Hilbert Spaces

Definition. A complex inner product space (or pre-Hilbert space) is a complex vector space $X$ together with an inner product: a function from $X \times X$ into $\mathbb{C}$ (denoted by $\langle y, x \rangle$) satisfying:

1. $(\forall x \in X) \quad \langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ iff $x = 0$.
2. $(\forall \alpha, \beta \in \mathbb{C}) \quad (\forall x, y, z \in X), \quad \langle z, \alpha x + \beta y \rangle = \alpha \langle z, x \rangle + \beta \langle z, y \rangle$.
3. $(\forall x, y \in X) \quad \langle y, x \rangle = \overline{\langle x, y \rangle}$

Remarks.

(2) says the inner product is linear in the second variable;
(3) says the inner product is sesquilinear;
(2) and (3) imply $\langle \alpha x + \beta y, z \rangle = \bar{\alpha} \langle x, z \rangle + \bar{\beta} \langle y, z \rangle$, so the inner product is conjugate linear in the first variable.

Definition. For $x \in X$, let $\|x\| = \sqrt{\langle x, x \rangle}$.

Cauchy-Schwarz Inequality. $(\forall x, y \in X) \quad |\langle y, x \rangle| \leq \|x\| \cdot \|y\|$, with equality iff $x$ and $y$ are linearly dependent.

Proof. The result is obvious if $\langle y, x \rangle = 0$. Suppose $\gamma \equiv \langle y, x \rangle \neq 0$. Then $x \neq 0$, $y \neq 0$. Let $z = \gamma |\gamma|^{-1} y$. Then $\langle z, x \rangle = \gamma |\gamma|^{-1} \langle y, x \rangle = |\gamma| > 0$. Let $v = x \|x\|^{-1}$, $w = z \|z\|^{-1}$. Then $\|v\| = \|w\| = 1$ and $\langle w, v \rangle > 0$. Since $0 \leq \|v - w\|^2 = \langle v, v \rangle - 2 \Re \langle w, v \rangle + \langle w, w \rangle$, it follows that $\langle w, v \rangle \leq 1$ (with equality iff $v = w$, which happens iff $x$ and $y$ are linearly dependent). So $|\langle y, x \rangle| = \langle z, x \rangle = \|x\| \cdot \|z\| \langle w, v \rangle \leq \|x\| \cdot \|z\| = \|x\| \cdot \|y\|$. $\square$

Facts.

1' $(\forall x \in X) \|x\| \geq 0; \|x\| = 0$ iff $x = 0$.
2' $(\forall \alpha \in \mathbb{C})(\forall x \in X) \quad \|\alpha x\| = |\alpha| \cdot \|x\|$.
3' $(\forall x, y \in X) \quad \|x + y\| \leq \|x\| + \|y\|$.
Proof of (3'):
\[\|x + y\|^2 = \|x\|^2 + 2\Re \langle y, x \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle y, x \rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2.\]

Hence \(\|\cdot\|\) is a norm on \(X\); called the norm induced by the inner product \(\langle \cdot, \cdot \rangle\).

**Definition.** An inner product space which is complete with respect to the norm induced by the inner product is called a **Hilbert space.**

**Example.** \(X = \mathbb{C}^n\). For \(x = (x_1, \ldots, x_n)\) and \(y = (y_1, \ldots, y_n) \in \mathbb{C}^n\), let \(\langle y, x \rangle = \sum_{j=1}^n \overline{y}_j x_j\).

Then \(\|x\| = \sqrt{\sum_{j=1}^n |x_j|^2}\) is the \(l^2\)-norm on \(\mathbb{C}^n\).

**Examples of Hilbert spaces:**
- any finite dimensional inner product space
- \(l^2 = \{(x_1, x_2, x_3, \ldots) : x_k \in \mathbb{C}, \sum_{k=1}^\infty |x_k|^2 < \infty\}\) with \(\langle y, x \rangle = \sum_{k=1}^\infty \overline{y}_k x_k\)
- \(L^2(A)\) for any measurable \(A \subset \mathbb{R}^n\), with inner product \(\langle g, f \rangle = \int_A g(x)f(x)dx\).

**Incomplete inner product space**
\[C([a, b])\) with \(\langle g, f \rangle = \int_a^b \overline{g(x)}f(x)dx\)

\(C([a, b])\) with this inner product is not complete; it is dense in \(L^2([a, b])\), which is complete.

**Parallelogram Law.** Let \(X\) be an inner product space. Then \((\forall x, y \in X)\)
\[\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).\]

**Proof.** \[\|x + y\|^2 + \|x - y\|^2 = \langle x + y, x + y \rangle + \langle x - y, x - y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle = 2(\langle x, x \rangle + \langle y, y \rangle) = 2(\|x\|^2 + \|y\|^2).\] \(\square\)

**Polarization Identity.** Let \(X\) be an inner product space. Then \((\forall x, y \in X)\)
\[\langle y, x \rangle = \frac{1}{4} \left( \|y + x\|^2 - \|y - x\|^2 - i\|y + ix\|^2 + i\|y - ix\|^2 \right).\]

**Proof.** Expanding out the implied inner products, one shows easily that
\[\|y + x\|^2 - \|y - x\|^2 = 4\Re \langle y, x \rangle\) and \(\|y + ix\|^2 - \|y - ix\|^2 = -4\Im \langle y, x \rangle.\] \(\square\)

Note: In a real inner product space, \(\langle y, x \rangle = \frac{1}{4} (\|y + x\|^2 - \|y - x\|^2).\)

**Remark.** In an inner product space, the inner product determines the norm. The polarization identity shows that the norm determines the inner product. But not every norm on a vector space \(X\) is induced by an inner product.
Orthogonality. If \( \langle y, x \rangle = 0 \), we say \( x \) and \( y \) are orthogonal and write \( x \perp y \). For any subset \( A \subset X \), define \( A^\perp = \{ x \in X : \langle y, x \rangle = 0 \ \forall y \in A \} \). Since the inner product is linear in the second component and continuous, \( A^\perp \) is a closed subspace of \( X \). Also \( (\text{span}(A))^\perp = A^\perp \), \( \overline{A}^\perp = A^\perp \), and \( (\overline{\text{span}(A)})^\perp = A^\perp \).

The Pythagorean Theorem. If \( x_1, \ldots, x_n \in X \) and \( x_j \perp x_k \) for \( j \neq k \), then
\[
\left\| \sum_{j=1}^{n} x_j \right\|^2 = \sum_{j=1}^{n} \| x_j \|^2.
\]

Proof. If \( x \perp y \) then \( \| x + y \|^2 = \| x \|^2 + 2Re\langle y, x \rangle + \| y \|^2 = \| x \|^2 + \| y \|^2 \). Now use induction. □

Convex Sets. A subset \( A \) of a vector space \( X \) is called convex if \( (\forall x, y \in A) \ (\forall t \in (0,1)) \ (1 - t)x + ty \in A \).

Examples.

1. Every subspace is convex.
2. In a normed linear space, \( B(x, \epsilon) \) is convex for \( \epsilon > 0 \) and \( x \in X \).
3. If \( A \) is convex and \( x \in X \), then \( A + x \equiv \{ y + x : y \in A \} \) is convex.

Theorem. Every nonempty closed convex subset \( A \) of a Hilbert space \( X \) has a unique element of smallest norm.

Proof. Let \( \delta = \inf\{ \| x \| : x \in A \} \). If \( x, y \in A \), then \( \frac{1}{2}(x + y) \in A \) by convexity, and by the parallelogram law,
\[
\| x - y \|^2 = 2(\| x \|^2 + \| y \|^2) - \| x + y \|^2 \leq 2(\| x \|^2 + \| y \|^2) - 4\delta^2.
\]
Clearly, if \( \|x\| = \|y\| = \delta \), then \( \|x - y\|^2 \leq 4\delta^2 - 4\delta^2 = 0 \), so \( x = y \). For existence, choose \( \{y_n\}_{n=1}^{\infty} \subset A \) for which \( \|y_n\| \rightarrow \delta \). As \( n, m \rightarrow \infty \),

\[
\|y_n - y_m\|^2 \leq 2(\|y_n\|^2 + \|y_m\|^2) - 4\delta^2 \rightarrow 0,
\]

so \( \{y_n\} \) is Cauchy. By completeness, \( \exists y \in X \) for which \( y_n \rightarrow y \), and since \( A \) is closed, \( y \in A \). Also \( \|y\| = \lim \|y_n\| = \delta \).

**Corollary.** If \( A \) is a nonempty closed convex set in a Hilbert space and \( x \in X \), then \( \exists \) a unique closest element of \( A \) to \( x \).

**Proof.** Let \( z \) be the unique smallest element of the nonempty closed convex set \( A - x = \{y - x : y \in A\} \), and let \( y = z + x \). Then \( y \in A \) is clearly the unique closest element of \( A \) to \( x \).

**Orthogonal Projections onto Closed Subspaces**

**The Projection Theorem.** Let \( M \) be a closed subspace of a Hilbert space \( X \).

1. For each \( x \in X \), \( \exists \) unique \( u \in M \), \( v \in M^\perp \) such that \( x = u + v \). (So as vector spaces, \( X = M \oplus M^\perp \).)

Define the operators \( P : X \rightarrow M \) and \( Q : X \rightarrow M^\perp \) by \( P : x \mapsto u \) and \( Q : x \mapsto v \).

2. If \( x \in M \), \( Px = x \) and \( Qx = 0 \); if \( x \in M^\perp \), \( Px = 0 \) and \( Qx = x \).

3. \( P^2 = P \), Range\((P) = M \), Null Space\((P) = M^\perp \); \( Q^2 = Q \), Range\((Q) = M^\perp \), Null Space\((Q) = M \).

4. \( P, Q \in \mathcal{B}(X,X) \). \( \|P\| = 0 \) if \( M = \{0\} \); otherwise \( \|P\| = 1 \). \( \|Q\| = 0 \) if \( M^\perp = \{0\} \); otherwise \( \|Q\| = 1 \).

5. \( Px \) is the unique closest element of \( M \) to \( x \), and \( Qx \) is the unique closest element of \( M^\perp \) to \( x \).

6. \( P + Q = I \) (obvious by the definition of \( P \) and \( Q \)).

**Proof Sketch.** Given \( x \in X \), \( x + M \) is a closed convex set. Define \( Qx \) to be the smallest element of \( x + M \), and let \( Px = x - Qx \). Since \( Qx \in x + M \), \( Px \in M \). Let \( z = Qx \). Suppose \( y \in M \) and \( \|y\| = 1 \), and let \( \alpha = \langle y, z \rangle \). Then \( z - \alpha y \in x + M \), so \( \|z\|^2 \leq \|z - \alpha y\|^2 = \|z\|^2 - \alpha \langle z, y \rangle - \bar{\alpha} \langle y, z \rangle + |\alpha|^2 = \|z\|^2 - |\alpha|^2 \). So \( \alpha = 0 \). Thus \( z \in M^\perp \). Since clearly \( M \cap M^\perp = \{0\} \), the uniqueness of \( u \) and \( v \) in (1) follows. (2) is immediate from the definition. (3) follows from (1) and (2). For \( x, y \in X \), \( \alpha x + \beta y = (\alpha Px + \beta Py) + (\alpha Qx + \beta Qy) \), so by uniqueness in (1), \( P(\alpha x + \beta y) = \alpha Px + \beta Py \) and \( Q(\alpha x + \beta y) = \alpha Qx + \beta Qy \). By the Pythagorean Theorem, \( \|x\|^2 = \|Px\|^2 + \|Qx\|^2 \), so \( P, Q \in \mathcal{B}(X,X) \) and \( \|P\|, \|Q\| \leq 1 \). The rest of (4) follows from (2). Fix \( x \in X \). If \( y \in X \), then \( \|x - y\|^2 = \|Px - Py\|^2 + \|Qx - Qy\|^2 \).
If \( y \in M \), then \( \|x - y\|^2 = \|P x - y\|^2 + \|Q x\|^2 \), which is clearly minimized by taking \( y = P x \).
If \( y \in M^\perp \), then \( \|x - y\|^2 = \|P x\|^2 + \|Q x - y\|^2 \), which is clearly minimized by taking \( y = Q x \).
\( \square \)

**Corollary.** If \( M \) is a closed subspace of a Hilbert space \( X \), then \((M^\perp)^\perp = M\). In general, for any \( A \subset X \), \((A^\perp)^\perp = \text{span}\{A\}\), which is the smallest closed subspace of \( X \) containing \( A \), often called the **closed linear span** of \( A \).

### Bounded Linear Functionals and Riesz Representation Theorem

**Proposition.** Let \( X \) be an inner product space, fix \( y \in X \), and define \( f_y : X \to C \) by \( f_y(x) = \langle y, x \rangle \). Then \( f_y \in X^* \) and \( \|f_y\| = \|y\| \).

**Proof.** \( |f_y(x)| = |\langle y, x \rangle| \leq \|x\| \cdot \|y\| \), so \( f_y \in X^* \) and \( \|f_y\| \leq \|y\| \). Since \( |f_y(y)| = |\langle y, y \rangle| = \|y\|^2 \), \( \|f_y\| \geq \|y\| \). So \( \|f_y\| = \|y\| \).
\( \square \)

**Theorem.** Let \( X \) be a Hilbert space.

1. If \( f \in X^* \), then \( \exists \) a unique \( y \in X \) \( \ni f = f_y \), i.e., \( f(x) = \langle y, x \rangle \ \forall \ x \in X \).

2. The map \( \psi : X \to X^* \) given by \( \psi : y \mapsto f_y \) is a conjugate linear isometry of \( X \) onto \( X^* \).

**Proof.**

1. If \( f \equiv 0 \), let \( y = 0 \). If \( f \in X^* \) and \( f \not\equiv 0 \), then \( M \equiv f^{-1}(\{0\}) \) is a proper closed subspace of \( X \), so \( \exists z \in M^\perp \ni \|z\| = 1 \). Let \( \alpha = \overline{f(z)} \) and \( y = \alpha z \). Given \( x \in X \),

\[
u \equiv f(x)z - f(z)x \in M, \text{ so } 0 = \langle z, u \rangle = f(x)\langle z, z \rangle - f(z)\langle z, x \rangle = f(x) - \langle \alpha z, x \rangle = f(x) - \langle y, x \rangle, \text{ i.e., } f(x) = \langle y, x \rangle.
\]

Uniqueness: if \( \langle y_1, x \rangle = \langle y_2, x \rangle \ \forall \ x \in X \), then (letting \( x = y_1 - y_2 \) \( \|y_1 - y_2\|^2 = 0 \), so \( y_1 = y_2 \).

(2) follows immediately from (1), the previous proposition, and the conjugate linearity of the inner product in the first variable.
\( \square \)

**Corollary.** \( X^* \) is a Hilbert space with the inner product \( \langle g, f \rangle = \overline{\langle \psi^{-1}(g), \psi^{-1}(f) \rangle} \) (i.e., \( \langle f_y, f_z \rangle = \langle y, x \rangle = \langle x, y \rangle \)).

**Proof.** Clearly \( \langle f, f \rangle \geq 0 \), \( \langle f, f \rangle = 0 \) iff \( \psi^{-1}(f) = 0 \) iff \( f = 0 \), and \( \overline{\langle g, f \rangle} = \langle g, f \rangle \).

Also \( \langle f_y, \alpha_1 f_{x_1} + \alpha_2 f_{x_2} \rangle = \langle f_y, \alpha_1 x_1 + \alpha_2 x_2 \rangle = \alpha_1 \langle y, x_1 \rangle + \alpha_2 \langle y, x_2 \rangle = \alpha_1 \langle f_y, f_{x_1} \rangle + \alpha_2 \langle f_y, f_{x_2} \rangle, \) so \( \langle \cdot, \cdot \rangle \) is an inner product on \( X^* \). Since \( \langle f_y, f \rangle = \langle y, g \rangle = \|y\|^2 = \|f_y\|^2, \langle \cdot, \cdot \rangle \) induces the norm on \( X^* \). Since \( X^* \) is complete, it is a Hilbert space.
\( \square \)

**Remark.** Part (1) of the Theorem above is often called [one of] the Riesz Representation Theorem[s].
**Strong convergence/Weak convergence**

Let $X$ be a Hilbert space. We say $x_n \to x$ **strongly** if $\|x_n - x\| \to 0$ as $n \to \infty$. This is the usual concept of convergence in the metric induced by the norm, and is also called convergence in norm. We say $x_n \to x$ **weakly** if $\langle y, x_n \rangle \to \langle y, x \rangle$ as $n \to \infty$. (Other common notations for weak convergence are $x_n \rightharpoonup x$, $x_n \wto x$.) The Cauchy-Schwarz inequality shows that strong convergence implies weak convergence. Also, if $x_n \to x$ strongly, then $\|x_n\| \to \|x\|$ since $|\|x_n\| - \|x\|| \leq \|x_n - x\|$.

**Example.** (Weak convergence does not imply strong convergence if dim $X = \infty$). Let $X = l^2$. For $k = 1, 2, \ldots$, let $e_k = (0, \ldots, 0, 1, 0, \ldots)$ (so $\{e_k : k = 1, 2, \ldots\}$ is an orthonormal set in $l^2$).

**Claim.** $e_k \to 0$ weakly as $k \to \infty$.

**Proof.** Fix $y \in l^2$. Then $\sum_{k=1}^{\infty} |y_k|^2 < \infty$, so $y_k \to 0$. So $\langle y, e_k \rangle = \overline{y_k} \to 0$. \[\square\]

Note that $\|e_k\| = 1$, so $e_k$ does not converge to zero strongly.

**Remark.** If dim $X < \infty$, then weak convergence $\Rightarrow$ strong convergence (exercise).

**Theorem.** Suppose $x_n \to x$ weakly in a Hilbert space $X$. Then

(a) $\|x\| \leq \liminf_{k \to \infty} \|x_k\|

(b) If $\|x_k\| \to \|x\|$, then $x_k \to x$ strongly (i.e., $\|x_k - x\| \to 0$).

**Proof.**

(a) $0 \leq \|x - x_k\|^2 = \|x\|^2 - 2\Re\langle x, x_k \rangle + \|x_k\|^2$. By hypothesis, $\langle x, x_k \rangle \to \langle x, x \rangle = \|x\|^2$. So taking lim inf above, $0 \leq \|x\|^2 - 2\|x\|^2 + \liminf \|x_k\|^2$, i.e. $\|x\|^2 \leq \liminf \|x_k\|^2$.

(b) If $x_k \to x$ weakly and $\|x_k\| \to \|x\|$, then $\|x - x_k\|^2 = \|x\|^2 - 2\Re\langle x, x_k \rangle + \|x_k\|^2 \to \|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0$. \[\square\]

**Remark.** The **Uniform Boundedness Principle** implies that if $x_k \to x$ weakly, then $\|x_k\|$ is bounded.

**Orthogonal Sets**

**Definition.** Let $X$ be an inner product space. Let $A$ be a set (not necessarily countable). A set $\{u_\alpha\}_{\alpha \in A} \subset X$ is called an **orthogonal** set if $\langle \forall \alpha \neq \beta \in A \rangle \langle u_\beta, u_\alpha \rangle = 0$. (Often it is also assumed that each $u_\alpha \neq 0$.)
Orthonormal Sets

**Definition.** Let $X$ be an inner product space. A set $\{u_\alpha\}_{\alpha \in A}$ is called an orthonormal set if it is orthogonal and $(\forall \alpha \in A) \|u_\alpha\| = 1$. For each $x \in X$, define a function $\hat{x} : A \to \mathbb{C}$ by $\hat{x}(\alpha) = \langle u_\alpha, x \rangle$. The $\hat{x}(\alpha)$’s are called the Fourier coefficients of $x$ with respect to the orthonormal set $\{u_\alpha\}_{\alpha \in A}$.

**Theorem.** If $\{u_1, \ldots, u_k\}$ is an orthonormal set in an inner product space $X$, and $x = \sum_{j=1}^{k} c_j u_j$, then $c_j = \langle u_j, x \rangle$ for $1 \leq j \leq k$ and $\|x\|^2 = \sum_{j=1}^{k} |c_j|^2$.

**Proof.** $\langle u_i, x \rangle = \sum c_j \langle u_i, u_j \rangle = c_i$. Now use the Pythagorean Theorem. □

**Corollary.** Every orthonormal set is linearly independent.

**Example.** If $A$ is finite, say $A = \{1, 2, \ldots, n\}$, then for any $x \in X$, we know that the closest element of $\text{span}\{u_1, \ldots, u_n\}$ to $x$ is $\sum_{k=1}^{n} \langle u_k, x \rangle u_k$.

**Theorem.** (Gram-Schmidt process) Let $V$ be a subspace of an inner product space $X$, and suppose $V$ has a finite or countable basis $\{x_n\}_{n \geq 1}$. Then $V$ has a basis $\{u_n\}_{n \geq 1}$ which is orthonormal (we reserve the term “orthonormal basis” to mean something else); moreover we can choose $\{u_n\}_{n \geq 1}$ so that for all $m \geq 1$, $\text{span}\{u_1, \ldots, u_m\} = \text{span}\{x_1, \ldots, x_m\}$.

**Proof Sketch.** Define $\{u_n\}$ inductively. Start with $u_1 = \frac{x_1}{\|x_1\|}$. Having defined $u_1, \ldots, u_{n-1}$, let $v_n = x_n - \sum_{j=1}^{n-1} \langle u_j, x_n \rangle u_j$ and $u_n = \frac{v_n}{\|v_n\|}$. □

**Definition.** Let $A$ be a nonempty set. For each $\alpha \in A$, let $y_\alpha$ be a nonnegative real number. Define $\sum_{\alpha \in A} y_\alpha = \sup\{\sum_{\alpha \in F} y_\alpha : F \subset A \text{ and } F \text{ is finite}\}$.

**Proposition.** If $\sum_{\alpha \in A} y_\alpha < \infty$, then $y_\alpha \neq 0$ for at most countably many $\alpha$.

**Proof.** For each $k$, it is clear that $A_k = \{\alpha : y_\alpha > k^{-1}\}$ is a finite set. But $\{\alpha : y_\alpha \neq 0\} = \bigcup_{k=1}^{\infty} A_k$. □

**Definition.** Let $A$ be a nonempty set. Define $l^2(A)$ to be the set of functions $f : A \to \mathbb{C}$ for which $\sum_{\alpha \in A} |f(\alpha)|^2 < \infty$. Then $l^2(A)$ is a Hilbert space with inner product $\langle g, f \rangle = \sum_{\alpha \in A} g(\alpha) f(\alpha)$ and norm $\|f\|_2 = \sqrt{\langle f, f \rangle}$.

**Bessel’s Inequality.** Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set in a Hilbert space $X$, let $x \in X$, and let $\hat{x}(\alpha) = \langle u_\alpha, x \rangle$. Then $\sum_{\alpha \in A} |\hat{x}(\alpha)|^2 \leq \|x\|^2$.

**Proof.** By the previous Theorem, this is true for every finite subset of $A$. Take the sup. □

**Corollary.** Let $\{u_\alpha\}_{\alpha \in A}$, $x$ be as above. Then
(1) $\hat{x} \in l^2(A)$ and $\|\hat{x}\|_2 \leq \|x\|

(2) $\{\alpha \in A : \hat{x}(\alpha) \neq 0\}$ is countable.

**Theorem.** Let $X$ be a Hilbert space and let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set. Define $F : X \to l^2(A)$ (F is for Fourier) by $F(x) = \hat{x}$ where $\hat{x}(\alpha) = \langle u_\alpha, x \rangle$. Then $F$ is a bounded linear operator with $\|F\| = 1$, which maps $X$ onto $l^2(A)$.

**Proof.** Clearly $F$ is linear. By (1) of the Corollary, $F$ is bounded and $\|F\| \leq 1$. If $x = u_\alpha$ for some $\alpha \in A$, $\|\hat{x}\|_2 = 1 = \|x\|$, so $\|F\| = 1$. Given $f \in l^2(A)$, $f(\alpha) \neq 0$ only for a countable set $A_f \subseteq A$; enumerate them $\alpha_1, \alpha_2, \alpha_3, \ldots$. Let $x_k = \sum_{j=1}^{k} f(\alpha_j)u_j$. Clearly $\hat{x}_k(\alpha) = f(\alpha)$ for $\alpha_1, \ldots, \alpha_k$ and $\hat{x}_k(\alpha) = 0$ otherwise. So

$$\|\hat{x}_k - f\|_2^2 = \sum_{j=k+1}^{\infty} |f(\alpha_j)|^2 \to 0 \quad \text{as} \quad k \to \infty.$$  

Thus $\hat{x}_k \to f$ in $l^2(A)$, and in particular $\hat{x}_k$ is a Cauchy sequence in $l^2(A)$. Since each $x_k$ is a finite linear combination of the $u_\alpha$’s, $\|x_j - x_k\| = \|\hat{x}_j - \hat{x}_k\|_2$, so $\{x_k\}$ is Cauchy in $X$, so $x_k \to x$ in $X$ for some $x \in X$. For each $\alpha \in A$,

$$\hat{x}(\alpha) = \langle u_\alpha, x \rangle = \lim_{k \to \infty} \langle u_\alpha, x_k \rangle = \lim_{k \to \infty} \hat{x}_k(\alpha) = f(\alpha).$$

So $F(x) = f$ and $F$ is onto. \hfill \Box

**Theorem.** Let $X$ be a Hilbert space. Every orthonormal set in $X$ is contained in a maximal orthonormal set (i.e., an orthonormal set not properly contained in any orthonormal set).

**Proof.** Zorn’s lemma. \hfill \Box

**Corollary.** Every Hilbert space has a maximal orthonormal set.

**Theorem.** Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set in a Hilbert space $X$. The following conditions are equivalent:

(a) $\{u_\alpha\}_{\alpha \in A}$ is a maximal orthonormal set.

(b) The set of finite linear combinations of the $u_\alpha$’s is dense in $X$.

(c) $(\forall x \in X) \|x\|^2 = \sum_{\alpha \in A} |\hat{x}(\alpha)|^2$ (Parseval’s relation).

(d) $(\forall x, y \in X) \langle y, x \rangle = \sum_{\alpha \in A} y(\alpha)\hat{x}(\alpha)$.

(e) $(\forall x \in X)$ if $(\forall \alpha \in A) \langle u_\alpha, x \rangle = 0$ then $x = 0$.

**Proof.**

(a) $\Rightarrow$ (b): Let $V = \text{span}\{u_\alpha : \alpha \in A\}$ and $M = \bar{V}$. Then $M$ is a closed subspace. Since $\{u_\alpha\}$ is maximal, $V^\perp = \{0\}$, so $M^\perp = \{0\}$, so $M = X$. 

(b) $\Rightarrow$ (a): Let $\tilde{V}$ be the linear span of the finite linear combinations of the $u_\alpha$’s. Since $\tilde{V}$ is dense in $X$, it contains the closures of its projections on $X$.

(c) $\Rightarrow$ (d): By Parseval’s relation, $\langle x, y \rangle = \sum_{\alpha \in A} \langle u_\alpha, x \rangle \langle u_\alpha, y \rangle$.

(d) $\Rightarrow$ (c): This is trivial.
Hilbert Spaces

Theorem. Norm Convergence of Fourier Series

or an

Corollary . complete orthonormal set

is an isometry

Caution

Parseval’s Equality holds for this

Hilbert Spaces

Definition. An orthonormal set \{u_\alpha\} in a Hilbert space \(X\) satisfying the conditions in the previous theorem is called a complete orthonormal set (or a complete orthonormal system) or an orthonormal basis in \(X\).

Caution. If \(X\) is infinite dimensional, an orthonormal basis is not a basis in the usual definition of a basis for a vector space (i.e., each \(x \in X\) has a unique representation as a finite linear combination of basis elements). Such a basis in this context is called a Hamel basis.

Definition. Let \(X\) and \(Y\) be inner product spaces. A map \(T : X \to Y\) which is linear, bijective, and preserves inner products (i.e., \((\forall x, y \in X) \langle x, y \rangle = \langle Tx, Ty \rangle\) — this implies \(T\) is an isometry \(\|x\| = \|Tx\|\)) is called a unitary isomorphism.

Corollary. If \(X\) is a Hilbert space and \(\{u_\alpha\}_{\alpha \in A}\) is an orthonormal basis of \(X\), then the map \(F : X \to l^2(A)\) mapping \(x \mapsto \widehat{x}\) (where \(\widehat{x}(\alpha) = \langle x, u_\alpha \rangle\)) is a unitary isomorphism.

Corollary. Every Hilbert space is unitarily isomorphic to \(l^2(A)\) for some \(A\).

Norm Convergence of Fourier Series

Theorem. Let \(X\) be a Hilbert space, \(\{u_\alpha\}_{\alpha \in A}\) be an orthonormal set in \(X\), and let \(x \in X\). Let \(\{\alpha_j\}_{j \geq 1}\) be any enumeration of \(\{\alpha \in A : \langle u_\alpha, x \rangle \neq 0\}\). Then \(\|x\|^2 = \sum_{j=1}^{\infty} |\langle u_\alpha, x \rangle|^2\) (i.e. Parseval’s Equality holds for this \(x\)) iff \(\lim_{n \to \infty} \left\|x - \sum_{j=1}^{n} \langle u_\alpha, x \rangle u_\alpha\right\| = 0\) (i.e. the Fourier series \(\sum_{j=1}^{\infty} \widehat{x}(\alpha_j) u_\alpha\) converges to \(x\) in norm).

Proof. Let \(M_n = \text{span}\{u_{\alpha_1}, \ldots, u_{\alpha_n}\}\) and let \(P_n\) be the orthogonal projection onto \(M_n\) (so \(I - P_n\) is the orthogonal projection onto \(M_n^\perp\)). Then \(P_n x = \sum_{j=1}^{n} \langle u_{\alpha_j}, x \rangle u_{\alpha_j}\) and \(\|P_n x\|^2 = \sum_{j=1}^{n} |\langle u_{\alpha_j}, x \rangle|^2\). Also \(\|x\|^2 = \|P_n x\|^2 + \|(I - P_n)x\|^2\), so \(\|x\|^2 - \|P_n x\|^2 = \|x - P_n x\|^2\). Hence \(\lim_{n \to \infty} \|P_n x\|^2 = \|x\|^2\) iff \(\lim_{n \to \infty} \|x - P_n x\|^2 = 0\), which is the desired conclusion. (Note: If \(\{\alpha \in A : \langle u_\alpha, x \rangle \neq 0\}\) is finite, say \(\{\alpha_1, \ldots, \alpha_n\}\), then Parseval holds iff \(\|P_n x\|^2 = \|x\|^2\) iff \(x = P_n x\), i.e., \(x = \sum_{j=1}^{n} \langle u_{\alpha_j}, x \rangle u_{\alpha_j} \in M_n\).

Corollary. Let \(\{u_\alpha\}_{\alpha \in A}\) be an orthonormal set in a Hilbert space \(X\). Then \(\{u_\alpha\}\) is an orthonormal basis iff for each \(x \in X\) and each enumeration \(\{\alpha_j\}_{j \geq 1}\) of \(\{\alpha \in A : \langle u_\alpha, x \rangle \neq 0\}\), \(\lim_{n \to \infty} \left\|x - \sum_{j=1}^{n} \langle u_{\alpha_j}, x \rangle u_{\alpha_j}\right\| = 0\).
Cardinality of Orthonormal Bases

**Proposition.** $l^2(A)$ is unitarily isomorphic to $l^2(B)$ iff card($A$) = card($B$).

**Proposition.** Any pair of orthonormal bases in a Hilbert space have the same cardinality.

**Proposition.** A Hilbert space $X$ is separable iff it has a countable orthonormal basis.

**Remark.** For a separable Hilbert space $X$, one can show directly without invoking Zorn’s lemma that $X$ has a countable complete orthonormal set.

**Proof.** Clear if dim $X < \infty$. Suppose dim $X = \infty$. Let $z_1, z_2, \ldots$ be a countable dense subset. Apply Gram-Schmidt (dropping zero vectors along the way) to get an orthonormal sequence $u_1, u_2, \ldots$ whose finite linear combinations include $z_1, z, \ldots$, and thus are dense. □

**Theorem.** (Orthogonal projection in terms of orthonormal bases.) Let $X$ be a Hilbert space, and let $M$ be a closed subspace of $X$. Let $\{v_\beta\}_{\beta \in \mathcal{B}}$ be a complete orthonormal set in $M$, and let $\{w_\gamma\}_{\gamma \in \mathcal{C}}$ be a complete orthonormal set in $M^\perp$. Then $\{v_\beta\} \cup \{w_\gamma\}$ is a complete orthonormal set in $X$. The orthogonal projection of $X$ onto $M$ is $P_X = \sum_{\beta \in \mathcal{B}} (v_\beta, x)v_\beta$, and the orthogonal projection of $X$ onto $M^\perp$ is $Q_X = \sum_{\gamma \in \mathcal{C}} (w_\gamma, x)w_\gamma$.

**Proof.** Follows directly from $X = M \oplus M^\perp$ and the projection theorem. □

**Example.** (Orthogonal Polynomials in weighted $L^2$ spaces.) Fix $a, b \in \mathbb{R}$ with $-\infty < a < b < \infty$. Let $w \in C(a, b)$ with $w(x) > 0$ on $(a, b)$ and $\int_a^b w(x)dx < \infty$. The function $w$ is called the weight function. For example, take $w(x) = (1 - x^2)^{-1/2}$ on $(-1, 1)$. Define

$$L_w^2(a, b) = \left\{ f : f \text{ is measurable on } (a, b) \text{ and } \int_a^b |f(x)|^2 w(x)dx < \infty \right\}$$

and define $\langle g, f \rangle_w = \int_a^b g(x)f(x)w(x)dx$ for $f, g \in L_w^2(a, b)$. Then (after identifying $f$ and $g$ when $f = g$ a.e.), $L_w^2(a, b)$ is a Hilbert space.

**Claim.** Polynomials are dense in $L_w^2(a, b)$.

**Proof.** First note that if $f \in L^\infty(a, b)$, then $f \in L_w^2(a, b)$ since $\int_a^b |f(x)|^2 w(x)dx \leq \|f\|^2_w \int_a^b w(x)dx$, and thus $\|f\|_w \leq M\|f\|_\infty$, where $M = \left(\int_a^b w(x)dx\right)^{1/2} < \infty$. Given $f \in L_w^2(a, b)$, $\exists g \in C[a, b]$ for which $\|f - g\|_w < \frac{1}{2}\epsilon$ (exercise). By the Weierstrass Approximation Theorem, polynomials are dense in $(C[a, b], \|\cdot\|_\infty)$, so $\exists$ a polynomial $p$ for which $\|g - p\|_\infty < (2M)^{-1}\epsilon$. Then $\|f - p\|_w \leq \|f - g\|_w + \|g - p\|_w < \frac{1}{2}\epsilon + M\|g - p\|_\infty < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon$. □

Applying the Gram-Schmidt process to the linearly independent set $\{1, x, x^2, \ldots\}$ using the inner product of $L_w^2(a, b)$ produces a sequence of polynomials which are orthonormal in $L_w^2(a, b)$.

**Theorem.** The orthogonal polynomials in $L_w^2(a, b)$ are a complete orthonormal set in $L_w^2(a, b)$.

**Proof.** Finite linear combinations are dense. □