Bundled Lessons in Vector Spaces

Part-III Vector Spaces with Inner Product

A. K. Kapoor
http://0space.org/users/kapoor
akkapoor@cmi.ac.in; akkhcu@gmail.com

Contents

Inner Product Spaces **2**
§0.1 Norm and Inner Product in a Vector Space 3
§0.2 Parallelogram and Polarization Identities 4
§0.3 Cauchy Schwarz and Triangle Inequalities 4
§0.4 Triangle Inequality 6

1 Orthogonality **7**
§1.1 Orthogonality 8
§1.2 Grahm Schmidt orthogonalization 10

2 Complete Orthonormal Set **12**
§2.1 Complete orthonormal set 13
§2.2 Representation in an o.n. basis 15
§2.3 Change of o.n. basis in Dirac notation 16

3 Dirac Bra-Ket Notation **18**
§3.1 Change of o.n. basis in Dirac notation 19

Lesson-1 Inner Product Spaces

Lesson Overview

Syllabus Vector spaces with inner product; Properties of inner product; Parallelogram and polarization identities; Cauchy Schwarz and triangle inequalities.

Lesson Objecives To define and give examples of inner product spaces; To prove the parallelogram and polarization identities, Cauchy Schwarz and triangle inequalities. To give few examples.

Prerequisites Understanding of the definition of vector spaces

References

1. Halmos P. R. *Finite Dimensional Vector Spaces* Springer Verlag, East West Edition (1974).

2. Fraleigh J. B. *A First Course in Abstract Algebra*, Pearson Education Limited, Essex, (2014).

---

§0.1 Norm and Inner Product in a Vector Space

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

Norm and scalar product

Definition 1 **Norm** *of a vector f in a vector space V is a real number $\|f\|$ satisfying the following properties.*

(N-1) $\|f\| \geq 0$, and, $\|f\| = 0$ if and only if $f = 0$.

(N-2) $\|\alpha f\| = |\alpha| \|f\|$

(N-3) $\|f + g\| \ \leq \|f\| + \|g\|$ (Triangle Inequality)

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

Definition 2 *A* **scalar product**, *(or inner 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*

(S-1) $(f, f) \geq 0, \text{and } (f, f) = 0 \quad$ iff $f = 0$

(S-2) $(f, g) = (g, f)^*$

(S-3) $(f, \alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 (f, g_2) + \alpha_2 (f, g_2)$

(S-4) $(\alpha_1 f_1 + \alpha_2 f_2, g) = \alpha_1^* (f_1, g) + \alpha_2^* (f, g)$

We shall not discuss real vector spaces with inner product.

Remark: The property (S-4) can be proved from properties (S-2) and (S-3). Thus we have

$$\begin{aligned} (\alpha_1 f_1 + \alpha_2 f_2, g) &= [(g, \alpha_1 f_1) + (g, \alpha_2 f_2)]^* && (1) \\ &= [\alpha_1 (g, f_1) + \alpha_2 (g, f_2)]^* && (2) \\ &= \alpha_1^* (g, f_1) + \alpha_2^* (g, f_2) && (3) \end{aligned}$$

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

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

§0.2 Parallelogram and Polarization Identities

We shall now prove two important identities.

Parallelogram Identity

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

The proof of the parallelogram identity is easy. We begin from the l.h.s.

$$\begin{aligned}
\|f+g\|^2 + \|f-g\|^2 &= (f+g, f+g) + (f-g, f-g) \\
&= [(f,f) + (f,g) + (g,f) + (g,g)] + [(f,f) - (f,g) - (g,f) + (g,g)] \\
&= 2\|f\|^2 + 2\|g\|^2 \qquad (5)
\end{aligned}$$

Polarization Identity

The polarization identity is given by

$$4(f,g) = \|f+g\|^2 - \|f-g\|^2 + i\|f-ig\|^2 - i\|f+ig\|^2 \tag{6}$$

Proof :

$$\begin{aligned}
\|f+g\|^2 &= (f+g, f+g) = (f,f) + (f,g) + (g,f) + (g,g) && (7) \\
\|f-g\|^2 &= (f-g, f-g) = (f,f) - (f,g) - (g,f) + (g,g) && (8) \\
\|f-ig\|^2 &= (f-ig, f-ig) = (f,f) - i(f,g) + i(g,f) + (g,g) && (9) \\
\|f+ig\|^2 &= (f+ig, f+ig) = (f,f) + i(f,g) - i(g,f) + (g,g) && (10)
\end{aligned}$$

Adding Eq.(7) and Eq.(8) gives the parallelogram identity. In a similar fashion taking Eq.(7) - Eq.(8) + $i \otimes$ Eq.(9) - $i \otimes$ Eq.(10) 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 separately.

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.

§0.3 Cauchy Schwarz and Triangle Inequalities

As a preparation we first prove an intermediate result.

Theorem 1 *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.

$$\begin{aligned} F(\lambda) &= \|f - \lambda g\|^2 & (11) \\ &= (f - \lambda g, f - \lambda g) & (12) \\ &= (f,f) - \lambda(f,g) - \lambda^*(g,f) + \|\lambda\|^2(g,g) & (13) \end{aligned}$$

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

$$-(f,g) - (g,f) + 2\alpha(g,g) = 0 \qquad (14)$$

$$i(g,f) - i(f,g) + 2\beta(g,g) = 0 \qquad (15)$$

$$(16)$$

hence

$$\alpha = [(f,g) + (g,f)]/2(g,g) \qquad (17)$$

$$\beta = i[(f,g) - (g,f)]/2(g.g) \qquad (18)$$

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

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

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 $\|x\| \geq 0$ for all values of λ. We find the minimum of $\|x\|$ and set it ≥ 0

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

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.

§0.4 Triangle Inequality

We are now in a position to prove the triangle inequalities

$$\begin{aligned} \|f + g\| &\leq \|f\| + \|g\| && (19) \\ \|f - g\| &\leq \|f\| + \|g\| && (20) \end{aligned}$$

Proof : Consider

$$\begin{aligned} \|f + g\|^2 &= (f + g, f + g) && (21) \\ &= (f, f) + (f, g) + (g, f) + (g, g) && (22) \\ &= (f, f) + 2\mathrm{Re}(f, g) + (g, g) && (23) \\ &\leq \|f\|^2 + \|g\|^2 + 2|(f, g)|, \qquad [\because \mathrm{Re} z \leq |z|]. && (24) \end{aligned}$$

Using the Cauchy Schwarz inequality we get

$$\begin{aligned} \|f\|^2 + \|g\|^2 + 2|(f, g)| &\leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 && (25) \\ &= [\|f\| + \|g\|]^2 && (26) \end{aligned}$$

$\therefore$ We get the desired inequality

$$\|f + g\| \leq \|f\| \, + \|g\|$$

1 Orthogonality

Lesson Overview

Syllabus Vector spaces with inner product; Properties of inner product; Parallelogram and polarization identities; Cauchy Schwarz and triangle inequalities.

Lesson Objecives To define the concept of orthogonality of vectors.Orthognalization of linearly independent vectors.

Prerequisites Understanding of the definition of vector spaces; concept of inner product.

References

1. Halmos P. R. *Finite Dimensional Vector Spaces* Springer Verlag, East West Edition (1974).

2. Fraleigh J. B. *A First Course in Abstract Algebra*, Pearson Education Limited, Essex, (2014).

---

§1.1 Orthogonality

Definition 3 *We say that two vectors f and g are* **orthogonal** *if* $(f, g) = 0$

Lemma: If $g \neq 0$ then the vector

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

is orthogonal to g.

Proof: Consider

$$\begin{aligned} (g, x) &= (g, f - \frac{(g, f)}{(g, g)} g) = (g, f) - \frac{(g, f)}{(g, g)} (g, g) && (27) \\ &= (g, f) - (g, f) = 0 && (28) \end{aligned}$$

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

Definition 4 *Two vectors f and g are* **orthogonal** *if* $(f, g) = 0$.

Definition 5 *A set of vectors $\mathcal{X}$ is an* **orthogonal set** *if $\forall$ pair $x, y \in \mathcal{X}$, we have $(x, y) = 0$.*

Definition 6 *A set of vectors $\mathcal{X}$ is called* **orthonormal set** *if*

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

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

Definition 7 *A set $\{x_1, x_2, ..., x_r\}$ is an* **orthonormal set** *iff $(x_i, x_j) = \delta_{ij}$.*

⟫⟪**Short Examples 1 (Orthonormal Sets)** *We will now give several examples of orthonormal sets.*

(1a) 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.

(1b) Consider the vector space $\mathbb{C}^n$ with inner product of two column vectors x, y defined by $x^\dagger y$ an o.n. set is given by

$$x_1 = \begin{pmatrix} 1 & 0 & 0... & 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 & 1 & 0... & 0 \end{pmatrix}, \quad .... \quad x_n = \begin{pmatrix} 0 & 0 & 0... & 1 \end{pmatrix} \tag{29}$$

is an o.n. set.

(1c) Consider the complex vector space of all polynomials $\mathbb{P}_n$. We then have following examples of orthonormal sets.

1. *[(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.*

(1d) In the vector space of square integrable functions, $\mathcal{L}^2(-\infty,\infty)$, 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.

(1e) In the vector space of all functions defined on interval $[-\pi, \pi]$ and satisfying

$$f(x+2\pi) = f(x)$$

and inner product

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

an orthonormal set is

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

Definition 8 *An orthonormal set is called a* **complete orthonormal set** *if it is not contained in any larger orthonormal set.*

Theorem 3 *An orthogonal set $\mathcal{X} = \{x_1, x_2, ..x_r\}$ of non-zero vectors is linearly independent.*

Proof : Consider

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_r x_r = 0 \tag{30}$$

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

$$\alpha_1(x_1, x_1) = 0 \Rightarrow \alpha_1 = 0 \tag{31}$$

$$(\because x_1 \neq 0 \Rightarrow (x_1, x_1) \neq 0). \tag{32}$$

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 4 *If $\mathcal{U} = \{u_1, u_2, ..., u_n\}$ 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 $\mathcal{U}$

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

Remark: One can show that all the o.n. sets given above are complete.**??**

§1.2 Grahm Schmidt orthogonalization

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

Proof: 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.

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

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \qquad \text{(Bessel Inequality)} \tag{33}$$

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

$$\begin{aligned}
(y, y) &= (x - \sum_k \lambda_k u_k, x - \sum_j \lambda_j u_j) && (34)\\
&= (x, x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_j \sum_k \lambda_k^* \lambda_j (u_j, u_k) && (35)\\
&= (x, x) - \sum_k \lambda_k^*(u_k, x) - \sum_j \lambda_j (x, u_j) + \sum_k \lambda_k^* \lambda_k && (36)
\end{aligned}$$

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

$$\begin{aligned}
(y, y) &= (x, x) - \sum (x, u_k)(u_k, x) - \sum (u_j, x)(x, u_j) + \sum (x, u_j)(u_j, x) && (37)\\
&= (x, x) - \sum_k (x, u_k)(u_k, x) && (38)\\
&= (x, x) - \sum_k |(u_k, x)|^2 && (39)
\end{aligned}$$

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

$$\sum_k |(u_k, x)|^2 \leq \|x\|^2 \tag{40}$$

2 Complete Orthonormal Set

Lesson Overview

Syllabus Representation in an o.n. basis. Dirac notation

Lesson Objecives To explain construction of representation in an o.n. basis; to describe the results on change of basis; to briefly explain the Dirac notation

Prerequisites Understanding ofconcept of inner product and orthogonality; properties of a complete orthonormal set.

References

1. Halmos P. R. *Finite Dimensional Vector Spaces* Springer Verlag, East West Edition (1974).

2. Fraleigh J. B. *A First Course in Abstract Algebra*, Pearson Education Limited, Essex, (2014).

---

§2.1 Complete orthonormal set

Theorem 5 (Orthonormal Sets) *If $\mathfrak{U} = \{u_1, u_2, ..., u_n\}$ is any finite o.n. set in an inner product space having finite dimension, the following conditions $(P1)$ to $(P6)$ on $\mathfrak{U}$ are equivalent to each other.*

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

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

(P3) The subspace spanned by $\mathfrak{U}$ 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 $\mathfrak{U} \cup f/\|f\|$ would be an orthonormal set containing $\mathfrak{U}$. But this is impossible because $\mathfrak{U}$ is a complete set. Therefore, $f = 0$.

$\underline{(P2) \Rightarrow (P3):}$ Assume that $(P3)$ is not true. If the subspace spanned by $\mathfrak{U}$ is not whole space, there would exist a vector $f \neq 0$ such that f is not a linear combination of elements in $\mathfrak{U}$. 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.

$\underline{(P3) \Rightarrow (P4):}$ We are given that the subspace spanned by $\mathcal{U}$ 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 $\mathcal{U}$ is o.n. set we get $\alpha_k = (u_k, f)$. $\therefore f = \sum_k (u_k, f) u_k$.

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

$$f = \sum_i \lambda_i u_i; \quad \text{with} \quad \lambda_i = (u_i, f) \tag{41}$$

$$g = \sum_j \mu_j u_j, \quad \text{where} \quad \mu_j = (u_j, g) \tag{42}$$

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

$$\begin{aligned}
(f, g) &= \Big(\sum_i \lambda_i u_i, \sum_j \mu_j u_j \Big) && (43)\\
&= \sum_i \sum_j \lambda_i^* \mu_j (u_i, u_j) && (44)\\
&= \sum_{ij} \lambda_i^* \mu_j \delta_{ij} && (45)\\
&= \sum_i \lambda_i^* \mu_i && (46)
\end{aligned}$$

where, in step (45), we have used the orthogonality property $(u_i, u_j) = \delta_{ij}$. This gives us the desired result

$$(f, g) = \sum_i (f, u_i)(u_i, g) \tag{47}$$

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

$\underline{(P6) \Rightarrow (P1):}$

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 $\mathcal{U}$

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 6 *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.

§2.2 Representation in an o.n. basis

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 vectors in a vector space are denoted by $|f\rangle$, called kets, or ket vectors. The linear functionals on the vector space are denoted as $\langle g|$, called bra vectors. The action of a linear functionals on a vector is written as a bracket $\langle g|f\rangle$. The names bra and ket are derived from the bra(c)ket. In the inner product spaces all 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 $\mathcal{E} = \{|e_1\rangle, |e_2\rangle, ..., |e_N\rangle\}$ be an o.n. basis. If a vector $|\psi\rangle$ is written as linear combination of the basis elements in $\mathcal{E}$,

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

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 $P_k, (\equiv |e_k\rangle\langle e_k|)$, acting on the vector $|\psi\rangle$ giving $|e_k\rangle\langle e_k|\psi\rangle$:

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

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

$$\sum_k |e_k\rangle\langle e_k| = I$$

This relation is referred to as *completeness relation.*

§2.3 Change of o.n. basis in Dirac notation

Let x be a vector in a vector space. Let $\mathcal{E}$ and $\mathcal{U}$ be two o.n. bases. Let $\underline{x}$ and $\underline{\underline{x}}$ denote the components of the vector x w.r.t. the bases $\mathcal{E}$ and $\mathcal{U}$ respectively. Similarly, let $[\underline{T}]$ denote the matrix representing an operator T w.r.t. the first basis $\mathcal{E}$ and $[\underline{\underline{T}}]$ be the matrix w.r.t. the second basis $\mathcal{U}$. Let us take the first o.n. basis as $\mathcal{E} = e_1, e_2, \ldots, e_N$ then we have the following expressions.

$$x_k = \langle e_k|x\rangle, \qquad [\underline{T}]_{jk} =< e_j|T|e_k > \tag{48}$$

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

$$\underline{x}_i = \langle u_i|x\rangle, \qquad [\underline{T}]_{jk} = \langle u_j|T|u_k\rangle \tag{49}$$

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

$$\underline{x}_i = \langle e_i|x\rangle = \langle e_i|I|x\rangle = \langle e_i|\{\sum_k |u_k\rangle\langle u_k|\}|x\rangle \tag{50}$$

$$= \sum_k \langle e_i|u_k\rangle\langle u_k|x\rangle = \sum_k \langle e_i|u_k\rangle\, \underline{\underline{x}}_k. \tag{51}$$

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

$$[\underline{T}]_{jk} = \langle e_j|T|e_k\rangle \tag{52}$$

$$= \langle e_j|\Big\{\sum_m |u_m >< u_m|\Big\}T\Big\{\sum_n |u_n\rangle\langle u_n|\Big\}|e_k\rangle \tag{53}$$

$$= \sum_m\sum_n \langle e_j|u_m\rangle\langle u_m|T|u_n\rangle\langle u_n|e_k\rangle \tag{54}$$

$$= \sum_m\sum_n \langle e_j|u_m\rangle[\underline{\underline{T}}]_{mn}\langle u_n|e_k\rangle \tag{55}$$

This gives the change of basis formula for the matrices representing the operators. *All these results are valid for finite dimensional vector spaces only. Their use in case of infinite dimensional vector spaces requires a separate detailed discussion.*

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 vectors in a vector space are denoted by $|f\rangle$, called kets, or ket vectors. The linear functionals on the vector space are denoted as $\langle g|$, called bra vectors. The action of a linear functionals on a vector is written as a bracket $\langle g|f\rangle$. The names bra and ket are derived from the bra(c)ket. In the inner product spaces all 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 $\mathcal{E} = \{|e_1\rangle, |e_2\rangle, ..., |e_N\rangle\}$ be an o.n. basis. If a vector $|\psi\rangle$ is written as linear combination of the basis elements in $\mathcal{E}$,

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

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 $P_k, (\equiv |e_k\rangle\langle e_k|)$, acting on the vector $|\psi\rangle$ giving $|e_k\rangle\langle e_k|\psi\rangle$:

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

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

$$\sum_k |e_k\rangle\langle e_k| = I$$

This relation is referred to as *completeness relation.*

§3.1 Change of o.n. basis in Dirac notation

Let x be a vector in a vector space. Let $\mathcal{E}$ and $\mathcal{U}$ be two o.n. bases. Let $\underline{x}$ and $\underline{\underline{x}}$ denote the components of the vector x w.r.t. the bases $\mathcal{E}$ and $\mathcal{U}$ respectively. Similarly, let $[\underline{T}]$ denote the matrix representing an operator T w.r.t. the first basis $\mathcal{E}$ and $[\underline{\underline{T}}]$ be the matrix w.r.t. the second basis $\mathcal{U}$. Let us take the first o.n. basis as $\mathcal{E} = e_1, e_2, \ldots, e_N$ then we have the following expressions.

$$x_k = \langle e_k|x\rangle, \qquad [\underline{T}]_{jk} =< e_j|T|e_k > \tag{56}$$

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

$$\underline{\underline{x}}_i = \langle u_i|x\rangle, \qquad [\underline{\underline{T}}]_{jk} = \langle u_j|T|u_k\rangle \tag{57}$$

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

$$\underline{x}_i = \langle e_i|x\rangle = \langle e_i|I|x\rangle = \langle e_i|\{\sum_k |u_k\rangle\langle u_k|\}|x\rangle \tag{58}$$

$$= \sum_k \langle e_i|u_k\rangle\langle u_k|x\rangle = \sum_k \langle e_i|u_k\rangle\, \underline{\underline{x}}_k. \tag{59}$$

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

$$[\underline{T}]_{jk} = \langle e_j|T|e_k\rangle \tag{60}$$

$$= \langle e_j|\Big\{\sum_m |u_m >< u_m|\Big\}T\Big\{\sum_n |u_n\rangle\langle u_n|\Big\}|e_k\rangle \tag{61}$$

$$= \sum_m\sum_n \langle e_j|u_m\rangle\langle u_m|T|u_n\rangle\langle u_n|e_k\rangle \tag{62}$$

$$= \sum_m\sum_n \langle e_j|u_m\rangle[\underline{\underline{T}}]_{mn}\langle u_n|e_k\rangle \tag{63}$$

This gives the change of basis formula for the matrices representing the operators. *All these results are valid for finite dimensional vector spaces only. Their use in case of infinite dimensional vector spaces requires a separate detailed discussion.*