## 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) $(\forall x \in X)\langle x, x\rangle \geq 0$ and $\langle x, x\rangle=0$ iff $x=0$.
(2) $(\forall \alpha, \beta \in \mathbb{C})(\forall x, y, z \in X),\langle z, \alpha x+\beta y\rangle=\alpha\langle z, x\rangle+\beta\langle z, y\rangle$.
(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 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. $(\forall x, y \in X)|\langle y, x\rangle| \leq\|x\| \cdot\|y\|$, 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 $\langle z, x\rangle=\bar{\gamma}|\gamma|^{-1}\langle y, x\rangle=|\gamma|>0$. Let $v=x\|x\|^{-1}, w=z\|z\|^{-1}$. Then $\|v\|=\|w\|=1$ and $\langle w, v\rangle>0$. Since $0 \leq\|v-w\|^{2}=\langle v, v\rangle-2 \mathcal{R} e\langle w, v\rangle+\langle w, w\rangle$, it follows that $\langle w, v\rangle \leq 1$ (with equality iff $v=w$, which happens iff $x$ and $y$ are linearly dependent). So $|\langle y, x\rangle|=\langle z, x\rangle=\|x\| \cdot\|z\|\langle w, v\rangle \leq\|x\| \cdot\|z\|=\|x\| \cdot\|y\|$.

## Facts.

(1') $(\forall x \in X)\|x\| \geq 0 ;\|x\|=0$ iff $x=0$.
$\left(2^{\prime}\right)(\forall \alpha \in \mathbb{C})(\forall x \in X)\|\alpha x\|=|\alpha| \cdot\|x\|$.
$\left(3^{\prime}\right)(\forall x, y \in X)\|x+y\| \leq\|x\|+\|y\|$.

## Proof of ( $3^{\prime}$ ):

$$
\|x+y\|^{2}=\|x\|^{2}+2 \mathcal{R} e\langle y, x\rangle+\|y\|^{2} \leq\|x\|^{2}+2|\langle y, x\rangle|+\|y\|^{2} \leq\|x\|^{2}+2\|x\| \cdot\|y\|+\|y\|^{2} .
$$

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 $x=\left(x_{1}, \ldots, x_{n}\right)$ and $y=\left(y_{1}, \ldots, y_{n}\right) \in \mathbb{C}^{n}$, let $\langle y, x\rangle=\sum_{j=1}^{n} \overline{y_{j}} x_{j}$. Then $\|x\|=\sqrt{\sum_{j=1}^{n}\left|x_{j}\right|^{2}}$ is the $l^{2}$-norm on $\mathbb{C}^{n}$.

## Examples of Hilbert spaces:

- any finite dimensional inner product space
- $l^{2}=\left\{\left(x_{1}, x_{2}, x_{3}, \ldots\right): x_{k} \in \mathbb{C}, \sum_{k=1}^{\infty}\left|x_{k}\right|^{2}<\infty\right\}$ 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 $\langle g, f\rangle=\int_{A} \overline{g(x)} f(x) d x$.


## Incomplete inner product space

$C([a, b])$ with $\langle g, f\rangle=\int_{a}^{b} \overline{g(x)} f(x) d x$
$C([a, b])$ 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 $(\forall x, y \in X)$

$$
\|x+y\|^{2}+\|x-y\|^{2}=2\left(\|x\|^{2}+\|y\|^{2}\right)
$$

Proof. $\|x+y\|^{2}+\|x-y\|^{2}=\langle x+y, x+y\rangle+\langle x-y, x-y\rangle=\langle x, x\rangle+\langle x, y\rangle+\langle y, x\rangle+$ $\langle y, y\rangle+\langle x, x\rangle-\langle x, y\rangle-\langle y, x\rangle+\langle y, y\rangle=2(\langle x, x\rangle+\langle y, y\rangle)=2\left(\|x\|^{2}+\|y\|^{2}\right)$.

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

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

Proof. Expanding out the implied inner products, one shows easily that

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

Note: In a real inner product space, $\langle y, x\rangle=\frac{1}{4}\left(\|x+y\|^{2}-\|x-y\|^{2}\right)$.
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 $\langle y, i x\rangle=i\langle y, x\rangle$. Use the parallelogram law to show $\langle z, x+y\rangle=\langle z, x\rangle+\langle z, y\rangle$. Then show $\langle y, \alpha x\rangle=\alpha\langle y, x\rangle$ successively for $\alpha \in \mathbb{N}$, $\alpha^{-1} \in \mathbb{N}, \alpha \in \mathbb{Q}, \alpha \in \mathbb{R}$, and finally $\alpha \in \mathbb{C}$.

Continuity of the Inner Product. Let $X$ be an inner product space with induced norm $\|\cdot\|$. Then $\langle\cdot, \cdot\rangle: X \times X \rightarrow \mathbb{C}$ is continuous.
Proof. Since $X \times X$ and $\mathbb{C}$ are metric spaces, it suffices to show sequential continuity. Suppose $x_{n} \rightarrow x$ and $y_{n} \rightarrow y$. Then by the Schwarz inequality,

$$
\left|\left\langle y_{n}, x_{n}\right\rangle-\langle y, x\rangle\right|=\left|\left\langle y_{n}, x_{n}-x\right\rangle+\left\langle y_{n}-y, x\right\rangle\right| \leq\left\|x_{n}-x\right\| \cdot\left\|y_{n}\right\|+\|x\| \cdot\left\|y_{n}-y\right\| \rightarrow 0
$$

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 $A^{\perp}=\{x \in X:\langle y, x\rangle=0 \quad \forall y \in A\}$. Since the inner product is linear in the second component and continuous, $A^{\perp}$ is a closed subspace of $X$. Also

$$
(\operatorname{span}(A))^{\perp}=A^{\perp}, \quad \bar{A}^{\perp}=A^{\perp}, \quad \text { and } \quad(\overline{\operatorname{span}(A)})^{\perp}=A^{\perp}
$$

The Pythagorean Theorem. If $x_{1}, \ldots, x_{n} \in X$ and $x_{j} \perp x_{k}$ for $j \neq k$, then

$$
\left\|\sum_{j=1}^{n} x_{j}\right\|^{2}=\sum_{j=1}^{n}\left\|x_{j}\right\|^{2}
$$

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

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

## 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,

$$
\|x-y\|^{2}=2\left(\|x\|^{2}+\|y\|^{2}\right)-\|x+y\|^{2} \leq 2\left(\|x\|^{2}+\|y\|^{2}\right)-4 \delta^{2} .
$$

Uniqueness follows: if $\|x\|=\|y\|=\delta$, then $\|x-y\|^{2} \leq 4 \delta^{2}-4 \delta^{2}=0$, so $x=y$. For existence, choose $\left\{y_{n}\right\}_{n=1}^{\infty} \subset A$ for which $\left\|y_{n}\right\| \rightarrow \delta$. As $n, m \rightarrow \infty$,

$$
\left\|y_{n}-y_{m}\right\|^{2} \leq 2\left(\left\|y_{n}\right\|^{2}+\left\|y_{m}\right\|^{2}\right)-4 \delta^{2} \rightarrow 0
$$

so $\left\{y_{n}\right\}$ is Cauchy. By completeness, $\exists y \in X$ for which $y_{n} \rightarrow y$, and since $A$ is closed, $y \in A$. Also $\|y\|=\lim \left\|y_{n}\right\|=\delta$.

Corollary. If $A$ is a nonempty closed convex set in a Hilbert space and $x \in X$, then $\exists \mathrm{a}$ 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 $x \in M, P x=x$ and $Q x=0$; if $x \in M^{\perp}, P x=0$ and $Q x=x$.
(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) $P x$ is the unique closest element of $M$ to $x$, and $Q x$ 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 $Q x$ to be the smallest element of $x+M$, and let $P x=x-Q x$. Since $Q x \in x+M, P x \in M$. Let $z=Q x$. 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 $x, y \in X, \alpha x+\beta y=(\alpha P x+\beta P y)+(\alpha Q x+\beta Q y)$, so by uniqueness in (1), $P(\alpha x+\beta y)=\alpha P x+\beta P y$ and $Q(\alpha x+\beta y)=\alpha Q x+\beta Q y$. By the Pythagorean Theorem, $\|x\|^{2}=\|P x\|^{2}+\|Q x\|^{2}$, 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 $\|x-y\|^{2}=\|P x-P y\|^{2}+\|Q x-Q y\|^{2}$.

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 $y=Q x$.

Corollary. If $M$ is a closed subspace of a Hilbert space $X$, then $\left(M^{\perp}\right)^{\perp}=M$. In general, for any $A \subset X,\left(A^{\perp}\right)^{\perp}=\overline{\operatorname{span}\{A\}}$, 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 $f_{y}: X \rightarrow C$ by $f_{y}(x)=\langle y, x\rangle$. Then $f_{y} \in X^{*}$ and $\left\|f_{y}\right\|=\|y\|$.

Proof. $\left|f_{y}(x)\right|=|\langle y, x\rangle| \leq\|x\| \cdot\|y\|$, so $f_{y} \in X^{*}$ and $\left\|f_{y}\right\| \leq\|y\|$. Since $\left|f_{y}(y)\right|=|\langle y, y\rangle|=$ $\|y\|^{2},\left\|f_{y}\right\| \geq\|y\|$. So $\left\|f_{y}\right\|=\|y\|$.

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., $f(x)=\langle y, x\rangle \forall x \in X$.
(2) The map $\psi: X \rightarrow X^{*}$ 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 $0=\langle z, u\rangle=f(x)\langle z, z\rangle-f(z)\langle z, x\rangle=f(x)-\langle\alpha z, x\rangle=$ $f(x)-\langle y, x\rangle$, i.e., $f(x)=\langle y, x\rangle$. Uniqueness: if $\left\langle y_{1}, x\right\rangle=\left\langle y_{2}, x\right\rangle \forall x \in X$, then (letting $\left.x=y_{1}-y_{2}\right)\left\|y_{1}-y_{2}\right\|^{2}=0$, 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.

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

Proof. Clearly $\langle f, f\rangle \geq 0,\langle f, f\rangle=0$ iff $\psi^{-1}(f)=0$ iff $f=0$, and $\overline{\langle g, f\rangle}=\langle f, g\rangle$. Also $\left\langle f_{y}, \alpha_{1} f_{x_{1}}+\alpha_{2} f_{x_{2}}\right\rangle=\left\langle f_{y}, f_{\bar{\alpha}_{1} x_{1}+\bar{\alpha}_{2} x_{2}}\right\rangle=\overline{\left\langle y, \bar{\alpha}_{1} x_{1}+\bar{\alpha}_{2} x_{2}\right\rangle}=\alpha_{1} \overline{\left\langle y, x_{1}\right\rangle}+\alpha_{2} \overline{\left\langle y, x_{2}\right\rangle}=$ $\alpha_{1}\left\langle f_{y}, f_{x_{1}}\right\rangle+\alpha_{2}\left\langle f_{y}, f_{x_{2}}\right\rangle$, so $\langle\cdot, \cdot\rangle$ is an inner product on $X^{*}$. Since $\left\langle f_{y}, f_{y}\right\rangle=\overline{\langle y, y\rangle}=\|y\|^{2}=$ $\left\|f_{y}\right\|^{2},\langle\cdot, \cdot\rangle$ induces the norm on $X^{*}$. Since $X^{*}$ is complete, it is a Hilbert space.

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 $x_{n} \rightarrow x$ strongly if $\left\|x_{n}-x\right\| \rightarrow 0$ as $n \rightarrow \infty$. This is the usual concept of convergence in the metric induced by the norm, and is also called convergence in norm. We say $x_{n} \rightarrow x$ weakly if $(\forall y \in X)\left\langle y, x_{n}\right\rangle \rightarrow\langle y, x\rangle$ as $n \rightarrow \infty$. (Other common notations for weak convergence are $x_{n} \rightharpoonup x, x_{n} \xrightarrow{w} x$.) The Cauchy-Schwarz inequality shows that strong convergence implies weak convergence. Also, if $x_{n} \rightarrow x$ strongly, then $\left\|x_{n}\right\| \rightarrow\|x\|$ since $\left|\left\|x_{n}\right\|-\|x\|\right| \leq\left\|x_{n}-x\right\|$.
Example. (Weak convergence does not imply strong convergence if $\operatorname{dim} X=\infty$ ). Let $\swarrow k^{\text {th }}$ entry
$X=l^{2}$. For $k=1,2, \ldots$, let $e_{k}=(0, \ldots, 0,1,0, \ldots) \quad$ (so $\left\{e_{k}: k=1,2, \ldots\right\}$ is an orthonormal set in $l^{2}$ ).

Claim. $e_{k} \rightarrow 0$ weakly as $k \rightarrow \infty$.
Proof. Fix $y \in l^{2}$. Then $\sum_{k=1}^{\infty}\left|y_{k}\right|^{2}<\infty$, so $y_{k} \rightarrow 0$. So $\left\langle y, e_{k}\right\rangle=\overline{y_{k}} \rightarrow 0$.
Note that $\left\|e_{k}\right\|=1$, so $e_{k}$ does not converge to zero strongly.
Remark. If $\operatorname{dim} X<\infty$, then weak convergence $\Rightarrow$ strong convergence (exercise).
Theorem. Suppose $x_{n} \rightarrow x$ weakly in a Hilbert space $X$. Then
(a) $\|x\| \leq \lim \inf _{k \rightarrow \infty}\left\|x_{k}\right\|$
(b) If $\left\|x_{k}\right\| \rightarrow\|x\|$, then $x_{k} \rightarrow x$ strongly (i.e., $\left\|x_{k}-x\right\| \rightarrow 0$ ).

## Proof.

(a) $0 \leq\left\|x-x_{k}\right\|^{2}=\|x\|^{2}-2 \mathcal{R} e\left\langle x, x_{k}\right\rangle+\left\|x_{k}\right\|^{2}$. By hypothesis, $\left\langle x, x_{k}\right\rangle \rightarrow\langle x, x\rangle=\|x\|^{2}$. So taking liminf above, $0 \leq\|x\|^{2}-2\|x\|^{2}+\lim \inf \left\|x_{k}\right\|^{2}$, i.e. $\|x\|^{2} \leq \lim \inf \left\|x_{k}\right\|^{2}$.
(b) If $x_{k} \rightarrow x$ weakly and $\left\|x_{k}\right\| \rightarrow\|x\|$, then $\left\|x-x_{k}\right\|^{2}=\|x\|^{2}-2 \mathcal{R} e\left\langle x_{k}, x\right\rangle+\left\|x_{k}\right\|^{2} \rightarrow$ $\|x\|^{2}-2\|x\|^{2}+\|x\|^{2}=0$.

Remark. The Uniform Boundedness Principle implies that if $x_{k} \rightarrow x$ weakly, then $\left\|x_{k}\right\|$ is bounded.

## Orthogonal Sets

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

## Orthonormal Sets

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

Theorem. If $\left\{u_{1}, \ldots, u_{k}\right\}$ is an orthonormal set in an inner product space $X$, and $x=$ $\sum_{j=1}^{k} c_{j} u_{j}$, then $c_{j}=\left\langle u_{j}, x\right\rangle$ for $1 \leq j \leq k$ and $\|x\|^{2}=\sum_{j=1}^{l}\left|c_{j}\right|^{2}$

Proof. $\left\langle u_{i}, x\right\rangle=\sum c_{j}\left\langle u_{i}, u_{j}\right\rangle=c_{i}$. 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 $\operatorname{span}\left\{u_{1}, \ldots, u_{n}\right\}$ to $x$ is $\sum_{k=1}^{n}\left\langle u_{k}, x\right\rangle u_{k}$.

Theorem. (Gram-Schmidt process) Let $V$ be a subspace of an inner product space $X$, and suppose $V$ has a finite or countable basis $\left\{x_{n}\right\}_{n \geq 1}$. Then $V$ has a basis $\left\{u_{n}\right\}_{n \geq 1}$ which is orthonormal (we reserve the term "orthonormal basis" to mean something else); moreover we can choose $\left\{u_{n}\right\}_{n \geq 1}$ so that for all $m \geq 1$, $\operatorname{span}\left\{u_{1}, \ldots, u_{m}\right\}=\operatorname{span}\left\{x_{1}, \ldots, x_{m}\right\}$.

Proof Sketch.. Define $\left\{u_{n}\right\}$ inductively. Start with $u_{1}=\frac{x_{1}}{\left\|x_{1}\right\|}$. Having defined $u_{1}, \ldots, u_{n-1}$, let $v_{n}=x_{n}-\sum_{j=1}^{n-1}\left\langle u_{j}, x_{n}\right\rangle u_{j}$ and $u_{n}=\frac{v_{n}}{\left\|v_{n}\right\|}$.

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

Definition. Let $A$ be a nonempty set. For each $\alpha \in A$, let $y_{\alpha}$ be a nonnegative real number. Define $\sum_{\alpha \in A} y_{\alpha}=\sup \left\{\sum_{\alpha \in F} y_{\alpha}: F \subset A\right.$ and $F$ 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 $A_{k} \equiv\left\{\alpha: y_{\alpha}>k^{-1}\right\}$ is a finite set. But $\left\{\alpha: y_{\alpha} \neq 0\right\}=$ $\cup_{k=1}^{\infty} A_{k}$.

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

Bessel's Inequality. Let $\left\{u_{\alpha}\right\}_{\alpha \in A}$ be an orthonormal set in a Hilbert space $X$, let $x \in X$, and let $\widehat{x}(\alpha)=\left\langle u_{\alpha}, x\right\rangle$. 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.
Corollary. Let $\left\{u_{\alpha}\right\}_{\alpha \in A}, x$ be as above. Then
(1) $\widehat{x} \in l^{2}(A)$ and $\|\widehat{x}\|_{2} \leq\|x\|$
(2) $\{\alpha \in A: \widehat{x}(\alpha) \neq 0\}$ is countable.

Theorem. Let $X$ be a Hilbert space and let $\left\{u_{\alpha}\right\}_{\alpha \in A}$ be an orthonormal set. Define $F: X \rightarrow l^{2}(A)\left(F\right.$ is for Fourier) by $F(x)=\widehat{x}$ where $\widehat{x}(\alpha)=\left\langle u_{\alpha}, x\right\rangle$. 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 $\alpha \in A,\|\widehat{x}\|_{2}=1=\|x\|$, 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 $x_{k}=\sum_{j=1}^{k} f\left(\alpha_{j}\right) u_{j}$. Clearly $\widehat{x}_{k}(\alpha)=f(\alpha)$ for $\alpha_{1}, \ldots, \alpha_{k}$ and $\widehat{x}_{k}(\alpha)=0$ otherwise. So

$$
\left\|\widehat{x}_{k}-f\right\|_{2}^{2}=\sum_{j=k+1}^{\infty}\left|f\left(\alpha_{j}\right)\right|^{2} \rightarrow 0 \quad \text { as } \quad k \rightarrow \infty
$$

Thus $\widehat{x}_{k} \rightarrow f$ in $l^{2}(A)$, and in particular $\widehat{x}_{k}$ is a Cauchy sequence in $l^{2}(A)$. Since each $x_{k}$ is a finite linear combination of the $u_{\alpha}$ 's, $\left\|x_{j}-x_{k}\right\|=\left\|\widehat{x}_{j}-\widehat{x}_{k}\right\|_{2}$, so $\left\{x_{k}\right\}$ is Cauchy in $X$, so $x_{k} \rightarrow x$ in $X$ for some $x \in X$. For each $\alpha \in A$,

$$
\widehat{x}(\alpha)=\left\langle u_{\alpha}, x\right\rangle=\lim _{k \rightarrow \infty}\left\langle u_{\alpha}, x_{k}\right\rangle=\lim _{k \rightarrow \infty} \widehat{x}_{k}(\alpha)=f(\alpha) .
$$

So $F(x)=f$ and $F$ is onto.
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.
Corollary. Every Hilbert space has a maximal orthonormal set.
Theorem. Let $\left\{u_{\alpha}\right\}_{\alpha \in A}$ be an orthonormal set in a Hilbert space $X$. The following conditions are equivalent:
(a) $\left\{u_{\alpha}\right\}_{\alpha \in A}$ is a maximal orthonormal set.
(b) The set of finite linear combinations of the $u_{\alpha}$ 's is dense in $X$.
(c) $(\forall x \in X)\|x\|^{2}=\sum_{\alpha \in A}|\widehat{x}(\alpha)|^{2}$ (Parseval's relation).
(d) $(\forall x, y \in X)\langle y, x\rangle=\sum_{\alpha \in A} \overline{\widehat{y}(\alpha)} \widehat{x}(\alpha)$.
(e) $(\forall x \in X)$ if $(\forall \alpha \in A)\left\langle u_{\alpha}, x\right\rangle=0$ then $x=0$.

## Proof.

(a) $\Rightarrow(\mathrm{b})$ : Let $V=\operatorname{span}\left\{u_{\alpha}: \alpha \in A\right\}$ and $M=\bar{V}$. Then $M$ is a closed subspace. Since $\left\{u_{\alpha}\right\}$ 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 $\epsilon>0$ (WLOG assume $\epsilon<\|x\|$ ), choose $y \in V \ni\|x-y\|<\epsilon$, say $y \in \operatorname{span}\left\{u_{\alpha_{1}}, \ldots, u_{\alpha_{k}}\right\}$. Let $z=\widehat{x}\left(\alpha_{1}\right) u_{\alpha_{1}}+\cdots+\widehat{x}\left(\alpha_{k}\right) u_{\alpha_{k}}$. Then $z$ minimizes $\|x-w\|$ over $w \in \operatorname{span}\left\{u_{\alpha_{1}}, \ldots, u_{\alpha_{k}}\right\}$ so $\|x-z\| \leq\|x-y\|<\epsilon$. Thus $\|x\|<\|z\|+\epsilon$, so $(\|x\|-\epsilon)^{2}<\|z\|^{2}$ and $\|z\|^{2}=\sum_{j=1}^{k}\left|\widehat{x}\left(\alpha_{j}\right)\right|^{2} \leq \sum_{\alpha \in A}|\widehat{x}(\alpha)|^{2}$. Let $\epsilon \rightarrow 0$ to get $\|x\|^{2} \leq \sum_{\alpha \in A}|\widehat{x}(\alpha)|^{2}$. The other inequality is Bessel's inequality.
$(c) \Rightarrow(d)$ : Use polarization. Only countably many terms in the sum are nonzero.
$(\mathrm{d}) \Rightarrow(\mathrm{e}):$ Suppose $(\forall \alpha \in A)\left\langle x, u_{\alpha}\right\rangle=0$. Then $\widehat{x}(\alpha) \equiv 0$, so $\|x\|^{2}=\langle x, x\rangle=0$, so $x=0$.
(e) $\Rightarrow$ (a): If $\left\{u_{\alpha}\right\}$ is not maximal, then $\exists x \neq 0 \ni\left\langle x, u_{\alpha}\right\rangle=0$ for all $\alpha \in A$.

Definition. An orthonormal set $\left\{u_{\alpha}\right\}$ 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 $T: X \rightarrow Y$ which is linear, bijective, and preserves inner products (i.e., $(\forall x, y \in X)\langle x, y\rangle=\langle T x, T y\rangle$ - this implies $T$ is an isometry $\|x\|=\|T x\|)$ is called a unitary isomorphism.

Corollary. If $X$ is a Hilbert space and $\left\{u_{\alpha}\right\}_{\alpha \in A}$ is an orthonormal basis of $X$, then the map $F: X \rightarrow l^{2}(A)$ mapping $x \mapsto \widehat{x}$ (where $\left.\widehat{x}(\alpha)=\left\langle x, u_{\alpha}\right\rangle\right)$ 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, $\left\{u_{\alpha}\right\}_{\alpha \in A}$ be an orthonormal set in $X$, and let $x \in X$. Let $\left\{\alpha_{j}\right\}_{j \geq 1}$ be any enumeration of $\left\{\alpha \in A:\left\langle u_{\alpha}, x\right\rangle \neq 0\right\}$. Then $\|x\|^{2}=\sum_{j \geq 1}\left|\left\langle u_{\alpha_{j}}, x\right\rangle\right|^{2}$ (i.e. Parseval's Equality holds for this $x$ ) iff $\lim _{n \rightarrow \infty}\left\|x-\sum_{j=1}^{n}\left\langle u_{\alpha_{j}}, x\right\rangle u_{\alpha_{j}}\right\|=0$ (i.e. the Fourier series $\sum_{j=1}^{\infty} \widehat{x}\left(\alpha_{j}\right) u_{\alpha_{j}}$ converges to $x$ in norm).
Proof. Let $M_{n}=\operatorname{span}\left\{u_{\alpha_{1}}, \ldots, u_{\alpha_{n}}\right\}$ and let $P_{n}$ be the orthogonal projection onto $M_{n}$ (so $I-P_{n}$ is the orthogonal projection onto $\left.M_{n}^{\perp}\right)$. Then $P_{n} x=\sum_{j=1}^{n}\left\langle u_{\alpha_{j}}, x\right\rangle u_{\alpha_{j}}$ and $\left\|P_{n} x\right\|^{2}=$ $\sum_{j=1}^{n}\left|\left\langle u_{\alpha_{j}}, x\right\rangle\right|^{2}$. Also $\|x\|^{2}=\left\|P_{n} x\right\|^{2}+\left\|\left(I-P_{n}\right) x\right\|^{2}$, so $\|x\|^{2}-\left\|P_{n} x\right\|^{2}=\left\|x-P_{n} x\right\|^{2}$. Hence $\|x\|^{2}=\sum_{j \geq 1}\left|\left\langle u_{\alpha_{j}}, x\right\rangle\right|^{2}$ iff $\lim _{n \rightarrow \infty}\left\|P_{n} x\right\|^{2}=\|x\|^{2}$ iff $\lim _{n \rightarrow \infty}\left\|x-P_{n} x\right\|^{2}=0$, which is the desired conclusion. (Note: If $\left\{\alpha \in A:\left\langle u_{\alpha}, x\right\rangle \neq 0\right\}$ is finite, say $\left\{\alpha_{1}, \ldots, \alpha_{n}\right\}$, then Parseval holds iff $\left\|P_{n} x\right\|^{2}=\|x\|^{2}$ iff $x=P_{n} x$, i.e., $x=\sum_{j=1}^{n}\left\langle u_{\alpha_{j}}, x\right\rangle u_{\alpha_{j}} \in M_{n}$.)
Corollary. Let $\left\{u_{\alpha}\right\}_{\alpha \in A}$ be an orthonormal set in a Hilbert space $X$. Then $\left\{u_{\alpha}\right\}$ is an orthonormal basis iff for each $x \in X$ and each enumeration $\left\{\alpha_{j}\right\}_{j \geq 1}$ of $\left\{\alpha \in A:\left\langle u_{\alpha}, x\right\rangle \neq 0\right\}$, $\lim _{n \rightarrow \infty}\left\|x-\sum_{j=1}^{n}\left\langle u_{\alpha_{j}}, x\right\rangle u_{\alpha_{j}}\right\|=0$.

## 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 $\operatorname{dim} X<\infty$. Suppose $\operatorname{dim} X=\infty$. 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 $\left\{v_{\beta}\right\}_{\beta \in \mathcal{B}}$ be a complete orthonormal set in $M$, and let $\left\{w_{\gamma}\right\}_{\gamma \in \mathcal{C}}$ be a complete orthonormal set in $M^{\perp}$. Then $\left\{v_{\beta}\right\} \cup\left\{w_{\gamma}\right\}$ is a complete orthonormal set in $X$. The orthogonal projection of $X$ onto $M$ is $P x=\sum_{\beta \in \mathcal{B}}\left\langle v_{\beta}, x\right\rangle v_{\beta}$, and the orthogonal projection of $X$ onto $M^{\perp}$ is $Q x=\sum_{\gamma \in \mathcal{C}}\left\langle w_{\gamma}, x\right\rangle w_{\gamma}$.

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 $\int_{a}^{b} w(x) d x<\infty$. The function $w$ is called the weight function. For example, take $w(x)=\left(1-x^{2}\right)^{-1 / 2}$ on $(-1,1)$. Define

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

and define $\langle g, f\rangle_{w}=\int_{a}^{b} \overline{g(x)} f(x) w(x) d x$ 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 $\int_{a}^{b}|f(x)|^{2} w(x) d x \leq$ $\|f\|_{\infty}^{2} \int_{a}^{b} w(x) d x$, and thus $\|f\|_{w} \leq M\|f\|_{\infty}$, where $M=\left(\int_{a}^{b} w(x) d x\right)^{\frac{1}{2}}<\infty$. 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 $\left(C[a, b],\|\cdot\|_{\infty}\right)$, so $\exists$ a polynomial $p$ for which $\|g-p\|_{\infty}<(2 M)^{-1} \epsilon$. 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 $\left\{1, x, x^{2}, \ldots\right\}$ 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.

