Hilbert Spaces I

Definition of Inner Product

A *scalar product* (or *inner product*) on a linear space X is a mapping from $X \times X$ to $\mathbb{K}$, denoted $(\cdot,\cdot)$, which satisfies the following axioms

$$
\begin{aligned}
&(a_1) \quad (x,y) = \overline{(y,x)} \quad \forall\, x,y \in X\,,\\
&(a_2) \quad (x+y,z) = (x,z) + (y,z) \quad \forall\, x,y,z \in X\,,\\
&(a_3) \quad (\alpha x,y) = \alpha(x,y) \quad \forall\, \alpha \in \mathbb{K},\ \forall\, x,y \in X\,,\\
&(a_4) \quad (x,x) \geq 0 \ \ \forall x \in X, \ \ \text{and} \ \ (x,x) = 0 \iff x = 0\,.
\end{aligned}
$$

We have denoted by $\overline{(y,x)}$ the complex conjugate of (y,x)

A space X together with such a product is called an *inner product space*.

We have $(x, \alpha y) = \overline{\alpha}(x,y)$ for all $\alpha \in \mathbb{K}$ and all $x, y \in X$.

Two vectors $x,\ y \in X$ are called *orthogonal* if their scalar product is equal to zero: $(x,y) = 0$.

We define the *length* of a vector $x \in X$ as $\|x\| = \sqrt{(x,x)}$.

mapping $x \to \|x\|$ satisfies

$$
\begin{aligned}
&(i) \quad \|x\| = 0 \iff x = 0\,;\\
&(ii) \quad \|\alpha x\| = |\alpha| \cdot \|x\| \quad \forall \alpha \in \mathbb{K},\ \forall\, x \in X\,;\\
&(iii) \quad \|x+y\| \leq \|x\| + \|y\| \quad \forall\, x,y \in X\,.
\end{aligned}
$$

The mapping $\|\cdot\| : X \to [0,\infty)$ is a *norm* on X, and X is a *normed space*.

Cauchy-Schwarz inequality:

$$|(x,y)| \leq \|x\| \cdot \|y\| \quad \forall x,y \in X \tag{1}$$

In fact, we have

$$0 \leq \|x+\alpha y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\,\overline{\alpha}(x,y) + |\alpha|^2\|y\|^2,$$

for all $\alpha \in \mathbb{K}$ and all $x, y \in X$. Taking

$$\alpha = -(x,y)/\|y\|^2$$

We get (1).

The usual scalar product of $X = \mathbb{R}^n$ is defined by

$$(x,y) = \sum_{i=1}^{n} x_i y_i \quad \forall\, x = (x_1, \ldots, x_n)^T,\ y = (y_1, \ldots, y_n)^T \in X,$$

and the corresponding norm is

$$\|x\| = \sqrt{(x,x)} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Parallelogram law

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \text{ for all } x,y \in X.$$

Lemma If V is an inner product space, then $x_n \to x$, $y_n \to y$ implies that $(x_n, y_n) \to (x,y)$.

In particular, if $x = \sum_{j=1}^{\infty} x_j$, then

$$(x,y) = \sum_{j=1}^{\infty} (x_j, y)$$

Proof. We have

$$|(x_n, y_n) - (x,y)| = |(x_n - x, y_n) + (x, y_n - y)|$$

$$\leq \|x_n - x\|\|y_n\| + \|y_n - y\|\|x\| \to 0,\ n \to \infty.$$

We define on $\mathbb{C}^n$ *an inner product* by

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i} \quad \forall\, x = (x_1, \ldots, x_n)^T,\ y = (y_1, \ldots, y_n)^T \in \mathbb{C}^n\,,$$

and the corresponding (Euclidean) norm is

$$\|x\| = \sqrt{(x, x)} = \sqrt{\sum_{i=1}^{n} |x_i|^2}\,.$$

The function $\mathrm{d}(x, y) = \sqrt{(x-y, x-y)}$

is called the distance function induced by the inner product (,).

If $||\cdot||$ is the norm induced by an inner product $(\cdot\,,\cdot)$, then if X is real,

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2;$$

if X is complex,

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2$$

Let X be a linear space over $\mathbb{K}$ equipped with a scalar (inner) product $(\cdot,\cdot)$.

Define the norm

$$\|x\| = \sqrt{(x,x)}, \quad x \in X\,.$$

If $(X, \|\cdot\|)$ is a Banach space (i.e., (X,d) is a complete metric space, where $d(x,y) = \|x-y\|$, $x,y \in X$), then X is said to be a **Hilbert space**. In other words, a Hilbert space is a Banach space $(X, \|\cdot\|)$ whose norm is given by a scalar product.

$\mathbb{R}^k$, $\mathbb{C}^k$, $L^2(\Omega)$ are Hilbert spaces equipped with the usual inner products:

$$(x,y) = \sum_{i=1}^{k} x_i y_i, \quad x = (x_1,\ldots,x_k)\,,\ y = (y_1,\ldots,y_k) \in \mathbb{R}^k\,,$$

$$(x,y) = \sum_{i=1}^{k} x_i \overline{y_i}\,, \quad x = (x_1,\ldots,x_k),\ y = (y_1,\ldots,y_k) \in \mathbb{C}^k\,,$$

$$(u,v)_{L^2(\Omega)} = \int_\Omega uv\,dx\,, \quad u,v \in L^2(\Omega)$$

The corresponding induced norms are

$$\|x\|^2 = \sum_{i=1}^{k} x_i^2, \quad x = (x_1, \ldots, x_k) \in \mathbb{R}^k\,,$$

$$\|x\|^2 = \sum_{i=1}^{k} |x_i|^2, \quad x = (x_1, \ldots, x_k) \in \mathbb{C}^k\,,$$

$$\|u\|^2_{L^2(\Omega)} = \int_\Omega u^2\, dx\,, \quad u \in L^2(\Omega).$$

Every Cauchy sequence in $\mathbb{R}^n$ is convergent since the corresponding coordinate sequences are Cauchy in $\mathbb{R}$, hence convergent in that space. So $\mathbb{R}^n$ equipped with the above scalar product and norm is a Hilbert space over $\mathbb{R}$.

An inner product space is also called a Pre-Hilbert space.

Theorem (Jordan–von Neumann). *Let $(H, \|\cdot\|)$ be a normed space. Then the norm $\|\cdot\|$ is given by a scalar product (i.e., there exists a scalar product $(\cdot,\cdot) : H \times H \to \mathbb{K}$ such that $\|x\| = \sqrt{(x,x)}$, $x \in H$) if and only if $\|\cdot\|$ satisfies the parallelogram law. (Hence, a Banach space $(H, \|\cdot\|)$ is Hilbert $\Longleftrightarrow$ its norm $\|\cdot\|$ satisfies the parallelogram law).*

Proof. Assuming that $\|\cdot\|$ is generated by a scalar product $(\cdot,\cdot)$,

we have for all $x, y \in H$

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= (x+y, x+y) + (x-y, x-y) \\ &= 2(\|x\|^2 + \|y\|^2)\,. \end{aligned} \tag{1}$$

i.e., the norm satisfies the parallelogram law.

Assume that the norm $\|\cdot\|$ of H satisfies (1) , we consider the

Case $\mathbb{K}=\mathbb{R}$. Define $f: H\times H\to\mathbb{R}$ by

$$f(x,y)=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2),\quad x,y\in H,$$

$$f(x,x)=\frac{1}{4}\|2x\|^2=\|x\|^2\quad\forall x\in H,$$

$$f(x,y)=f(y,x)\quad\forall x,y\in H,$$

$$f(x,0)=0\quad\forall x\in H.$$

$$f(x_1+x_2,y)=\frac{1}{4}(\|x_1+x_2+y\|^2-\|x_1+x_2-y\|^2)$$

$$f(x_1-x_2,y)=\frac{1}{4}(\|x_1-x_2+y\|^2-\|x_1-x_2-y\|^2)$$

$$\begin{aligned}&f(x_1+x_2,y)+f(x_1-x_2,y)\\=\ &\frac{1}{2}(\|x_1+y\|^2-\|x_1-y\|^2)\\=\ &2f(x_1,y).\end{aligned}$$

In the special case $x_1=x_2=x$ we have

$$f(2x,y)=2f(x,y)\quad\forall x,y\in H.$$

Taking $x_1+x_2=x$ and $x_1-x_2=x'$

we get
$$\begin{aligned}f(x,y)+f(x',y)&=2f\left(\frac{x+x'}{2},y\right)\\&=f(x+x',y).\end{aligned}$$

Thus
$$f(nx,y)=nf(x,y)\text{ for all }n\in\mathbb{N}$$

and so
$$f(nx,y)=nf(x,y)\quad\forall x,y\in H,\ \forall n\in\mathbb{Z}$$

Now for a rational number $r=m/n$, $m,n\in\mathbb{Z}$, $n\neq 0$, we have

$$f\left(\frac{m}{n}x, y\right) = mf\left(\frac{1}{n}x, y\right) = \frac{m}{n}f(x, y),$$

so

$$f(rx, y) = rf(x, y) \quad \forall x, y \in H, \ \forall r \in \mathbb{Q}.$$

Since f is continuous on $H \times H$, this extends to $r \in \mathbb{R}$, i.e.,

$$f(rx, y) = rf(x, y) \quad \forall x, y \in H, \ \forall r \in \mathbb{R}.$$

Summarizing, we see that H satisfies the conditions for an inner product.

Sufficiency in the complex case $\mathbb{K} = \mathbb{C}$ can be treated similarly, with $f : H \times H \to \mathbb{C}$ defined by

$$f(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2)$$

the scalar product generating a norm is unique. Indeed, if $\langle\cdot,\cdot\rangle$ and $(\cdot,\cdot)$ are two scalar products $(x,x) = \langle x,x\rangle = \|x\|^2,\, x\in H,$ we have from

$$(x+y,x+y) = \langle x+y,x+y\rangle \quad \forall x,y\in H\,,$$

that

$$\mathrm{Re}(x,y) = \mathrm{Re}\langle x,y\rangle \quad \forall x,y\in H\,, \tag{2}$$

and this completes the proof in the real case. If $\mathbb{K} = \mathbb{C}$, then by replacing y by iy in (2), we also get

$$\mathrm{Im}(x,y) = \mathrm{Im}\langle x,y\rangle \quad \forall x,y\in H\,.$$

Example The norm $\|\cdot\|_\infty$ on C([a,b]) does not come from an inner product.

In fact, taking $f(x) = 1,\quad g(x) = \dfrac{x-a}{b-a}$, we have

$$\|f\|_\infty = \|g\|_\infty = 1, \qquad \|f+g\|_\infty = 2, \qquad \|f-g\|_\infty = 1.$$

The Parallelogram law does not hold.

Example When $p \geq 1, p \neq 2$, the norm $\|\cdot\|_{l^p}$ on $l^p(\mathbb{K})$ does not come from an inner product.

In fact, for $x = (1,1,0,\ldots), y = (1,-1,0,\ldots) \in l^p$,

We have $if\ p \neq \infty$

$$\|x\| = \|y\| = 2^{1/p}, \|x+y\| = \|x-y\| = 2, \quad \|x+y\|^2 + \|x-y\|^2 = 8 \neq 4 \times 2^{2/p} = 2(\|x\|^2 + \|y\|^2).$$

On the other hand, if $u = (1,1,0,\ldots)$, $v = (1,0,\ldots)$, then

$$\|u\|_\infty = \|v\|_\infty = 1, \quad \|u+v\|_\infty = 2, \|u-v\|_\infty = 1.$$

The Parallelogram law does not hold.

Definition. A normed space is said to be *uniformly convex* if

$$\forall\varepsilon > 0 \;\; \exists\delta > 0 \text{ such that}$$

$$\left[x, y \in E, \|x\| \le 1, \|y\| \le 1 \text{ and } \|x - y\| > \varepsilon\right] \Rightarrow \left[\left\|\frac{x+y}{2}\right\| < 1 - \delta\right].$$

Lemma Every inner product space is uniformly convex.

Proof

For any $\varepsilon > 0$, taking $\delta = 1 - \sqrt{1 - \frac{\varepsilon^2}{4}}$, we have from

$$\left\|\frac{x+y}{2}\right\|^2 = \frac{1}{2}\left(\|x\|^2 + \|y\|^2\right) - \left\|\frac{x-y}{2}\right\|^2$$ that

$\|x\| \le 1, \|y\| \le 1, \|x - y\| > \varepsilon \;\rightarrow \left\|\frac{x+y}{2}\right\|^2 < 1 - \frac{\varepsilon^2}{4}$, that is, $\left\|\frac{x+y}{2}\right\| < 1 - \delta.$

Schauder Bases in Normed Spaces

Definition A countable set $\{e_j\}_{j=1}^{\infty}$ is a Schauder basis for a normed space X if every $x \in X$ can be written uniquely as

$$x = \sum_{i=1}^{\infty} \alpha_i e_i \qquad \text{for some } \alpha_i \in \mathbb{K}$$

Example The collection $\{\mathrm{e}^{(j)}\}_{j=1}^{\infty}$ is a Schauder basis for l^p for every $1 \leq p < \infty$, but is not a Schauder basis for l^{∞}.

In fact, we have showed that $\forall\, \varepsilon > 0, \exists\, N > 0\ s.t.$ for all $n \geq N$, we have

$$\left\| x - \sum_{j=1}^{n} x_j \mathrm{e}^{(j)} \right\|_{l^p} < \varepsilon$$

So we can write $x = \Sigma_{j=1}^{\infty} x_j \mathrm{e}^{(j)}$ as an equality in l^p, in the sense that the sum converges in l^p.

Consider $x \in l^{\infty}$ with $x_j = 1$ for every j. The equality

$$\sum_{j=1}^{\infty} \alpha_j \mathrm{e}^{(j)} = x$$

would mean that the partial sums converge to x in l^{∞}; but for any finite n we have

$$\left\| \sum_{j=1}^{n} \alpha_j \mathrm{e}^{(j)} - x \right\|_{l^{\infty}} = \|(\alpha_1 - 1, \dots, \alpha_n - 1, 1, \dots, 1)\|_{l^{\infty}} \geq 1,$$

and so the partial sums cannot converge whatever our choice of coefficients $\{\alpha_j\}$.

Definition Two elements x and y of an inner-product space V are said to be orthogonal if $(x, y) = 0$

A set E in an inner-product space is orthonormal if $\|\mathrm{e}\| = 1$ for every e ∈ E and $(e_1, e_2) = 0$ for any $e_1, e_2 \in E$ with $e_1 \neq e_2$.

Lemma If $\{e_1, \dots, e_n\}$ is an orthonormal set in an inner product space V, then for any $\{\alpha_j\}_{\mathrm{i}=1}^{n} \epsilon\ \mathbb{K}$

$$\left\|\sum_{\mathrm{j}=1}^{n} \alpha_j e_j\right\|^2 = \sum_{\mathrm{j}=1}^{n} |\alpha_j|^2$$

Proposition (Gram-Schmidt) Suppose that V is an inner-product space and E = $(e_j)_{j\in J}$ ∈ V, with J = {1,... , n} or J = N, is a linearly independent sequence. Then there exists an orthonormal sequence $\tilde{E} = (\tilde{e_j})_{j\in J}$ such that $\mathrm{Span}(e_1, \cdots, e_k) = \mathrm{Span}(\tilde{e}_1, \dots, \tilde{e}_k)$ for every $k \in J$ *and so* $clin(\tilde{E}) = clin(E)$.

In a finite-dimensional inner-product space this process guarantees the existence of an orthonormal basis, i.e. a basis of orthonormal elements: starting with any basis we use the Gram-Schmidt process to find an orthonormal basis. A similar result holds in any Hilbert space.

Proposition (Gram–Schmidt orthonormalisation) *Suppose that V is an inner-product space and $E = (e_j)_{j \in \mathcal{J}} \in H$, with $\mathcal{J} = \{1, \ldots, n\}$ or $\mathcal{J} = \mathbb{N}$, is a linearly independent sequence. Then there exists an orthonormal sequence $\tilde{E} = (\tilde{e}_j)_{j \in \mathcal{J}}$ such that*

$$\mathrm{Span}(e_1, \ldots, e_k) = \mathrm{Span}(\tilde{e}_1, \ldots, \tilde{e}_k) \qquad \text{(A)}$$

for every $k \in \mathcal{J}$, and so $\mathrm{clin}(\tilde{E}) = \mathrm{clin}(E)$.

Proof We proceed by induction, starting with $\tilde{e}_1 = e_1/\|e_1\|$.

Suppose that we already have an orthonormal set $(\tilde{e}_1, \ldots, \tilde{e}_n)$ whose linear span is the same as $(e_1, \ldots, e_n)$. Then we can define $\tilde{e}_{n+1}$ by setting

$$e'_{n+1} = e_{n+1} - \sum_{i=1}^{n} (e_{n+1}, \tilde{e}_i)\tilde{e}_i \qquad \text{and} \qquad \tilde{e}_{n+1} = \frac{e'_{n+1}}{\|e'_{n+1}\|}.$$

The span of $(\tilde{e}_1, \ldots, \tilde{e}_{n+1})$ is the same as the span of $(\tilde{e}_1, \ldots, \tilde{e}_n, e_{n+1})$, which is the same as the span of $(e_1, \ldots, e_n, e_{n+1})$ using the induction hypothesis. Clearly $\|\tilde{e}_{n+1}\| = 1$ and for $m \le n$ we have

$$(\tilde{e}_{n+1}, \tilde{e}_m) = \frac{1}{\|\mathrm{e}'_{n+1}\|}\Big((e_{n+1}, \tilde{e}_m) - \sum_{i=1}^{n} (e_{n+1}, \tilde{e}_i)(\tilde{e}_i, \tilde{e}_m)\Big) = 0$$

since $(\tilde{e}_1, \ldots, \tilde{e}_n)$ are orthonormal.

That the closed linear spans of $\tilde{E}$ and E coincide is a consequence of (A) any element in $\mathrm{clin}(E)$ can be approximated arbitrarily closed by finite linear combinations of the $\{e_j\}$, and hence by finite linear combinations of the $\{\tilde{e}_j\}$, so is an element of $\mathrm{clin}(\tilde{E})$. The same argument in reverse yields the equality of the closed linear spans. $\square$

Convergence of Orthogonal Series

Suppose that $\{e_j\}_{j=1}^{\infty}$ is an orthonormal set in an inner-product space V. If the series $\sum_{j=1}^{\infty} \alpha_j e_j$ converges to some x ∈ V, then, taking the inner product with some e_k, we obtain

$$(x, e_k) = \left(\sum_{j=1}^{\infty} \alpha_j e_j, e_k\right) = \sum_{j=1}^{\infty} \alpha_j (e_j, e_k) = \alpha_k,$$

Which shows that the coefficients α_j are completely determined, with $\alpha_j = (x, e_j)$.

(Bessel's inequality) Let V be an inner-product space and $(e_j)_{j=1}^{\infty}$ an orthonormal set in V. Then for any x ∈ V we have $\sum_{j=1}^{\infty} |(x, e_j)|^2 \le \|x\|^2$.

Proof. Let $x_k = \sum_{j=1}^{k} (x, e_j) e_j$; then $\|x_k\|^2 = \sum_{j=1}^{k} |(x, e_j)|^2$ and so

$$\|x - x_k\|^2 = \|x\|^2 - (x, x_k) - (x_k, x) + \|x_k\|^2 = \|x\|^2 - \sum_{j=1}^{k} (x, e_j)(e_j, x) - \sum_{j=1}^{k} \overline{(x, e_j)}(x, e_j) + \|x_k\|^2 = \|x\|^2 - \|x_k\|^2.$$ Thus

$$\sum_{j=1}^{k} |(x, e_j)|^2 = \|x\|^2 - \|x - x_k\|^2 \leqslant \|x\|^2.$$ Taking $k \to \infty$, we obtain the inequality.

Lemma Let H be a Hilbert space and $\{e_n\}_{n=1}^{\infty}$ an orthonormal set in H.

$\sum_{n=1}^{\infty} \alpha_n e_n$ converges $\Leftrightarrow$ $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$ and then

$$\|\sum_{n=1}^{\infty} \alpha_n e_n\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2 \qquad (*)$$

Proof If $\sum_{j=1}^{n} \alpha_j e_j \to x$ as $n \to \infty$, then $\left\|\sum_{j=1}^{n} \alpha_j e_j\right\|^2 = \sum_{j=1}^{n} |\alpha_j|^2 \to \|x\|^2$, $n \to \infty$. Thus $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$.

Conversely, if $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$, then $\left(\sum_{j=1}^{n} |\alpha_j|^2\right)_{n=1}^{\infty}$ is a Cauchy sequence. Setting

$x_n = \sum_{j=1}^{n} \alpha_j e_j$, we have, taking $m > n$

$\|x_m - x_n\|^2 = \left\|\sum_{j=n+1}^{m} \alpha_j e_j\right\|^2 = \sum_{j=n+1}^{m} |\alpha_j|^2$. Hence (x_n) is Cauchy and so converges

to some $x \in H$. The equality (*) follows as before.

Corollary Let H be a Hilbert space and $\{e_n\}_{n=1}^{\infty}$ an orthonormal set in H. Then

$\sum_{n=1}^{\infty} (x, e_n) e_n$ converges for every x $\in$ H.

Proposition *Let* $E = \{e_j\}_{j=1}^{\infty}$ *be an orthonormal set in a Hilbert space* H*. Then the following statements are equivalent:*

(a) E *is a basis for* H*;*

(b) *for any* x *we have*

$$x = \sum_{j=1}^{\infty} (x, e_j) e_j \qquad \textit{for all} \qquad x \in H;$$

(c) *Parseval's identity holds:*

$$\|x\|^2 = \sum_{j=1}^{\infty} |(x, e_j)|^2 \qquad \textit{for all} \qquad x \in H;$$

(d) $(x, e_j) = 0$ *for all* j *implies that* $x = 0$*; and*

(e) $\mathrm{clin}(E) = H$.

Example The sequence $(\boldsymbol{e}^{(j)})_{j=1}^{\infty}$ defined by $\boldsymbol{e}^{(j)} = (0, 0, \ldots, 1, 0, \ldots)$, is an orthonormal basis for for ℓ^2, since it is clear that if $(\boldsymbol{x}, \boldsymbol{e}^{(j)}) = x_j = 0$ for all j then $\boldsymbol{x} = \boldsymbol{0}$.

Part (e) means that the linear span of E is dense in H, i.e. for any $x \in H$ and any $\epsilon > 0$ there exists an $n \in \mathbb{N}$ and $\alpha_j \in \mathbb{K}$ such that

$$\left\| x - \sum_{j=1}^{n} \alpha_j e_j \right\| < \epsilon.$$

See Exercise 9.8 for an example showing that if E is linearly independent but not orthonormal, then $\operatorname{clin}(E) = H$ does not necessarily imply that E is a basis for H.

Proof First we show (a)⇔(b). If E is an orthonormal basis for H, then we can write

$$x = \sum_{j=1}^{\infty} \alpha_j e_j, \text{ i.e. } x = \lim_{n\to\infty} \sum_{j=1}^{n} \alpha_j e_j.$$

Clearly if $k \leq n$ we have

$$\left(\sum_{j=1}^{n} \alpha_j e_j, e_k \right) = \alpha_k;$$

taking the limit $n \to \infty$ it follows that $\alpha_k = (x, e_k)$ and hence (a) holds. The same argument shows that if we assume (b), then this expansion is unique, and so E is a basis.

We show that $(b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (b)$, and then that $(b) \Rightarrow (e)$ and $(e) \Rightarrow (d)$. (b) $\Rightarrow$ (c) is immediate from (*). (c) $\Rightarrow$ (d) is immediate since $\|x\| = 0$ implies that $x = 0$. (d) $\Rightarrow$ (b) Take $x \in H$ and let $y = x - \sum_{j=1}^{\infty}(x, e_j)e_j$. For each $m \in \mathbb{N}$ we have

$$(y, e_m) = (x, e_m) - \lim_{n\to\infty} \left(\sum_{j=1}^{n} (x, e_j)e_j, e_m \right) = 0$$

since eventually $n \geq m$. It follows from (d) that $y = 0$, i.e. that

$$x = \sum_{j=1}^{\infty} (x, e_j)e_j$$

as required. (b)⇒ (e) is clear, since given any x and $\epsilon > 0$ there exists an n such that

$$\left\| \sum_{j=1}^{n} \alpha_j e_j - x \right\| < \epsilon.$$

(e)⇒ (d) Suppose that $x \in H$ with $(x, e_j) = 0$ for every j. Choose x_n contained in the linear span of E such that $x_n \to x$. Then $\|x\|^2 = (x, x) = \lim_{n\to\infty}(x_n, x) = 0$, since x_n is a (finite) linear combination of the e_j. So $x = 0$.

Proposition *An infinite-dimensional Hilbert space is separable if and only if it has a countable orthonormal basis.*

Proof If a Hilbert space has a countable basis, then we can construct a count- able dense set by taking finite linear combinations of the basis elements with rational coefficients, and so it is separable.

If H is separable, let $E' = (x_n)_{n=1}^{\infty}$ be a countable dense subset. In par- ticular, the closed linear span of E' is the whole of H. Remove from E' any element x_n that can be written as a linear combination of $\{x_1, \cdots, x_{n-1}\}$, to give a new set E whose linear span is still dense but that is linearly independent. Now use the Gram–Schmidt process to obtain a countable orthonormal set whose closed linear span is all of H. The above Proposition guarantees that E is therefore a countable orthonormal basis.

Theorem *Any infinite-dimensional separable Hilbert s-pace H over $\mathbb{K}$ is isometrically isomorphic to $l^2(\mathbb{K})$, i.e. $H \equiv l^2(\mathbb{K})$.*

Proof H has a countable orthonormal basis $\{e_j\}_{j=1}^{\infty}$ since it is separable. Define a linear map $\phi : H \to l^2$ by setting

$$\phi(u) := ((u, e_1), (u, e_2), \cdots, (u, e_n), \cdots);$$

clearly the inverse map, given by $\phi^{-1}(\alpha) = \sum_{j=1}^{\infty} \alpha_j e_j$ is also linear, where

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n, ...),$$

Also, ϕ is bijective and we have $\|u\|_H = \|\phi(u)\|_2$, so ϕ is an isometry.

Projections in Hilbert Spaces

Theorem *Let H be a Hilbert space with scalar product $(\cdot,\cdot)$ and induced norm $\|\cdot\|$, and let C be a nonempty, convex, closed subset of H. Then for all $x \in H$ there exists a unique $y \in C$ such that*

$$\|x - y\| = d(x, C) := \inf_{v \in C} \|x - v\| .$$

Moreover, for any $\tilde{y} \in C$,

$$Re\,(\tilde{y} - y, x - y) \leq 0.$$

Proof.

Assume $x \in H \setminus C$. Denote $\rho = d(x, C)$. By the definition of inf, for all $n \in \mathbb{N}$ there exists $y_n \in C$ such that

$$\rho \leq \|x - y_n\| < \rho + \frac{1}{n} ,$$

which gives

$$\lim_{n \to \infty} \|x - y_n\| = \rho .$$

Apply the parallelogram law to $x - y_n$ and $x - y_m$ to get

$$\|2x - (y_n + y_m)\|^2 + \|y_n - y_m\|^2 = 2\left(\|x - y_n\|^2 + \|x - y_m\|^2\right), \quad \text{for all m, n.}$$

Since $\frac{1}{2}(y_n + y_m)$ is in C, we have

$$\|2x - (y_n + y_m)\|^2 = 4\left\|x - \frac{1}{2}(y_n + y_m)\right\|^2 \geq 4\rho^2.$$

Thus

$$\|y_n - y_m\|^2 \leq 2(\|x - y_m\|^2 + \|x - y_n\|^2) - 4\rho^2$$

and so (y_n) is a Cauchy sequence. Let $y_n \to y$. We have $y \in C$ because C is closed.

It follows that

$$\|x - y\| = \rho.$$

We now prove uniqueness. Suppose $\|x-y\| = \rho = \|x-y'\|$ for some $y, y' \in C$. We use the parallelogram law for $x-y, x-y'$ to obtain

$$\|2x-(y+y')\|^2 + \|y-y'\|^2 = 2\big(\|x-y\|^2 + \|x-y'\|^2\big)$$

which implies

$$4\|x-(1/2)(y+y')\|^2 + \|y-y'\|^2 = 4\rho^2 .$$

$(1/2)(y+y') \in C$ since it is a convex combination, therefore

$$4\|x-(1/2)(y+y')\|^2 \geq 4\rho^2$$

yielding

$$\|y-y'\|^2 \leq 4\rho^2 - 4\rho^2 = 0\,,$$

and thus $y = y'$.

If $\tilde{y} \in C, then\ since\ C\ is\ convex,\ we\ have\ \ for\ any\ \ t \in (0,1),\ that\ (1-t)y + \tilde{y} \in C,\ and\ so$

$$\begin{aligned}\|x-y\|^2 &\leq \big\|x-\big((1-t)y+t\hat{y}\big)\big\|^2\\ &= \|x-y\|^2 - 2t\,\mathrm{Re}(x-y, \tilde{y}-y) + t^2\|\tilde{y}-y\|^2\end{aligned}$$

Thus

$$\mathrm{Re}(x-y, \tilde{y}-y) \leq 0.$$

Linear Subspaces and Orthogonal Complements

If X is a subset of a Hilbert space H, then the *orthogonal complement of X in H* is

$$X^{\perp} = \{u \in H : \ (u, x) = 0 \qquad \text{for all} \qquad x \in X\}.$$

Clearly, if $Y \subseteq X$, then $X^{\perp} \subseteq Y^{\perp}$. Note also that $X \cap X^{\perp} \subset \{0\}$.

If X is a subset of H, then $X^{\perp}$ is a closed linear subspace of H.

Observe that $\{e_j\}_{j=1}^{\infty}$ is a basis for H if and only if $\left(\{e_j\}_{j=1}^{\infty}\right)^{\perp} = \{0\}$.

We now show that given any closed linear subspace U of H, any $x \in H$ has a unique decomposition in the form $x = u + v$, where $u \in U$ and $v \in U^{\perp}$: we say that H is the direct sum of U and $U^{\perp}$ and write $H = U \oplus U^{\perp}$.

Proposition *If U is a closed linear subspace of a Hilbert space H, then any $x \in H$ can be written uniquely as*

$$x = u + v \qquad \textit{with} \qquad u \in U, \quad v \in U^{\perp},$$

i.e. $H = U \oplus U^{\perp}$. The map $P_U : H \to U$ defined by

$$P_U x := u$$

is called the orthogonal projection of x onto U, *and satisfies*

$$P_U^2 x = P_U x \qquad \textit{and} \qquad \|P_U x\| \leq \|x\| \qquad \textit{for all } x \in H.$$

Proof If U is a closed linear subspace, then U is closed and convex. For a given $x \in H$ there is a unique closest point $u \in U$. It is now simple to show that $x - u \in U^{\perp}$ and then such a decomposition is unique. In fact, given any $v \in U$, we have $u \pm v \in U$, so

$$\mathrm{Re}(x - u, \pm v) \leq 0,$$

which shows that $\mathrm{Re}(x - u, \pm v) = 0$. Choosing instead $u \pm \mathrm{i}v \in U$ we obtain

$$\mathrm{Im}(x - u, v) = 0, \text{ and so } (x - u, v) = 0 \text{ for every } v \in U, \text{ i.e. } x - u \in U^{\perp}.$$

Finally, the uniqueness follows easily: if $x = u_1 + v_1 = u_2 + v_2$, then

$$u_1 - u_2 = v_2 - v_1, \text{ and so}$$

$$\|u_1 - u_2\|^2 = (u_1 - u_2, v_2 - v_1) = 0,$$

since $u_1 - u_2 \in U$ and $v_2 - v_1 \in U^{\perp}$.

If $P_U x$ denotes the closest point to x in U, then clearly $P_U^2 = P_U$, and it follows from the fact that $(u, x - u) = 0$ that

$$\|x\|^2 = \|u\|^2 + \|x - u\|^2,$$

and so

$$\|P_U x\| \leq \|x\|,$$

i.e. the projection can only decrease the norm.

Lemma If $X \subseteq H$, then $X \subseteq (X^{\perp})^{\perp}$ with equality if and only if X is a closed linear subspace of H.

Proof Any $x \in X$ satisfies

$$(x, z) = 0 \quad \text{for every } z \in X^{\perp};$$

so $X \subseteq (X^{\perp})^{\perp}$.

Now suppose that z ∈ $(X^{\perp})^{\perp}$, so that (z, y) = 0 for every y ∈ $X^{\perp}$. If X is a closed linear subspace, then we can write z = x + y, where x ∈ X and y ∈ $X^{\perp}$. But then (since z ∈ $(X^{\perp})^{\perp}$) we have 0 = (z, y) = (x + y, y) = $\|y\|^2$, so in fact y = 0 and therefore z ∈ X.

Finally, if then $X = (X^{\perp})^{\perp}$, then X must be a closed linear subspace.

Theorem *Let $E = \{e_j\}_{j\in\mathcal{J}}$ be an orthonormal set, where $\mathcal{J} = \mathbb{N}$ or $(1, 2, \ldots, n)$. Then for any $x \in H$, the orthogonal projection of x onto* $\operatorname{clin}(E)$, *which is the closest point to x in* $\operatorname{clin}(E)$, *is given by*

$$P_E x := \sum_{j\in\mathcal{J}} (x, e_j) e_j.$$

Proof Consider $x - \sum_{j\in\mathcal{J}} \alpha_j e_j$. Then

$$\begin{aligned}
\left\| x - \sum_{j\in\mathcal{J}} \alpha_j e_j \right\|^2 &= \|x\|^2 - \sum_{j\in\mathcal{J}} (x, \alpha_j e_j) - \sum_{j\in\mathcal{J}} (\alpha_j e_j, x) + \sum_{j\in\mathcal{J}} |\alpha_j|^2 \\
&= \|x\|^2 - \sum_{j\in\mathcal{J}} \overline{\alpha_j}(x, e_j) - \sum_{j\in\mathcal{J}} \alpha_j \overline{(x, e_j)} + \sum_{j\in\mathcal{J}} |\alpha_j|^2 \\
&= \|x\|^2 - \sum_{j\in\mathcal{J}} |(x, e_j)|^2 \\
&\quad + \sum_{j\in\mathcal{J}} \left[|(x, e_j)|^2 - \overline{\alpha_j}(x, e_j) - \alpha_j \overline{(x, e_j)} + |\alpha_j|^2 \right] \\
&= \|x\|^2 - \sum_{j\in\mathcal{J}} |(x, e_j)|^2 + \sum_{j\in\mathcal{J}} |(x, e_j) - \alpha_j|^2,
\end{aligned}$$

and so the minimum occurs when $\alpha_j = (x, e_j)$ for all $j \in \mathcal{J}$. □

Definition. *A system of mutually orthogonal vectors of unit length in a* inner product *space X is called orthonormal.*

Such a system $\{e_\alpha\}$ is called an orthonormal basis in X if, for every $x \in X$, there exists an at most countable subsystem $\{e_{\alpha_n}\} \subset \{e_\alpha\}$ and a finite or countable collection of scalars $\{c_n\}$ for which $x = \sum_{n=1}^{\infty} c_n e_{\alpha_n}$, where the series converges in X.

A system of vectors is called *complete* if its linear span is dense. An orthonormal basis is a complete system.

Theorem. *Every nonzero Hilbert space possesses an orthonormal basis.*

PROOF. Let $\mathcal{B}$ be the set of all orthonormal systems in a Hilbert space X partially ordered by inclusion. Every chain $\mathcal{B}_0 \subset \mathcal{B}$ has an upper bound: we can take the union $\mathcal{V}$ of all vectors belonging to the families in $\mathcal{B}_0$. Any two different vectors x and y in $\mathcal{V}$ are orthogonal, since $x \in \mathcal{V}_1 \in \mathcal{B}_0$, $y \in \mathcal{V}_2 \in \mathcal{B}_0$, and either $\mathcal{V}_1 \subset \mathcal{V}_2$ or $\mathcal{V}_2 \subset \mathcal{V}_1$ by the linear ordering of $\mathcal{B}_0$. By Zorn's lemma there is a maximal element in $\mathcal{B}$, i.e., an orthonormal family $\{e_\alpha\}$ that is not a part of a larger orthonormal system. This means that there is no nonzero vector orthogonal to all e_α. From X=clin$(\{e_\alpha\}) \oplus$ clin$(\{e_\alpha\})^\perp$ = clin$(\{e_\alpha\})$, we conclude that the linear span of $\{e_\alpha\}$ is dense in X. Hence every vector x is the limit of a sequence of linear combinations of some countable subfamily $\{e_{\alpha_n}\}$. We have

$$x = \sum_{n=1}^{\infty} (x, e_{\alpha_n}) e_{\alpha_n}.$$

Let P be a set with a (partial) order relation $\leq$. We say that a subset $Q \subset P$ is *totally ordered* if for any pair (a, b) in Q either $a \leq b$ or $b \leq a$ (or both!). Let $Q \subset P$ be a subset of P; we say that $c \in P$ is an *upper bound* for Q if $a \leq c$ for every $a \in Q$. We say that $m \in P$ is a *maximal* element of P if there is *no* element $x \in P$ such that $m \leq x$, except for $x = m$. Note that a maximal element of P need not be an upper bound for P. A totally ordered subset is called a chain.

We say that P is *inductive* if every totally ordered subset Q in P has an upper bound.

Lemma (Zorn). *Every nonempty ordered set that is inductive has a maximal element.*