MP-Lecs Finite Dimensional Vector Spaces

Lecture Notes Based on Mathematical Physics Courses

Offered at University of Hyderabad and IIT Bhubaneswar

A. K. Kapoor
email:akkhcu@gmail.com

(Ver 2 , September 19, 2015)

Contents

Preface

These lecture notes are based on Mathematical Physics courses taught at University of Hyderabad and IIT Bhubaneswar. The material included here can be covered in about twelve to fifteen lectures. The primary reference has been the book by Halmos [1].

A companion set of notes of solved examples and exercises is planned as a separate unit.

I would like to thank the students of Hyderabad University of different batches who attended my course in Mathematical Physics. In particular I thank the M.Sc. students of I.I.T. Bhubaneswar whose help has been crucial in finalizing the lecture notes and making it available to a larger audience.

Several important topics could not be covered due to constraints on number of lecture hours available. References for these have been provided in the end.

References

[1] Paul R. Halmos, *Finite Dimensional Vector Spaces*

Part I

Basics concepts

Lecture 1

Groups, Fields and Vector Spaces

Definition 1 *To every ordered pair* $\langle a, b\rangle$ *of elements of a set* $\mathcal{X}$ *a* **binary operation** *assigns an element, denoted by* $a*b$*, of the set* $\mathcal{X}$*. For a binary opeartion to be a valid one it must be defined for all pairs and the* $a*b$ *must belong to the set and the result of binary operation must be unique.*

Definition 2 *A* **group** *is a pair* $\langle \mathcal{G}, *\rangle$ *with a binary operation* $*$ *defined on a set* $\mathcal{G}$ *such that the following properties.*

(G-1) Associative property : $a*(b*c) = (a*b)*c \qquad \forall\ a, b, c \in \mathcal{G}$

(G-2) Existence of identity : $\exists$ an element $e \in \mathcal{G}$ such that

$$e*a = a*e = a \qquad \forall\ a \in \mathcal{G}.$$

(G-3) Existence of inverse : $\forall\ a \in \mathcal{G}$ there exists an element a' such that

$$a*a' = a'*a = e$$

EXAMPLES OF GROUPS

1. The set of all real numbers $\mathbb{R}$ forms a group with addition as the binary operation.
2. The set of all complex numbers $\mathbb{C}$ is a group with addition as group operation.
3. The set of all positive, non-zero, real numbers $\mathbb{R}^+$ is a group with respect to multiplication as group operation.
4. The set of all $N \times N$ real (or complex) matrices form a group under matrix addition.
5. The group of all $N \times N$ real (or complex) matrices with determinant $\neq 0$ form a group under the matrix multiplication.

§1 Fields

Definition 3 *A* **field** $\mathcal{F}$ *is a triple* $\langle \mathcal{F}, +, \cdot \rangle$*, where,* $\cdot$ *and* $+$ *are two binary operations defined on a set* $\mathcal{F}$ *such that the axioms (F-I) to (F-III), given below, are satisfied. The elements of the field will be called* **scalars** *and will be denoted by greek letters* $\alpha, \beta, \gamma, \ldots$.

(F-1) To every pair α, β the scalar $\alpha + \beta$ is called the **sum** of α, β which satsifies the following axioms $\forall \alpha, \beta, \gamma \in \mathcal{F}$

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

(ii) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

(iii) $\exists$ a unique scalar 0 such that $0 + \alpha = \alpha = \alpha + 0$

(iv) $\forall \alpha \in \mathcal{F}$ $\exists$ a unique scalar $(-\alpha) \in \mathcal{F}$ we have $\alpha + (-\alpha) = 0$

These properties imply that $\mathcal{F}$ is a group with $+$ as binary operation.

(F-2) The scalar $\alpha \cdot \beta$ will be called the **product** of α, β and has the following properties.

(i) Commutative Property : $\alpha \cdot \beta = \beta \cdot \alpha$

(ii) Associative Property : $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

(iii) Existence of multiplicative identity : $\exists$ a unique scalar 1 such that

$$\alpha.1 = 1.\alpha = \alpha$$

(iv) $\forall \alpha \neq 0$ $\exists$ a scalar denoted by α^{-1} such that $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$

(F-3) The sum and the product obey the distributive property :

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$$

EXAMPLES OF FIELDS

1. Set of all rational numbers $\mathbb{Q}$ is a field with usual addition and multiplication as the two binary operations.

2. Set of all real numbers $\mathbb{R}$ is a field with usual addition and multiplication as the two binary operations.

3. Set of all complex numbers $\mathbb{C}$ is a field with usual addition and multiplication as the two binary operations.

4. The set $\mathbb{Z}^+$ of all positive integers is not a field with the usual addition and multiplication as two binary operations (give all possible reasons).

5. The set $\mathbb{Z}$ of all integers is not a field with the usual addition and multiplication. (Give one reason).

§2 Vector Spaces

Definition 4 *Let $\mathcal{F}$ be a field and $+$ be a binary operation defined on a set $\mathcal{V}$. The triple $\langle \mathcal{V}, +, \mathbb{F} \rangle$ is a* **vector space** *on a field $\mathcal{F}$ if the following properties are satisfied.*

(V-1) To every pair of vectors $f, g \in \mathcal{V}$, there corresponds a vector $f + g \in \mathcal{V}$ called the sum of f and g such that

(i) $f + g = g + f \qquad \forall\ f, g \in \mathcal{V}$

(ii) $f + (g + h) = (f + g) + h \qquad \forall\ f, g, h \in \mathcal{V}$

(iii) $\exists$ a unique vector $0 \in \mathcal{V}$ such that

$$f + 0 = f \qquad \forall\ f \in \mathcal{V}$$

(iv) To every vector $f \in \mathcal{V}$, there corresponds a vector $-f \in \mathcal{V}$ such that

$$f + (-f) = 0$$

(V-2) $\forall\ \alpha \in \mathcal{F}$ and $f \in \mathcal{V}$ there corresponds a unique vector $\alpha f \in \mathcal{V}$ such that

$$\alpha(\beta f) = (\alpha\beta) f \qquad \forall\ \alpha, \beta \in \mathcal{F}$$

and

$$1.f = f \qquad \forall f \in \mathcal{V}$$

(V-3) $\forall \alpha, \beta \in \mathcal{F}$ and $\forall f, g \in \mathcal{V}$ we have

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

and

$$\alpha(f + g) = \alpha f + \alpha g$$

EXAMPLES OF VECTOR SPACES

(I)
1. Every field $\mathcal{F}$ is also a vector space over $\mathcal{F}$ as field of scalars. Thus we have the following important special examples of vector spaces.
2. Set of all complex numbers $\mathbb{C}$ is a complex vector space with $\mathbb{C}$ as the field of scalars.
3. Set of all real numbers $\mathbb{R}$ is a real vector space with $\mathbb{R}$ as the field of scalars.
4. Set of all rational numbers $\mathbb{Q}$ is a rational vector space with $\mathbb{Q}$ as the field of scalars.

(II) Set of all n-tuples $(\alpha_1, \alpha_2, ..., \alpha_n)$ where $\alpha_k \in \mathcal{F}$ is denoted by $\mathcal{F}^n$. This set is vector space with $\mathcal{F}$ as field of scalars. Thus

1. $\mathbb{C}^n$ is a complex vector space over $\mathbb{C}$ as the field of scalars.

2. $\mathbb{R}^n$ is a real vector space over $\mathbb{R}$ as the field of scalars.
3. $\mathbb{Q}^n$ is a rational vector space over $\mathbb{Q}$ as the field of scalars.

(III) 1. All polynomials in a variable t, with coefficients in any field $\mathcal{F}$ is vector space $\mathscr{P}$.

$$\mathscr{P} = \{p(t) | p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t + ... + \alpha_n t^n + ...\text{and } \alpha_j \in \mathcal{F}\}$$

Here $\mathcal{F}$ can be any of the fields such as $\mathbb{C}, \mathbb{R}, \mathbb{Q}, \ldots$

2. Consider the set $\mathscr{P}$ of all polynomials in a variable t, with coefficients in any field $\mathscr{F}$ and consider the subset $\mathscr{P}_N$ consisting of all polynomials of degree $\leq N$. Then $\mathscr{P}_N$ is a vector space.

(IV) 1. Let $\mathcal{F}$ be set of all functions defined on an interval $[a, b]$ and having complex values. With any one of the fields $\mathbb{C}, \mathbb{R}$, or $\mathscr{Q}$, $\mathcal{F}$ is a vector space.

2. Let $\mathcal{F}$ be as in (IV-1) and $\mathscr{C}^{(0)}$ be the subset of all continuous functions. Then $C^{(0)}$ is a vector space.

3. Let $\mathcal{F}$ be as in (IV-1) and $C^{(r)}$ be the subset of all functions for which $r-$ derivatives exist and are continuous on $[a, b]$. The $C^{(r)}$ is a vector space.

4. Let $\mathscr{C}^{(0)}$ be as in (IV-2). Let $\mathcal{S}$ be a subset of $\mathscr{C}^{(0)}$ consisting of those functions which vanish at a given point x_0.Then $\mathcal{S}$ is vector space. In general, if one can takes all functions which vanish at $x_1, x_2, \ldots, x_n$ then also we get a vector space.

(V) Let $\mathbb{M}_N$ be the set of all $N \times N$ matrices whose element are scalars from a field $\mathcal{F}$. With standard matrix addition as vector addition $\mathbb{M}_N$ is a vector space over the same field F

(VI) The set of all functions f on an interval $[a, b]$, for which $\int_a^b |f(x)|^p\,dx$ is finite, is a vector space denoted by $\mathscr{L}^p[a, b]$. That addition of two functions in $\mathscr{L}^p[a, b]$ gives back a function in the same space will not be proved here. The space $\mathscr{L}^p[a, b]$, for $p = 2$, is the set of all square integrable functions on the interval $[a, b]$.

(VII) The set of all infinite sequences $(\alpha_1, \alpha_2, \ldots, ..)$, such that the infinite series

$$\sum_{k=1}^{\infty} |\alpha_k|^p$$

converges, is a vector space denoted by ℓ^p. That the sum of two sequences, $\alpha, \beta \in \ell^p$ is also in ℓ^p, space requires a proof which will not be given here.

(VIII) A set $\{0\}$, consisting of only one element, the null vector, is a vector space over any field.

§3 Subspace

Definition 5 *Let $\mathcal{V}$ be a vector space over a field $\mathcal{F}$. Let $\mathscr{S}$ be a subset of $\mathcal{V}$. Let the vector addition in $\mathscr{S}$ be defined in the same way as in $\mathcal{V}$. If $\mathscr{S}$ is also vector space over the same field $\mathcal{F}$, we say that $\mathscr{S}$ is* **subspace** *of $\mathcal{V}$.*

EXAMPLES OF SUBSPACES

1. Every vector space $\mathcal{V}$ is subspace of itself.

2. The subset having only the null vector, 0, is a subspace of every vector space.

3. Let $\mathcal{V}_1$ be the vector space of complex numbers over the field of real numbers. Let $\mathcal{V}_2$ be the vector space of all real numbers with $\mathbb{R}$ as the field of scalars. The $\mathcal{V}_2$ is a subspace of $\mathcal{V}_1$.

4. The set $C^{(1)}$ of functions with continuous first derivative is a subspace of the vector space of all continuous functions with the same field of scalars.

5. Let $C^{(0)}$[a,b] be the set of all continuous complex valued functions on the interval $[a, b]$. This set is a vector space and we have

   (a) the subset consisting of of all functions which vanish at a given point x_0 is a subspace.

   (b) the subset of $C^{(0)}$ consisting of all functions having value 1/2 at a point x_0 is not a subspace.

   (c) The set of all solutions of a linear differential equation

   $$a_0(x)\frac{d^n y}{dx^n} + a_1(x)\frac{d^{n-1}y}{dx^{n-1}} + \ldots + y(x) = 0$$

   is a vector space.

6. Consider the set of all vectors in three dimensions, $\mathbb{R}^3$ which is real vector space. The subset S_1 of all vectors which are multiples of a fixed vector $\vec{A}$ and the subset S_2 of all vectors in a given fixed plane passing through the origin, and are two examples of subspaces of $\mathbb{R}^3$.

It is easy to see that intersection of two subspaces of a vector space is again a subspace.

Lecture 2

Linear Independence, Basis and Dimension

§1 Linear Independence

Definition 6 *A set of vectors* $\mathscr{S} = \{f_1, f_2, ...f_n\}$ *is called* **linearly dependent set** *if* $\exists$ *a set of scalars* $\alpha_1, \alpha_2, \ldots$ *such that not all* α*'s are zero and*

$$\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n = 0$$

Definition 7 *A set of vectors* $\mathscr{S} = \{f_1, f_2, ...f_n\}$ *is called* **linearly independent set** *if it is not a linearly dependent set. This means that a set* X *is linearly independent if*

$$\alpha_1 f_2 + \alpha_2 f_2 + \cdots + \alpha_n f_n = 0$$

implies $\alpha_1 = \alpha_2 = \cdots = 0$.

Definition 8 *Let* $\{f_1, f_2, ..., f_m\}$ *be a finite set of vectors in vector space* V*. Let* $\alpha_1, \alpha_2, \ldots, \alpha_m$ *be a set of scalars and* $f \in \mathcal{V}$ *be such that*

$$f = \alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_m f_m$$

Then we say that f *is* **linear combination** *of the vectors* $f_1, f_2, \ldots, f_m$.

Properties Of Linear Combination

1. If $f \in \mathcal{V}$ is a linear combination of $\{f_1, f_2, \ldots\}$, then the scalars α_i in
$$f = \sum \alpha_i f_i$$
are uniquely determined if and only if $\{f_1, f_2, \ldots\}$ is a independent set.
2. If $\{f_i\}$ is a linearly independent set, a necessary and sufficient condition that $f \in \mathcal{V}$ be a linear combination of $\{f_i\}$ is that the set $\{f, f_i\}$ be linearly dependent.
3. Every set of vectors containing a linearly dependent set is also linearly dependent.

Definition 9 *A vector space is called* **finite dimensional** *if* $\exists$ *an integer* N *such that every set containing more than* N *elements is a linearly dependent set.*

§2 Basis

Definition 10 *A set of vectors $\mathscr{X}$ is called a* **basis** *in a vector space $\mathcal{V}$ if the following two properties are satisfied.*

- the set $\mathscr{X}$ is a linearly independent set, and
- every vector $f \in \mathcal{V}$ is a linear combination of vectors in $\mathscr{X}$, i.e.,

$$f = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$$

  where $x_k \in \mathscr{X}$ for all $k = 1, 2, \ldots, n$.

Examples Of Basis

1. Vectors $\{\vec{i}, \vec{j}, \vec{k}\}$ form a basi s for the set of all vectors in three dimension.
2. Any three vectors
3. Any three vectors which are not coplanar form a basis in the space of vectors in three dimension.
4. $\{1, x, x^2, ..., x^N\}$ is a basis in the space of all polynomials of degree N.
5. The set $\bigcup_n \{\cos nx, \sin nx\}$, where $n = 1, 2, 3, ...$, is a basis in space of all periodic functions on $[-\pi, \pi]$ with period 2π.
6. The vectors $\mathscr{E} = e_1, e_2, ..., e_N$ where
$$e_1 = (1, 0, 0, ..., 0)\ ; e_2 = (0, 1, 0, ...0)\ ; ....\ e_N = (0, 0, 0, ..., 1)$$
form a basis in the vector space $\mathbb{C}^N$. This basis will be called the canonical basis or the standard basis.
7. The vectors $\mathscr{E} = \{e_1, e_2, ..., e_N\}$ also form a basis in R^N and in $\mathbb{Q}^N$.

Theorem)‖(**1 (Number of Elements in a Basis)** *The number of elements in any one basis is equal to number of elements in every other basis.*

Definition 11 *For a finite dimensional space the number of elements in a basis is defined to be the* **dimension** *of the vector space.*

Summary Of Properties Of Bases

Given that a vector space $\mathcal{V}$ has dimension N we have the following properties.

1. Every set containing $N+1$ or more vectors is a linearly dependent set.

2. A set of N vectors is a basis if and only if it is linearly independent.

3. A set of N vectors $\mathscr{X}$ is a basis iff every vector in $\mathcal{V}$ is linear combination of vectors in the set

   $\mathscr{X}$.

Definition 12 *Let $S = \{f_1, f_2, \ldots, f_m\}$ be subset of a vector space. The* **linear span** *of S is the set of all vectors f such that f is linear combination of vectors $f_1, f_2, \ldots, f_m \in S$. Linear Span of $S = \{f | f = \sum_{k=1}^{m} \alpha_k f_k\}$ and $f_k \in V$ and $\alpha_k \in V\}$*

§3 Linear Functional

Definition 13 *A* **linear functional** *on a vector space is a mapping from the vector space to the field of scalars :*

$$\Psi : f \mapsto \Psi(f) \in F$$

such that Ψ is linear:

$$\Psi(\alpha f + \beta g) = \alpha \Psi(f) + \beta \Psi(g)$$

Definition 14 *The set of all linear functional on a vector space $\mathcal{V}$ forms a vector space by itself if the addition of two functionals, ϕ and ψ, is defined by*

$$(\psi + \phi)(f) = \psi(f) + \phi(g)$$

This vector space is called **vector space dual** *to $\mathcal{V}$ and is denoted by $\tilde{\mathcal{V}}$.*

Examples of Functionals

1. The functional which assigns $0 \in \mathcal{F}$ to every vector is a linear functional.

2. In $\mathscr{C}^n$ let $x = (\xi_1, \xi_2, ..., \xi_n)$ and $\Psi(x)$ defined by

   $$\Psi(x) = \alpha_1 \xi_1 + \alpha_2 \xi_2 + \ldots + \alpha_n \xi_n$$

   is a linear functional where $\alpha_k \in \mathscr{C}$.

3. In the function space $\mathscr{L}^2[a,b]$ given $f L^2[a,b]$ define a functional Ψ by

   $$\Psi(f) = \int_a^b g^*(x) f(x) dx,$$

   for a fixed $g \in \mathscr{L}^2[a,b]$, then Ψ is linear functional.

4. For $f \in R^n$ the functional Φ_1 and Φ_2 defined below are not linear functional. Let $f = (\alpha_1, \alpha_2, ..., \alpha_n)$ and

$$\Phi_1(f) = \sum_k |\alpha_k|; \qquad \Phi_2(f) = \sum_k |\alpha_k|^2$$

Lecture 3

Linear Functionals

§1 Linear Functional

Definition 15 *A* **linear functional** *on a vector space is a mapping from the vector space to the field of scalars :*

$$\Psi : f \mapsto \Psi(f) \in \mathscr{F}$$

such that Ψ is linear:

$$\Psi(\alpha f + \beta g) = \alpha\Psi(f) + \beta\Psi(g)$$

Examples Of Functionals

1. The functional which assigns $0 \in \mathcal{F}$ to every vector is a linear functional.

2. In $\mathscr{C}^n$ let $x = (\xi_1, \xi_2, ..., \xi_n)$ and $\Psi(x)$ defined by

   $$\Psi(x) = \alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_n\xi_n$$

   is a linear functional where $\alpha_k \in \mathscr{C}$.

3. In the function space $\mathscr{L}^2[a, b]$ given $f\mathscr{L}^2[a, b]$ define a functional Ψ by

   $$\Psi(f) = \int_a^b g^*(x)f(x)dx,$$

   for a fixed $g \in \mathscr{L}^2[a, b]$, then Ψ is linear functional.

4. For $f \in R^n$ the functional Φ_1 and Φ_2 defined below are not linear functionals. Let $f = (\alpha_1, \alpha_2, ..., \alpha_n)$ and

   $$\Phi_1(f) = \sum_k |\alpha_k|; \qquad \Phi_2(f) = \sum_k |\alpha_k|^2$$

5. Given a vector x_0, we define a functional Ψ_0 by

   $$\Psi_{x_0}(f) = \begin{cases} 1 & \text{if } f = x_0 \\ 0 & \text{if } f \neq x_0 \end{cases} \tag{3.1}$$

   It is easy to check that the functional $\Psi(x_0)$ is a linear functional.

6. In the vector space, $\mathbb{R}^3$, of all real vectors in 3 dimension a linear functional can be defined as follows.

   Choose a vector $\vec{X} \in \mathbb{R}^3$ and define a functional Ψ_X by

$$\Psi_X(\vec{A}) = \vec{X} \cdot \vec{A}, \qquad \forall \vec{A} \in \mathbb{R}^3$$

Equality of two functionals: Two linear functionals Ψ and Φ on a vector space $\mathscr{V}$ are said to be equal if $\Psi(f) = \Phi(f), \quad \forall f \in \mathscr{V}$.

§2 Dual Vector Space

We will now define twin operations involving linear functionals and scalars
(i) addition of two linear functionals and
(ii) multiplication of a linear functional by a scalars

Given two arbitrary linear functionals Ψ and Φ, their sum $\Psi + \Phi$ is defined by giving its action on an arbitrary vector $f \in \mathscr{V}$ as

$$(\Psi + \Phi)f = \Psi(f) + \Phi(f).$$

The sum of two linear functionals is again a linear functional. Given a scalar α and a linear functional Ψ, multiplication of linear functional by scalar α, $(\alpha\Psi)$ is defined by

$$(\alpha\Psi)(f) = \alpha\Psi(f)$$

and the product $\alpha\Psi$ is again a linear functional. Then we have the following result:

Theorem)|(**2** *The set of all linear functional on a vector space $\mathcal{V}$ forms a vector space. This vector space is called* **vector space dual** *to $\mathcal{V}$ and is denoted by $\widetilde{\mathcal{V}}$.*

The proof is easy. We need to verify that sum of two linear functionals and multiplication of a linear functional by a scalar result in linear functionals. The proof is obvious. Still let us write it down.

Checking Linearity of $\Psi_1 + \Psi_2$: Let Ψ_1, Ψ_2 be two linear functionals. Consider $\Phi = \alpha_1\Psi_1 + \alpha_2\Psi_2$, then for all $f \in \mathscr{V}$

$$(\alpha_1\Psi_1 + \alpha_2\Psi_2)f = (\alpha_1\Psi_1)(f) + (\alpha_2\Psi_2)(f) \tag{3.2}$$

$$= \alpha_1\Psi_1(f) + \alpha_2\Psi_2(f). \tag{3.3}$$

It can be proved that the dimension of a vector space dual to $\mathscr{V}$ is equal to the dimension of the vector $\mathscr{V}$ space itself. Remember that every vector space has a basis. So we can ask for a basis for the dual vector space. A useful construction of a basis in the dual space starts with a basis $\mathscr{B}$ in the vector space, and the basis obtained will be called basis dual to $\mathscr{B}$.

Definition 16 Dual Basis: *Let* $\mathscr{B} = x_1, x_2, \ldots, x_N \subset \mathscr{V}$ *be a basis. Let linear functionals* $\Psi_1, \Psi_2, \ldots$ *be defined, (as in Eq.(3.1)), by*

$$\Psi_1(f) = \begin{cases} 1 & \text{if } f = x_1 \\ 0 & \text{if } f \neq x_1 \end{cases} \tag{3.4}$$

$$\Psi_2(f) = \begin{cases} 1 & \text{if } f = x_2 \\ 0 & \text{if } f \neq x_2 \end{cases} \tag{3.5}$$

etc. In general, we have

$$\Psi_k(f) = \begin{cases} 1 & \text{if } f = x_k \\ 0 & \text{if } f \neq x_k \end{cases} \tag{3.6}$$

where $k = 1, 2, \cdots, N$*. Then you can check that* Ψ_k *is a basis in the dual vector space. It is called basis dual to the chosen basis* $\mathscr{B}$*.*

The definition of dual basis can be summarized as in the table given below.

	x_1	x_2	...	x_k	...	x_N
Ψ_1	1	0	...	0	...	0
Ψ_2	0	1	...	0	...	0
...	0	0	...	0	...	0
Ψ_k	0	0	...	1	...	0
...	0	0	...	0	...	0
Ψ_N	0	0	...	0	...	1

(3.7)

Prove it now Show that the dual basis as defined as above is in fact a basis by proving the two properties, linear independence and spanning the whole space, that a basis must have.

)|(**Short Examples 1** (Dual Basis) *In the vector space* $\mathbb{R}^3$*, given a basis* $\{\vec{A}, \vec{B}, \vec{C}\}$ *the set of vectors*

$$\vec{a} = \frac{\vec{b} \times \vec{c}}{|\vec{b} \times \vec{c}|}, \qquad \vec{b} = \frac{\vec{c} \times \vec{a}}{|\vec{c} \times \vec{a}|}, \qquad \vec{c} = \frac{\vec{a} \times \vec{b}}{|\vec{a} \times \vec{b}|}. \tag{3.8}$$

define linear functionals in the sense of <Example 5> on page 11.

Remark: The dual of dual of vector space $\mathscr{V}$ is the vector space $\mathscr{V}$ itself.

vs-dial-03001

Question: *Give examples of functionals on $\mathbb{R}^3$. This means given a vector $\vec{A} = (A_1, A_2, A_3)$, give a rule for constructing a real number from the three components*

(i) $\Psi_1(\vec{A}) = A_1 + A_2 + A_3$

(ii) $\Psi_2(\vec{A}) = A_1\, A_2/A_3$

(iii) $\Psi_3(\vec{A}) = A_1\, A_2 + A_2\, A_3 + A_3\, A_1$

(iv) $\Psi_4(\vec{A}) = A_1\, A_2\, A_3$

(v) $\Psi_5(\vec{A}) = \log(A_1 + A_2 + A_3)$

(vi) $\Psi_6(\vec{A}) = \sqrt{A_1^2 + A_2^2 + A_3^2}$

Question: *Which of the functionals in the above example are linear and which are not linear?*

(i) Ψ_2 and Ψ_5 are not functionals because they are not defined for all vectors. For example Ψ_2 is not defined for vectors which have zero third component.
Ψ_4 is not defined if $(A_1 + A_2 + A_3)$ is zero negative.

(ii) $\Psi_1(\vec{A})$ is a linear functional. This can be seen as follows.

> **Thinking:** In order to prove that $\Psi_1(\vec{A})$ is a linear I must show that
>
> $$\Psi_1(\alpha\vec{A} + \beta\vec{B}) = \alpha\Psi_1(\vec{A}) + \beta\Psi_1(\vec{B})$$

Let us assume that $\vec{A} = (A_1, A_2, A_3), \vec{B} = (B_1, B_2, B_3)$, then

$$\alpha\vec{A} + \beta\vec{B} = (\alpha A_1 + \beta B_1, \alpha A_2 + \beta B_2, \alpha A_3 + \beta B_3)$$

Now calculate

$$\begin{aligned}
\Psi_1(\vec{A}) &= A_1 + A_2 + A_3 && (3.9)\\
\Psi_1(\vec{B}) &= B_1 + B_2 + B_3 && (3.10)\\
\Psi_1(\alpha\vec{A} + \beta\vec{B}) &= (\alpha A_1 + \beta B_1) + (\alpha A_2 + \beta B_2) + (\alpha A_3 + \beta B_3) && (3.11)\\
&= \alpha(A_1 + A_2 + A_3) + \beta(B_1 + B_2 + B_3) && (3.12)\\
\therefore\quad \Psi_1(\alpha\vec{A} + \beta\vec{B}) &= \alpha\Psi_1(\vec{A}) + \beta\Psi_1(\vec{B}) && (3.13)
\end{aligned}$$

(iii) You can convince yourself that Ψ_2, Ψ_4, Ψ_6 are not linear functionals.

Lecture 4

Linear Operators-I

§1 Vector Space of Linear Operators

Definition 17 *An* **operator***, T, on a vector space $\mathcal{V}$ is a mapping*

$$T : \mathcal{V} \to \mathcal{V}$$

from the vector space $\mathcal{V}$ into itself. In other words, to an arbitrary vector f from the vector space, an operator, T, assigns a unique vector, Tf, in the vector space $\mathcal{V}$.

$$T : f \mapsto Tf \in \mathcal{V}$$

Definition 18 *An operator, T on a vector space is a* **linear operator** *if it satisfies the property*

$$T(\alpha + \beta g) = \alpha\ T + \beta\ Tg$$

$\forall \alpha, \beta \in \mathcal{F}$ and $\forall\ g \in \mathcal{V}$. Equivalently an operator T is linear if

$$T(f + g) = Tf + Tg \text{ and } T(\alpha f) = \alpha Tf$$

It is, therefore, seen that for an operator T to be linear it is necessary that $Tf = 0$ if $f = 0$.

Definition 19 *Given two linear operators A and B we can define their* **sum** *, $A+B$, by means of the following rule for its action on an arbitrary vector.*

$$(A + B)f = Af + Bf$$

The sum of two linear operators is again a linear operator.

Definition 20 **Multiplication of a linear operator** *T by a scalar α is a linear operator defined by*

$$(\alpha T)f = \alpha(Tf)$$

Theorem **3** *With addition of linear operators and scalar multiplication defined as above, the set of all linear operators on a vector space $\mathcal{V}$ is again a vector space. If the dimension of the vector space $\mathcal{V}$ is n, the dimension of the vector space of all the operators on$\mathcal{V}$ is n^2.*

Definition 21 *With sum, product and multiplication by a scalar defined for operators, the following expression defines* **polynomial** *in a linear operator A.*

$$p(A) = \alpha_0 I + \alpha_1 A + \alpha_2 A^2 + +\alpha_n A^n$$

The operator $p(A)$ is again a linear operator.

§2 Product and Commutator

Definition 22 **Product of two operators**, *A and B, is defined as in the case of mappings.*

$$(AB)f = A(Bf)$$

When A and B are linear operators, the product AB is also a linear operator.

Definition 23 *The* **commutator** *of two operators is defined to be*

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

Definition 24 *The* **anticommutator** *is defined by*

$$[A, B]_+ = AB + BA$$

PROPERTIES OF COMMUTATOR

The commutator satisfies the following properties.

$$[A, B] = -[B, A] \tag{4.1}$$

$$[\alpha_1 A_1 + \alpha_2 A_2, B] = \alpha_1 [A_1, B] + \alpha_2 [A_2, B] \tag{4.2}$$

$$[A, \beta_1 B_1 + \beta_2 B_2] = \beta_1 [A, B_1] + \beta_2 [A, B_2] \tag{4.3}$$

$$[A, BC] = B[A, C] + [A, B]C \tag{4.4}$$

$$[AB, C] = A[B, C] + [A, C]B \tag{4.5}$$

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{4.6}$$

The last relation is known as the Jacobi identity.

§3 Inverse of an Operator

Definition 25 *Let T be an operator on a vector space. We say T is* **one to one** *if action of T on two* distinct vectors *gives* distinct answers.

$$x_1 \neq x_2 \implies Tx_1 \neq Tx_2$$

This is equivalent to the condition

$$Tx_1 = Tx_2 \Longrightarrow x_1 = x_2$$

Definition 26 *An operator T on a vector space is called* **onto** *if $\forall\ y \in \mathcal{V}$ we can find at least one $x \in \mathcal{V}$ such that $Tx = y$. This x may, in general, not be unique.*

Definition 27 *An operator is called* **invertible** *if it is both one to one and onto.*

Definition 28 *Let T be an operator which is both one to one and onto. We define* **inverse** *of T by giving its action on an arbitrary vector $u \in \mathcal{V}$.*

Because T is onto, we can find a vector u such that $Tu = v$. Since T is one to one it follows that u satisfying $Tu = v$ is uniquely determined once the vector v is specified. We define **inverse** of T, to be denoted by T^{-1}, by the equation

$$T^{-1}v = u$$

This definition coincides with the defintion of the inverse for a mapping. The inverse satisfies

$$(AB)^{-1} = B^{-1}A^{-1},$$

$$(\alpha A)^{-1} = (1/\alpha)A^{-1}, \alpha \neq 0.$$

Definition 29 *Let T be a linear operator on a vector space $\mathcal{V}$. The* **range** *$\mathscr{R}(T)$ is the set of vectors obtained by applying T on all vectors $f \in \mathcal{V}$.*

$$\mathscr{R}(T) = \{g | g = Tf, \in \mathcal{V}\}$$

Definition 30 *Also the* **null space** *of an operator, $\mathscr{N}(T)$, is the set of all those vectors x for which $Tx = 0$.*

$$\mathscr{N}(\ T) = \{x | x \in \mathcal{V} \text{ and } Tx = 0\}$$

Both $\mathscr{R}(T)$ and $\mathscr{N}(T)$ are subspaces of the vector space $\mathcal{V}$. (Proof ?!)

Definition 31 *The dimension of $\mathscr{R}(T)$ for an operator is called the rank of the operator T. Obviously rank $(T) \leq \dim \mathcal{V}$.*

Clearly rank of an operator is the maximum number of linearly independent vectors that can be selected from Tf when varies over the entire vector space $\mathcal{V}$.

§4 Eigen-values and Eigen-vectors

Definition 32 *A subspace $\mathscr{M} \subset \mathcal{V}$ is said to be an* **invariant subspace** *of an linear operator X if $\forall f \in \mathscr{M}\, Xf \in \mathscr{M}$.*

Definition 33 *Let T be linear operator. If f is a non-zero vector satisfying*

$$Tf = \lambda f$$

or some scalar λ , we say that f is an **eigen-vector** *of operator T and λ is the corresponding* **eigen-value.**

Note that $f = 0$ will always satisfy the equation $Tf = \lambda f$ for an arbitrary λ. Therefore, null vector is, by definition, excluded from being an eigen-vector.

It is possible that for a given λ there are more than one eigen-vectors satisfying the eigen-value equation $Tf = \lambda$. Therefore, we define

Definition 34 *Let λ be an eigen-value of an operator T. Let $\nu(\lambda)$ denote the number of linearly independent eigen-vectors $Tx = \lambda x$. If $\nu(\lambda) = 1$ we say that the eigen-value λ is* **non-degenerate**. *When $\nu(\lambda) > 1$, we say that the eigen-value λ is* **degenerate** *and the* **degeneracy** *of the eigen-value λ is defined to be equal to the number of linearly independent eigen-vectors with eigen-value λ.*

Lecture 5

Linear Operators-II

§1 Properties of Operators

An operator is defined to be invertible if it is one to one and onto In case of vector spaces of finite dimension either one of these two properties is sufficient for a linear operator to be invertible. In this lecture this and other issues related to existence of inverse are discussed.

Theorem)‖(**4** *Let T be a linear operator on a finite dimensional vector space $\mathcal{V}$ $(\dim V = N)$. Then the following statements are equivalent.*

(L1) $Tx = 0$ implies $x = 0$.

(L2) T is one to one.

(L3) If $\mathscr{X} = \{x_1, x_2, \ldots\}$ is a lineraly independent set then $T\mathscr{X} \equiv \{Tx_1, Tx_2, \ldots\}$ is also linearly independent set.

(L4) If $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ is a basis then $T\mathscr{X} \equiv \{Tx_1, Tx_2, \ldots, Tx_N\}$ is also a basis.

(L5) T is an onto operator.

(L6) Let $\mathscr{X} = \{x_1, x_2, \ldots, \}$ and $\mathscr{Y} = \{y_1, y_2, \ldots, \}$ be sets of vectors such that $y_j = Tx_j$. If $\mathscr{Y}$ is a linearly independent set of vectors, then $\mathscr{X}$ is also linearly independent.

(L7) If $\mathscr{X}$ and $\mathscr{Y}$ are as in (6) above and if $\mathscr{Y}$ is a basis, then $\mathscr{X}$ is also a basis.

Proof

: We shall prove that
$(1) \Longrightarrow (2) \Longrightarrow (3) \Longrightarrow (4) \Longrightarrow (5) \Longrightarrow (6) \Longrightarrow (7) \Longrightarrow (1)$.

Proof of $(L1) \Longrightarrow (L2)$ We are given (L1) :$Tx = 0 \Longrightarrow x = 0$. Consider $Tx_1 = Tx_2$; Then $Tx_1 - Tx_2 = 0$ Using linearity we get $T(x_1 - x_2) = 0$. Using (L1) we get $x_1 = x_2$. Thus $Tx_1 = Tx_2 \rightarrow$ This means that T is one to one.

Proof of $(I2) \Longrightarrow (I3)$ To prove $\{Tx_1, Tx_2, ...\}$is linearly independent consider

$$\alpha_1(Tx_1) + \alpha_2(Tx_2) + ... = 0 = T0$$

$$T(\alpha_1 x_1 + \alpha_2 x_2 + ...) = T0$$

Using (L2),T is one to one gives $\alpha_1 x_2 + \alpha_2 x_2 + ... = 0$ It is given that $x_1, x_2, ...$ is linearly independent, hence we get $\alpha_1 = \alpha_2 = ... = 0$ This proves that $\{Tx_1, Tx_2, ...\}$ is linearly independent.

Proof of $(L3) \Longrightarrow (L4)$ Let $B = \{e_1, e_2, ..., e_N\}$ be a set of basis vectors. $\therefore \{e_1, e_2, ..., e_N\}$ is linearly independent. Hence using (L3) we see that $\{Te_1, Te_2, ..., Te_N\}$ is also linearly independent set. The number of elements, N, in this set is equal to the dimension of the vector space, hence the set is also a basis set.

Proof of $(L4) \Longrightarrow (L5)$ To prove (5),*i.e.*, T is onto we must show that $\forall$ vectors $y \in \mathcal{V}$, we can find a vector x such that $Tx = y$. Let$\{e_1, e_2, ..., e_N\}$ be a basis set then (4) gives that $\{Te_1, Te_2, ..., Te_N\}$ is also a basis.
$\therefore$ given an arbitrary vector $y \in \mathcal{V}$, we can expand y in terms of the vectors $\{Te_1, Te_2, ..., Te_N\}$:

$$y = \alpha_1 Te_1 + \alpha_2 Te_2 + ... + \alpha_N Te_N$$

or , using linearity we have

$$y = T(\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_N e_N)$$

Therefore, we have shown that y can be written as Tx, where x is given by

$$x = (\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_N e_N)$$

Proof of $(L5) \Longrightarrow (L6)$ Given T is onto means that for every $y \in \mathcal{V}$ we can find at least one vector $x \in \mathcal{V}$ such that $y = Tx$. Hence, starting from $\{y_1, y_2, ...\}$ we can form the set $\{x_1, x_2, ...\}$ such that $Tx_k = y_k$. Assume, as given in (L6), that $\{y_1, y_2, ...\}$ is LI . To prove that $\{x_1, x_2, ...\}$ is LI consider

$$\alpha_1 x_1 + \alpha_2 x_2 + ... = 0$$

Applying T on the above equation we get

$$T(\alpha_1 x_1 + \alpha_2 x_2 + ...) = 0$$

or, using the linearity

$$\alpha_1 Tx_1 + \alpha_2 Tx_2 + ... = 0$$

or,

$$\alpha_1 y_1 + \alpha_2 y_2 + ... = 0$$

As $\{y_1, y_2, ...\}$ is given to be linearly independent the above relation can be satisfied only when

$$\alpha_1 = \alpha_2 = ... = 0$$

This proves that the set $\{x_1, x_2, ...\}$ is LI .

Proof of $(L6) \Longrightarrow (L7)$ $\mathscr{Y} = \{y_1, y_2, ..., y_N\}$ is a basis set and is, therefore, LI. (L6) gives that the set $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ is also LI. Linear independence of the set $\mathscr{X}$ along with the fact that the number of elements in $\mathscr{X}$ is equal to the dimension of the vector space $\mathcal{V}$ proves that the set $\mathscr{X}$ is a basis set.

Proof of $(L7) \Longrightarrow (L1)$ Let $\mathscr{Y} = \{y_1, y_2, ...y_N\}$ be a basis and let the set $\mathscr{X} = \{x_1, x_2, ..., x_N\}$ be as in the statement (L7) then $\mathscr{X}$ is a basis.

Assume x is such that $Tx = 0$, we have to show that $x = 0$. Expand x in terms of vectors in the basis set $\mathscr{X}$:

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_N x_N$$

Applying T on the above equation, and using $Tx = 0$, we get

$$T(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_N x_N) = 0$$

$$\alpha_1 Tx_1 + \alpha_2 Tx_2 + \ldots + \alpha_N Tx_N = 0$$

$$\alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_N y_N = 0$$

The last equation above can hold only when

$$\alpha_1 = \alpha_2 = \ldots = \alpha_N = 0$$

This gives the desired result $x = 0$ proving (L1).

§2 Invertibilty

We have defined an operator to be invertible if it is one to one and onto. For linear operators on *finite* dimensional vector spaces the property of being one to one is equivalent to the property of being onto. Thus a linear operator T on a finite dimensional vector space has an inverse if any one of the following two statements holds.

(*a*) the operator is one to one. (*b*) the operator be onto.

Now we will summarize, with some repetitions, a set of statements, each of which is necessary and sufficient so that a linear operator in a finite dimensional vector space may be invertible.

Theorem ‖5 **(Conditions for an Operator to be Invertible)** *Let T be a linear operator on a finite dimensional vector space. Then the following eleven statements are equivalent.*

(I1) Inverse of T exists.

(I2) T is one to one.

(I3) T is onto.

(I4) The operator T takes a set of basis vectors to basis vectors. This means that $\mathscr{X}$ is a set of basis vectors implies that $T\mathscr{X}$ is also a basis set.

(I5) Let $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ be a subset of the vector space $\mathcal{V}$ and let $y_k = Tx_k$. If the set $\mathscr{Y} = \{y_1, y_2, \ldots, y_N\}$ is a basis then the set $\mathscr{X}$ is also a basis.

(I6) $Tx = 0$ implies $x = 0$

(I7) Range space of T is entire vector space $T\mathcal{V} = \mathcal{V}$.

(I8) The null space of $\mathscr{N}(T) = \{0\}$.

(I9) $\dim(T\mathcal{V}) = \dim(\mathcal{V})$.

(I10) Rank $(T) = \dim V$.

(I11) $\dim \mathscr{N}(T) = 0$

Proofs

: The proof that (I1) $\Longleftrightarrow$ and (I2) and (I1) $\Longleftrightarrow$ (I3) are repetitions of statements (L1) to (L7).

(I4), (I5), and (I6) are equivalent to (I2) and to (I3) has already been proved.

The property (I7),i.e., range space of T is entire vector space, $T\mathcal{V} = \mathcal{V}$, coincides with T being onto and is therefore same as (I3).

The property (I8),i.e., the null space $\mathscr{N}(T) = 0$, is just a restatement of (I6), *i.e.*, $Tx = 0$ implies $x = 0$.

Let $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ is a basis set for $\mathcal{V}$. Now we the following implications hold.

T has inverse
$\Longleftrightarrow T$ is one to one
$\Longleftrightarrow$ set $T\mathscr{X} = \{Tx_1, Tx_2, ..., Tx_N\}$ is a basis
set (using (I4) $\Longleftrightarrow$ linear span of $T\mathscr{X} = \mathcal{V}$
$\Longleftrightarrow T\mathcal{V} = \mathcal{V}$.
The last implication follows from the fact that $\mathscr{X} \subseteq \mathcal{V}$ implies $T\mathscr{X} \subseteq \mathcal{V}$ which together with $TX = V$ gives $T\mathcal{V} = \mathcal{V}$.

For proving invertibility being equivalent to (I10) recall rank$(T) = \dim(T\mathcal{V})$. Using (I9) we get the desired result that the operator T has an inverse if and only if rank $(T) = \dim \mathcal{V}$. Thus item 10. is true if and only if (I1) is true.

The property (I10) gives that a linear operator T is invertible if and only if rank $(T) = \dim(\mathcal{V})$. Now $\dim \mathscr{N}(T) + \text{rank}(T) = \dim(\mathcal{V})$. Hence T is invertible if and only if $\dim \mathscr{N}(T) = 0$. This proves T has inverse $\iff$ I11.

Theorem **6** *Let T, S, R be arbitrary (linearity is not demanded) operators on a vector space such that*

$$TS = RT = I \tag{5.1}$$

where I is the identity operator. Then T is invertible and

$$R = S = T^{-1}$$

Proof

We begin with noting the linearity is not demanded as a condition on the operators. Let R and S exist such that Eq.(5.1) is satisfied. Then

$$T(Sx) = (TS)x = x, \forall x \in \mathcal{V}.$$

Thus T is onto because given any vector $x \in V$ there exists a vector ($y = Sx$) such that $Ty = x$.

Next using $RT = I$ we will show that T is one to one. To show that T is one to one, we must prove that $Tx_1 = Tx_2 \iff x_1 = x_2$. Let $Tx_1 = Tx_2$ apply R on both sides. This gives $R(Tx_1) = R(Tx_2)$, or, $(RT)x_1 = (RT)x_2$ using the given property $RT = I$ we get the desired result that $x_1 = x_2$. Thus $Tx_1 = Tx_2 \iff x_1 = x_2$ .Therefore T is one to one. Thus T is invertible because T is both one to one and onto.

Conversely if T^{-1} exists , the given relations are satisfied for $R = S = T^{-1}$. There (II.12) is necessary and sufficient for an operator to have an inverse. Note that when the operator T is linear, each *one* of the two conditions (a) $TS = I$ (b) $RT = I$ is *separately* sufficient for T to have an inverse.

Lecture 6

Matrix Representation

We shall now discuss a way of representing vectors in a vector space, having dimension N, by N component column vectors and linear operators by $N \times N$ matrices .

Let $\mathscr{X} = x_1, x_2, ..., x_N$ be a basis. Every vector $x \in \mathcal{V}$ can be expanded in terms of the basis vectors $x_1, x_2, ..., x_N$. Thus

$$x = \xi_1 x_1 + \xi_2 x_2 + ... + \xi_N x_N. \tag{6.1}$$

The scalars $\xi_1, \xi_2, ..., \xi_N$ will be called the components of the vector x with respect to the basis $\mathscr{X}$. Knowing the vector x the scalars $\xi_1, \xi_2, ..., \xi_N$ are uniquely fixed and conversely if the scalars $\xi_1, \xi_2, ..., \xi_N$ are given, the vectors in the basis $\mathscr{X}$ can be used to get the vector. We shall assemble the components $\xi_1, \xi_2, ..., \xi_N$ in form of an N- component column denoted by x.

$$x \mapsto \mathsf{x} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ ... \\ \xi_N \end{pmatrix} \tag{6.2}$$

Let $x \in \mathcal{V}$ be a vector and T be a linear operator. The answers for the column representing a vector and for matrix representing an operator depends on chioce of basis, and will change when a new basis is selected.

To find the matrix representing a linear operator T, we note that the knowledge of the action of a linear operator on a set of basis vectors is sufficient to know the action of an operator on any vector. Thus we consider the basis $\mathscr{X} = x_1, x_2, ..., x_N$ and apply the operator T on every element to obtain the set $T\mathscr{X} = \{Tx_1, Tx_2, Tx_3, ...Tx_N\}$. Next we expand the vectors in the set Tx_k so obtained in terms of the basis vectors.

$$Tx_k = \sum_j t_{jk} x_j, \qquad k = 1, 2, ..., \tag{6.3}$$

The m^{th} row and n^{th} column of the matrix T is given by t_{mn}. We write the above N equations for $k = 1, 2, ..., N$ as

$$Tx_1 = t_{11}x_1 + t_{21}x_2 + t_{31}x_3 + ... + t_{N1}x_N \tag{6.4}$$

$$Tx_2 = t_{12}x_1 + t_{22}x_2 + t_{32}x_3 + ... + t_{N2}x_N \tag{6.5}$$

$$Tx_3 = t_{13}x_1 + t_{23}x_2 + t_{33}x_3 + ... + t_{N3}x_N \tag{6.6}$$

$$Tx_N = t_{1N}x_1 + t_{2N}x_2 + t_{3N}x_3 + ... + t_{NN}x_N \tag{6.7}$$

$$\tag{6.8}$$

The rule for constructing the matrix T for the operator T is to collect the coefficients appearing in the above equations as a matrix and take its transpose. Thus we have

$$T \to \mathsf{T} = \text{Transpose of} \begin{bmatrix} t_{11} & t_{21} & t_{31} & \dots & t_{N1} \\ t_{12} & t_{22} & t_{32} & \dots & t_{N2} \\ t_{13} & t_{23} & t_{33} & \dots & t_{N3} \\ \dots & \dots & \dots & \dots & \dots \\ t_{1N} & t_{2N} & t_{3N} & \dots & t_{NN} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1N} \\ t_{21} & t_{22} & t_{23} & \dots & t_{2N} \\ t_{31} & t_{32} & t_{33} & \dots & t_{3N} \\ \dots & \dots & \dots & \dots & \dots \\ t_{N1} & t_{N2} & t_{N3} & \dots & t_{NN} \end{bmatrix} \tag{6.9}$$

Thus $\mathsf{T}_{ij} = \mathsf{t}_{ij}$. With this every vector space of dimension N becomes isomorphic to $\mathbb{F}^N$. Every relation between vectors and operators is equivalent to a relation between $N-$ component columns and $N \times N$ matrices. For example

If $y = Tx$ we have $\mathsf{y}= \mathsf{T}\ \mathsf{x}$; Similarly, if $AB = C$ then $\mathsf{A} = \mathsf{BC}$ where A, B, C are operators and $x, y, ..$ are vectors in $\mathcal{V}$.

§1 An Example

- Let $e_1 = (1, 1, 0), e_2 = (0, 1, 1), e_3 = (1, 0, 1)$ be a basis in $\mathbb{R}^3$. Find components of a vector $f = (x, y, z)$ and represent it w.r.t. the basis $\{e_1, e_2, e_3\}$
- Let an operator T be defined as $Te_1 = (1, 0, 0), Te_2 = (0, 1, 0)$ and $Te_3 = (1, 1, 1)$ Knowledge of action of an operator T on a basis is sufficient to find its action on any vector. Given e_1, e_2, e_3 as above, find the vector $g = Tf$. 3. Find the representatives of two vectors f, g and the matrix T w.r.t. the basis (e_1, e_2, e_3) and verify that $\mathsf{T}\ \mathsf{f} = \mathsf{g}$

SOLUTION :

(1) Let $f = (x, y, z)$ be written as a linear combination of the vectors e_1, e_2, e_3 :

$$f = ae_1 + be_2 + ce_3 \tag{6.10}$$

$$(x, y, z) = a(1, 1, 0) + b(0, 1, 1) + c(1, 0, 1) \tag{6.11}$$

$$= (a + c, a + b, b + c) \tag{6.12}$$

$$\text{or} a + c = x; \qquad a + b = y; \qquad b + c = z \tag{6.13}$$

$$\tag{6.14}$$

This is easily solved to give

$$a = (x + y - z)/2, b = (y + z - x)/2, c = (z + x - y)/2.$$

$$(x, y, z) = \frac{(x + y - z)}{2}e_1 + \frac{(y + z - x)}{2}e_2 + \frac{(z + x - y)}{2}e_3$$

Thus $f \to \mathsf{f}$ where

$$f \mapsto \mathsf{f} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad (6.15)$$

$$\mathsf{f} = \begin{pmatrix} \frac{(x+y-z)}{2} \\ \frac{(y+z-x)}{2} \\ \frac{(z+x-y)}{2} \end{pmatrix} \quad (6.16)$$

(2) To find how T acts on a general vector $h = ae_1 + be_2 + ce_3$, we compute $g = Th$. Using the linearity property we get

$$g = T[ae_1 + be_2 + ce_3] \quad (6.17)$$

$$= aTe_1 + bTe_2 + cTe_3 \quad (6.18)$$

$$= a(1, 0, 0) + b(0, 1, 0) + c(1, 1, 1) \quad (6.19)$$

$$= (a + c, b + c, c) \quad (6.20)$$

$$\therefore Tf = \left(x, z, \frac{(z + x - y)}{2}\right) \quad (6.21)$$

For later use we remark that the vector g written as linear combination of $\{e_1, e_2, e_3\}$ becomes

$$g = \frac{(x + y + z)}{4}e_1 + \frac{(3z - x - y)}{4}e_2 + \frac{(3x - y - z)}{4}e_3.$$

(3) We construct the matrix for the operator w.r.t. the basis (e_1, e_2, e_3). For this purpose we must express Te_1, Te_2 and Te_3 as linear combinations of e_1, e_2, e_3.

$$Te_1 = (1, 0, 0) = \frac{1}{2}e_1 - \frac{1}{2}e_2 + \frac{1}{2}e_3 \quad (6.22)$$

$$Te_2 = (0, 1, 0) = \frac{1}{2}e_1 + \frac{1}{2}e_2 - \frac{1}{2}e_3 \quad (6.23)$$

$$Te_3 = (1, 1, 1) = \frac{1}{2}e_1 + \frac{1}{2}e_2 + \frac{1}{2}e_3 \quad (6.24)$$

$$(6.25)$$

Therefore, the matrix, T, representing the operator T w.r.t. the basis $\{e_1, e_2, e_3\}$ is given by

$$\mathsf{T} = \text{Transpose of} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad (6.26)$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad (6.27)$$

A second way of computing the vector g is to apply T on f to get g as follows.

$$\mathsf{g} = \mathsf{T}\mathsf{f} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \times \begin{bmatrix} (x+y-z)/2 \\ (y+z-x)/2 \\ (z+x-y)/2 \end{bmatrix} \tag{6.28}$$

or

$$g = \begin{bmatrix} (x+y-z)/4 \\ (3z-x-y)/4 \\ (3x-y-z)/4 \end{bmatrix} \tag{6.29}$$

This gives the components of the vector g w.r.t. the basis $\{e_1, e_2, e_3\}$. To get back the vector we reconstruct the vector g as

$$\begin{aligned} g &= \frac{(x+y+z)}{4}e_1 + \frac{(3z-x-y)}{4}e_2 + \frac{(3x-y-z)}{4}e_3 && (6.30) \\ &= \left(\frac{(x+y+z)}{4}, \frac{(x+y+z)}{4}, 0\right) && (6.31) \\ &\quad + \left(0, \frac{(3z-x-y)}{4}, \frac{(3z-x-y)}{4}\right) && (6.32) \\ &\quad + \left(\frac{(3x-y-z)}{4}, 0, \frac{(3x-y-z)}{4}\right) && (6.33) \\ &= \left(x, z, \frac{(z+x-y)}{2}\right) && (6.34) \end{aligned}$$

which agrees with the result obtained above.

§2 Change of Basis

The elements of columns x representing a vector $x \in \mathcal{V}$ are just the components of the vector x with respect to a given basis. These would change when a different basis is used. Similarly the values of the matrix elements of T for a given operator depend on the choice of basis. We shall now give formulas on relating the components of a vector w.r.t. two different bases and on relation between the matrices representing an operator w.r.t. two two bases.

Let x be a vector $\in \mathcal{V}$ and T be a linear operator. Let $\mathscr{X} = \{x_k | k = 1, 2, ...N\}$ and $\mathscr{Y} = \{y_k | k = 1, 2, .., N\}$ be two bases. It is useful to define an operator S which takes elements in basis $\mathscr{X}$ to the elements in the basis $\mathscr{Y}$. Let

$$Sx_k = y_k,$$

thus $\mathscr{Y} = S\mathscr{X}$. The operator S will be an invertible operator.

Let the columns $\underline{\mathsf{x}}$ and $\underline{\underline{\mathsf{x}}}$ represent the vector x w.r.t. the two bases $\mathscr{X}$ and $\mathscr{Y}$ respectively. Similarly, let the matrices $\underline{\mathsf{T}}$ and $\underline{\underline{\mathsf{T}}}$ represent the operator T w.r.t. the two bases $\mathscr{X}$ and$\mathscr{Y}$ respectively.

With respect to the first basis $\mathscr{X}$ let the representation be given by

$$x \mapsto \underline{\mathsf{x}}; \qquad T \mapsto \underline{\mathsf{T}}$$

and the representation with respect to the second basis $\mathscr{Y}$

$$x \mapsto \underline{\underline{\mathsf{x}}},; \qquad T \mapsto \underline{\underline{\mathsf{T}}}$$

In order to exhibit the relation between the two sets of representatives we make use of the linear operator S defined above by means of equations $Sx_k = y_k$. Note that the operator S is invertible. Let S be the matrix representing the operator S w.r.t. basis $\mathscr{X}$. Then we have the following relations.

$$\underline{\underline{\mathsf{x}}} = \mathsf{S}^{-1}\underline{\mathsf{x}}; \qquad \text{and} \qquad \underline{\underline{\mathsf{T}}} = \mathsf{S}^{-1}\underline{\mathsf{T}}\mathsf{S}$$

We shall skip the proof of these relations [See Halmos].

Lecture 7

Topics for Further Study

The most important topics not covered in these lecture notes and suggested references are as follows.

1. Jordan Canonical forms [1]
2. Direct product Spaces [1]
3. Multilinear functionals [1]
4. Quotient and Tensor Product Spaces [2]
5. Vectors and Tensors [2]

References

[1] Paul R. Halmos, *Finite Dimensional Vector Spaces*

[2] Pannkaj Sharan, *Space Time and Gravitation*

Part II

Inner Product Spaces

Lecture 8

Norm and Inner Product

Important: *From now onwards all the vector spaces we deal are complex vector spaces of finite dimension unless mentioned otherwise.*

§1 Norm and scalar product

Definition 35 **Norm** *of a vector f in a vector space V is a real number $\|f\|$ satisfying the following properties.*

1. $\|f\| \geq 0$, and, $\|f\| = 0$ if and only if $f = 0$.
2. $\|\alpha f\| = |\alpha| \|f\|$
3. $\|f + g\| \leq \|f\| + \|g\|$ (Triangle Inequality)

Quick Question: *Is norm a linear functional !? WHY ?*

Definition 36 *An* **inner product***, (or scalar product), denoted by (f, g), in a complex vector space $\mathcal{V}$ is a complex valued function of the ordered pair of vectors $f, g \in \mathcal{V}$ such that*

1. $(f, f) \geq 0, \text{and } (f, f) = 0$ iff $f = 0$
2. $(f, g) = (g, f)^*$
3. $(f, \alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 (f, g_2) + \alpha_2 (f, g_2)$
4. $(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1^* (f_1, g) + \alpha_2^* (f, g)$

Remark: We shall not discuss real vector spaces with inner product.

)|(**Short Examples 2** *Some examples of scalar product are given. That the properties of the scalar product are satisfied can be easily checked.*

(2a) In $\mathbb{C}^N$ a vector is a N-component column vector with complex numbers as entries. For two vectors f, g:

$$f = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_N \end{pmatrix} \qquad g = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_N \end{pmatrix}$$

define

$$(f, g) = \sum_k \alpha_k^* \beta_k \tag{8.1}$$

(2b) In the vector space $\mathscr{L}^2(-\infty, \infty)$ of all square integrable functions, scalar product between two functions $\psi(x), \phi(x)$ is defined as

$$(\psi(x), \phi(x)) = \int_{-\infty}^{\infty} \psi^*(x)\phi(x), dx \tag{8.2}$$

(2c) In the vector space of all polynomials $\mathscr{P}$ with complex coefficients, a scalar product of two polynomials $p(t), q(t)$ can be defined as

$$(p, q) = \int_{-1}^{1} p^*(t)q(t)\, dt \tag{8.3}$$

oslash In general several different scalar products can be defined in a given vector space; the choice of scalar product for a vector space is not unique.

(2d) In the vector space of all polynomials a few examples of other scalar products are

$$(p, q) = \int_{-\infty} \infty e^{-t^2} p^*(t)q(t)\, dt \tag{8.4}$$

$$(p, q) = \int_0^{\infty} e^{-t} p^*(t)q(t)\, dt \tag{8.5}$$

$$(p, q) = \int_a^b w(t)p^*(t)q(t)\, dt \tag{8.6}$$

where in the last example $w(t)$ is any nonsingular positive function defined in the range $a \leq t \leq b$.

Examples Of Properties Of Inner Product The property (4) can be proved from properties (2) and (3). Thus we have

$$\begin{aligned} (\alpha_1 f_1 + \alpha_2 f_2, g) &= [(g, \alpha_1 f_1) + (g, \alpha_2 f_2)]^* && (8.7) \\ &= [\alpha_1 (g, f_1) + \alpha_2 (g, f_2)]^* && (8.8) \\ &= \alpha_1^*(g, f_1) + \alpha_2^*(g, f_2) && (8.9) \end{aligned}$$

Using the property (2) once again we get the desired result:

$$(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1(f_1, g) + \alpha_2(f_:2, g)$$

We shall now prove two important identities.

Parallelogram Identity

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2) \tag{8.10}$$

Polarization Identity

$$4(f,g) = \|f+g\|^2 - \|f-g\|^2 + i\|f-ig\|^2 - i\|f+ig\|^2 \tag{8.11}$$

PROOF :

$$\|f+g\|^2 = (f+g, f+g) = (f,f) + (f,g) + (g,f) + (g,g) \tag{8.12}$$

$$\|f-g\|^2 = (f-g, f-g) = (f,f) - (f,g) - (g,f) + (g,g) \tag{8.13}$$

$$\|f-ig\|^2 = (f-ig, f-ig) = (f,f) - i(f,g) + i(g,f) + (g,g) \tag{8.14}$$

$$\|f+ig\|^2 = (f+ig, f+ig) = (f,f) + i(f,g) - i(g,f) + (g,g) \tag{8.15}$$

Adding Eq.(8.12) and Eq.(8.13) gives the parallelogram identity. In a similar fashion taking Eq.(8.12) - Eq.(8.13) + $i\otimes$ Eq.(8.14) - $i\otimes$ Eq.(8.15) gives the polarization identity.

Defining norm from an inner product In a vector space with an inner product if we define

$$\|f\| = \sqrt{(f,f)},$$

then $\|f\|$ has all the properties of the norm. The two properties (1) and (2) of the norm are automatically satisfied. The third property, viz., the triangle inequality will be proved below after the proof of Cauchy Schwarz inequality.

Conversely, if a norm is defined in a complex vector space we ask: "can we introduce a norm such that the relation is maintained?" The answer is YES if and only if the norm satisfies the parallelogram identity. The right hand side of the polarization identity can then be taken as the definition of inner product. The result will satisfy all the axioms for the inner product.

§2 Cauchy Schwarz inequality

As a preparation we first prove an intermediate result.

Theorem)‖(**7** *If f is a given vector and $g \neq 0$ be any vector $\|f-\lambda g\|$ is minimum when $\lambda = \lambda_0$ where*

$$\lambda_0 = \frac{(f,g)^*}{\|g\|^2} = \frac{(g,f)}{(g,g)}$$

and the minimum value of $\|f-\lambda g\|$ is given by

$$\|f-\lambda g\|_{min} = \|f\|^2 - |(f,g)|^2/\|g\|^2$$

Proof:

Let $F(\lambda) = \|f-\lambda g\|^2$. We compute $F(\lambda)$, write it as function of the real and imaginary parts of $\lambda(\equiv \alpha + i\beta)$ and minimize $F(\lambda)$ w.r.t. α and β.

$$F(\lambda) = \|f-\lambda g\|^2 \tag{8.16}$$

$$= (f-\lambda g, f-\lambda g) \tag{8.17}$$

$$= (f,f) - \lambda(f,g) - \lambda^*(g,f) + \|\lambda\|^2(g,g) \tag{8.18}$$

Substituting $\lambda = \alpha + i\beta$ we get

$$F(\lambda) = (f,f) - \alpha[(f,g) + (g,f)] + i\beta[(g,f) - (f,g)] + (\alpha^2 + \beta^2)(g,g)$$

Note that the right hand side has to be real. WHY ?! Setting

$$\frac{\partial F}{\partial \alpha} = 0, \qquad \text{and} \qquad \frac{\partial F}{\partial \beta} = 0$$

we get

$$-(f,g) - (g,f) + 2\alpha(g,g) = 0 \tag{8.19}$$

$$i(g,f) - i(f,g) + 2\beta(g,g) = 0 \tag{8.20}$$

hence

$$\alpha = [(f,g) + (g,f)]/2(g,g) \tag{8.21}$$

$$\beta = i[(f,g) - (g,f)]/2(g.g) \tag{8.22}$$

This gives the desired value λ_0 corresponding to the minimum of $F(\lambda)$ as

$$\lambda_0 = \alpha + i\beta = \frac{(g,f)}{(g,g)} = \frac{(f,g)^*}{(g,g)}$$

and the minimum value of $F(\lambda_0)$ is then computed to be

$$F(\lambda)|_{\min} = (f,f) - |(f,g)|^2/(g,g)$$

Theorem)‖(**8 (Cauchy Schwarz Inequality)** *Let $f, g \in \mathcal{V}$. Then*

$$|(f,g)| \le \|f\|\,\|g\|$$

The equality holds if and only if f and g are linearly dependent.

Proof :If $f = 0$ or $g = 0$, the equality holds trivially and there is nothing to prove because both sides are zero. Therefore, we assume $g \neq 0$. Consider $x = f - \lambda g$. Then we have $\|x\| \ge 0$ for all values of λ. We find the minimum of $\|x\|$ and set it ≥ 0

$$\min \|x\|^2 \ge 0$$

Using the previous result $\|x\|^2 = \|f - \lambda g\|^2$ is minimum when λ is equal to $\frac{(g,f)}{(g,g)}(\equiv \lambda_0)$ and minimum value of $\|x\|^2 = \|f - \lambda g\|^2$ is given by

$$\min \|x\|^2 = (f,f) - \frac{|(f,g)|^2}{(g,g)}$$

Thus we get

$$(f, f) - \frac{|(f, g)|^2}{(g, g)} \geq 0$$

or

$$(f, f)(g, g) \geq |(f, g)|^2$$

which is just the desired Cauchy Schwarz inequality

$$|(f, g)| \leq \|f\| \, \|g\|.$$

Note that when the Cauchy Schwarz inequality becomes equality $\min\|x\|^2 = 0$. This is possible if and only if $x = 0$ for $\lambda = \lambda_0$. This gives $f - \lambda_0 g = 0$ which means that f and g are linearly dependent.

§3 Triangle Inequality

We are now in a position to prove the triangle inequalities

$$\|f + g\| \quad \leq \quad \|f\| + \|g\| \tag{8.23}$$

Proof : Consider

$$\begin{aligned}
\|f + g\|^2 &= (f + g, f + g) && (8.24)\\
&= (f, f) + (f, g) + (g, f) + (g, g) && (8.25)\\
&= (f, f) + (f, g) + (f, g)^* + (g, g) && (8.26)\\
&= (f, f) + 2\text{Re } (f, g) + (g, g), \quad [\because z + z^* = 2\text{Re } z] && (8.27)\\
&\leq \|f\|^2 + \|g\|^2 + 2|(f, g)|, \quad [\because \text{Re } z \leq |z| \,]. && (8.28)
\end{aligned}$$

Using the Cauchy Schwarz inequality, $|(f, g)| \leq \|f\|\|g\|$, we get

$$\begin{aligned}
\|f\|^2 + \|g\|^2 + 2|(f, g)| &\leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 && (8.29)\\
&= [\|f\|^2 + \|g\|]^2 && (8.30)
\end{aligned}$$

$\therefore$ We get the desired inequality:

$$\|f + g\| \leq \|f\| \; + \|g\| \tag{8.31}$$

it has the interpretation that *length of any side of a triangle is less than the sum of the lengths of the other two sides.* See figure below.

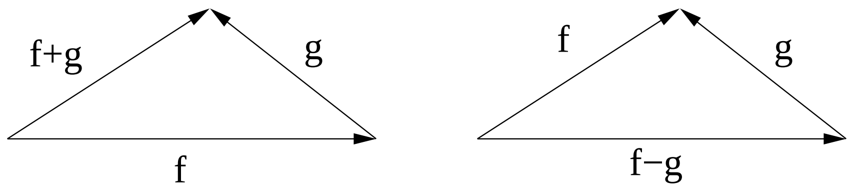

Fig. **1** Interpretation of Triangle Inequalities

The proof of a second triangle inequality, which says that *length of any side of a triangle is greater than the difference of the lengths of other two sides of the triangle.* This is translated into mathematical form as

$$\|f+g\| \geq \big|\|f\| - \|g\|\big|, \tag{8.32}$$

and also as $\|f-g\| \geq \big|\|f\| - \|g\|\big|$. The proof of this inequality is left as an exercise for the reader.

Lecture 9

Orthogonality

§1 Orthogonality

Definition 37 *We say that two vectors f and g are* **orthogonal** *if* $(f, g) = 0$

LEMMA : If $g \neq 0$, the vector

$$x = f - \frac{(g, f)}{(g, g)} g$$

is orthogonal to g.

Proof : Consider

$$(g, x) = \left(g, f - \frac{(g, f)}{(g, g)} g\right) = (g, f) - \frac{(g, f)}{(g, g)}(g, g) \tag{9.1}$$

$$= (g, f) - (g, f) = 0 \tag{9.2}$$

Therefore, g is orthogonal to $x = f - \frac{(g, f)}{(g, g)} g$.

Definition 38 *Two vectors f and g are said to be* **orthonormal**, *if f, g are orthogonal and* $\|f\| = \|g\| = 1$.

Definition 39 *A set of vectors $\mathscr{X}$ is an* **orthogonal set** *if $\forall$ pair $x, y \in \mathscr{X}$, we have $(x, y) = 0$.*

Definition 40 *A set of vectors $\mathscr{X}$ is called* **orthonormal set** *if*

(a) *for every pair $x, y \in \mathscr{X}$ we have $(x, y) = 0$ and*

(b) *for every $x \in \mathscr{X}$ we have $\|x\| = 1$.*

Definition 41 *A set $\{x_1, x_2, ..., x_r\}$ is an* **orthonormal set** *iff $(x_i, x_j) = \delta_{ij}$.*

)|(**Short Examples 3** *(Orthonormal Sets)*

We will now give several examples of orthonormal sets.

(3a) In the vector space $\mathbb{R}^3$, the set of unit vectors $\{\vec{i}, \vec{j}, \vec{k}\}$ along the three coordinate axes is an orthonormal set. In fact, if the coordinate axes are rotated the unit vectors along the new axes will again from an o.n. set.

(3b) Consider the vector space $\mathbb{C}^n$, with inner product of two column vectors x, y defined by $x^\dagger y$, the set of vectors $\{x_1, x_2, \ldots\}$ given by

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \qquad ...., \qquad x_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}, \tag{9.3}$$

is an o.n. set.

(3c) Consider the complex vector space of all polynomials $\mathbb{P}_n$. We then have following examples of orthonormal sets.

(i) With inner product of two polynomials $p(t), q(t)$ defined as

$$(p, q) = \int_{-\infty}^{\infty} p^*(t) q(t) e^{-t^2} \, dt$$

The set of all Hermite polynomials $\{H_0(t), H_1(t), \ldots, H_n(t), \ldots\}$ is an o.n. set.

(ii) With inner product of two polynomials $p(t), q(t)$ defined as

$$(p, q) = \int_{-1}^{1} p^*(t) q(t) \, dt$$

The set of all Legendre polynomials $\{P_0(t), P_1(t), \ldots, P_n(t), \ldots\}$ is an o.n. set.

(iii) With inner product of two polynomials $p(t), q(t)$ defined as

$$(p, q) = \int_{0}^{\infty} p^*(t) q(t) t^\nu e^{-t} \, dt$$

The set of all Laguerre polynomials $\{L_0^\nu(t), L_1^\nu(t), \ldots, L_n^\nu(t), \ldots\}$ is an o.n. set.

(iv) The set of monomials $\{1, t, t^2, t^3, \ldots\}$ is not an orthonormal set with any of the above three inner products.

⊘ *The above examples clearly show that a set being o.n. set depends on the choice scalar product.*

(3d) In the vector space of square integrable functions, $\mathscr{L}^2(-\infty, \infty)$, the scalar product of two functions $\psi(x), \phi(x)$ is defined to be

$$(\psi, \phi) = \int_{-\infty}^{\infty} \psi^*(x) \phi(x) \, dx$$

In this space the harmonic oscillator wave functions form an. o.n. set.

(3e) In the vector space of all functions defined on interval $[-\pi, \pi]$ and satisfying

$$f(-\pi) = f(\pi)$$

and having inner product defined by

$$(f, g) = \int_{-\pi}^{\pi} f^*(x) g(x) \, dx$$

an orthonormal set is

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots \cos nx, \sin nx, \ldots\}$$

Remark:

The orthonormal property of a set $\mathscr{E} = \{x_k | k = 1, 2, \ldots, r\}$ can be written as $(x_k, x_\ell) = \delta_{k\ell}$, where $\delta_{k\ell}$ is Kronecker delta defined by

$$\delta_{k\ell} = \begin{cases} 1 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell. \end{cases} \tag{9.4}$$

Definition 42 *An orthonormal set is called a* **complete orthonormal set** *if it is not contained in any larger orthonormal set.*

Theorem)‖(**9** *An orthogonal set $\mathscr{X} = \{x_1, x_2, ..x_r\}$ of non-zero vectors is linearly independent.*

Proof : Consider

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_r x_r = 0 \tag{9.5}$$

Taking scalar product with x_1 gives zero for all terms except the first one. Thus

$$\alpha_1 (x_1, x_1) = 0 \Rightarrow \alpha_1 = 0 \tag{9.6}$$

$$(\because x_1 \neq 0 \Rightarrow (x_1, x_1) \neq 0). \tag{9.7}$$

Remark : Earlier we have seen that the vector $h = f - \lambda g$ is orthogonal to the vector g if λ is taken to be $(g, f)/(g, g)$. The following theorem generalizes this result to orthogonal sets.

Theorem)‖(**10** *If $\mathscr{U} = u_1, u_2, ..., u_n$ is any finite orthogonal set containing non zero vectors of an inner product space and if $\lambda_k = (u_k, x)/(u_k, u_k)$, then the vector h defined by*

$$h = f - \lambda_1 u_1 - \lambda_2 u_2 - ... - \lambda_k u_k$$

is orthogonal to every element u_k in the set $\mathscr{U}$

The result follows easily by taking the scalar products (h, u_k) for different k.

§2 Grahm Schmidt Orthogonalization Procedure

Let $\mathscr{X} = \{x_1, x_2, \ldots, x_r\}$ be a linearly independent set. Then one can construct a set of vectors $\mathscr{E} = \{e_1, e_2, ...e_r\}$ such that the vectors e_k are linear combinations of the vectors in $\mathscr{X}$ and the set $\mathscr{E}$ is an orthonormal set.

Proof: Recursively define

$$\begin{aligned} u_1 &= x_1, & e_1 &= u_1/\|u_1\| \\ u_2 &= x_2 - (e_1, x_2)e_2, & e_2 &= u_2/\|u_2\| \\ u_3 &= x_3 - (e_1, x_3)e_3 - (e_2, x_3)e_2, & e_3 &= u_3/\|u_3\| \\ u_r &= x_r - \sum_{k=1}^{r-1} (e_k, x_r)e_k, & e_r &= u_r/\|u_r\| \end{aligned}$$

It is easily verified that $\{e_1, e_2, ...\}$ is an o.n. set.

§3 Bessel's Inequality

If $\mathscr{U} = u_1, u_2, ..., u_r$ is any finite orthonormal set in an inner product space then for all $x \in \mathcal{V}$ we have

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \qquad \text{(Bessel Inequality)} \tag{9.8}$$

Proof : For every vector y , we have $(y, y) \geq 0$. Therefore, taking y to be

$$y = x - \sum_k \lambda_k u_k \qquad \text{with } u_k = (u_k, x).$$

we get

$$
\begin{aligned}
(y, y) &= (x - \sum_k \lambda_k u_k, x - \sum_j \lambda_j u_j) && (9.9)\\
&= (x, x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_j \sum_k \lambda_k^* \lambda_j (u_j, u_k) && (9.10)\\
&= (x, x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_k \lambda_k^* \lambda_k && (9.11)
\end{aligned}
$$

One of two the summations in the last term has been done using $(u_j, u_k) = \delta_{jk}$. Substituting $\lambda_j = (u_j, x)$ we get

$$
\begin{aligned}
(y, y) &= (x, x) - \sum (x, u_k)(u_k, x) - \sum (u_j, x)(x, u_j) + \sum (x, u_j)(u_j, x) && (9.12)\\
&= (x, x) - \sum_k (x, u_k)(u_k, x) && (9.13)\\
&= (x, x) - \sum_k |(u_k, x)|^2 && (9.14)
\end{aligned}
$$

Using $(y, y) \geq 0$ we get the desired Bessel's inequality.

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \tag{9.15}$$

Lecture 10

Complete Orthonormal Sets

§1 Complete Orthonormal Sets

Theorem **11 (Orthonormal Sets)** *If $\mathscr{U} = \{u_1, u_2, ..., u_n\}$ is any finite o.n. set in an inner product space having finite dimension, the following conditions $(P1)$ to $(P6)$ on $\mathscr{U}$ are equivalent to each other.*

(P1) The set $\mathscr{U}$ is complete.

(P2) If $(x, u_k) = 0 \quad \forall k$ then $x = 0$.

(P3) The subspace spanned by $\mathscr{U}$ is whole space.

(P4) If $f \in \mathcal{V}$ then

$$f = \sum_k (u_k, f)u_k$$

(P5) If f and g are in $\mathcal{V}$ then

$$(f, g) = \sum (f, u_k)(u_k, g)$$

(P6) If $x \in \mathcal{V}$ then,

$$\|x\|^2 = \sum_k |(u_k, x)|^2$$

PROOFS We shall prove that

$$(P1) \Rightarrow (P2) \Rightarrow (P3) \Rightarrow (P4) \Rightarrow (P5) \Rightarrow (P6) \Rightarrow (P1).$$

$\underline{(P1) \Rightarrow (P2):}$ If $\exists$ a vector f such that $(u_k, f) = 0 \forall k$ and $f \neq 0$. Then the set $\mathscr{U} \cup f/\|f\|$ would be an orthonormal set containing $\mathscr{U}$. But this is impossible because $\mathscr{U}$ is a complete set. Therefore, $f = 0$.

$\underline{(P2) \Rightarrow (P3):}$ Assume that $(P3)$ is not true. If the subspace spanned by $\mathscr{U}$ is not whole space, there would exist a vector $f \neq 0$ such that f is not a linear combination of elements in $\mathscr{U}$. Hence $g = f - \sum_k (f, u_k)u_k$ is different from zero and is, by construction, orthogonal to all u_k, this contradicts $(P2)$. Thus we have proved $\sim(P3) \Rightarrow\ \sim(P2)$ giving us the required result $(P2) \Rightarrow (P3)$..

> **Remember** One of the ways to a statement $A \Rightarrow B$ is to start from negation of statement B and prove negation of statement A. This method, $\sim B \Rightarrow\sim A$, is what has been used in the above two cases to write the proof. For more, see

$(P3) \Rightarrow (P4)$: We are given that the subspace spanned by $\mathscr{U}$ is whole space. Hence every vector is a linear combination of $\{f_i\}$:

$$f = \sum_i \alpha_i u_i.$$

Taking scalar product with u_k and using the fact that $\mathscr{U}$ is o.n. set we get $\alpha_k = (u_k, f)$. $\therefore$ $f = \sum_k (u_k, f) u_k$.

$(P4) \Rightarrow (P5)$**:** Let f, g be two arbitrary vectors in the vector space. The result (4) applied to two vectors f and g gives

$$f = \sum_i \lambda_i u_i; \quad \text{with} \quad \lambda_i = (u_i, f) \tag{10.1}$$

$$g = \sum_j \mu_j u_j, \quad \text{where} \quad \mu_j = (u_j, g) \tag{10.2}$$

The result $(P5)$ follows by computing (f, g) using the orthogonality properties of u_k.

$$(f,g) = \Big(\sum_i \lambda_i u_i, \sum_j \mu_j u_j\Big) \tag{10.3}$$

$$= \sum_i \sum_j \lambda_i^* \mu_j (u_i, u_j) \tag{10.4}$$

$$= \sum_{ij} \lambda_i^* \mu_j \delta_{ij} \tag{10.5}$$

$$= \sum_i \lambda_i^* \mu_i \tag{10.6}$$

where, in step (10.5), we have used the orthogonality property $(u_i, u_j) = \delta_{ij}$. This gives us the desired result

$$(f, g) = \sum_i (f, u_i)(u_i, g) \tag{10.7}$$

$(P5) \Rightarrow (P6)$**:** If we set $f = g = x$ in the result of $(P5)$, we get $(P6)$.

$(P6) \Rightarrow (P1)$:

> **Recall** One of the methods, known as *reduction ad absurdum*, of proving $A \Rightarrow B$ is to assume that B is not true and to derive a contradiction. This is what will be used to write this part of the the proof.

To obtain a contradiction let us assume that $(P1)$ is not true the set $\mathscr{U}$ is not complete. Then there exists a vector $h \neq 0$ which is orthogonal to all $u_k, (h, u_k) = 0$. We apply $(P5)$

$$\|x\|^2 = \sum_k |(u_k, x)|^2$$

to $x = h$. The left hand side is non-zero while the r.h.s. is zero, hence a contradiction.

This proves $(P6) \Rightarrow (P1)$.

Theorem **12** *If $\mathcal{V}$ is vector space with inner product, then there exists a complete o.n. sets in $\mathcal{V}$, and every o.n. set contains exactly n elements.*

If we start from a basis set and apply the Gram-Schmidt orthogonalization procedure we would get a complete o.n. set. We skip the proofs and discussion.

So, for example starting from a basis of monomials $\{1, t, t^2, \ldots\}$ and taking the scalar product of two polynomials $p(t), q(t)$ to be

$$(p, q) = \int_{\infty}^{\infty} e^{-t^2} p(t) q(t)\, dt$$

we would get Hermite polynomials as the o.n. basis in the space of all polynomials.

Lecture 11

Linear Operators in Inner Product Spaces

Theorem **13** *Let T be a linear operator in an inner product space. Let f and g be arbitrary vectors then*

$$(f, Tg) = (x_1, Tx_1) - (x_2, Tx_2) + i(x_3, Tx_3) - i(x_4, Tx_4)$$

where

$$x_1 = f + g; \quad x_2 = f - g; \quad x_3 = f - ig; \quad x_4 = f + ig$$

Proof follows by proceeding in a way similar to the proof of polarization identity. Use linearity of the operator T and expand the right hand side of the above identity to be proved.

Theorem **14 (When is a linear operator zero ?)**

(Z1) If $(f, Tg) = 0$ holds for all f and g, then $T = 0$.

(Z2) If $(f, Tf) = 0$ is true for all $f \in \mathcal{V}$, then $T = 0$.

(Z3) If $(x_i, Tx_j) = 0$ holds for all elements x_i, x_j in a basis $\mathscr{X} = \{x_1, x_2, \ldots, x_N\}$ then $T = 0$.

Proof of (Z1) Given that $(g, Tf) = 0$ holds for all $g, f \in \mathcal{V}$. Therefore, we take $g = Tf$. This gives $(Tf, Tf) = 0$ which in turn implies $Tf = 0$ for all $f \in \mathcal{V}$. Therefore, $T = 0$.

Proof of (Z2) Using linearity of T we have already proved the result that

$$4(f, Tg) = (x_1, Tx_1) - (x_2, Tx_2) - i(x_3, Tx_3) + i(x_4, Tx_4)$$

where

$$x_1 = f + g; x_2 = f - g; x_3 = f - ig; x_4 = f + ig$$

Since $(f, Tf) = 0$ for all vectors $f \in \mathcal{V}$, the right hand side is zero. Hence we get $(f, Tg) = 0$ for all f, g in the vector space . Hence using part (Z2) we get $T = 0$.

Proof of (Z3) Since an arbitrary vector f can be written as a linear combination, $f = \sum \alpha_k x_k$, we can prove (2) by using linearity of T. Hence the result $T = 0$ follows.

§1 Adjoint of an Operator

Representation Theorem For Linear Functionals It can be proved that for every linear functional $\psi : f \longrightarrow \psi(f)$ on a complex inner product space $\mathcal{V}$ of finite dimension, there exists a vector $g \in \mathcal{V}$ such that $<< Proof?? >>$

$$\psi(f) = (g, f) \tag{11.1}$$

Definition 43

Given a linear operator T on an inner product space of finite dimension now we shall define **adjoint** of T, to be denoted by $T^\dagger$. The adjoint $T^\dagger$ will be defined once its action on an arbitrary vector f is specified. The functional ϕ, defined by

$$\phi : g \to \phi(g) = (f, Tg), \tag{11.2}$$

is a linear functional on $\mathcal{V}$, and hence there exists a vector unique vector $h \in \mathcal{V}$ such that

$$(h, g) = \phi(g) = (f, Tg) \tag{11.3}$$

We, then, define $T^\dagger f = h$. Thus the operator $T^\dagger$ has the property

$$(f, Tg) = (T^\dagger f, g), \qquad \forall f, g \in \mathcal{V}. \tag{11.4}$$

Remark If we find $(f, Tg) = (Xf, g)$ holds for all f and g we can conclude immediately that $X = T^\dagger$. WHY? The proof is left as an exercise for you.

Properties of Adjoint Let A be a linear operator in a complex inner product space.

(A1) $A^\dagger$, the adjoint of a linear operator is again a linear operator.

(A2) $(A^\dagger)^\dagger = A$

(A2) $(\alpha A)^\dagger = \alpha^* A^\dagger$

(A3) $(A + B)^\dagger = A^\dagger + B^\dagger$

(A4) $(AB)^\dagger = B^\dagger A^\dagger$

(A5) If A is invertible, $A^\dagger$ is invertible and $(A^\dagger)^{-1} = (A^{-1})^\dagger$.

)|(**Short Examples 4** *In vector space $\mathbb{C}^n$, an operator is represented by $n \times n$ complex matrix. The adjoint of the corresponding operator is obtained by interchanging rows and columns and taking complex conjugate. This is the familiar definition of adjoint of matrix. The proof is left* *as an exercise for you.*

☑ **Problem 1:** *In the Hilbert space op square integrable functions $\mathscr{L}^2(-\infty, infty)$ find adjoint of an operator X defined by*

$$T\psi(x) = \psi(ax+b), \quad a \neq 0, \text{ and } a, b \in \mathbb{R}$$

Solution: *To find $X^\dagger$, we must define its action on arbitrary square integrable function $\psi(x)$ starting from known action of X. The starting point is the relation*

$$(T\phi(x), \psi(x)) = (\phi(x), T^\dagger\psi(x)) \tag{11.5}$$

The left hand side of this equation in the vector space $\mathbb{L}^2$ takes the form

$$\int_{-\infty}^{\infty} (T\phi(x))^* \psi(x)\, dx = \int_{-\infty}^{\infty} \phi(ax+b)^* \psi(x)\, dx \tag{11.6}$$

$$\text{we should try to write it as } \int_{-\infty}^{\infty} \phi(x)^* (???)\, dx \tag{11.7}$$

The right hand side of Eq.(11.5) is

$$\int_{-\infty}^{\infty} \phi(x)^* (T^\dagger \psi(x))\, dx \tag{11.8}$$

comparing the above expressions (12.8) *and* (11.7), *we will get the expression (???) which will give us the result of action of $T^\dagger$ on $\psi(x)$.*

Question: How can we manipulate Eq.(11.6) , and rewrite it in the desired form as in (11.7)?

(i) We can change the variable from x to new variable $t = ax + b$

(ii) That is correct, so let us then proceed.

Changing the integration variable we have

$$\int_{-\infty}^{\infty} (T\phi(x))^* \psi(x)\, dx = \int_{-\infty}^{\infty} \phi(ax+b)^* \psi(x)\, dx$$

$$= \int_{-\infty}^{\infty} \phi(t)^* \psi((t-b)/a\, (dt/a) \tag{11.9}$$

$$= \int_{-\infty}^{\infty} \phi(x)^* (1/a) \psi((x-b)/a)\, dx \tag{11.10}$$

This gives us the desired answer

$$T^\dagger \psi(x) = (1/a)\psi\big((t-b)/a\big). \tag{11.11}$$

Lecture 12

Hermitian and Unitary Operators

§1 Hermitian Operators

Definition 44 *An operator A is* **hermitian** *if $A^\dagger = A$.*

Definition 45 *A linear operator U is called* **unitary** *if $U^\dagger = U^{-1}$. In case of a finite dimensional vector space, it is equivalent to demanding*

$$UU^\dagger = I (or U^\dagger U = I).$$

Theorem **15 (When is an operator hermitian ?)** *Each of the following two statements give condition for hermiticity of an operator.*

(H1) An operator T is hermitian if and only if $(Tg, f) = (g, Tf)$ holds for all $f, g \in \mathcal{V}$.

(H2) In a finite dimensional vector space an operator T is self adjoint if and only if (f, Tf) is real $\forall f \in \mathcal{V}$.

Proof of (H1) : Let T be a hermitian operator. Using the definition of adjoint we have

$$(g, Tf) = (T^\dagger g, f)$$

or

$$(g, Tf) = (Tg, f) \qquad (\because T = T^\dagger)$$

Let $(g, Tf) = (Tg, f)$ for all f and g in the vector space. Then we get

$$(g, Tf) = (Tg, f) \qquad \text{(given)} \tag{12.1}$$
$$(g, Tf) = (g, T^\dagger f) \quad \text{(Use def of } T^\dagger) \tag{12.2}$$
$$(g, (T - T^\dagger)f) = 0 \tag{12.3}$$

holds $\forall g$ and f. Select $g = (T - T^\dagger)f$. This gives $\|(T - T^\dagger)f\| = 0$. Therefore,

$$(T - T^\dagger)f = 0 \qquad \forall f \in \mathcal{V}. \tag{12.4}$$

Hence $T = T^\dagger$

Proof of (H2) : Let (f, Tf) be real. Then

$$\begin{aligned}(f,Tf) &= (f,Tf)^* \quad \text{given} && (12.5)\\ &= (Tf,f) \quad (\text{ property of inner product }) && (12.6)\\ &= (f,T^\dagger f) \quad (\text{ def of adjoint }) && (12.7)\end{aligned}$$

Thus $(f, Tf) = (f, T^\dagger f)$ holds $\forall f \in \mathcal{V}$. This implies $(f, (T - T^\dagger)f) = 0$, hence $T = T^\dagger$. Therefore, T is hermitian.

Theorem)‖(**16** *If X is any operator we may write, $X = A + iB$, where A and B are hermitian operators.*

The proof is easy. We write $A = (X + X^\dagger)/2; \qquad B = (X - X^\dagger)/2i$ It is straight forward to verify that A and B are hermitian and that $X = A + B$.

§2 Unitary Operators

Theorem)‖(**17 (When is an Operator Unitary ?)** *In a finite dimensional vector space, $\mathcal{V}$, the following conditions on an operator X are equivalent.*

(U1) X *is unitary.*

(U2) $(Xf, Xg) = (f, g) \qquad \forall f, g \in \mathcal{V}$

(U3) $\|Xf\| = \|f\| \qquad \forall f \in \mathcal{V}$.

Proof of $(U1) \Rightarrow (U2)$: $(Xf, Xg) = (f, X^\dagger Xg) = (f, g)$

Proof of $(U2) \Rightarrow (U3)$: (U3) follows from (U2) by setting $g = f$ in (U2).

Proof of $(U3) \Rightarrow (U1)$: Given that $\|Xf\| = \|f\| \qquad \forall f \in \mathcal{V}$ we have $(Xf, Xf) = (f, f)$. This in turn gives $(f, X^\dagger Xf) = (f, f)$ or

$$(f, (X^\dagger X - I)f) = 0 \forall f \in \mathcal{V}$$

Thus $(X \dagger X - I) = 0$. This means that $X^\dagger X = I$. In finite dimensional spaces we then have the result that X is unitary.

)⦀(**Short Examples 5** *Consider the Hilbert space $\mathscr{L}^2(-\infty, \infty)$ of square integrable functions. We give some examples without a detailed discussion.*

(5a) Let $\hat{X}$ be defined as

$$\hat{X}\psi(x) = e^{ikx}\psi(x), \quad k \in \mathbb{R}$$

then

$$\hat{X}^\dagger\psi(x) = e^{-ikx}\psi(x), \quad k \in \mathbb{R}$$

(5b) The parity operator $\hat{P}$ defined by

$$\hat{P}\psi(\vec{x}) = \psi(-vecx)$$

is hermitian, $\hat{P}^\dagger = \hat{P}$.

(5c) The adjoint of translation operator $\hat{T}$ defined by

$$\hat{T}\psi(x) = \psi(x+a)$$

is given by .

$$\hat{T}^\dagger\psi(x) = \psi(x-a)$$

(5d) The translation operator $\hat{T}$ defined above is a unitary operator.

§3 Properties of Eigenvalues and Eigenvectors

Theorem)⦀(**18 (Eigenvalues and Eigenvectors of Hermitian Operators)** *Two important properties of hermitian operators are given below.*

(E1) The eigenvalues of a hermitian operators are real.

(E2) The eigenvectors of a hermitian operator corresponding to two distinct eigenvalues are orthogonal.

Proof of (E1) : Let λ be an eigenvalue and f be eigenvector of T with $Tf = \lambda f$. Since T is a hermitian operator we have

$$(x, Ty) = (Tx, y), \qquad \forall\, x, y \in \mathscr{V}. \tag{12.8}$$

Therefore setting $x = y = f$ in (12.8), we get $(f, Tf) = (Tf, f)$ we get

$$(Tf, f) = (f, Tf) \tag{12.9}$$

$$\Rightarrow \qquad (\lambda f, f) = (f, \lambda f) \tag{12.10}$$

$$\therefore \qquad (\lambda^* - \lambda)(f, f) = 0 \tag{12.11}$$

As $f \neq 0, (f, f) \neq 0$ and hence we must have $\lambda^* - \lambda = 0$ Therefore the eigenvalues of a hermitian operator are real.

Proof of (E2) :To prove that two eigenvectors corresponding to a different eigenvalues are orthogonal. Let $Tf = \lambda f$ and $Tg = \mu g$. and T be a hermitian operator $T\dagger = T$ and $\lambda \neq \mu$. Then proceeding as in proof of (E1)

$$(f, Tg) = (Tf, g) \qquad (\text{Since } T^\dagger = T)$$

We have

$$(f, \mu g) = (\lambda f, g)$$

or

$$\mu(f, g) = \lambda^*(f, g) = \lambda(f, g)$$

because the eigenvalues λ, μ are real. For $\lambda \neq \mu$ the above equation implies that $(f, g) = 0$. Hence f and g are orthogonal.

Theorem ⟧⟦**19 (Eigenvalues and Eigenvectors of Unitary Operators)** *(E3) If λ as an eigenvalue of a unitary operator then $|\lambda| = 1$. It can be equivalently written in several forms such as $\lambda^*\lambda = 1$, or $\lambda^* = 1/\lambda$, or $\lambda = e^{i\alpha}$ with $\alpha \in \mathbb{R}$.*

(E4) The eigenvectors of a unitary operator corresponding to two distinct eigenvalues are orthogonal.

Proof of (E3) : Let U be a unitary operator having λ as an eigenvalue and f as an eigenvector.

$$Uf = \lambda f \tag{12.12}$$

$$\Rightarrow \quad (Uf, Uf) = (f, f), \quad (\text{ since } U \text{ is unitary }) \tag{12.13}$$

$$\Rightarrow (\lambda f, \lambda f) = (f, f) \tag{12.14}$$

$$\Rightarrow \lambda * \lambda(f, f) = (f, f) \tag{12.15}$$

$$\Rightarrow \big(|\lambda|^2 - 1\big)(f, f) = 0 \tag{12.16}$$

$$\Rightarrow |\lambda|^2 = 1, \qquad \because (f, f) \neq 0 \tag{12.17}$$

Therefore, $|\lambda| = 1$. This means that λ is phase and $\lambda = \exp(i\alpha)$.

Proof of (E4) : Let λ and μ be two distinct eigenvalues of a unitary operator U and let f and g be the corresponding eigenvectors. Thus

$$Uf = \lambda f, \qquad Ug = \mu g, \qquad \text{and } \lambda \neq \mu. \tag{12.18}$$

Since U is unitary

$$(Uf, Ug) = (f, g) \tag{12.19}$$

$$\Rightarrow (\lambda f, \mu g) = (f, g) \tag{12.20}$$

$$\Rightarrow (\lambda^* \mu)(f, g) = (f, g) \tag{12.21}$$

Since $|\lambda| = 1$, we have $\lambda^*\lambda = 1$, or $\lambda^* = 1/\lambda$. The above equation then gives

$$[(\mu/\lambda) - 1](f, g) = 0 \tag{12.22}$$

$$\therefore (f, g) = 0, \qquad \because \mu \neq \lambda \text{ and } (\mu/\lambda - 1) \neq 0. \tag{12.23}$$

Lecture 13

Change of Orthonormal Basis and Dirac Notion

---

§1 Representation in an Orthonormal Basis

Let $\mathscr{E} = \{e_1, e_2, ..., e_N\}$ be an o.n. basis. A vector $f \in \mathcal{V}$ can be represented by its components with respect to a basis obtained by expanding f in terms of the basis:

$$f = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_3 e_3. \tag{13.1}$$

Thus the vector f is represented by a column

$$f \mapsto \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ .. \\ \alpha_n \end{bmatrix}. \tag{13.2}$$

The matrix representing an operator T has elements t_{jk} which can be obtained by expanding Te_k in terms of the basis elements $e_1, e_2, ..., e_N$ again.

$$Te_k = \sum_j t_{jk} e_j. \tag{13.3}$$

When $\mathscr{E}$ is an o.n. basis, the numbers α_k appearing in Eq.(13.2) are easily obtained by taking the scalar product of Eq.(13.1) with e_k. Thus we get

$$\alpha_1 = (e_1, f); \quad \alpha_2 = (e_2, f); \quad .....; \quad \alpha_k = (e_k, f). \tag{13.4}$$

Similarly, t_{jk} appearing in Eq.(13.3) are obtained taking scalar product of Eq.(13.3) with e_j and using the fact that $\mathscr{E}$ is o.n. set. This gives

$$t_{jk} = (e_j, Te_k). \tag{13.5}$$

Thus a vector f and an operator T are represented, respectively, by a column and a matrix as follows.

$$f \mapsto \begin{bmatrix} (e_1, f) \\ (e_2, f) \\ \ldots \\ (e_n, f) \end{bmatrix}, \tag{13.6}$$

$$T \mapsto \mathsf{T} = \begin{pmatrix} (e_1, Te_1) & (e_1, Te_2) & \cdots & \cdots & (e_1, Te_N) \\ (e_2, Te_1) & (e_2, Te_2) & \cdots & \cdots & (e_2, Te_N) \\ \cdots & \cdots & (e_j, Te_k) & \cdots & \cdots \\ (e_N, Te_1) & (e_N, Te_2) & \cdots & \cdots & (e_N, Te_N) \end{pmatrix}. \tag{13.7}$$

§2 Change of o.n. Basis

Thus representatives are easy to construct and expressions easy to remember for the case when the basis in an o.n. basis. Next we shall discuss how the elements of the column and the matrix change when a different basis is selected.

Theorem)|(**20** *Let* $\mathscr{E} = \{e_1, e_2, ..., e_N\}$ *and* $\mathscr{U} = \{u_1, u_2, ..., u_N\}$ *be two complete o.n. sets. Define an operator* X *by*

$$Xe_j = u_j.$$

The operator X *which takes an o.n. set* $\mathscr{E}$ *to another o.n. basis* $\mathscr{U}$ *is a unitary operator.*

Proof :

Because both the sets $\mathscr{E}$ and $\mathscr{U}$ are o.n. sets we have

$$(Xe_i, Xe_j) = (u_i, u_j) = \delta_{ij} = (e_i, e_j)$$

or

$$(e_i, X^\dagger X e_j) - (e_i, e_j) = 0 \qquad \Rightarrow (e_i, (X^\dagger X - I)e_j) = 0 \qquad \forall \quad i \text{ and } j.$$

This implies that $(f, (X^\dagger X - I)g) = 0$ for all f and g in the vector space. To see this expand $f = \sum_i \alpha_i e_i$, and $g = \sum_j \beta_j e_j$ and consider

$$(f, (X^\dagger X - I)g) = \left(\sum_i \alpha_i e_i), (X^\dagger X - I) \sum_j \beta_j e_j \right) \tag{13.8}$$

$$= \sum_i \sum_j \alpha_{ij} (e_i, (X^\dagger X - I)e_j) \tag{13.9}$$

$$= 0. \tag{13.10}$$

Hence $X^\dagger X - I = 0$, or X is unitary.

§3 Dirac Bra-Ket Notation

In this lecture I will explain the Dirac bra-ket notation for vector spaces with inner product. This notation is extremely useful for quantum mechanics. When an o.n. basis is selected in the vector space, Dirac notation is very convenient and several formulas such concerning representations and change of basis become simple and easy to remember.

The vector in a vector space are denoted by $|f\rangle$, called kets. The linear functionals on the vector space are denoted as $\langle g|$, called bra. The action of a linear functionals on a vector is written as a bracket $< g|f >$. The names bra and ket are derived from the bra(c)ket. A bracket is split into two parts which are named as the *bra and ket*

$$(.) \to (|) \to \langle|, \quad |\rangle.$$

In the inner product spaces a linear functionals ψ can be viewed as coming from some vector j so that

$$\psi(f) = (j, f)$$

and the distinction between the vectors and linear functionals can be dropped, if we take note of the correspondence of linear functional ψ with the vector j. We shall not talk about the linear functionals any more.

The scalar product of two vectors $|\psi\rangle$ and $|\phi\rangle$ is thus denoted by $\langle\phi|\psi\rangle$.

Let $\mathscr{E} = \{|e_1\rangle, |e_2\rangle, ..., |e_N\rangle\}$ be an o.n. basis. If a vector $|\psi\rangle$ is written as linear combination of the basis elements in $\mathscr{E}$,

$$|\psi\rangle = \sum \alpha_k |e_k\rangle$$

the coefficients will be given by the scalar products $\alpha_k = \langle e_k|\psi\rangle$. Substituting the value of α we can write the expansion of $|\psi\rangle$ as

$$|\psi\rangle = \sum_k |e_k\rangle\langle e_k|\psi\rangle.$$

The vector $|e_k\rangle\langle e_k|\psi\rangle$ appearing inside the sum can be thought of as a linear operator T_k, $(\equiv |e_k\rangle\langle e_k|)$, which on the vector $|\psi\rangle$ gives $|e_k\rangle\langle e_k|\psi\rangle$.

$$T_k|\psi\rangle = \Big(|e_k\rangle\langle e_k|\Big)|\psi\rangle.$$

The relation can be viewed as a statement that the relation $|\psi\rangle = \sum T_k|\psi\rangle$ holds for every vector $|\psi\rangle$. Thus $\sum T$ must be equal to identity operator. Hence we get

$$\sum_k |e_k\rangle\langle e_k| = \hat{I}.$$

This relation is referred to as *completeness relation.*

§4 Change Of O.N. Basis

Let x be a vector in a vector space. Let $\mathscr{E}$ and $\mathscr{U}$ be two o.n. bases. Let $\underline{\mathsf{x}}$ and $\underline{\underline{\mathsf{x}}}$ denote the components of the vector x w.r.t. the bases $\mathscr{E}$ and $\mathscr{U}$ respectively. Similarly let $\underline{\mathsf{T}}$ denote the matrix representing an operator T w.r.t. the first basis $\mathscr{E}$. Let $\underline{\underline{\mathsf{T}}}$ be the matrix w.r.t. the second basis $\mathscr{U}$.

Let us take the first o.n. basis as $\mathscr{E} = \{e_1, e_2, \ldots, e_N\}$ then we have the following expressions.

$$\underline{\mathsf{x}}_k = \langle e_k|x\rangle, \qquad [\underline{\mathsf{T}}]_{jk} = \langle e_j|T|e_k\rangle. \tag{13.11}$$

If we take the second o.n. basis as $\mathscr{U} = \{u_1, u_2, ..., u_N\}$ then we have the following expressions.

$$\underline{\underline{\mathsf{x}}}_i = \langle u_i|x\rangle, \qquad [\underline{\underline{\mathsf{T}}}]_{jk} =< u_j|T|u_k > \tag{13.12}$$

We want to find relations between
(i) components of $\underline{\mathsf{x}}$ and $\underline{\underline{\mathsf{x}}}$,
(ii) elements of the matrices $\underline{\mathsf{T}}$ and $\underline{\underline{\mathsf{T}}}$. The change of basis can be achieved by using the completeness relation. For example,

$$\begin{aligned} \underline{\mathsf{x}}_i &= \langle e_i|x\rangle = \langle e_i|I\rangle x = \langle e_i|\Big(\sum_k |u_k\rangle\langle u_k|\Big)|x\rangle && (13.13)\\ &= \sum_k \langle e_i|u_k\rangle\langle u_k|x\rangle = \sum_k \langle e_i|u_k\rangle \underline{\underline{\mathsf{x}}}_k. && (13.14) \end{aligned}$$

This gives the required relation between the components of the vector x w.r.t. the two basis sets $\mathscr{E}$ and $\mathscr{U}$. Similarly,

$$\begin{aligned} [\underline{\mathsf{T}}]_{jk} &= \langle e_j|T|e_k\rangle && (13.15)\\ &= \langle e_j|\Big(\sum_m |u_m\rangle\langle u_m|\Big)T\Big(\sum_n |u_n\rangle\langle u_n|\Big)|e_k\rangle && (13.16)\\ &= \sum_m\sum_n \langle e_j|u_m\rangle\langle u_m|T|u_n\rangle\langle u_n|e_k\rangle && (13.17)\\ &= \sum_m\sum_n \langle e_j|u_m\rangle\, [\underline{\underline{\mathsf{T}}}]_{mn}\, \langle u_n|e_k\rangle. && (13.18) \end{aligned}$$

This gives the change of basis formula for the matrices representing the operators. If we introduce a matrix S whose elements are given by

$$\mathsf{S}_{jm} = \langle e_j|u_m\rangle,$$

the last equation is seen to be like matrix multiplication

$$\underline{\mathsf{T}} = \mathsf{S}\,\underline{\underline{\mathsf{T}}}\mathsf{S}^\dagger. \tag{13.19}$$

Notice that the matrix S is unitary.

$$\mathsf{S}\mathsf{S}^\dagger = I = \mathsf{S}^\dagger\mathsf{S}. \tag{13.20}$$

To see this consider mn element of $\mathsf{S}^\dagger\mathsf{S}$

$$
\begin{aligned}
(\mathsf{S}^\dagger\mathsf{S})_{mn} &= \sum_k \mathsf{S}^\dagger_{mk}\mathsf{S}_{kn} = \sum_k \mathsf{S}^*_{km}\mathsf{S}_{kn} && (13.21)\\
&= \sum_k \langle e_k|u_m\rangle^* \langle e_k|u_n\rangle = \sum_k \langle u_m|e_k\rangle\langle e_k|u_n\rangle && (13.22)\\
&= \langle u_m| \left(\textstyle\sum_k |e_k\rangle\langle e_k| \right) |u_n\rangle && (13.23)\\
&= \big(\langle u_m|\big)\ \hat{I}\ \big(|u_n\rangle\big) = \langle u_m|u_n\rangle, \qquad \text{use completeness relation} && (13.24)\\
&= \delta_{mn}. && (13.25)
\end{aligned}
$$

Therefore, the matrix $\underline{\mathsf{S}}$ is a unitary matrix. This is just a statement of the fact that two orthonormal basis set are related by a unitary operator.

Question for you: Verify that the matrix $\underline{\mathsf{S}}$ is just the matrix for the linear operator $\hat{S}$ defined by

$$S|e_k\rangle = |u_k\rangle \tag{13.26}$$

in the basis set $\{|e_k\rangle | k = 1, 2, \ldots, N\}$.

Remark: ALL THESE RESULTS ARE VALID FOR FINITE DIMENSIONAL VECTOR SPACES ONLY. THIER USE IN CASE OF INFINITE DIMENSIONAL VECTOR SPACES REQUIRES A SEPARATE DETAILED DISCUSSION.

Topics and References for Further Study

The first part of the lecture notes should be supplemented by a study of the topics listed below.

- Jordan canonical forms, Halmos [1], Sec. 56-58
- Direct Sum and Quotient Spaces, Halmos [1], Sec 18-22.
- Bilinear and Multilinear Forms Halmos [1], Sec. 23, 29-31
- Tensor Product of Vector spaces and Tensors Pankaj Sharan[3], Ch 5, Tulsi Dass and S.K. Sharma[4], Ch. 3. Halmos Sec. 24-25

In continuation of the second part of the lecture note a study of the following topics is recommended.

- Orthogonal sum vector spaces, Halmos [1].
- Spectral Theorem, Sadri Hassani[2], Ch 4
- Infinite Dimensional Vector Spaces, von Neumann [5] and [6], Ch II.

.

References

[1] Paul R. Halmos, *Finite-Dimensional Vector Spaces*, Springer New York (1974).

[2] Sadri Hassani, *Mathematical Physics, Amodern Introduction to Its Foundations*, Springer, New York (1999)

[3] Pankaj Sharan, *Space Time and Gravitation*, Hindustan Book Agency, India (2009),

[4] Tulsi Dass and S. K. Sharma, *Mathematical Physics in Classical and Quantum Physics*, Universities Press, India (1999).

[5] John von Neumann, *Functional Operators Vol-II, The Geometry of Orthogonal Spaces.* Princeton University Press (1950).

[6] John von Neumann, Mathematical Foundations of Quantum Mechnaics, Princeton University Press (1955).