CHAPTER 5 [Hilbert Spaces]

Def. (5.1): (1) An inner product space is also called a pre-Hilbert space.
(2) An inner product space (is a pre-Hilbert space) X is called a Hilbert space if it is complete in the sense of a metric space.

Notation. We shall use $H$ to represent a Hilbert space.

Example (5.1): (1) $\mathbb{R}^n$ is a Hilbert space with inner product defined by: $(x, y) = \sum_{i=1}^{n} x_i y_i$; $x, y \in \mathbb{R}^n$.

Because we have proved that $\mathbb{R}^n$ is an inner product space and we know from analysis that $\mathbb{R}^n$ is also complete.
(2) $\mathbb{C}^n$ is a Hilbert space with inner product space defined by: $(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}$; $x, y \in \mathbb{C}^n$.
(3) The space $l^2$ of all complex sequences $x = \{x_i\}$ such that $\sum_{i=1}^{\infty} |x_i|^2 < \infty$ is an inner product space under the inner product defined by: $(x, y) = \sum_{i=1}^{\infty} x_i \overline{y_i}$; $y = \{y_i\} \in l^2$.

We also know that $l^2$ is complete, hence $l^2$ is a Hilbert space.
(4) Every finite dimensional inner product space is a Hilbert space.

Because every finite dimensional inner product space is a finite dimensional n.b.s. and we have proved in this that every finite dimensional n.b.s. is complete.
Theorem (5.4) [Schwarz's Inequality]

If \( x \) and \( y \) are two vectors in a Hilbert space, then

\[
|(x,y)| \leq \sqrt{(x,x)} \cdot \sqrt{(y,y)} = \|x\| \|y\| \rightarrow 0.
\]

and equality holds in (1) if and only if \( x \) and \( y \) are linearly dependent.

Pf: See corresponding proof in Inner product spaces, only replacing Inner product spaces by Hilbert spaces.

Theorem (5.5): The Inner product in a Hilbert space \( H \) is jointly continuous if it is a continuous function.

Proof: See corresponding Pf in Inner product spaces.

Theorem (5.6) [Parallelogram Law]

If \( x \) and \( y \) are any vectors in a Hilbert space \( H \), then

\[
\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)
\]

Pf: Same as for Inner product spaces.

Exercise (5.7): (1) Let \( X = l_p \), \( p > 1 \), \( p \neq 2 \) is a norm linear space but not a Hilbert space, because we have proved in Ch#4i that it is not an Inner product space.

(2) The space \( X = C[a,b] \) is not a Hilbert space, because we have proved that it is not an Inner product space.

Available at
www.mathcity.org
Recall: The line segment joining two given elements \( x \) and \( y \) of a space \( X \) is defined to be the set of all \( z \in X \) of the form: \( z = tx + (1-t)y \) for every real no. \( t \) such that \( 0 \leq t \leq 1 \).

A subset \( M \) of \( X \) is said to be Convex if for every \( x, y \in M \) the line segment joining \( x \) and \( y \) is contained in \( M \), i.e. \( z = tx + (1-t)y \in M \) for every \( t \) such that \( 0 \leq t \leq 1 \).

\[
\begin{array}{c}
\text{(Convex set)} \\
M \\
\end{array}
\]

(Not Convex)

Definition (5.8): If \( C \) is any non-empty subset of a Hilbert space \( H \), we define \( d(x, C) \) (the distance from \( x \) to \( C \)) by

\[
d(x, C) = \inf_{y \in C} \| x - y \|
\]

Theorem: A closed convex subset \( C \) of a Hilbert space \( H \) contains a unique vector of smallest norm.

Proof: Let \( C \) be a closed convex subset in \( H \). We show that it contains a unique vector of the smallest norm.

Since \( C \) is convex, so by above definition, it is non-empty, and contains \( \frac{1}{2} (x+y) \), wherever it contains \( x \) and \( y \).

Let \( d = \inf \{ \| x \| : x \in C \} \), then by def. of an infimum, there exists a sequence \( \{ x_n \} \subset C \)

\[
d = \inf \{ \| x \| : x \in C \}
\]
Such that $x_n \to d$. (Because of a result).

and by the convexity of $C$; $\frac{1}{2} (x_n + x_m)$ is in $C$.

and $\|x_n + x_m\| \geq d$ (by def. of $d$)

$\implies \|x_n + x_m\| \geq 2d$.

Now using Parallelogram Law, we have

$\|x_n + x_m\|^2 + \|x_n - x_m\|^2 = 2 (\|x_n\|^2 + \|x_m\|^2).

$\implies \|x_n - x_m\|^2 = 2 (\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2.

= 2 \|x_n\|^2 + 2 \|x_m\|^2 - \|x_n + x_m\|^2

\leq 2 \|x_n\|^2 + 2 \|x_m\|^2 - (2d)^2 \leq 2 \|x_n\|^2 + 2 \|x_m\|^2 - 4d^2

\implies 2d^2 + 2d^2 - 4d^2 = 0 \text{ as } m, n \to \infty \text{ (by above, } x_n, x_m \text{)}

$\implies \|x_n - x_m\|^2 \to 0 \text{ as } m, n \to \infty$

This shows that $\{x_n\}$ is a Cauchy sequence in $C$.

Now since $H$ is complete and $C$ is a closed subspace of $H$, so $C$ is complete, so the Cauchy sequence $\{x_n\}$ converges in $C$ i.e. $x_n \to x \in C$ (say); Then $x = \lim x_n$.

$\implies \|x\| = \|\lim x_n\| = \lim \|x_n\| (\because \text{ norm is a continuous function})

= d \quad (\because \lim n \to d)$

is also vector in $C$ with smallest norm.

Now we show that $x$ is unique. For this let us suppose $x'$ is another vector in $C$ with $x' \neq x$, which also has norm $d$ i.e. $\|x\| = d$.

Now $x, x'$ are in $C$ and $C$ is convex, so that $\frac{1}{2} (x+x')$ is also in $C$ and by applying Parallelogram law, we have

$\|\frac{x+x'}{2}\|^2 + \|\frac{x-x'}{2}\|^2 = 2 (\|\frac{x}{2}\|^2 + \|\frac{x'}{2}\|^2)$.
\[ \| x + x' \|_2^2 = 2 \left( \| x \|_2^2 \right) - \| x - x' \|_2^2 \]

\[ \leq \| x \|_2^2 + \| x' \|_2^2 - \| x - x' \|_2^2 \]

\[ = \frac{d^2}{2} + \frac{d^2}{2} = d^2 \]

ie \[ \| x + x' \|_2 \leq d \], which is a contradiction to the definition of \( d \) (\( d = \inf \{ \| x \| : x \in C \} \)). This contradiction arises due to our wrong supposition that \( x \neq x' \). Hence \( x = x' \) ie \( x \) is unique.

This completes the proof.

---

**Theorem (5.10):** In any Hilbert space, the inner product is related to the norm by the following identity, called Polarization identity, which is:

\[ \langle x, y \rangle = \frac{1}{2} \left( \| x + y \|_2^2 - \| x - y \|_2^2 + \| x + iy \|_2^2 - \| x - iy \|_2^2 \right) \]

\[ = \sum_{k=0}^{\infty} \frac{i^k}{k!} \| x + iy \|_2^2. \]

**Proof:** Same as the proof in inner product spaces.

---

**Theorem (Assignment):** If \( B \) is a complex Banach space whose norm obeys the parallelogram law and if an inner product is defined on \( B \) by the polarization identity, then \( B \) is a Hilbert space.

**Proof:**
Definitions (5.11):

1. Two vectors \( x \) and \( y \) in a Hilbert space \( H \) are said to be orthogonal if \( (x, y) = 0 \). We express symbolically the orthogonal vectors \( x \) and \( y \) by \( x \perp y \).
2. A vector \( x \) is said to be orthogonal to a non-empty set \( A \) if \( (x, y) = 0 \) for every \( y \) in \( A \); we write it as \( x \perp A \).
3. Two non-empty sets \( A \) and \( B \) in a Hilbert space \( H \) are said to be orthogonal if \( (x, y) = 0 \) for every \( x \) in \( A \) and every \( y \) in \( B \); we write it as \( A \perp B \).
4. A set \( A \) is said to be orthogonal if for every pair of elements \( x, y \) in \( A \) with \( x \neq y \), we have \( (x, y) = 0 \).

Remark: 1. \( x \perp 0 \) for every \( x \) in a Hilbert space \( H \).
2. If \( x \perp y \), then \( y \perp x \).
3. 0 is only vector orthogonal to itself.
4. If \( x \perp y \), \( x \perp z \), then \( x \perp y+z \) and \( x \perp ax \) for any scalar \( a \).
5. If \( x \perp y_n \), where \( y_n \to y \); then \( x \perp y \).

Proof: See proof in Inner product spaces.

Theorem (5.12) [Pythagorean Theorem]

1. If \( x \) and \( y \) are orthogonal vectors in a Hilbert space \( H \), then \( \|x+y\|^2 = \|x-y\|^2 = \|x\|^2 + \|y\|^2 \).
2. Generalized Pythagorean Thm:

If \( \{x_1, x_2, \ldots, x_n\} \) is an orthogonal set in a Hilbert space \( H \), then \( \|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2 \).

Proof: See proof in Inner product spaces.
Definition (5.14): 
If $M$ is any subset of a Hilbert space $H$, then the orthogonal complement of $M$, denoted by $M^\perp$, is defined as:

$$M^\perp = \{x \in H : (x,y) = 0 \text{ for every } y \in M\}$$

and also $M^{\perp\perp} = (M^\perp)^\perp = \{x \in H : (x,y) = 0 \text{ for every } y \in M^\perp\} = \{x \in H : x \perp M^\perp\}$.

Remark: From the above definition, it is clear that:

1. $[0]^\perp = H$
2. $H^\perp = [0]$

Theorem (5.14): Let $M_1$, $M_2$ be subsets of a Hilbert space $H$, then prove the following:

(I) $M_1 \subseteq M_1^{\perp\perp}$ is any subset of $H$ is contained in its double orthogonal complement.

(II) If $M_1 \subseteq M_2$, then $M_1^\perp \subseteq M_2^\perp$.

(III) $(M_1 \cup M_2)^\perp = M_1^\perp \cap M_2^\perp$ and $(M_1 \cap M_2)^\perp \supseteq M_1^\perp \cup M_2^\perp$.

(IV) $M_1^{\perp\perp} = M_1^{\perp\perp}$

(V) $M_1 \cap M_1^{\perp\perp} = [0]$

(VI) $M_1^\perp$ is a closed linear space.

Proof: (I)

(II) Let $x \in M_1^{\perp\perp}$.

$\Rightarrow (x,y) = 0$ for every $y \in M_1$ (by def.)

$\Rightarrow (x,y) = 0$ for every $y \in M_1$ ($\because M_1 \subseteq M_2$).

$\Rightarrow x \perp M_1$

$\Rightarrow x \in M_1^\perp$ (by def.)

So that $M_1^{\perp\perp} \subseteq M_1^\perp$. 
(II) Since $M_i \subseteq M_i \cup M_i$ and $M_i \subseteq M_i \cup M_i$ (always true)

$\implies (M_i \cup M_i)^+ \subseteq M_i^+$ and $(M_i \cup M_i)^+ \subseteq M_i^+$ (by (I)).

$\implies (M_i \cup M_i)^+ \subseteq M_i^+ \cap M_i^+ \implies (4)$

Now let $x \in M_i^+ \cap M_i^+ \implies x \in M_i$ and $x \in M_i^+$.

So by def: $(x, u) = 0$ for every $u \in M_i^+$ and $(x, v) = 0$ for every $v \in M_i^+$.

and so $(x, u) = 0$ for every $u \in M_i \cup M_i$.

$\implies x \in (M_i \cup M_i)^+$

so that $M_i^+ \cap M_i^+ \subseteq (M_i \cup M_i)^+ \implies (5)$

From (4) & (5); we get: $(M_i \cup M_i)^+ = M_i^+ \cap M_i^+$.

Next we show that $(M_i \cap M_i)^+ \supseteq M_i^+ \cup M_i^+$.

For this since $M_i \cap M_i \subseteq M_i$ and $M_i \cap M_i \subseteq M_i$,

$\implies M_i \subseteq (M_i \cap M_i)^+ \text{ and } M_i \subseteq (M_i \cap M_i)^+$ (by (II)).

$\implies M_i^+ \cup M_i^+ \subseteq (M_i \cap M_i)^+ \text{ as required.}$

(IV) By part (I), we have: $M_i \subseteq M_i^{++}$

and so part (II), we have: $(M_i^{++})^+ \subseteq M_i^+$

ie $M_i^{+++} \subseteq M_i^+ \implies (IV)$

Also by part (I), $M_i \subseteq (M_i^{++})^+$ ie $M_i \subseteq M_i^{+++} \implies (V)$

From (V) and (IV); we have $M_i = M_i^{+++}$.

(V) If $M_i \cap M_i^{++} = \emptyset$, then clearly $M_i \cap M_i^{++} = \emptyset \subseteq \{0\}$

ie $M_i \cap M_i^{++} \subseteq \{0\}$

if $M_i \cap M_i^{++} = \emptyset$, then let $x \in M_i \cap M_i^{++}$ implies $x \in M_i$ and $x \in M_i^{++}$.

Now since $x \in M_i^{++} \implies (x, x) = 0$ ie $\|x\| = 0$ ie $\|x\| = 0$ ie $x = 0 \subseteq \{0\}$

ie $x \subseteq \{0\}$. Therefore $M_i \cap M_i^{++} \subseteq \{0\}$. 

We show that $M^+$ is a closed linear subspace.

For this we first recall that "A subset $M$ of a linear space $X$ is a subspace of $X$ if for any $x,y \in M$ and any scalars $\alpha, \beta$, we have $\alpha x + \beta y \in M$.

Now let $x,y$ be any two elements in $M^+$ and $\alpha, \beta$ be any scalars. Then for any $u \in M$, we have:

$$(x,u) = 0 \text{ and } (y,u) = 0 \text{ and therefore:}$$

$$(\alpha x + \beta y, u) = (\alpha x, u) + (\beta y, u) = \alpha (x, u) + \beta (y, u) = \alpha \cdot 0 + \beta \cdot 0 = 0$$

Thus $(\alpha x + \beta y, u) = 0$ for any $u \in M$.

$\Rightarrow \alpha x + \beta y \in M^+$, which shows that $M^+$ is a subspace of $H$.

To complete the proof, it remains to show that $M^+$ is closed and in order to prove this, it is enough to show that if \{ $x_n$ \} is any convergent sequence in $M^+$ converging to a point $x$ (say) \( x_n \to x \), then \( x \in M^+ \).

Now for any $u \in M$, we can write:

$$(x,u) = \lim_{n \to \infty} (x_n,u) \quad \text{[as $x_n \to x$]}$$

$$= \lim_{n \to \infty} (x_n,u) \quad \text{[as inner product is continuous]}$$

$$= 0, \quad \text{because $x_n \in M^+$ in as $(x_n)$ is a seq. in $M^+$}$$

$$(x,u) = 0 \quad \text{for any $u \in M^+ \Rightarrow x \perp M$.}$$

$\Rightarrow x \in M^+$. Thus $M^+$ is closed linear subspace of $H$. Thus completing the proof.
Theorem 6.64: Let $M$ be a (closed) linear subspace of a Hilbert space $H$, then $M \cap M^\perp = \{0\}$.

**Proof:** Let $x \in M \cap M^\perp$. Then $x \in M$ and $x \in M^\perp$, hence $(x, y) = 0$ for every $y \in M$.

$$\Rightarrow (x, x) = 0$$

Because $x \in M^\perp$.

$$\Rightarrow \|x\|^2 = 0 \Rightarrow \|x\| = 0 \Rightarrow x = 0.$$

This shows that $0 \in M \cap M^\perp \Rightarrow \{0\} \subseteq M \cap M^\perp$.

But we know that part (c) of previous Theorem, $M \cap M^\perp = \{0\}$.

Hence $M \cap M^\perp = \{0\}$.

**Remark:** For sets $M$ and $M^\perp$, $M \cap M^\perp = \{0\}$ and for subspaces $M$ and $M^\perp$, $M \cap M^\perp = \{0\}$. The reason is that it is not necessary for $0$ to present in any subset but every subspace contains $0$.

**Recall:**

1. Any subspace of a linear space $X$ is convex.
2. For any subspace $M$ of a linear space $X$ and $x \in X$, the set $x + M = \{x + m : m \in M\}$ is convex.

Theorem 6.65: Let $M$ be a closed linear subspace of a Hilbert space $H$. Let $x$ be a vector in $M$ and let $d = d(x, M)$. Then there exists a unique vector $y_0 \in M$ such that $\|x - y_0\| = d$.

**Proof:** Let us set $C = x + M$ (the translation of $M$ by $x$). Then the set $C = x + M$ is a closed convex set and $d$ is the distance from the origin to $C$ (see figure).
So by Thm. (5.9), there exists a unique vector \( \mathbf{y_0} \) in \( \mathbf{C} \) such that \( \| \mathbf{y_0} \| = d = \inf \{ \| \mathbf{x} \| : \mathbf{x} \in \mathbf{C} \} \).

Since \( \mathbf{x} \in \mathbf{C} \), so by def. of \( \mathbf{C} \), \( \mathbf{x_0} = \mathbf{x} + \mathbf{y} \) for some \( \mathbf{y} \).

Let us put \( \mathbf{x} - \mathbf{x_0} = \mathbf{y_0} \), then the vector \( \mathbf{y_0} = \mathbf{x} - \mathbf{x_0} \) is easily seen to be in \( \mathbf{M} \) (\( \mathbf{x_0} = \mathbf{x} + \mathbf{y}, \mathbf{y_0} = \mathbf{x} - \mathbf{x_0} = \mathbf{x} - (\mathbf{x} + \mathbf{y}) = -\mathbf{y} \)) and \( \| \mathbf{y_0} \| = \| \mathbf{x} - \mathbf{y} \| \) (\( (\mathbf{x}) = \mathbf{x} - \mathbf{x_0} \)).

12. \( \| \mathbf{x} - \mathbf{y} \| = \| \mathbf{y_0} \| = d \) (from above)

Thus there exists a vector \( \mathbf{y_0} \) in \( \mathbf{M} \) such that \( \| \mathbf{x} - \mathbf{y} \| = d \).

It remains to prove the uniqueness of the vector \( \mathbf{y_0} \).

For this let \( \mathbf{y_1} \) be another vector in \( \mathbf{M} \) such that \( \| \mathbf{x} - \mathbf{y_1} \| = d \).

Then \( \mathbf{x} = \mathbf{y_0} + \mathbf{y_1} \) is a vector in \( \mathbf{C} \) (\( \mathbf{C} = \mathbf{x + M} \)) such that \( \mathbf{x} = \mathbf{y_0} + \mathbf{y_1} \) and \( \| \mathbf{x} \| = \| \mathbf{x} - \mathbf{y_0} \| = d \) is \( \| \mathbf{x} \| = d \),

which is a contradiction to the fact that there is a unique vector \( \mathbf{y_0} \) in \( \mathbf{C} \) such that \( \| \mathbf{x} \| = d \). This is because of our wrong supposition. Hence there exists a unique vector \( \mathbf{y_0} \) in \( \mathbf{M} \) such that \( \| \mathbf{x} - \mathbf{y_0} \| = d \).

**Theorem (5.17)** If \( \mathbf{M} \) is a proper closed linear subspace of a Hilbert space \( \mathbf{H} \), then there exists a non-zero vector \( \mathbf{z_0} \) in \( \mathbf{H} \) such that \( \mathbf{z_0} \perp \mathbf{M} \).

**Proof:** Let \( \mathbf{x} \) be a vector not in \( \mathbf{M} \) and let \( d = d(\mathbf{x}, \mathbf{M}) \), then by above Thm., there exists a unique vector \( \mathbf{y_0} \) in \( \mathbf{M} \) such that \( \| \mathbf{x} - \mathbf{y_0} \| = d \).
Let us take \( z_0 = x - y \) and observe that since
\( d > 0 \), \( z_0 \) is a non-zero vector in \( H \) \((||z_0|| = ||x - y|| = d > 0)\).

In order to show that \( z_0 \perp M \), it is enough to show
that \( z_0 \perp y \) for every \( y \in M \).

For this let \( \lambda \in \mathbb{C} \), then we have:

\[
||z_0 - \lambda y|| = ||x - y - \lambda y|| = ||x - (y + \lambda y)||
\]

\[
\geq d \quad \text{[by def of } d]\]

\[
= ||z_0|| \quad \text{[} ||z_0|| = ||x - y|| = d\}
\]

ie \( ||z_0 - \lambda y|| \geq ||z_0|| \)

\[
\Rightarrow ||z_0 - \lambda y||^2 \geq ||z_0||^2
\]

Then from this, we have:

\[
(z_0 - \lambda y, z_0 - \lambda y) \geq (z_0, z_0)
\]

\[
\Rightarrow (z_0, z_0) - (z_0, \lambda y) - (\lambda y, z_0) + (\lambda y, \lambda y) \geq (z_0, z_0)
\]

\[
\Rightarrow -\lambda (z_0, y) - \lambda (y, z_0) + \lambda \bar{y} (y, y) \geq 0
\]

\[
\Rightarrow -\lambda (z_0, y) - \lambda (\bar{z}_0, \bar{y}) + \lambda \bar{y} (y, y) \geq 0
\]

\[
\Rightarrow -\lambda (z_0, y) - \lambda (\bar{z}_0, \bar{y}) + |\lambda|^2 (y, y) \geq 0 \quad \Rightarrow 0
\]

Put \( \lambda = \bar{\mu}(x, y) \) for an arbitrary real number \( \mu \),

then \( 0 \) becomes:

\[
-\mu (z_0, y) - \mu (z_0, \bar{y}) + |\mu (z_0, y)|^2 (y, y) \geq 0.
\]

\[
\Rightarrow -\mu |(z_0, y)|^2 - \mu |(z_0, \bar{y})|^2 + \mu^2 |(z_0, y)|^2 (y, y) \geq 0. \quad \text{[since \( \mu \neq 0 \)]}
\]

\[
\Rightarrow -2 \mu |(z_0, y)|^2 + \mu^2 |(z_0, y)|^2 \geq 0 \quad \Rightarrow \quad 2
\]

Now put \( a = |(z_0, y)|^2 \) and \( b = ||y||^2 \), then from \( 0 \), we obtain:
-2 \mu a + \mu^2 ab \geq 0, \forall \text{ real nos: } \mu

\Rightarrow \mu a (\mu b - 2) \geq 0, \forall \text{ real nos: } \mu \rightarrow 0

However if \( a > 0 \), then 0 is impossible. For all sufficiently small positive \( \mu \) e.g. for \( a=1, b=1, \mu=1 \)
we get \(-1 > 0\), which is not possible.

We see from this that \( a=0 \) is only possibility. which means \( |(x_0, y)| = 0 \) \( \Rightarrow a = |(x_0, y)| \)

\[ |(x_0, y)| = 0 \Rightarrow (x_0, y) = 0 \Rightarrow x_0 \perp y \text{ for all } y \in M \]
\Rightarrow x_0 \perp M, \text{ which completes the proof.}

\text{Definition: (5.18)}
\text{let } M \text{ and } N \text{ be two subspaces of a linear space } L \text{. we define } M + N = \{x+y : x \in M, y \in N\}

Since \( M \) and \( N \) are subspaces, it is easy to see that \( M+N \) is also a subspace spanned (generated) by all vectors in \( M \) and \( N \) together i.e. \( M+N = [MUN] \).

\text{Definition: (5.19)}
\text{if } M+N = L, \text{ then we say that } L \text{ is the sum of the subspaces } M \text{ and } N.

This means that \( \forall \) vector in \( L \) is expressible as the sum of a vector in \( M \) and a vector in \( N \) i.e. \( \forall z \in L, \text{ then } z = x+y \text{ where } x \in M \land y \in N \).

if each vector \( z \) in \( L \) is expressible uniquely in the form of \( z = x+y \text{ with } x \in M \land y \in N \); then we say that \( L \) is the direct sum of the subspaces \( M \) and \( N \). Symbolically we write it as \( L = M \oplus N \).
Theorem: Let $M$ and $N$ be subspaces of a linear space $L$. Then $L = M \oplus N$ if and only if $M \cap N = \{0\}$.

Remark: The condition in this above theorem that the subspaces $M$ and $N$ have only the origin in common is often expressed by saying that $M$ and $N$ are disjoint.

Remark: Two non-empty sets $S_1, S_2$ of a Hilbert space $H$ are said to be orthogonal (written as $S_1 \perp S_2$) if $x \perp y$ for all $x \in S_1$ and for all $y \in S_2$.

Theorem (5.20): If $M$ and $N$ are closed linear subspaces of a Hilbert space $H$ such that $M \perp N$, then the linear subspace $M + N$ is closed.

Proof: To show that $M + N$ is closed, we need to show that all the limit points of $M + N$ are in $M + N$.

Let $\{z_n\}$ be a sequence in $M + N$ converging to a limit point $z$. It is enough to show that $z$ is in $M + N$.

Since $M \perp N$, we see that $M$ and $N$ are disjoint (i.e., $M \cap N = \{0\}$).

So by above Thm, the sum $M + N$ can be strengthened to the direct sum $M \oplus N$ and thus each $z_n$ can be expressed uniquely in the form $z_n = x_n + y_n$, where $x_n$ is in $M$ and $y_n$ is in $N$.

Since $x_n$ and $y_n$ are orthogonal ($\perp$ M \cap N), so by Pythagorean theorem, we have:

$$\|z_n - z_m\|^2 = \| (x_n + y_n) - (x_m + y_m) \|^2 = \| (x_n - x_m) \|^2 + \| (y_n - y_m) \|^2 = \| x_n - x_m \|^2 + \| y_n - y_m \|^2 \quad \text{(by Pyth.: Thm.)}$$
\[ \| x_n - x_m \|^2 + \| y_n - y_m \|^2 = \| x_n - x_m \|^2 < \varepsilon^2 \text{ for } m, n \geq N. \]

So from above, we have:

\[ \| x_n - x_m \|^2 + \| y_n - y_m \|^2 < \varepsilon^2 \text{ for } m, n \geq N. \]

\[ \| x_n - x_m \|^2 < \varepsilon^2 \text{ and } \| y_n - y_m \|^2 < \varepsilon^2 \text{ for } m, n \geq N. \]

which shows that \{x_n\} and \{y_n\} are Cauchy sequences in M and N respectively.

Also M and N are closed subspaces of the complete space (Hilbert space) H, so M and N are complete.

So by the completeness, there exists vectors \(x \in M\) and \(y \in N\) such that \(x_n \rightarrow x\) and \(y_n \rightarrow y\).

Since \(x + y\) is a vector in \(M+N\), so we have:

\[ x = \lim x_n = \lim (x_n + y_n) = \lim x_n + \lim y_n = x + y \in M+N \]

\[ x \in M+N. \text{ Thus } M+N \text{ is closed. Thus completing the proof.} \]
Theorem (5.21): [Projection Theorem]

Statement: If \( M \) is a closed linear subspace of a Hilbert space \( H \), then \( H = M \oplus M^\perp \).

Proof: Let \( M \) be a closed linear subspace of \( H \), then \( M^\perp \) is also a closed linear subspace of \( H \) (proved already).

Also \( M \) and \( M^\perp \) are orthogonal, because if \( x \in M \), then by def: \( (x,y)=0 \) for all \( y \in M \). Since \( x \) was chosen arbitrary in \( M \), so \( (x,y)=0 \) for every \( y \in M^\perp \) and every \( x \in M^\perp \) so that \( M^\perp \perp M \).

Thus \( M \) and \( M^\perp \) are orthogonal closed linear subspaces of \( H \), therefore \( M + M^\perp \) is also a closed linear subspace of \( H \) (by previous result).

We need to show that \( H = M \oplus M^\perp \).

First we show that \( H = M + M^\perp \).

On the contrary, assume that \( H = M + M^\perp \), then \( M + M^\perp \) is a proper closed linear subspace of \( H \). Then by Theorem (5.17), there exists a non-zero vector \( z_0 \) such that \( z_0 \perp (M + M^\perp) \). So \( z_0 \in (M + M^\perp)^\perp \) [by def. of orthogonal complement of a set].

Now \( M \subseteq M + M^\perp \Rightarrow (M + M^\perp)^\perp \subseteq M^\perp \)

And \( M^\perp \subseteq M + M^\perp \Rightarrow (M + M^\perp)^\perp \subseteq M^\perp \)

So that \( (M + M^\perp)^\perp \subseteq M^\perp \cap M^\perp = \{0\} \) \( (\because \{0\} \subseteq M^\perp \cap M^\perp) \)

Hence \( z_0 \in (M + M^\perp)^\perp \subseteq \{0\} \Rightarrow z_0 \in \{0\} \Rightarrow z_0 = 0 \) is a zero vector, which is a contradiction to the fact that \( z_0 \neq 0 \).

So our supposition was wrong and hence \( H = M + M^\perp \).
To complete the proof, it is enough to observe that
since $M$ and $M^\perp$ are orthogonal, so $M \cap M^\perp = \{0\}$
thus by Theorem (5.15), the statement $H = M + M^\perp$ can
be strengthened to the $H = M \oplus M^\perp$.
This completes the required result.

Orthogonal sets in Hilbert spaces:

**Def. (5.22):** An orthogonal set in a Hilbert space $H$ is a
non-empty subset of $H$ which consists of mutually
orthogonal unit vectors. That is, it is a non-empty subset
$\{e_i\}$ of $H$ with the following properties.

(i) $(e_i, e_j) = 0$ if $i \neq j$
(ii) $(e_i, e_i) = 1$ if $i = j$.

**Examples (5.23):** See examples following the definition of
orthogonal sets in Inner product spaces.

**Remark (5.24):** If $H = \{0\}$ i.e. $H$ contains only the
zero element, then it has no orthogonal set.
If $H$ contains a non-zero vector $x$, then we can construct
$e$ by normalizing $x$, that is: $e = \frac{x}{\|x\|}$. Then the single
element set $\{e\}$ is clearly an orthogonal orthogonal set
because $(e, e) = \|e\|^2 = \left\|\frac{x}{\|x\|}\right\|^2 = \frac{\|x\|^2}{\|x\|^2} = 1$.

Generally speaking, if $\{x_i\}$ is a non-empty set of
mutually orthogonal non-zero vectors in $H$, and if the
\( x_i \) are normalized by replacing each of them by
\[ e_i = \frac{x_i}{\|x_i\|}, \]
Then the resulting set \( \{e_i\} \) is orthonormal set.

**Theorem (5.25):** Let \( \mathcal{B} = \{e_1, e_2, \ldots, e_n\} \) be an orthonormal set in a Hilbert space \( H \). If \( x \) is any vector in \( H \), then
\[ \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \leq \|x\|^2. \] (Bessel's inequality)

and \( x = \sum_{i=1}^{n} (x, e_i)e_i \) for each \( j \)
\[ \text{i.e.} \quad x = \sum_{i=1}^{n} (x, e_i)e_i \in \mathcal{B}. \]

**Proof:** We have:
\[ 0 \leq \|x - \sum_{i=1}^{n} (x, e_i)e_i\| = (x - \sum_{i=1}^{n} (x, e_i)e_i, x - \sum_{i=1}^{n} (x, e_i)e_i). \]
\[ = (x, x) - \sum_{i=1}^{n} (x, e_i)(x, e_i) - \sum_{i=1}^{n} (x, e_i)(e_i, e_i) \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)(x, e_j)(e_i, e_j). \]
\[ = (x, x) - \sum_{j=1}^{n} \|x, e_j\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)(x, e_j)(e_i, e_j). \]
\[ = (x, x) - \sum_{i=1}^{n} |(x, e_i)|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)(x, e_j)(e_i, e_j). \]
\[ = (x, x) - \sum_{i=1}^{n} |(x, e_i)|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)(x, e_j)(e_i, e_j). \]
\[ = (x, x) - \sum_{j=1}^{n} \|x, e_j\|^2 - \sum_{i=1}^{n} |(x, e_j)|^2 + \sum_{i=1}^{n} |(x, e_i)|^2. \]
\[ \implies 0 \leq \sum_{j=1}^{n} \|x, e_j\|^2 \leq \|x\|^2, \] which is equivalent to \( \mathcal{B} \).
In order to show that \( x = \sum_{i=1}^{n} (x, e_i) e_i + s \), consider any \( e_j \in S \) where \( j = 1, 2, \ldots, n \).

Then \( x - \sum_{i=1}^{n} (x, e_i) e_i = (x, e_j) - \sum_{i=1}^{n} (x, e_i) (e_i, e_j) \)

\[ = (x, e_j) - \sum_{i=1}^{n} (x, e_i) (e_i, e_j) \]

\[ = (x, e_j) - (x, e_j) (e_j, e_j) \]

For all \( i \) the second term on the right-hand side is zero because \( e_i \) is an o.n. set,

\[ = (x, e_j) - (x, e_j) \cdot 1 \]

\[ = 0 \]

This shows that \( x = \sum_{i=1}^{n} (x, e_i) e_i \) for each \( j \)

\[ \Rightarrow x = \sum_{i=1}^{n} (x, e_i) e_i + s \].

Thus completing the proof.

Theorem (5.26): If \( \{e_i\} \) is an orthonormal set in a Hilbert space \( H \) and if \( x \) is any vector in \( H \), the the set \( S^c = \{e_i : (x, e_i) \neq 0\} \) is either empty or countable.

Proof:
Theorem (5.27) [Generalization of Bessel's Inequality]

If \( \{e_i\} \) is an orthonormal set in a Hilbert space \( H \), then
\[
\sum |(x,e_i)|^2 \leq \|x\|^2
\]
for every vector \( x \) in \( H \).

Proof: Let us define a set \( \mathcal{S} \) as:
\[
\mathcal{S} = \{ e_i : (x,e_i) \neq 0 \}
\]
Then by Thm (5.26), \( \mathcal{S} \) is either empty or Countable.

If \( \mathcal{S} \) is empty, then \( (x,e_i) = 0 \), so \( \sum |(x,e_i)|^2 \) is zero and so in this case (1) reduces to \( 0 \leq \|x\|^2 \) which is obviously true.

If \( \mathcal{S} \) is Countable, then \( \mathcal{S} \) is finite or Countably infinite.

When \( \mathcal{S} \) is finite, let it can be written in the form
\[
\mathcal{S} = \{ e_1, e_2, \ldots, e_n \}
\]
for some positive integer \( n \).

In this case, we denote \( \sum |(x,e_i)|^2 \) to be \( \sum |(x,e_i)|^2 \) which is clearly independent of the order in which the vectors of \( \mathcal{S} \) are arranged. So Inequality (1) reduces to
\[
\sum |(x,e_i)|^2 \leq \|x\|^2,
\]
which is the Bessel Inequality when \( \{e_i\} \) is finite orthonormal set and it has been proved already in Theorem (5.25).

When \( \mathcal{S} \) is Countably infinite: let the vectors in \( \mathcal{S} \) be arranged in some definite order as \( \mathcal{S} = \{ e_1, e_2, \ldots, e_n, \ldots \} \).

Now by the theory of "Absolutely Convergent Series" we know that if \( \sum |(x,e_i)|^2 \) converges, then every series obtained from this series by re-arranging its terms also converges and all such series have the same sum.
So we therefore can define: \( \sum_{i=1}^{n} |x_i e_i|^2 \) to be \( \sum_{i=1}^{n} |(x_i, e_i)|^2 \).

and it follows from the above remark that \( \sum_{i=1}^{n} |x_i e_i|^2 \) is a non-negative extended real number which depends only on \( S \) and not on the arrangement of vectors in \( S \). So in this case (1) reduces to:

\[
\sum_{i=1}^{n} |x_i e_i|^2 \leq \|x\|^2 \quad \Rightarrow \quad (2)
\]

Now from Bessel's inequality for finite case, we have:

\[
\sum_{i=1}^{n} |x_i e_i|^2 \leq \|x\|^2
\]

It follows that no partial sum of the series on the left of (2) can exceed \( \|x\|^2 \) and so it is clear that (2) (a) is true

\[
\Rightarrow \quad \sum_{i=1}^{n} |x_i e_i|^2 \leq \|x\|^2 \Rightarrow \sum_{i=1}^{n} |(x_i, e_i)|^2 \leq \|x\|^2
\]

This completes the proof.

**Recall:** (1) Let \( P \) be a set of elements. Suppose there is a binary relation defined between certain pairs \( a, b \) of \( P \) expressed symbolically by \( a \leq b \), with the properties

(i) If \( a \leq b \) and \( b \leq c \), then \( a \leq c \) (Transitivity)

(ii) If \( a \in P \), then \( a \leq a \) (Reflexivity)

(iii) If \( a \leq b \) and \( b \leq a \), then \( a = b \) (Antisymmetry)

Then \( P \) is said to be partially ordered set.

For example, if \( P \) is the set of all subsets of a given set \( X \), the set inclusion \( (A \subseteq B) \) gives a partial ordering of \( P \).

(2) If \( P \) is a partially ordered set, moreover if for any pair \( a, b \) in \( P \) either \( a \leq b \) or \( b \leq a \), then \( P \) is said to be completely (totally, linearly, simply) ordered set.
A completely ordered set is called a chain.

eg: The real numbers are completely ordered by the relation "a is less than or equal to b i.e. a ≤ b".

Zorn's Lemma: (only recall)

Let P be a non-empty partially ordered set with the property that every completely ordered subset of P has an upper bound. Then P contains at least one maximal element.

Theorem: Every non-zero Hilbert space H contains a complete orthonormal set.

Proof: Let \( H \neq \{0\} \) and \( M \) be the set of all subsets of \( H \) which are orthonormal. We define a partially ordering in \( M \) by the usual set inclusion, so that \( M \) is a partially ordered set.

Since \( H \neq \{0\} \), therefore \( y + e \) is a vector in \( H \), \( y \in M \) where \( y = \frac{x}{\|x\|} \). (by previous remark (M is an orthonormal set).

So the set \( M \) of all orthonormal sets is non-empty.

Now let \( C = \{ E_i \mid i \in \Delta \} \) be an increasing chain of orthonormal subsets in \( M \) (ie. \( E_i \subseteq E_j \)).

Then \( \bigcup E_i \) is the upper bound of \( C \).

Now \( M \) is a partially ordered set and every chain in \( M \) has its upper bound, so by "Zorn's Lemma", there exists a maximal element in \( M \). Let \( \Lambda \) be that element that is the set which is maximal in \( M \) so that \( H \) contains a complete orthonormal set.

Available at
www.mathcity.org
Let \( \{e_i\} \) be an orthonormal set in a Hilbert space \( H \) and let \( x \) be a vector in \( H \). Then

\[
x = \sum (x, e_i) e_i + \{e_i\}\]
Theorem (5.31): Let $H$ be a Hilbert space and let $\{e_i\}$ be an orthonormal set in $H$, then the following are equivalent.
(a) $\{e_i\}$ is complete.
(b) $x \perp [e_i] \Rightarrow x = 0$
(c) if $x$ is an arbitrary vector in $H$, then $x = \sum (x,e_i)e_i$.
(d) if $x$ is an arbitrary vector in $H$, then $\|x\|^2 = \sum (x,e_i)^2$.

Proof: (a) $\Rightarrow$ (b)
Suppose (a) is true i.e $\{e_i\}$ is complete $\Rightarrow \{e_i\}$ is maximal.
On contrary suppose that (b) is not true, then there exists a vector $x \neq 0$ such that $x \perp [e_i]$.
Define $e = \frac{x}{\|x\|}$ (Normalization of $x$), then the set $\{e_i,e\}$ is an orthonormal set, which properly contains $\{e_i\}$, but this contradicts the completeness of $\{e_i\}$. Hence (b) is true.
(b) $\Rightarrow$ (c)
Suppose that (b) is true i.e $x \perp [e_i] \Rightarrow x = 0$.
Now by (5.30), we have $x = \sum (x,e_i)e_i$ is orthogonal to $\{e_i\}$

$$x = \sum (x,e_i)e_i \perp [e_i]$$

So by (b), we get: $x = \sum (x,e_i)e_i = 0$

or $x = \sum (x,e_i)e_i$: For any vector $x$ in $H$. Hence (c) is true.
(c) $\Rightarrow$ (d) Suppose that (c) is true i.e $x = \sum (x,e_i)e_i$ for any vector $x$ in $H$.

Now $x = \sum (x,e_i)e_i = \sum_{i=1}^{\infty} (x,e_i)e_i$

Then $\|x\|^2 = \langle x,x \rangle = \langle x, \sum_{i=1}^{\infty} (x,e_i)e_i \rangle$

$$= \langle x, \lim_{n \to \infty} \sum_{i=1}^{n} (x,e_i)e_i \rangle$$

$$= \lim_{n \to \infty} \langle x, e_i \rangle (x,e_i) \quad \text{[Inner Product is Continuous]}$$
\[ ||x||^2 = \lim_{n \to \infty} \sum_{i=1}^{n} |(x, e_i)|^2 \]

Using \( \sum (x, e_i) e_i \) in place of \( \sum (x, e_i) e_i \), we get

\[ ||x||^2 = \sum |(x, e_i)|^2 \] Hence (d) is true.

Finally (d) \( \Rightarrow \) (a)

Suppose that (d) is true i.e. \( ||x||^2 = \sum |(x, e_i)|^2 \)

we show that (a) is true. On the contrary assume that (a) is not true i.e. \( \{e_i\} \) is not complete, then it is properly contained in an orthonormal set \( \{e, e'\} \).

so by definition of orthonormal set, we can say that \( e \) is orthogonal to \( e_i \).

Now \( ||e||^2 = \sum |(e, e_i)|^2 \) (\( \ast \) by (d))

\[ = \sum ||0||^2 \] (\( \ast \) is a vector, therefore we take norm)

\[ = ||0|| \]

\[ = 0 \]

i.e. \( ||e|| = 0 \)

and this contradicts the fact that \( ||e|| = 1 \)

so our supposition was wrong and hence \( \{e_i\} \) is complete.

Hence (a) is true.

This completes the required proof.

Remark(5.xx): let \( \{e_i\} \) be a complete orthonormal set and let \( x \) be an arbitrary vector in a Hilbert space \( H \).

Then the numbers \( (x, e_i) \) are called the Fourier coefficients of \( x \). The expression \( (x, e_i) e_i \) is called the Fourier expansion of \( x \) and the equation \( ||x||^2 = \sum |(x, e_i)|^2 \) is called Parseval's equation or formula. - all w.r.t. the particular complete orthonormal set \( \{e_i\} \) under consideration.
The Gram-Schmidt Orthogonalization Process:

It is a constructive procedure for converting a linearly independent set \( \{x_1, x_2, \ldots, x_n, \ldots\} \) into a corresponding orthonormal set \( \{e_1, e_2, \ldots, e_n, \ldots\} \) with the property that for each \( n \), the linear subspace spanned by \( \{e_1, e_2, \ldots, e_n\} \) is the same as that spanned by \( \{x_1, x_2, \ldots, x_n\} \).

We state this process in the form of the following theorem.

**Theorem (5.33):** Suppose that \( \{x_1, x_2, \ldots, x_n, \ldots\} \) is a linearly independent set in a Hilbert space \( H \), then there exists an orthonormal set \( \{e_1, e_2, \ldots, e_n, \ldots\} \) with the property that for each \( n \), the linear subspace spanned by \( \{e_1, e_2, \ldots, e_n\} \) is the same as that spanned by \( \{x_1, x_2, \ldots, x_n\} \).

**Proof:** Certainly \( x_1 = 0 \), because the set \( \{x_1, x_2, \ldots, x_n, \ldots\} \) is linearly independent.

We define \( y_1, y_2, \ldots \) and \( e_1, e_2, \ldots \) recursively as follows:

\[
y_1 = x_1 \quad e_1 = \frac{y_1}{\|y_1\|}
\]

Clearly the subspace spanned by \( x_1 \) and \( e_1 \) are the same.

\[
y_2 = x_2 - (x_2, e_1)e_1 \quad e_2 = \frac{y_2}{\|y_2\|}
\]

\[
y_3 = x_3 - (x_3, e_1)e_1 - (x_3, e_2)e_2 \quad e_3 = \frac{y_3}{\|y_3\|}
\]

\[
y_n = x_n - (x_n, e_1)e_1 - \cdots - (x_n, e_{n-1})e_{n-1} \quad e_n = \frac{y_n}{\|y_n\|}
\]

\[
y_{n+1} = x_{n+1} - \sum_{i=1}^{n} (x_{n+1}, e_i)e_i \quad e_{n+1} = \frac{y_{n+1}}{\|y_{n+1}\|}
\]
The process terminates if \( \{ x_n \} \) is a finite set, otherwise it continues indefinitely.

Also note that \( y_n \to 0 \) because \( y_n, x_2, \ldots, x_n \) are l.i. Thus \( e_n \) is well-defined if the definition of \( e_n \) is valid. From the construction, it is clear that \( e_1, e_2, \ldots, e_n \) is a linear combination of \( x_1, x_2 \) and \( x_1 \neq x_2 \) is a linear combination of \( e_1, e_2 \).

Similarly \( x_2 \) is a linear combination of \( e_1, e_2, e_3 \) and \( e_2 \) is a linear combination of \( x_1, x_2, x_3 \).

So by induction each \( x_n \) is a linear combination of \( e_1, e_2, \ldots, e_n \) and each \( e_n \) is a linear combination of \( x_1, x_2, \ldots, x_n \).

Thus the linear subspace spanned by the \( x \)'s is the same as that spanned by the \( e \)'s.

Now it remains to show that the set of \( e \)'s is an orthonormal set i.e. \( \{ e_1, e_2, \ldots, e_n, \ldots \} \) is orthonormal.

Now since \( e_i = \frac{y_i}{\| y_i \|} \)

\[ \Rightarrow \| e_i \| = \frac{\| y_i \|}{\| y_i \|} = 1 \]

by induction.

Now we show that \( (e_i, e_j) = 0 \) for \( i \neq j, i, j = 1, 2, \ldots \).

Consider \( (e_i, e_i) = (e_i, \frac{y_i}{\| y_i \|}) = \frac{1}{\| y_i \|} (e_i, y_i) \)

\[ = \frac{1}{\| y_i \|} (e_i, x_2 - (x_2, e_1) e_1) \]

\[ = \frac{1}{\| y_i \|} \left[ (e_i, x_2) - (e_i, (x_2, e_1) e_1) \right] \]

\[ = \frac{1}{\| y_i \|} \left[ (e_i, x_2) - \frac{(x_2, e_1) (e_i, e_1)}{\| y_i \|} \right] \]
\[
\Rightarrow (e_1, e_2) = \frac{1}{\|y\|} \left[ (x_1, x_2) - (e_1, x_2) \right] (\because \|e_2\| = 1)
\]
\[
= 0
\]
\[
\Rightarrow (e_1, e_2) = 0.
\]
Suppose that \((e_i, e_j) = \rho \quad \text{for} \quad i, j = 1, 2, \ldots, n-1.\)

Now \((e_n, e_j) = \left( \frac{y_n}{\|y\|}, e_j \right)\)
\[
= \frac{1}{\|y\|} (y_n, e_j)
\]
\[
= \frac{1}{\|y\|} \left( x_n - \sum_{i=1}^{n-1} (x_n, e_i) e_i \right) e_j
\]
\[
= \frac{1}{\|y\|} \left[ (x_n, e_j) - \sum_{i=1}^{n-1} (x_n, e_i) (e_i, e_j) \right]
\]
\[
= \frac{1}{\|y\|} \left[ (x_n, e_j) - \sum_{i=1}^{n-1} (x_n, e_i) (e_i, e_j) \right]
\]
\[
= \frac{1}{\|y\|} \left[ (x_n, e_j) - (x_n, e_j) (e_j, e_j) \right]
\]
\[
= \frac{1}{\|y\|} \left[ (x_n, e_j) - (x_n, e_j) \right]
\]
\[
= 0
\]

Hence by induction \(\{e_1, e_2, \ldots, e_n\}\) form a orthonormal set. Hence the result follows.

The Conjugate Space of a Hilbert Space \(H\):

Let \(H\) be a Hilbert Space. By \(H^*\), we denote the Conjugate Space of \(H\) (i.e., the set of all continuous linear transformations of \(H\) into \(\mathbb{C}\)). The elements of \(H^*\) are called Continuous Linear Functionals or briefly Functional.

One of the fundamental properties of a Hilbert Space \(H\) is the fact that there is a natural correspondence between the vectors in \(H\) and the functionals in \(H^*\) as we shall see below.
If \( y \) is a vector in Hilbert space \( H \), then the complex function \( F_y \) defined by \( F_y(x) = (x, y) \) for \( x \in H \) is linear, because for any \( x_1, x_2 \in H \) and scalar \( \alpha \), we have

\[
F_y(x_1 + x_2) = (x_1 + x_2, y) \quad [\text{by def. of } F_y(x)]
\]

\[
= (x_1, y) + (x_2, y)
\]

\[
= F_y(x_1) + F_y(x_2)
\]

and \( F_y(\alpha x) = (\alpha x, y) \)

\[
= \alpha (x, y)
\]

\[
= \alpha F_y(x).
\]

Moreover,

\[
|F_y(x)| = |(x, y)|
\]

\[
\leq ||x|| ||y|| \quad [\text{by Schwarz's inequality}].
\]

For all \( x \in H \). This inequality shows that \( F_y \) is bounded (criterion \( M = ||y|| \)) and hence continuous and is therefore a functional on \( H \) i.e \( F_y \in H^* \).

Since \( |F_y(x)| \leq ||x|| ||y|| \) (by above).

Thus we have: \( ||F_y|| \leq ||y|| \) (Taking \( x \) over \( x \) with \( ||x|| = 1 \)).

Even more equality is attained here i.e \( ||F_y|| = ||y|| \), because

This is clear when \( y = 0 \) (if \( y = 0 \) then \( ||F_y|| = 0 \) because norm is non-negative)

and if \( y \neq 0 \), then

\[
||x|| = (x, y) = F_y(0) \quad (\because F_y(x) = (x, y))
\]

\[
\leq |F_y(0)|
\]

\[
\leq ||F_y|| ||y||
\]

\[
\Rightarrow ||x|| \leq ||F_y|| ||y|| \Rightarrow ||y|| \leq ||F_y||
\]

So that \( ||F_y|| = ||y|| \)

We see that for every \( y \in H \), there exists a functional \( F_y \) in \( H^* \) such that \( ||F_y|| = ||y|| \)

In such case, we say that \( y \rightarrow F_y : H \rightarrow H^* \) is a norm preserving mapping of \( H \) into \( H^* \).
[If \( T : X \to Y \) is a linear mapping from a n.e.s. \( X \) into a n.e.s. \( Y \), then \( \|
abla T \| \) is called norm preserving mapping if \( \|
abla T \| \leq 1 \) for all \( x \in X \).]

Theorem (5.34) [Riesz Representation Theorem]

Let \( H \) be a Hilbert space and let \( \Phi \) be an arbitrary functional in \( H^* \), then there exists a unique vector \( y \in H \) such that \( \Phi(x) = \langle x, y \rangle \) for every \( x \in H \) and \( \|
abla \Phi \| = \|y\| \).

Proof: Let \( M \) be the null space (kernel) of \( \Phi \), that is \( M = \{x \in H : \Phi(x) = 0\} \).

Since \( \Phi \) is continuous (\( \Phi \) is functional), so by the continuity of \( \Phi \), the null space \( M \) of \( \Phi \) is a closed subspace of \( H \), by a result saying that "the null space of a non-zero continuous linear operator is a closed subspace".

If \( M = H \), then \( \Phi(x) = 0 * (b) \) (by def. of \( M \))

\[ = \langle x, \bar{y} \rangle \text{ for all } x \in H \text{ and the theorem is proved.} \]

If \( M \neq H \), then \( M \) is a proper closed subspace of \( H \) and so there exists a non-zero vector \( y_0 \) in \( H \) which is orthogonal to \( M \) i.e. \( y_0 \perp M \) (by 5.17).

Since \( y_0 \) is not in \( M \), then \( \Phi(y_0) \neq 0 \). [by def. of \( M \)].

For any vector \( x \) in \( H \), the vector \( z = x - \frac{\Phi(x)}{\Phi(y_0)} \cdot y_0 \) is in \( M \), because \( \Phi(z) = \Phi(x) - \frac{\Phi(x)}{\Phi(y_0)} \cdot \Phi(y_0) = 0 \).

Also since \( y_0 \perp M \), so that \( y_0 \bot z \) (\( \because z \in M \))

\[ \Rightarrow \langle \bar{z}, y_0 \rangle = 0 \Rightarrow \langle x - \frac{\Phi(x)}{\Phi(y_0)} \cdot y_0, y_0 \rangle = 0 \]

\[ \Rightarrow \langle x, y_0 \rangle - \left( \frac{\Phi(x)}{\Phi(y_0)} \cdot y_0, y_0 \right) = 0 \Rightarrow \langle x, y_0 \rangle - \frac{\Phi(x)}{\Phi(y_0)} \cdot \langle y_0, y_0 \rangle = 0 \]

\[ \Rightarrow \frac{\Phi(x)}{\Phi(y_0)} \cdot \langle y_0, y_0 \rangle = \langle x, y_0 \rangle \Rightarrow \Phi(x) = \frac{\Phi(y_0)}{\Phi(y_0)} \cdot \langle x, y_0 \rangle \]

\[ \Rightarrow \Phi(x) = \left( x, \frac{\Phi(y_0)}{\Phi(y_0)} \cdot y_0 \right) = \left( x, \frac{\Phi(y_0)}{\Phi(y_0)} \cdot y_0 \right) \]
Let \( y = \frac{F(x)}{\|x\|} \) ; Then from we have:

\[
F(x) = (x, y) \quad \text{for all } x \in \mathbb{R}.
\]

To complete the proof, it remains to show that \( y \) is unique.

For this, if we also have \( F(x) = (x, y') \) for all \( x \), then

\[
(x, y) = (x, y')
\]

\[
\Rightarrow (x, y) - (x, y') = 0
\]

\[
\Rightarrow (x, y - y') = 0 \quad \text{for all } x \in \mathbb{R}.
\]

For particular \( x = y - y' \), we get:

\[
(x, y - y') = 0 \Rightarrow \|y - y'\|^2 = 0 \Rightarrow y - y' = 0
\]

\[
\Rightarrow y = y'. \text{ Hence } y \text{ is unique.}
\]

Next we show that \( \|F\| = \|y\| \)

we have: \( F(x) = (x, y) \)

\[
|F(x)| = |(x, y)|
\]

\[
\leq \|x\| \|y\| \quad (\text{Schwartz inequality})
\]

and thus it follows that

\[
\|F\| \leq \|y\| \quad \text{Taking } \sup \text{ over both sides}
\]

Also \( \|y\|^2 = (y, y) = F(y) \)

\[
\leq |F(y)|
\]

\[
\leq \|F\| \|y\|
\]

\[
\Rightarrow \|y\| \leq \|F\|
\]

so that \( \|F\| = \|y\| \). This completes the proof.

Available at

www.mathcity.org