} = \frac{s^2-2s-1}{(s-1)(s+1)(s^2+1)} = -\frac{1}{2}.\frac{1}{s-1} - \frac{1}{2}.\frac{1}{s+1} + \frac{s}{s^2+1} + \frac{1}{s^2+1}$

d) $x = \frac{1}{2}e^t - \frac{1}{2}e^{-t} + \cos t + \sin t = \sin t + \cos t - \cosh t$
Eliminating $\bar{x}$ in (b) we have
$\Rightarrow\ \bar{y} = \frac{s^2+2s-1}{(s-1)(s+1)(s^2+1)} = \frac{1}{2}.\frac{1}{s-1} + \frac{1}{2}.\frac{1}{s+1} - \frac{s}{s^2+1} + \frac{1}{s^2+1}$
$\Rightarrow\ \ y = \frac{1}{2}e^t + \frac{1}{2}e^{-t} - \cos t + \sin t = \sin t - \cos t + \cosh t$
So the results are:
$x = \sin t + \cos t - \cosh t$
$y = \sin t - \cos t + \cosh t$

4.2 The Dirac Delta Function – the impulse function

It represents an extremely large force acting for a minutely small interval of time. Consider a single rectangular pulse of width b and height $\frac{1}{b}$ ocurring at $t = a$, as shown in Figs. 1(a) and (b).
If we reduce the width of the pulse to $\frac{b}{2}$ and keep the area of the pulse constant (1 unit) the height of the pulse will be $\frac{2}{b}$. If we continue reducing the width of the pulse while maintaining an area of unity, then as $b \to 0$, the height $\frac{1}{b} \to \infty$ and we have the Dirac delta function. It is denoted by $\delta(t-a)$.
Graphically it is represented by a rectangular pulse of zero width and infinite height, Fig. 2.

If the Dirac delta function is at the origin, $a = 0$ and so it is denoted by $\delta(t)$

4.2.1 Integration involving the impulse function

From the definition of $\delta(t-a)$

$$\int_p^q \delta(t-a)dt = 1 \quad \text{for} \begin{cases} \text{(i)} & p < t < a, \quad \delta(t-a) = 0 \\ \text{(ii)} & t = a \text{ area of pulse } = 1 \\ \text{(iii)} & a < t < q, \quad \delta(t-a) = 0 \end{cases}$$

Now consider $\int_p^q f(t)\delta(t-a)dt$ since $f(t)\delta(t-a)$ is zero for all values of t within the interval $[p,q]$ except at the point $t = a$, $f(t)$ may be regarded as a constant $f(a)$, so that $\int_p^q f(t)\delta(t-a)dt = f(a)\int_p^q \delta(t-a)dt = f(a)$

Examples

Evaluate $\int_1^3 (t^2+4)\cdot\delta(t-2)dt$. Here $a = 2$ $f(t) = t^2+4 \quad \Rightarrow f(a) = f(2) = 2^2+4 = 8$

Evaluate

1. $\int_0^6 5\cdot\delta(t-3)dt$
2. $\int_2^5 e^{-2t}\cdot\delta(t-4)dt$

Laplace transform of $\delta(t-a)$

Recall that $\int_p^q f(t)\cdot\delta(t-a)dt = f(a), \quad p < a < q$

$\Rightarrow$ if $p = 0$ and $q = \infty$ then $\int_0^\infty f(t)\cdot\delta(t-a)dt = f(a)$

Hence, if $f(t) = e^{-st}$, this becomes

$\int_0^\infty e^{-st}\cdot\delta(t-a)dt = \mathcal{L}\{\delta t - a\} = e^{-as}$

Similarly $\mathcal{L}\{f(t)\cdot\delta(t-a)\} = \int_0^\infty e^{-st}\cdot f(t)\cdot\delta t - adt = f(a)e^{-as}$

4.2.2 Differential equations involving the impulse function

Example

Solve the equation $\ddot{x} + 4\dot{x} + 13x = 2\delta(t)$ where, at $t = 0$, $x = 2$ and $\dot{x} = 0$

a) Expressing in LT, we have $(s^2\bar{x} - sx_0 - x_1) + 4(s\bar{x} - x_0) + 13\bar{x} = 2\times 1$

b) Inserting the initial conditions and simplifying we have $\bar{x} = \frac{2s+10}{s^2+4s+13}$

c) Rearranging the denominator by completing the square, this can be written as $\bar{x} = \frac{2(s+2)}{(s+2)^2+9} + \frac{6}{(s+2)^2+9}$

d) The inverse LT is $x = 2e^{-2t}\cos 3t + 2e^{-2t}\sin 3t = 2e^{-2t}(\cos 3t + \sin 3t)$

4.3 Fourier Series

You have seen from Maclaurin's and Taylor's series that an infinitely differentiable function can be expressed in the form of an infinite series in x. Fourier series on the other hand, enables us to represent a *periodic function* as an infinite trigonometrical series in sine and cosine terms. We can use Fourier series to represent a function containing discontinuities unlike Maclaurin's and Taylor's series.

Periodic function: A function $f(t)$ is periodic iff
$f(t) = f(t + nT)$, $n = 0, \pm 1, \pm 2 \ldots$
T is called the period. For sine and cosine the period $T = 2\pi$ so that $\sin t = \sin(t + 2\pi n)$ and $\cos t = \cos(t + 2\pi n)$

Analytical description of a periodic function

Many periodic functions are non-sinusoidal

Examples

1. $f(t) = 3 \quad 0 < t < 4$
   $f(t) = 0 \quad 4 < t < 6$
   $f(t) = f(t + 6)$ i.e. the period is 6

2. $f(t) = \frac{5}{8}t \quad 0 < t < 8$
   $f(t) = f(t + 8)$

Sketch the following periodic functions

1. $f(t) = 4 \quad 0 < t < 5$
   $f(t) = 0 \quad 5 < t < 8$
   $f(t) = f(t + 8)$

2. $f(t) = 3t - t^2 \quad 0 < t < 3$
   $f(t) = f(t + 3)$

4.3.1 Fourier series of functions of period 2π

Any periodic function $f(x) = f(x + 2\pi n)$ can be written in Fourier series as

$$\begin{aligned} f(x) &= \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \\ &= \frac{1}{2}a_0 + a_1 \cos x + a_2 \cos 2x + \ldots + b_1 \sin x + b_2 \sin 2x + \ldots \end{aligned}$$

(where $a_0, a_n, b_n,\ n = 1, 2, 3\ldots$ are Fourier coefficients) or as

$$f(x) = \frac{1}{2}a_0 + c_1 \sin(x + \alpha_1) + c_2 \sin(2x + \alpha_2) + \ldots$$

where $c_i = \sqrt{a_i^2 + b_i^2}$ and $\alpha_i = \arctan(\frac{b_i}{a_i})$.
$c_1 \sin(x + \alpha_1)$ is the first harmonic or fundamental
$c_2 \sin(2x + \alpha_2)$ is the second harmonic
$c_n \sin(nx + \alpha_n)$ is the n^{th} harmonic.
For the Fourier series to accurately represent $f(x)$ it should be such that if we put $x = x_1$ in the series the answer should be approximately equal to the value of $f(x_1)$ i.e. the value should converge to $f(x_1)$ as more and more terms of the series are evaluated. For this to happen $f(x)$ must satisfy the following **Dirichlet conditions**:

a) $f(x)$ must be defined and single-valued.

b) $f(x)$ must be continuous or have a finite number of discontinuities within a periodic interval.

c) $f(x)$ and $f'(x)$ must be piecewise continuous in the periodic interval.

If these conditions are met the series converges fairly quickly to $f(x_1)$ if $x = x_1$, and only the first few terms are required to give a good approximation of the function $f(x)$
Fourier coefficients: The Fourier coefficients above are given by
$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx$
$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx$
$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx$
Odd and even functions
a) *Even functions*: A function $f(x)$ is said to be even if $f(-x) = f(x)$. The graph of an even function is, therefore, *symmetrical about the y-axis*. e.g.

$f(x) = x^2$ $f(x) = \cos x$

b) *Odd functions*: A function $f(x)$ is said to be odd if $f(-x) = -f(x)$; the graph of an odd function is thus symmetrical about the origin. e.g.
$f(x) = x^3$ $f(x) = \sin x$

Products of odd and even functions

(even)×(even)=even
(odd)×(odd)=even
(neither)×(odd)=neither
(neither)×(even)=neither

Theorem 1: If $f(x)$ is defined over the internal $-\pi < x < \pi$ and $f(x)$ is even, then the Fourier series for $f(x)$ contains cosine terms only. Here a_0 is included.

<u>*Example*</u>

$$\begin{aligned} f(x) &= 0 \quad -\pi < x < -\tfrac{\pi}{2} \\ f(x) &= 4 \quad -\tfrac{\pi}{2} < x < \tfrac{\pi}{2} \\ f(x) &= 0 \quad \tfrac{\pi}{2} < x < \pi \\ f(x) &= f(x + 2\pi) \end{aligned}$$

The waveform is symmetrical about the y-axis, therefore, it is even.

$\Rightarrow \quad f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx$

a) $a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{2}{\pi}\int_0^{\pi} f(x)dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} 4dx = \frac{2}{\pi}[4x]_0^{\frac{\pi}{2}} = 4$

b) $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx = \frac{2}{\pi}\int_0^{\pi} f(x)\cos nx dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} 4\cos nx dx = \frac{8}{\pi}\left[\frac{\sin nx}{n}\right]_0^{\frac{\pi}{2}} = \frac{8}{\pi n}\sin\frac{n\pi}{2}$

But $\sin\frac{n\pi}{2} = \begin{cases} 0 \text{ for } n \text{ even} \\ 1 \text{ for } n = 1, 5, 9 \dots \\ -1 \text{ for } n = 3, 7, 11 \dots \end{cases}$

$$\Rightarrow \quad f(x) = 2 + \frac{8}{\pi}\left\{\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \frac{1}{7}\cos 7x + \dots\right\}$$

Theorem 2: If $f(x)$ is an odd function defined over the interval $-\pi < x < \pi$, then the Fourier series for $f(x)$ contains sine terms only. Here $a_0 = a_n = 0$

Example

$$f(x) = -6 \quad -\pi < x < 0$$
$$f(x) = 6 \quad 0 < x < \pi$$
$$f(x) = f(x + 2\pi)$$

This is an odd function so $f(x)$ contains only the sine terms
i.e. $f(x) = \sum_{n=1}^{\infty} b_n \sin nx$
and $b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx$
$f(x) \sin nx$ is even since it is a product of two odd functions.
$\Rightarrow \; b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx dx = \frac{2}{\pi} \int_0^{\pi} 6 \sin nx dx = \frac{12}{\pi} \left[\frac{-cosnx}{n} \right]_0^{\pi} = \frac{12}{\pi n}(1 - \cos n\pi)$

$$f(x) = \frac{24}{\pi} \left\{ \sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \ldots \right\}$$

If $f(x)$ is neither even nor odd we must obtain expressions for a_0, a_n and b_n in full

Examples

Determine the Fourier series of the function shown.

$$f(x) = \frac{2x}{\pi} \quad 0 < x < \pi$$
$$f(x) = 2 \quad \pi < x < 2\pi$$
$$f(x) = f(x + 2\pi)$$

This is neither odd nor even,

$$\Rightarrow \quad f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \{a_n \cos nx + b_n \sin nx\}$$

a) $a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx = \frac{1}{\pi} \left\{ \int_0^{\pi} \frac{2}{\pi} x dx + \int_{\pi}^{2\pi} 2 dx \right\}$
$= \frac{1}{\pi} \left\{ \left[\frac{x^2}{\pi} \right]_0^{\pi} + [2x]_{\pi}^{2\pi} \right\} = \frac{1}{\pi} \{\pi + 4\pi - 2\pi\} = 3$
$\Rightarrow \quad a_0 = 3$

b) $a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx dx$
$= \frac{1}{\pi} \left\{ \int_0^{\pi} \left(\frac{2x}{\pi} \right) \cos nx dx + \int_{\pi}^{2\pi} 2 \cos nx dx \right\}$
$= \frac{2}{\pi} \left\{ \frac{1}{\pi} \left[\frac{x \sin nx}{n} \right]_0^{\pi} - \frac{1}{\pi n} \int_0^{\pi} \sin nx dx + \int_{\pi}^{2\pi} \cos nx dx \right\}$
$= \frac{2}{\pi} \left\{ \frac{1}{\pi n} (\pi \sin n\pi x) + \frac{1}{\pi n} \left[\frac{\cos nx}{n} \right]_0^{\pi} + \left[\frac{\sin nx}{n} \right]_{\pi}^{2\pi} \right\}$
$= \frac{2}{\pi} \left\{ \frac{1}{n} \sin n\pi x + \frac{1}{\pi n^2} (\cos \pi nx - 1) + \frac{1}{n} (\sin 2\pi nx - \sin n\pi x) \right\}$
$= \frac{2}{\pi} \left\{ \frac{1}{\pi n^2} (\cos \pi nx - 1) + \frac{1}{n} \sin 2n\pi x \right\}$
$a_n = 0$ (n even); $a_n = \frac{-4}{\pi^2 n^2}$ (n odd)

c) $b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx dx$
$= \frac{1}{\pi} \left\{ \int_0^{\pi} \left(\frac{2x}{\pi} \right) \sin nx dx + \int_{\pi}^{2\pi} 2 \sin nx dx \right\}$

$=\frac{2}{\pi}\left\{\frac{1}{\pi}\left[\frac{-x\cos nx}{n}\right]_0^\pi - \frac{1}{\pi n}\int_0^\pi \cos nx dx + \int_\pi^{2\pi}\sin nx dx\right\}$
$=\frac{2}{\pi}\left\{\frac{1}{\pi n}(-\pi\cos n\pi x) + \frac{1}{\pi n x}\left[\frac{\sin nx}{n}\right]_0^\pi + \left[\frac{-\cos nx}{n}\right]_\pi^{2\pi}\right\}$
$=\frac{2}{\pi}\left\{-\frac{1}{n}\cos n\pi x + (0-0) - \frac{1}{n}(\cos 2\pi n x - \cos n\pi x)\right\}$
$=\frac{2}{\pi}\left\{-\frac{1}{n}\cos 2n\pi x\right\} = -\frac{2}{\pi n}\cos 2n\pi x$
But $\cos 2n\pi = 1 \quad \Rightarrow \; b_n = -\frac{2}{\pi n}$

$$f(x) = \frac{3}{2} - \frac{4}{\pi^2}\left\{\cos x + \frac{1}{9}\cos 3x + \frac{1}{25}\cos 5x + \ldots\right\}$$
$$-\frac{2}{\pi}\left\{\sin x + \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x + \frac{1}{4}\sin 4x \ldots\right\}$$

4.3.2 Half-range series

Sometimes a function of period 2π is defined over the range 0 to π instead of the normal $-\pi$ to π, or 0 to 2π. In this case one can choose to obtain a half-range cosine series by assuming that the function is part of an even function or a sine series by assuming that the function is part of an odd function.
Example

$f(x) = 2x \quad 0 < x < \pi$
$f(x) = f(x + 2\pi)$
from fig. 1

To obtain a half-range cosine series we assume an even function as in fig. 2.
$a_0 = \frac{2}{\pi}\int_0^\pi f(x)dx = \frac{2}{\pi}\int_0^\pi = \frac{2}{\pi}\left[x^2\right]_0^\pi = 2\pi$
$a_n = \frac{2}{\pi}\int_0^\pi 2x\cos nx dx = \frac{4}{\pi}\left\{\left[\frac{x\sin nx}{n}\right]_0^\pi - \frac{1}{n}\int_0^\pi \sin nx dx\right\}$
Simplifying, $a_n = 0$ for n even and $a_n = \frac{-8}{\pi n^2}$ for n odd. In this $b_0 = 0$ and so $f(x) = \pi - \frac{8}{\pi}\left\{\cos x + \frac{1}{9}\cos 3x + \frac{1}{25}\cos 5x + \ldots\right\}$
Obtain a half-range sine series for $f(x)$.

4.3.3 Functions with arbitrary period T

i.e. $f(t) = f(t+T)$, frequency $f = \frac{1}{T}$ and angular frequency $\omega = 2\pi f$ $\Rightarrow \omega = \frac{2\pi}{T}$ and $T = \frac{2\pi}{\omega}$. The angle $x = \omega t$ and the Fourier series is $f(t) = \frac{1}{2}a_0 + \sum_{n=1}^\infty\{a_n\cos n\omega t + b_n \sin n\omega t\}$

$= \frac{1}{2}a_0 + \Sigma_{n=1}^{\infty}\{a_n \cos\frac{2\pi nt}{T} + b_n \sin\frac{2\pi nt}{T}\}$
where
$a_0 = \frac{2}{T}\int_0^T f(t)dt = \frac{\omega}{\pi}\int_0^{\frac{2\pi}{\omega}} f(t)dt$
$a_n = \frac{2}{T}\int_0^T f(t)\cos n\omega t dt = \frac{\omega}{\pi}\int_0^{\frac{2\pi}{\omega}} f(t)\cos n\omega t dt$
$b_n = \frac{2}{T}\int_0^T f(t)\sin n\omega t dt = \frac{\omega}{\pi}\int_0^{\frac{2\pi}{\omega}} f(t)\sin n\omega t dt$
Example
Determine the Fourier series for a periodic function defined by
$f(t) = 2(1+t) \quad -1 < t < 0$
$f(t) = 0 \quad 0 < t < 1$
$f(t) = f(t+2) \quad 0 < t < 1$
Answer:

$$\begin{aligned} f(t) &= \frac{1}{2} + \frac{4}{\omega^2}\left\{\cos\omega t + \frac{1}{9}\cos 3\omega t + \frac{1}{25}\cos 5\omega t \ldots\right\} \\ &\quad -\frac{2}{\omega}\left\{\sin\omega t + \frac{1}{2}\sin 2\omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{4}\sin 4\omega t\right\} \end{aligned}$$

4.3.4 Sum of a Fourier series at a point of finite discontinuity

At $x = x_1$ the series converges to the value $f(x_1)$ as the number of terms included increases to infinity. But if there is a 'jump' at x_1 as shown in Fig.
$f(x_1 - 0) = y_1$ (approaching x_1 from below) $f(x_1 + 0) = y_2$ (approaching x_1 from above)
If we sustitute $x = x_1$ in the Fourier series for $f(x)$, it can be shown that the series converges to the value $\frac{1}{2}\{f(x_1 - 0) + f(x_1 + 0)\}$ i.e. $\frac{1}{2}(y_1 + y_2)$, the average of y_1 and y_2.

4.4 Fourier Integrals

4.4.1 The Fourier integral

: While Fourier series is for periodic functions Fourier integral is for non-periodic function. If a non-periodic $f(x)$ (i) satisfies the Dirichlet conditions in every finite interval $(-a, a)$ and (ii) is absolutely integrable in$(-\infty, \infty)$, i.e. $\int_{-\infty}^{\infty}|f(x)|dx$ converges, then $f(x)$ can be represented by a Fourier's integral as follows:

$$f(x) = \int_0^\infty \{A(k)\cos kx + B(k)\sin kx\}dx \qquad (1)$$

where $A(k) = \int_{-\infty}^{\infty} f(x)\cos kx dx$ (2)

$B(k) = \int_{-\infty}^{\infty} f(x)\sin kx\}dx$ (3)

If x is a point of discontinuity, then $f(x)$ must be replaced by $\left(\frac{f(x+0)+f(x-0)}{2}\right)$ as in the case of Fourier series. This can, in other words, be expressed by the following theorem.

Theorem 1: If $f(x)$ is piecewise continuous in every finite interval and has a right-hand derivative and a left-hand derivative at every point and if $\int_{-\infty}^{\infty} |f(x)|dx$ exists, then $f(x)$ can be represented by a Fourier integral. At a point where $f(x)$ is discontinuous the value of the Fourier integral equals the average of the left- and right-hand limits of $f(x)$ at that point.

Examples

Find the Fourier integral representation of the function in fig. below

$$f(x) = \begin{cases} 1 \text{ if } |x| < 1 \\ 0 \text{ if } |x| > 1 \end{cases}$$

Solution: From (2) and (3) we have

$A(k) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\cos kx dx = \frac{1}{\pi}\int_{-1}^{1} f(x)\cos kx dx = \left.\frac{\sin kx}{\pi k}\right|_{-1}^{1} = \frac{\sin k}{\pi k}$

$B(k) = \frac{1}{\pi}\int_{-1}^{1} \sin kx dx = 0$

and (1) gives the answer

$$f(x) = \frac{2}{\pi}\int_0^\infty \frac{\cos kx \sin k}{k} dk \qquad (4)$$

The average of the left- right-hand limits of $f(x)$ at $x = 1$ is equal to $(1+0)/2$, that is, $1/2$.

Furthermore, from (4) and Theorem 1 we obtain

$$\int_0^\infty \frac{\cos kx \sin k}{k} = \frac{\pi}{2} f(x) = \begin{cases} \frac{\pi}{2} & \text{if } \ 0 \le x < 1 \\ \frac{\pi}{4} & \text{if } \ x = 1 \\ 0 & \text{if } \ x > 1 \end{cases}$$

This integral is called *Dirichlet's discontinuous factor*. If $x = 0$, then

$$\int_0^\infty \frac{\sin k}{k} dk = \frac{\pi}{2}.$$

This integral is the limit of the so-called *sine integral*

$$Si(u) = \int_0^u \frac{\sin k}{k} dk$$

as $u \to \infty$
In the case of a Fourier series the graghs of the partial sums are approximation curves of the periodic function represented by the series. Similarly, in the case of the Fourier integral, approximations are obtained by replacing ∞ by numbers a. Hence the integral

$$\int_0^a \frac{\cos kx \sin k}{k} dk$$

approximates the integral in (4) and therefore $f(x)$