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 $\langle x, y\rangle$) satisfying:

(1) $(\forall\, x \in X)\ \langle x, x\rangle \geq 0$; $\langle x, x\rangle = 0$ iff $x = 0$.

(2) $(\forall\, \alpha, \beta \in \mathbb{C})\ (\forall\, x, y, z \in X)$, $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$.

(3) $(\forall\, x, y \in X)\ \langle y, x\rangle = \overline{\langle x, y\rangle}$

Remarks.

(2) says the inner product is linear in the first variable;

(3) says the inner product is conjugate symmetric;

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

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

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

Proof. The result is obvious if $\langle x, y\rangle = 0$. Suppose $\gamma \equiv \langle x, y\rangle \neq 0$. Then $x \neq 0 \neq y$. Let $z = \frac{\gamma}{|\gamma|}y$. Then $\langle x, z\rangle = \frac{\bar{\gamma}}{|\gamma|}\langle x, y\rangle = |\gamma| > 0$. Let $v = \frac{x}{\|x\|}$, $w = \frac{z}{\|z\|}$. Then $\|v\| = \|w\| = 1$ and $\langle v, w\rangle > 0$. Since $0 \leq \|v - w\|^2 = \langle v, v\rangle - 2\mathcal{R}e\langle v, w\rangle + \langle w, w\rangle$, $\langle v, w\rangle \leq 1$ (with equality iff $v = w$, which happens iff x and y are lin. dep.) So $|\langle x, y\rangle| = \langle x, z\rangle = \|x\| \cdot \|z\|\langle v, w\rangle \leq \|x\| \cdot \|z\| = \|x\| \cdot \|y\|$. $\square$

Facts.

(1′) $(\forall\, x \in X)\|x\| \geq 0$; $\|x\| = 0$ iff $x = 0$.

(2′) $(\forall\, \alpha \in \mathbb{C})(\forall\, x \in X)\ \|\alpha x\| = |\alpha| \cdot \|x\|$.

(3′) $(\forall\, x, y \in X)\ \|x + y\| \leq \|x\| + \|y\|$.

(Proof of (3′): $\|x+y\|^2 = \|x\|^2 + 2\mathcal{R}e\langle x,y\rangle + \|y\|^2 \le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \le \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$.) 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 $\langle x,y\rangle = \sum_{j=1}^m x_j \overline{y_j}$. 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 $\langle x,y\rangle = \sum_{k=1}^\infty x_k \overline{y_k}$
- for any $A^{\text{meas}} \subset \mathbb{R}^n$, $L^2(A)$ with $\langle f,g\rangle = \int_A f(x)\overline{g(x)}dx$.

Incomplete inner product space

$C[a,b]$ with $\langle f,g\rangle = \int_a^b f(x)\overline{g(x)}dx$

$C[a,b]$ with this inner product is *not* complete; it is dense in $L^2[a,b]$ with this inner product, 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: $\|x+y\|^2 + \|x-y\|^2 = \langle x+y, x+y\rangle + \langle x-y, x-y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle + \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle = 2(\langle x,x\rangle + \langle y,y\rangle) = 2(\|x\|^2 + \|y\|^2)$.)

Polarization Identity. Let X be an inner product space. Then $(\forall\, x,y \in X)$

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right).$$

(Sketch: Expanding out the implied inner products, one shows easily that

$$\|x+y\|^2 - \|x-y\|^2 = 4\mathcal{R}e\langle x,y\rangle \text{ and } \|x+iy\|^2 - \|x-iy\|^2 = 4\Im\langle x,y\rangle.)$$

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

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.

Theorem. Suppose $(X, \|\cdot\|)$ is a normed linear space. The norm $\|\cdot\|$ is induced by an inner product iff the parallelogram law holds in $(X, \|\cdot\|)$.

Proof Sketch. ($\Rightarrow$) see above. ($\Leftarrow$) Use the polarization identity to define $\langle \cdot, \cdot \rangle$. Then immediately $\langle x, x\rangle = \|x\|^2$, $\langle y, x\rangle = \overline{\langle x, y\rangle}$, and $\langle ix, y\rangle = i\langle x, y\rangle$. Use the parallelogram law to show $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$. Then show $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ successively for $\alpha \in \mathbb{N}$, $\frac{1}{2} \in \mathbb{N}$, $\alpha \in \mathbb{Q}$, $\alpha \in \mathbb{R}$, and finally $\alpha \in \mathbb{C}$. $\square$

Continuity of the Inner Product. Let X be an inner product space with induced norm $\|\cdot\|$. Then $\langle \cdot, \cdot\rangle : X \times X \to \mathbb{C}$ is continuous.

Proof. Since $X \times X$ and $\mathbb{C}$ are metric spaces, it suffices to show sequential continuity. Suppose $x_n \to x$ and $y_n \to y$. Then by the Schwarz inequality,

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| = |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \leq \|x_n - x\| \cdot \|y_n\| + \|x\| \cdot \|y_n - y\| \to 0.$$

$\square$

Orthogonality. If $\langle x, y\rangle = 0$, we say x and y are *orthogonal* and write $x \perp y$. For any subset $A \subset X$, define $A^{\perp} = \{x \in X : \langle x, y\rangle = 0 \,\forall\, y \in A\}$. Since the inner product is linear in the first component and continuous, $A^{\perp}$ is a *closed subspace*. Also $(\text{span}\{A\})^{\perp} = A^{\perp}$, $(\bar{A})^{\perp} = A^{\perp}$, 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 $j \neq k$, $\left\|\sum_{j=1}^{n} x_j\right\|^2 = \sum_{j=1}^{n} \|x_j\|^2$.

Proof. If $x \perp y$ then $\|x + y\|^2 = \|x\|^2 + 2\mathcal{R}e\langle x, y\rangle^0 = \|x\|^2$. Apply induction. $\square$

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, for $\varepsilon > 0$ and $x \in X$, $B(\varepsilon, x)$ is convex.

(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 $\frac{x+y}{2} \in A$ by convexity, and by the parallelogram law, $\|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) - \|x + y\|^2 \leq 2(\|x\|^2 + \|y\|^2) - 4\delta^2$. Uniqueness follows: 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 $\|y_n\| \to \delta$. As $n, m \to \infty$, $\|y_n - y_m\|^2 \leq 2(\|y_n\|^2 + \|y_m\|^2) - 4\delta^2 \to 0$, so $\{y_n\}$ is Cauchy. By completeness, $\exists\, y \in X$ for which $y_n \to y$, and since A is closed, $y \in A$. Also $\|y\| = \lim \|y_n\| = \delta$. $\square$

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.

Sketch. 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 $u \in M$, $v \in M^{\perp} \ni x = u + v$. (So as vector spaces, $X = M \oplus M^{\perp}$.)

Define the operators $P : X \to M$ and $Q : X \to M^{\perp}$ 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).

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 $\|y\| = 1$. Let $\alpha = \langle x, y\rangle$. Then $z - \alpha y \in x + M$, so $\|z\|^2 \le \|z - \alpha y\|^2 = \|z\|^2 - \alpha\langle y, z\rangle - \bar{\alpha}\langle z, y\rangle + |\alpha|^2 = \|z\|^2 - |\alpha|^2$. 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 obvious by uniqueness. (3) follows from (1) and (2). For $x, y \in X$, $\alpha x + \beta y = \overset{\swarrow \in M}{(\alpha Px + \beta Py)} + \overset{\swarrow \in M^{\perp}}{(\alpha Qx + \beta Qy)}$, 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 $P, Q \in L(X, X)$ and $\|P\|, \|Q\| \le 1$. 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 $\|x - y\|^2 = \|Px - y\|^2 + \|Qx\|^2$, which is clearly min iff $y = Px$. If $y \in M^{\perp}$, then $\|x - y\|^2 = \|Px\|^2 + \|Qx - y\|^2$, which is clearly min iff $y = Qx$. □

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

Bounded Linear Functionals and the Riesz Representation Theorem

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

Proof. $|f_y(x)| = |\langle x, y\rangle| \le \|x\| \cdot \|y\|$, so $f_y \in X^*$ and $\|f_y\| \le \|y\|$. Since $|f_y(y)| = |\langle y, y\rangle| = \|y\|^2$. $\|f_y\| \ge \|y\|$. So $\|f_y\| = \|y\|$. □

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., $\ni f(x) = \langle x, y\rangle \forall\, x \in X$.

(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 $x \in X$, $u \equiv f(x)z - f(z)x \in M$, so $0 = \langle u, z\rangle = f(x)\langle z, z\rangle - f(z)\langle x, z\rangle = f(x) - \langle x, \alpha z\rangle = f(x) - \langle x, y\rangle$, i.e., $f(x) = \langle x, y\rangle$. Uniqueness: if $\langle x, y_1\rangle = \langle x, y_2\rangle \,\forall\, x \in X$, 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 second variable.

$\square$

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

Proof. Clearly $\langle f, f\rangle \geq 0$, $\langle f, f\rangle = 0$ iff $\psi^{-1}(f) = 0$ iff $f = 0$, and $\overline{\langle f, g\rangle} = \langle g, f\rangle$. Also $\langle \alpha_1 f_{x_1} + \alpha_2 f_{x_2}\rangle = \langle f_{\bar\alpha_1 x_1 + \bar\alpha_2 x_2}, f_y\rangle = \overline{\langle \bar\alpha_1 x_1 + \bar\alpha_2 x_2, y\rangle} = \alpha_1 \overline{\langle x_1, y\rangle} + \alpha_2 \overline{\langle x_2, y\rangle} = \alpha_1\langle f_{x_1}, f_y\rangle + \alpha_2\langle f_{x_2}, f_y\rangle$, so $\langle\cdot,\cdot\rangle$ is an inner product on X^*. Since $\langle f_y, f_y\rangle = \overline{\langle y, y\rangle} = \|y\|^2 = \|f_y\|^2$, $\langle\cdot,\cdot\rangle$ 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, also called convergence in norm. We say $x_n \to x$ *weakly* if $(\forall\, y \in X)$ $\langle x_n, y\rangle \to \langle x, y\rangle$ as $n \to \infty$. (Other common notations for weak convergence: $x_n \rightharpoonup x$, $x_n \xrightarrow{w} x$.)

Example. (Weak convergence $\not\Rightarrow$ strong convergence if $\dim X = \infty$). Let $X \in l^2$. For $k = 1, 2, \ldots$, let $e_k = (0, \ldots, 0, \underset{\swarrow k^{\text{th}} \text{ entry}}{1}, 0 \ldots)$ (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 $\langle e_k, y\rangle = \overline{y_k} \to 0$. $\square$

Note that $\|e_k\| = 1$, so $e_k \not\to 0$ 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) $0 \leq \|x - x_k\|^2 = \|x\|^2 - 2\mathcal{R}e\langle x, x_k\rangle + \|x_k\|^2$. 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 $\|x - x_k\|^2 = \|x\|^2 - 2\mathcal{R}e\langle x, x_k\rangle + \|x_k\|^2 \to \|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0$.

$\square$

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 $(\forall\, \alpha \neq \beta \in A)$ $\langle u_\alpha, u_\beta\rangle = 0$. (often include also that $u_\alpha \neq 0$).

Orthonormal Sets

Definition. Let X be an inner product space. A set $\{u_\alpha\}_{\alpha\in A}$ is called an *orthonormal set* if $(\forall\, \alpha \neq \beta \in A)\langle u_\alpha, u_\beta\rangle = 0$ (ortho-) and $(\forall\, \alpha \in A)\|u_\alpha\| = 1$ (normal). For each $x \in X$, define a function $\widehat{x} : A \to \mathbb{C}$ by $\widehat{x}(\alpha) = \langle x, u_\alpha\rangle$. The $\widehat{x}(\alpha)$'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 $c_j = \langle x, u_j\rangle$ for $1 \leq j \leq k$ and $\|x\|^2 = \sum_{j=1}^{l} |c_j|^2$ ($\langle x, u_i\rangle = \sum c_j\langle u_j, u_i\rangle = c_i$ and then 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 span$\{u_1, \ldots, u_n\}$ to x is $\sum_{k=1}^{n}\langle x, u_k\rangle u_k$.

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 $m \geq 1$, span$\{u_1, \ldots, u_m\}$ = span$\{x_1, \ldots, x_m\}$.

Sketch. Define $\{u_n\}$ by induction: $u_1 = \frac{x_1}{\|x_1\|}$. Having defined $u_1, \ldots, u_{n-1}$, let $v_n = x_n - \sum_{j=1}^{n-1}\langle x_n, u_j\rangle u_j$ and $u_n = \frac{v_n}{\|v_n\|}$. $\square$

Theorem. Let V be a finite dimensional subspace of a Hilbert space X. Let $\{u_1, \ldots, u_n\}$ be a basis for V which is orthonormal, and let P be the orthogonal projection of X onto V. Then $Px = \sum_{j-1}^{n} \langle x, u_j \rangle u_j$ and $\|x\|^2 = \|Px\|^2 + \|Qx\|^2 = \sum_{j=1}^{n} |\langle x, u_j \rangle|^2 + \|Qx\|^2$.

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

Remark. This definition is equivalent to the integral of nonnegative functions $f \in L^+(\mu)$ where μ is counting measure on A (defined on $\mathcal{P}(A)$): if $f(\alpha) = y_\alpha$, then $\sum_{\alpha \in A} y_\alpha = \int_A f d\mu$.

Definition. Let A be a nonempty set. Define $l^2(A) = L^2_{\mathbb{C}}(\mu)$ (i.e., functions $f : A \to \mathbb{C}$ for which $\sum_{\alpha \in A} |f(\alpha)|^2 < \infty$) where μ is counting measure on A. Then $l^2(A)$ is a Hilbert space with inner product $\langle f, g \rangle = \sum_{\alpha \in A} f(\alpha)\overline{g(\alpha)}$ $\left(= \int_A f \bar{g} d\mu\right)$ and norm $\|f\|_2 = \sqrt{\langle f, f \rangle}$. (Since $\mu E = 0 \Rightarrow E = \emptyset$, $f = g$ a.e. $\Rightarrow f = g$ everywhere,so no quotient is needed.)

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

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

Corollary.

(1) $\widehat{x} \in l^2(A)$ and $\|\widehat{x}\|_2 \leq \|x\|$ so

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

Theorem. Define $F : x \to l^2(A)$ (where X is a Hilbert space; F is for Fourier) by $F : x \mapsto \widehat{x}$ where $\widehat{x}(\alpha) = \langle x, u_\alpha \rangle$ (where $\{u_\alpha\}_{\alpha \in A}$ is an orthonormal set). Then F is a bounded linear operator with $\|F\| = 1$ mapping 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 $\alpha \in A$, $\|\widehat{x}\|_2 = 1 = \|x\|$, so $\|F\| = 1$. Given $f \in l^2(A)$, $f(\alpha) \neq 0$ only for a countable set $A_f \subset A$; enumerate them $\alpha_1, \alpha_2, \alpha_3, \ldots$. Let $x_k = \sum_{j=1}^{k} f(\alpha_j) u_j$. Clearly $\widehat{x}_k(\alpha) = f(\alpha)$ for $\alpha_1, \ldots, \alpha_k$ and $\widehat{x}_k(\alpha) = 0$ otherwise. So $\widehat{x}_k(\alpha) \to f(\alpha)$ pointwise on A, and since $|\widehat{x}_k(\alpha) - f(\alpha)|^2 \leq |f(\alpha)|^2 \in L'(\mu)$, $\widehat{x}_k \to f$ in $l^2(A)$ by the Dominated Convergence Theorem. Since each x_k is a *finite* linear combination of the u_α's, $\|x_j - x_k\| = \|\widehat{x}_j - \widehat{x}_k\|_2$, so $\{x_k\}$ is Cauchy in X, so $x_k \to x$ in X for some $x \in X$. For each $\alpha \in A$,

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

So F is onto. $\square$

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

Proof. Zorn's lemma. $\square$

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_α's is dense in X.

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

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

(e) $(\forall\, x \in X)$ if $(\forall\, \alpha \in A)\ \langle x, u_\alpha\rangle = 0$ 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) $\Rightarrow$ (c): Clear if $x = 0$. Given $x \neq 0$, and given $\varepsilon > 0$ (WLOG assume $\varepsilon < \|x\|$), choose $y \in V \ni \|x - y\| < \varepsilon$, say $y \in \text{span}\{u_{\alpha_1}, \ldots, u_{\alpha_k}\}$. Let $z = \widehat{x}(\alpha_1)u_{\alpha_1} + \cdots + \widehat{x}(\alpha_k)u_{\alpha_k}$. Then z minimizes $\|x - w\|$ over $w \in \text{span}\{u_{\alpha_1}, \ldots, u_{\alpha_k}\}$ so $\|x - z\| \le \|x - y\| < \varepsilon$. Thus $\|x\| < \|z\| + \varepsilon$, so $(\|x\| - \varepsilon)^2 < \|z\|^2$ and $\|z\|^2 = \sum_{j=1}^k |\widehat{x}(\alpha_j)|^2 \le \sum_{\alpha\in A} |\widehat{x}(\alpha)|^2$. So $\|x\|^2 \le \sum_{\alpha\in A} |\widehat{x}(\alpha)|^2$. the other inequality is Bessel's inequality. (c) $\Rightarrow$ (d): Use polarization. (d) $\Rightarrow$ (e): Suppose $(\forall\, \alpha \in A)\langle x, u_\alpha\rangle = 0$. Then $\widehat{x}(\alpha) \equiv 0$, so $\|x\|^2 = \langle x, x\rangle = 0$, so $x = 0$. (e) $\Rightarrow$ (a): If $\{u_\alpha\}$ is not maximal, then $\exists\, x \neq 0 \ni \langle x, u_\alpha\rangle = 0$ for all $\alpha \in A$. $\square$

Notation. An orthonormal set $\{u_\alpha\}$ in a Hilbert space X satisfying the conditions in the previous theorem is called a *complete orthonormal set* (or 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.

Convergence of Fourier Series (in norm)

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 $\{\alpha \in A : \langle x, u_\alpha\rangle \neq 0\}$. Then $\|x\|^2 = \sum_{j\geq 1} |\langle x, u_{\alpha_j}\rangle|^2$ (i.e. Parseval's Equality holds for this x) iff $\lim_{n\to\infty} \left\| x - \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j} \right\| = 0$ (i.e. the Fourier series $\sum_{j=1}^\infty \widehat{x}(\alpha_j) u_{\alpha_j}$ converges to x).

Proof. Let $M_n = \operatorname{span}\{u_{\alpha_1}, \ldots, u_{\alpha_n}\}$ and let P_n be the orthogonal projection onto M_n (so $I - P_n$ is the orthogonal projection onto $M_n^\perp$). Then $P_n x = \sum_{j=1}^n \langle x, u_{\alpha_h j}\rangle u_{\alpha_j}$, $\|P_n x\|^2 = \sum_{j=1}^n |\langle x, u_{\alpha_j}\rangle|^2$, and $\|x\|^2 = \|P_n x\|^2 + \|(I-P_n)x\|^2$, so $\|x\|^2 - \|P_n x\|^2 = \|x - P_n x\|^2$. Hence $\|x\|^2 = \sum_{j\geq 1} |\langle x, u_{\alpha_j}\rangle|^2$ iff $\lim_{n\to\infty} \|P_n x\|^2 = \|x\|^2$ iff $\lim_{n\to\infty} \|x - P_n x\|^2 = 0$. (Note: If $\{\alpha \in A : \langle x, u_\alpha\rangle \neq 0\}$ 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., $x = \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j} \in M_n$.) $\square$

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 $\{\alpha \in A : \langle x, u_\alpha\rangle \neq 0\}$, $\lim_{n\to\infty} \|x - \sum_{j=1}^n \langle x, u_{\alpha_j}\rangle u_{\alpha_j}\| = 0$.

Cardinality of Orthonormal Bases

Proposition. $l^2(A)$ is unitarily isomorphic to $l^2(B)$ iff $\operatorname{card}(A) = \operatorname{card}(B)$.

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. $\square$

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 $Px = \sum_{\beta\in\mathcal{B}} \langle x, v_\beta\rangle v_\beta$, and the orthog. proj. of X onto $M^\perp$ is $Qx = \sum_{\gamma\in\mathcal{C}} \langle x, w_\gamma\rangle w_\gamma$.

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

Example. (Orthogonal Polynomials in weighted L^2 spaces). Fix $a, b \in \mathbb{R}$ with $-\infty < a < b < \infty$. Let $w(x) \in C(a,b)$ with $w(x) > 0$ on (a,b) and $\int_a^b w(x)dx < \infty$. ($w(x)$ is called the weight function, e.g., $w(x) = \frac{1}{\sqrt{1-x^2}}$ 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 $\langle f, g\rangle w = \int_a^b f(x)\overline{g(x)}w(x)dx$ for $f, g \in L^2_w(a,b)$. Then (after identifying f and g when $f = g$ a.e.), $L^2_w(a,b)$ is a Hilbert space.

Claim. Polynomials are dense in $L^2_w(a,b)$.

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

Theorem. The orthogonal polynomials in $L^2_w(a,b)$ (the result of Gram-Schmidt on $\{1, x, x^2, \ldots\}$) are a complete o.n. set in $L^2_w(a,b)$.

Proof. Finite lin. comb. are dense. $\square$