MA3422 2016–17

Chapter 1: Hilbert Spaces

Contents

1.1 Definition and examples

1.1.1 Definition. An *inner product space* (also known as a pre-Hilbert space) is a vector space V over $\mathbb{K}$ (= $\mathbb{R}$ or $\mathbb{C}$) together with a map

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{K}$$

satisfying (for $x, y, z \in V$ and $\lambda \in \mathbb{K}$):

(i) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

(ii) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$

(iii) $\langle y, x\rangle = \overline{\langle x, y\rangle}$

(iv) $\langle x, x\rangle \geq 0$

(v) $\langle x, x\rangle = 0 \Rightarrow x = 0$

Note that it follows from the first 3 properties that:

(i)' $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$

(ii)' $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$

An inner product on V gives rise to a norm

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

If the inner product space is complete in this norm (or in other words, if it is complete in the metric arising from the norm, or if it is a Banach space with this norm) then we call it a *Hilbert space*.

Another way to put it is that a Hilbert space is a Banach space where the norm arises from some inner product, but to make that claim we need to know that the norm determines the inner product. (We will see that this is so later. The issue here is that an inner product space is two things, a vector space together with an inner product. A normed space is also two things, a vector space together with a norm. If a vector space H has two inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ giving rise to the same norm $\|x\| = \sqrt{\langle x, x \rangle_1} = \sqrt{\langle x, x \rangle_2}$, then in fact $\langle x, y \rangle_1 = \langle x, y \rangle_2$ holds for all $x, y \in H$. The inner product is uniquely determined by the norm, if there is an inner product. We will eventually get to proving this in Theorem 1.3.1.)

1.1.2 Examples. (i) $\mathbb{C}^n$ with the inner product $\langle z, w \rangle = \sum_{j=1}^n z_j \overline{w_j}$ is a Hilbert space (over $\mathbb{K} = \mathbb{C}$). (Here we mean that $z = (z_1, z_2, \ldots, z_n)$ and $w = (w_1, w_2, \ldots, w_n)$.)

We know that $\mathbb{C}^n$ is complete (in the standard norm, which is the one arising from the inner product just given, but also in any other norm) and so $\mathbb{C}^n$ is a Hilbert space.

(ii) $\mathbb{R}^n$ with the inner product $\langle x, y \rangle = \sum_{j=1}^n x_j y_j$ is a Hilbert space over $\mathbb{R}$.

(iii) ℓ^2 with the inner product

$$\langle a, b \rangle = \sum_{j=1}^{\infty} a_j \overline{b_j}$$

is a Hilbert space over $\mathbb{K}$ (where we mean that $a = \{a_j\}_{j=1}^{\infty}$, $b = \{b_j\}_{j=1}^{\infty}$). The fact that the series for $\langle a, b \rangle$ always converges is a consequence of Hölder's inequality with $p = q = 2$. The properties that an inner product must satisfy are easy to verify here. The norm that comes from the inner product is the norm $\| \cdot \|_2$ we had already on ℓ^2.

(iv) $L^2[0, 1]$, $L^2[a, b]$ and $L^2(\mathbb{R})$ are all Hilbert spaces with respect to the inner product

$$\langle f, g \rangle = \int f \overline{g}$$

(the integral to be taken over the appropriate domain). [Since $|f(x)||g(x)| \leq (1/2)(|f(x)|^2 + |g(x)|^2)$ it is quite easy to see that $f\overline{g}$ is integrable for f and g in the same L^2 space.]

1.1.3 Remarks. (i) The triangle inequality holds on any inner product and this is proved via the Cauchy-Schwarz inequality:

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|$$

(for the norm arising from inner product). Equality holds in this inequality if and only if x and y are linearly dependent.

(ii) One can use Cauchy-Schwarz to show that the inner product map $\langle\cdot,\cdot\rangle \colon V \times V \to \mathbb{K}$ is always continuous on $V \times V$ (for any inner product space, and where we take the product topology on $V \times V$). If we take a sequence (x_n, y_n) converging in $V \times V$ to a limit (x, y), then $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$ in V and so

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &\leq |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &= |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leq \|x_n\|\,\|y_n - y\| + \|x_n - x\|\,\|y\| \\ &\to 0 \text{ as } n \to \infty \end{aligned}$$

(In the last step we are using the fact that $(x_n)_{n=1}^\infty$ is bounded, that is $\sup_n \|x_n\| < \infty$, because $\lim_{n\to\infty} x_n$ exists.)

1.2 Orthonormal bases

1.2.1 Notation. By default we will use the letter H to denote a Hilbert space.

Two elements x and y of an inner product space are called *orthogonal* if $\langle x, y\rangle = 0$.

A subset $S \subset H$ of a Hilbert space (or of an inner product space) is called *orthogonal* if

$$x, y \in S, x \neq y \Rightarrow \langle x, y\rangle = 0.$$

S is called *orthonormal* if it is an orthogonal subset and if in addition $\|x\| = 1$ for each $x \in S$.

Observe that these definitions are phrased so that they apply to both finite and infinite subsets S.

Note also that if S is orthogonal, then $\{x/\|x\| : x \in S \setminus \{0\}\}$ is orthonormal.

1.2.2 Proposition. *If $S \subset H$ is any orthonormal subset of an inner product space H and if $x \in H$, then*

(i) $\langle x, \phi \rangle$ *is nonzero for at most a countable number of* $\phi \in S$.

(ii)
$$\sum_{\phi \in S} |\langle x, \phi \rangle|^2 \leq \|x\|^2 \qquad \textit{(Bessel's inequality).} \tag{1.2.1}$$

Observe that (i) implies that we can list those $\phi \in S$ for which $\langle x, \phi \rangle \neq 0$ as a finite or infinite list $\phi_1, \phi_2, \ldots$ and then (ii) means that

$$\sum_n |\langle x, \phi_n \rangle|^2 \leq \|x\|^2. \tag{1.2.2}$$

The sum is independent of the order in which the $\phi_1, \phi_2, \ldots$ are listed (because it is a series of positive terms).

Proof. Suppose $\phi_1, \phi_2, \ldots, \phi_n$ are n distinct elements of S and $x \in H$. Let $a_j = \langle x, \phi_j \rangle$ for $1 \leq j \leq n$. Then

$$\begin{aligned}
0 &\leq \left\| x - \sum_{j=1}^n a_j \phi_j \right\|^2 \\
&= \left\langle x - \sum_{j=1}^n a_j \phi_j, x - \sum_{k=1}^n a_k \phi_k \right\rangle \\
&= \langle x, x \rangle - \sum_{k=1}^n \bar{a}_k \langle x, \phi_k \rangle - \sum_{j=1}^n a_j \langle \phi_j, x \rangle + \sum_{j,k=1}^n a_j \bar{a}_k \langle \phi_j, \phi_k \rangle \\
&= \langle x, x \rangle - \sum_{k=1}^n \bar{a}_k a_k - \sum_{j=1}^n a_j \bar{a}_j + \sum_{j=1}^n a_j \bar{a}_j \\
&\qquad \text{since } \langle \phi_j, \phi_k \rangle = 0 \text{ if } j \neq k \\
&= \langle x, x \rangle - \sum_{k=1}^n |a_k|^2
\end{aligned}$$

Therefore
$$\sum_{k=1}^n |a_k|^2 = \sum_{k=1}^n |\langle x, \phi_k \rangle|^2 \leq \langle x, x \rangle = \|x\|^2. \tag{1.2.3}$$

Now we can finish the proof by making use of this finite version of Bessel's inequality.

To show that there are only a countable number of $\phi \in S$ with $\langle x, \phi \rangle \neq 0$, consider the set of all such ϕ:

$$S_x = \{\phi \in S : \langle x, \phi \rangle \neq 0\} = \bigcup_{n=1}^{\infty} S_x^n$$

where

$$S_x^n = \{\phi \in S : |\langle x, \phi \rangle| \geq 1/n\}.$$

Now each S_x^n is finite because if we could find N elements $\phi_1, \phi_2, \ldots, \phi_N$ in S_x^n then by (1.2.3)

$$\|x\|^2 \geq \sum_{j=1}^{N} |\langle x, \phi_j \rangle|^2 \geq N \left(\frac{1}{n}\right)^2$$

and so $N \leq n^2 \|x\|^2$.

Thus S_x^n is finite and S_x is a countable union of finite sets, hence countable.

If we list the elements of S_x in a finite or infinite list $\{\phi_1, \phi_2, \ldots\}$ then we can let $n \to \infty$ in (1.2.3) to get Bessel's inequality in the form (1.2.2). □

1.2.3 Lemma. *If V is a vector space (over a field $\mathbb{K}$) and $S \subseteq V$ is any subset, then there is a smallest vector subspace $\operatorname{span}(S)$ of V that contains S (and $\operatorname{span}(S)$ is referred to as the* linear span *of S).*

Proof. An abstract way to establish this is to take the intersection of all vector subspaces $E \subseteq V$ with $S \subseteq E$, then show that this intersection is a vector subspace.

A more constrictive way is to take $\operatorname{span}(S)$ to be the collection of all finite linear combinations $\sum_{j=1}^{n} a_j \phi_j$ with $a_j \in \mathbb{K}$ and $\phi_j \in S$ for $1 \leq j \leq n$. We should allow $n = 0$, with the empty sum interpreted as $0 \in V$ so that $0 \in \operatorname{span}(S)$ even if $S = \emptyset$. Clearly any vector subspace that contains S must contain these linear combinations, but we would still need to check that the collection of all such linear combinations is indeed a vector subspace (which is fairly easy to check). □

1.2.4 Proposition. *Let H be an inner product space $x \in H$ and $\phi_1, \phi_2, \ldots, \phi_n$ orthonormal elements of H. Let $E = \operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_n\}$.*

Then $y = \sum_{j=1}^{n} \langle x, \phi_j \rangle \phi_j$ is the unique element of E closest to x.

(In other words, $\|x - y\| < \|x - z\|$ holds for all $z \in E$ other than $z = y$.)

Proof. Note that $y \in E$ and

$$\begin{aligned}
\langle x-y, \phi_k\rangle &= \langle x, \phi_k\rangle - \left\langle \sum_{j=1}^n \langle x, \phi_j\rangle \phi_j, \phi_k \right\rangle \\
&= \langle x, \phi_k\rangle - \sum_{j=1}^n \langle x, \phi_j\rangle \langle \phi_j, \phi_k\rangle \\
&= \langle x, \phi_k\rangle - \langle x, \phi_k\rangle \\
&= 0
\end{aligned}$$

Thus for any $z \in E$, say $z = \sum_{k=1}^n a_k \phi_k$ (with $a_1, a_2, \ldots, a_n \in \mathbb{K}$) we have

$$\langle x-y, z\rangle = \sum_{k=1}^n \bar{a}_k \langle x-y, \phi_k\rangle = 0.$$

(We would usually say that $x-y$ is orthogonal to E, meaning orthogonal to all elements of E.)

So if $z \in E$ we have $x-z = (x-y)+(y-z)$ and $y-z \in E$. Hence

$$\begin{aligned}
\|x-z\|^2 &= \langle x-z, x-z\rangle \\
&= \langle (x-y)+(y-z), (x-y)+(y-z)\rangle \\
&= \langle x-y, x-y\rangle + \langle y-z, y-z\rangle \\
&\quad + \langle x-y, y-z\rangle + \langle y-z, x-y\rangle \\
&= \|x-y\|^2 + \|y-z\|^2 + 0 \\
&\geq \|x-y\|^2
\end{aligned}$$

This no $z \in E$ is closer to x than y is and we have strict inequality unless $y = z$. □

1.2.5 Notation. In the situation of Proposition 1.2.4, the element y is usually referred to as the *orthogonal projection of* x *on* E (because $y \in E$ and $x = y + (x-y)$ with $x-y$ orthogonal to (all elements of) E.

1.2.6 Definition. If H is an inner product space then a finite sequence $\phi_1, \phi_2, \ldots, \phi_n \in H$ of orthonormal vectors is called an *orthonormal basis* for H if $x = \sum_{j=1}^n \langle x, \phi_j\rangle \phi_j$ holds for each $x \in H$.

An infinite orthonormal sequence $\phi_1, \phi_2, \ldots \in H$ is also called an *orthonormal basis* for H if $x = \sum_{j=1}^\infty \langle x, \phi_j\rangle \phi_j$ holds for each $x \in H$.

1.2.7 Remark. Since it only involves finite sums, Proposition 1.2.4 implies that a finite orthonormal sequence $\phi_1, \phi_2, \dots, \phi_n \in H$ is an orthonormal basis for H if and only if $\operatorname{span}\{\phi_1, \phi_2, \dots, \phi_n\} = H$.

Moreover, it is easy to see that if $\phi_1, \phi_2, \dots, \phi_n$ are orthonormal then they must be linearly independent (hint: for $\sum_{j=1}^n a_j \phi_j = 0$, consider $\left\langle \sum_{j=1}^n a_j \phi_j, \phi_k \right\rangle$ for $1 \leq k \leq n$) and so form a basis for their linear span (with the linear algebra meaning for the word 'basis').

1.2.8 Example. In $H = \ell^2$, let e_n denote the sequence where all the terms are 0 except the n^{th} term, which is 1. It may be more helpful to write

$$e_n = (0, 0, \dots, 0, 1, 0, 0, \dots)$$

(with 1 in the n^{th} position).

Let

$$B = \{e_1, e_2, \dots\} = \{e_n : n \in \mathbb{N}\}.$$

Then B is orthonormal because $\|e_n\|_2 = 1$ for each n and if $n \neq m$ then $\langle e_n, e_m \rangle = 0$.

Also if $x = (x_j)_{j=1}^\infty = (x_1, x_2, \dots) \in \ell^2$, it is easy to see that $\langle x, e_n \rangle = x_n$. The idea of an orthonormal basis is that we can express x (any $x \in \ell^2$)

$$x = \sum_{j=1}^\infty \langle x, e_j \rangle e_j = \lim_{n \to \infty} \sum_{j=1}^n \langle x, e_j \rangle e_j$$

and we can check that quite easily in this case. We have

$$\sum_{j=1}^n \langle x, e_j \rangle e_j = \sum_{j=1}^n x_j e_j = (x_1, x_2, \dots, x_n, 0, 0, \dots)$$

and

$$x - \sum_{j=1}^n \langle x, e_j \rangle e_j = (0, 0, \dots, 0, x_{n+1}, x_{n+1}, \dots).$$

So

$$\left\| x - \sum_{j=1}^n \langle x, e_j \rangle e_j \right\|_2 = \sqrt{\sum_{j=n+1}^\infty |x_j|^2}.$$

Recall that $x \in \ell^2$ means that $\sum_{j=1}^{\infty} |x_j|^2 < \infty$ and so it follows that

$$\sum_{j=n+1}^{\infty} |x_j|^2 \to 0 \text{ as } n \to \infty.$$

We can see then that

$$\lim_{n\to\infty} \left\| x - \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\|_2 = 0$$

and that is what it means to say $x = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$.

1.2.9 Remark. Our aim is to consider the notion of an orthonormal basis for infinite dimensional Hilbert spaces like the above basis for ℓ^2 (known as the 'standard basis' because it is a rather obvious extension to the infinite case of the standard basis for $\mathbb{K}^n$). Since we use infinite sums in the infinite dimensional case, we are no longer dealing with pure linear algebra when we do this.

To make life simpler we restrict ourselves to separable Hilbert spaces, though it is possible to extend the notion of an orthonormal basis to an uncountable one and then to include general (non-separable) Hilbert spaces.

1.2.10 Proposition. *Let H be an inner product space and $S = \{\phi_1, \phi_2, \ldots\}$ an orthonormal sequence of vectors in H such that* $\operatorname{span}(S)$ *is dense in H. Then S is an orthonormal basis for H.*

Proof. Fix $x \in H$. Since $\operatorname{span}(S)$ is dense in S, there must be a sequence $y_1, y_2, \ldots \in \operatorname{span}(S)$ with $\lim_{n\to\infty} y_n = x$ (limit in the norm of H).

Now, each y_n is a finite linear combination

$$y_n = \sum_{j=1}^{m_n} a_{n,j} \phi_j$$

of elements of S (we can take the elements of S to be those from ϕ_1 up to the largest ϕ_{m_n} in the linear combination so as to get sums of this form).

Because of Proposition 1.2.4,

$$\left\| x - \sum_{j=1}^{m_n} \langle x, \phi_j \rangle \phi_j \right\| \leq \|x - y_n\|$$

and so

$$\lim_{n\to\infty} \left\| x - \sum_{j=1}^{m_n} \langle x, \phi_j \rangle \phi_j \right\| = 0.$$

Also because of Proposition 1.2.4,

$$\left\| x - \sum_{j=1}^{N} \langle x, \phi_j \rangle \phi_j \right\|$$

is monotone decreasing with N. (Because increasing $\operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_N\}$ to $\operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_{N+1}\}$ has to make the closest element in the span closer to x, or at least not farther away.)

So $\lim_{N\to\infty} \left\| x - \sum_{j=1}^{N} \langle x, \phi_j \rangle \phi_j \right\| = 0$, or in other words $x = \sum_{j=1}^{\infty} \langle x, \phi_j \rangle \phi_j$. □

1.2.11 Theorem. *For an inner product space H, the following are equivalent properties of H:*

(a) H is separable (that is, has a countable dense subset)

(b) H has a countable orthonormal basis (finite basis or infinite sequence).

Proof. $(a) \Rightarrow$ (b): Let $D = \{d_1, d_2, \ldots\} \subset H$ be a countable dense subset. We can then construct a finite or infinite sequence $\psi_1, \psi_2, \ldots$ which has

i) $\psi_1, \psi_2, \ldots, \psi_n$ is linearly independent (for each n in the infinite case, or for n as large as possible in the finite case)

ii) $D \subseteq \operatorname{span}\{\psi_1, \psi_2, \ldots\}$

To do this let $\psi_1 = d_1$ unless $d_1 = 0$, and if $d_1 = 0$ let $\psi_1 = d_j$ for the smallest j with $d_j \neq 0$. (In the case $d_j = 0$ always, then H is the zero space and the empty set counts as an orthonormal basis for H.) Next, continue inductively. Once $\psi_1, \psi_1, \ldots, \psi_n$ have been chosen let $\psi_{n+1} = d_j$ for the smallest j such that $d_j \notin \operatorname{span}\{\psi_1, \psi_2, \ldots, \psi_n\}$. If there is no such j, we stop at ψ_n.

Now use the Gram-Schmidt orhonormalisation procedure on $\psi_1, \psi_2, \ldots$ to find orthonormal vectors $\phi_1, \phi_2, \ldots$ such that

$$\operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_n\} = \operatorname{span}\{\psi_1, \psi_2, \ldots, \psi_n\}$$

(Recall that $\phi_1 = \psi_1/\|\psi_1\|$ and ψ_{n+1} is obtained by dividing

$$\psi_{n+1} - \sum_{j=1}^{n} \langle \psi_{n+1}, \phi_j \rangle \phi_j$$

— the difference between ϕ_{n+1} and its orthogonal projection onto the span of the previous ones — by its norm. See Theorem A.0.1 for a recap of Gram-Schmidt.)

In the case of a finite number of $\phi_1, \phi_2, \ldots, \phi_n$ we end up with an orthonormal basis for the finite dimensional H, but in the infinite case we have

$$\begin{aligned} D &\subset \operatorname{span}\{\psi_1, \psi_2, \ldots\} = \bigcup_{n=1}^{\infty} \operatorname{span}\{\psi_1, \psi_2, \ldots, \psi_n\} \\ &= \bigcup_{n=1}^{\infty} \operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_n\} = \operatorname{span}\{\phi_1, \phi_2, \ldots\} \end{aligned}$$

Since D is dense, Proposition 1.2.10 implies that $\{\phi_1, \phi_2, \ldots\}$ is an orthonormal basis for H.

$(b) \Rightarrow$ (a): Since the finite dimensional case is relatively easy, let $S = \{\phi_1, \phi_2, \ldots\}$ be an orthonormal basis for H. We look at the case $\mathbb{K} = \mathbb{R}$ (H a real Hilbert space) first. It is quite easy to check that the sets

$$D_n = \left\{ \sum_{j=1}^{n} q_j \phi_j : q_j \in \mathbb{Q} \text{ for } 1 \leq j \leq n \right\}$$

are each countable and that their union $D = \bigcup_{n=1}^{\infty} D_n$ is countable and dense in H. The closure of each D_n is easily seen to be the $\mathbb{R}$-linear span of $\phi_1, \phi_2, \ldots, \phi_n$ and so the closure of D includes all finite linear combinations $\sum_{j=1}^{n} x_j \phi_j$. But, each $x \in H$ is a limit of such finite linear combinations (by the definition of an orthonormal basis). Hence the closure of D is all of H. As D is countable, this shows that H must be separable.

In the complex case (H a Hilbert space over $\mathbb{K} = \mathbb{C}$) we must take $q_j \in \mathbb{Q} + i\mathbb{Q}$ instead, so that we can get all finite $\mathbb{C}$-linear combinations of the ϕ_j in the closure of D (and there is no other difference in the proof).

In (the easier) finite case we could just take $n = \dim H$ (no need to take unions). $\square$

1.2.12 Remark. If we had considered the non-separable case we would have been able to show (using completeness) that every Hilbert space has an orthonormal basis (possibly uncountable). But then if H is a non-separable Hilbert space, it has to be an orthonormal basis in the more general sense that allows uncountable orthonormal bases because Theorem 1.2.11 implies that only separable Hilbert spaces have countable orthonormal bases.

1.2.13 Theorem. *If H is a separable inner product space and $S = \{\phi_1, \phi_2, \ldots\} \subset H$ is an orthonormal basis for H, then for each $x, y \in H$ we have*

$$\langle x, y\rangle = \sum_{\phi\in S} \langle x, \phi\rangle \overline{\langle y, \phi\rangle} = \sum_{j\geq 1} \langle x, \phi_j\rangle \overline{\langle y, \phi_j\rangle}$$

(where the latter sum has a finite range of j in the case where S is finite) and

$$\|x\|^2 = \sum_{\phi\in S} |\langle x, \phi\rangle|^2.$$

Proof. The case where S is finite being rather easy, we consider the infinite case only. We have

$$x = \lim_{n\to\infty} \sum_{j=1}^{n} \langle x, \phi_j\rangle \phi_j$$

and

$$y = \lim_{n\to\infty} \sum_{k=1}^{n} \langle y, \phi_k\rangle \phi_k.$$

By continuity of the inner product

$$\begin{aligned}
\langle x, y\rangle &= \lim_{n\to\infty} \left\langle \sum_{j=1}^{n} \langle x, \phi_j\rangle \phi_j, \sum_{k=1}^{n} \langle y, \phi_k\rangle \phi_k \right\rangle \\
&= \lim_{n\to\infty} \sum_{j,k=1}^{n} \langle x, \phi_j\rangle \overline{\langle y, \phi_k\rangle} \langle \phi_j, \phi_k\rangle \\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \langle x, \phi_j\rangle \overline{\langle y, \phi_j\rangle} \\
&= \sum_{n=1}^{\infty} \langle x, \phi_n\rangle \overline{\langle y, \phi_n\rangle}
\end{aligned}$$

If we do this for $x = y$ we find the second part of the statement. □

1.2.14 Theorem. *Every separable Hilbert space H over $\mathbb{K}$ is isometrically isomorphic to either $\mathbb{K}^n$ (if H has finite dimension n) or to ℓ^2. The isometric isomorphism preserves the inner product.*

Proof. The finite dimensional case is just linear algebra and we treat this as known.

If H is infinite dimensional and separable, then it has a countably infinite orthonormal basis $S = \{\phi_1, \phi_2, \ldots\}$. We can define a map

$$\begin{aligned} T\colon H &\to \ell^2 \\ \text{by } Tx &= (\langle x, \phi_n \rangle)_{n=1}^\infty \end{aligned}$$

By Theorem 1.2.13, the map T is well defined (actually maps into ℓ^2) and preserves the inner product and the norm. That is $\langle x, y \rangle = \langle Tx, Ty \rangle$ and $\|x\|_H = \|Tx\|_2$ for $x, y \in H$. Moreover, it is easy to see that T is a linear map.

From $\|Tx\| = \|x\|$ we can see that the kernel of T is just $\{0\}$ and so T is injective and what remains to be seen is that T is surjective.

To show that, consider any $a = \{a_n\}_{n=1}^\infty \in \ell^2$. Then one can quite easily verify that $\sum_{n=1}^\infty a_n \phi_n$ satisfies the Cauchy criterion for series (equivalently that the sequence of its partial sums $s_n = \sum_{j=1}^n a_j \phi_j$ is a Cauchy sequence in H, or equivalently that for $n < m$, $\|s_n - s_m\| = \left\| \sum_{j=n+1}^m a_j \phi_j \right\|$ is small as long as n is big enough) because a calculation with inner products shows that

$$\left\| \sum_{j=n}^m a_j \phi_j \right\|^2 = \sum_{j=n}^m |a_j|^2.$$

Now $a \in \ell^2 \Rightarrow \sum_{n=1}^\infty |a_n|^2 < \infty \Rightarrow \sum_{n=1}^\infty |a_n|^2$ is Cauchy and so it follows $\sum_{n=1}^\infty a_n \phi_n$ is Cauchy in H. Take $x \in H$ to be the sum of this series (which exists since H is complete) and then an argument using continuity of the inner product shows that

$$\langle x, \phi_n \rangle = a_n$$

for each n. Thus $Tx = a$ and T is surjective. □

1.2.15 Remark. We recall then that in finite dimensions we can say that all vector spaces of dimension n (over the same field) are isomorphic (choose ordered bases for each and match up coefficients with respect to the two bases). For inner product spaces we can choose orthonormal bases and this makes the correspondence more useful.

Nevertheless it is not necessarily a good idea to fix an orthonormal basis in advance. It can be advantageous to choose one to suit the problem. There is no single best basis for all problems and making an arbitrary choice may be a bad step.

For infinite dimensional separable Hilbert spaces we now have a similar statement, that all are the same. Again it depends on a choice of basis.

We don't actually have examples of further separable Hilbert spaces than ℓ^2 (in infinite dimensions). So we will repair that omission now.

1.2.16 Example. (a) $L^2[0,1]$ is a separable Hilbert space.

Unfortunately, we will only give a rough outline of the steps involved in checking this.

i) The (almost everywhere equivalence classes of) continuous functions are dense in $L^2[0,1]$. (This implies that $L^2[0,1]$ is the completion of $(C[0,1], \|\cdot\|_2)$.)

   In fact the (equivalence classes of) continuous functions $f\colon [0,1] \to \mathbb{K}$ with the extra restriction that $f(0) = f(1)$ are also dense. We write $\mathrm{CP}[0,1] = \{f \in C[0,1] : f(0) = f(1)\}$ (CP for continuous periodic).

   This is usually proved by an appeal to a theorem called Lusin's theorem.

ii) $\mathrm{CP}[0,1]$ is the closure of the trigonometric polynomials

$$p(x) = \sum_{-N}^{N} a_n e^{2\pi i n x}$$

   in the uniform norm $\|\cdot\|_\infty$. (This is a version of a theorem known as the Weierstrass theorem, or a consequence of a general version called the Stone-Weierstrass theorem.)

   Since $\|f\|_2 \leq \|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$, it follows that the closure of the trigonometric polynomials in $L^2[0,1]$ contains $\mathrm{CP}[0,1]$, and hence is all of $L^2[0,1]$.

iii) $L^2[0,1]$ is a separable Hilbert space.

   The idea is that trigonometric polynomials $p(x) = \sum_{-N}^{N} a_n e^{2\pi i n x}$ with each a_n rational (that is $a_n \in \mathbb{Q}$ when $\mathbb{K} = \mathbb{R}$, or each $a_n \in \mathbb{Q} + i\mathbb{Q}$ if $\mathbb{K} = \mathbb{C}$) form a countable dense subset of $\mathrm{CP}[0,1]$, hence of $L^2[0,1]$.

(b) There are a number of other commonly used variants of the previous example.

For instance, if $a < b$ then $L^2[a,b]$ is also a separable Hilbert space.

One way to prove this is to show that the map $T\colon L^2[a,b] \to L^2[0,1]$ given by

$$Tf(x) = \sqrt{b-a}\, f(a + (b-a)x) \qquad (f \in L^2[a,b], x \in [0,1])$$

is a linear isomorphism of inner product spaces (that is a vector space isomorphism with $\langle Tf, Tg\rangle_{L^2[0,1]} = \langle f, g\rangle_{L^2[a,b]}$ for $f, g \in L^2[a, b]$.

Another way is to follow a proof like we used for $L^2[0, 1]$ but replacing trigonometric polynomials by ordinary polynomials. (Then we need a different form of the Weierstrass theorem along with Lusin's theorem.)

(c) $L^2(\mathbb{R})$ is separable.

Here we can consider $\bigcup_{n=1}^\infty L^2[-n, n]$ as dense in $L^2(\mathbb{R})$. (To do this extend functions defined on $[-n, n]$ to the whole of $\mathbb{R}$ by making them 0 outside $[-n, n]$. Or we can think of the functions in $L^2(\mathbb{R})$ that are almost everywhere 0 on $\mathbb{R} \setminus [-n, n]$. If $f \in L^2(\mathbb{R})$ then one can show that $f\chi_{[-n,n]} \to f$ in $\|\cdot\|_2$ by using the dominated convergence theorem.)

Each $L^2[-n, n]$ has its own countable dense subset as we know. If $D_n \subset L^2[-n, n]$ is countable and dense in $L^2[-n, n]$, then $\bigcup_{n=1}^\infty D_n$ will be countable and dense in $L^2(\mathbb{R})$.

1.2.17 Example. An important and non-trivial example of an orthonormal basis is $H = L^2[0, 2\pi]$ with

$$S = \left\{ \phi_n(t) = \frac{1}{\sqrt{2\pi}} e^{int} : n \in \mathbb{Z} \right\}.$$

This fact then includes the information that $f \in L^2[0, 2\pi]$ implies

$$f = \sum_{n\in\mathbb{Z}} \langle f, \phi_n\rangle \phi_n.$$

This series is known as the Fourier series for f and the Hilbert space theory tells us that it converges to f in the norm of $L^2[0, 2\pi]$. This means that the partial sums

$$S_n f = \sum_{j=-n}^{n} \langle f, \phi_j\rangle \phi_j \to f$$

in the sense that

$$\|S_n f - f\|_2 = \sqrt{\int_0^{2\pi} |S_n f(t) - f(t)|^2 \, dt} \to 0$$

as $n \to \infty$. Thus the Fourier series converges to the function in a root-mean-square sense, but that is not the same as pointwise convergence. In fact, at any given point $t \in [0, 2\pi]$, there is no guarantee that $\lim_{n\to\infty} S_n f(t)$ exists or equals $f(t)$ if it does exist. When we recall that elements of $L^2[0, 2\pi]$ are not exactly functions, but rather equivalence classes of functions up to almost everywhere equality, we should not be surprised that we cannot pin down the Fourier series at any specific point of $[0, 2\pi]$.

Of course, it requires proof that this is indeed an example of an orthonormal basis. By integration we can easily check that S is orthonormal and then, according to the general theory, it would be enough to show that $\mathrm{span}(S)$ is a dense in $L^2(\mathbb{R})$.

This follows by the same logic as we used to show that $L^2[0, 1]$ is separable (just changing the notion of trigonometric polynomial from those with periods $n \in \mathbb{N}$ to those with periods $2\pi n$).

But maybe an easier way is to consider the isomorphism we gave previously in Examples 1.2.16 (b), that is $T\colon L^2[0, 2\pi] \to L^2[0, 1]$ with $Tf(x) = \sqrt{2\pi} f(2\pi x)$. When one does Fourier series om $[0, 1]$ one has the advantage of fewer places where the factor $1/\sqrt{2\pi}$ is needed. $T\phi_n$ is just $e^{2\pi i n x}$.

Another approach that is sometimes used is to put $1/(2\pi)$ in front of the integral when defining the inner product (and norm) on $L^2[0, 2\pi]$. With that approach the $\phi_n(t)$ would not have any $1/\sqrt{2\pi}$ factor.

1.2.18 Example. Using the fact (mentioned above without proof in Examples 1.2.16 (b)) that the linear combinations of $\phi_n(x) = x^n$ ($n = 0, 1, 2, \ldots$) are dense in $L^2[-1, 1]$. That is the polynomials are dense in $L^2[-1, 1]$.

Applying the Gram-Schmidt process to these functions ϕ_n yields an orthonormal basis for $L^2[-1, 1]$ that is related to the Legendre polynomials. The first few iterations of Gram-Schmidt yield $p_0(x) = 1/\sqrt{2}$, $p_1(x) = \sqrt{3/2}x$, $p_2(x) = \sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2 - \frac{1}{2}\right)$, $p_3(x) = \sqrt{\frac{7}{2}}\left(\frac{5}{2}x^3 - \frac{3}{2}x\right)$. The Legendre polynomials $P_n(x)$ are related to the p_n by $P_n(x) = \sqrt{\frac{2}{2n+1}} p_n(x)$. The Legendre polynomials are normalised by $P_n(1) = 1$ rather than

$$\|p_n\|_2 = \sqrt{\int_{-1}^{1} |p_n(x)|^2\,dx} = 1.$$

From the fact that the p_n form an orthonormal basis we have

$$f = \sum_{n=0}^{\infty} \langle f, p_n\rangle p_n = \sum_{n=0}^{\infty} \left(\int_{-1}^{1} f(x) p_n(x)\,dx\right) p_n$$

for each $f \in L^2[-1, 1]$.

1.3 Parallelogram Identity

The next theorem explains an earlier remark that there is only one inner product for a given Hilbert norm. This result could have been give earlier as it does not use any complicated theory.

1.3.1 Theorem (Parallelogram Identity)**.** *Let E be a normed space. Then there is an inner product on E which gives rise to the norm if and only if the parallelogram identity*

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

is satisfied by all $x, y \in E$. Moreover the inner product is uniquely determined by the norm.

Proof. $\Rightarrow$: This is a simple calculation with inner products.

$$\begin{aligned}
\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle + \\
&\quad \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle \\
&= 2\langle x,x\rangle + 2\langle y,y\rangle
\end{aligned}$$

To show that the inner product is uniquely determined by the norm, we can subtract rather than add.

$$\begin{aligned}
\|x+y\|^2 - \|x-y\|^2 &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle + \\
&\quad -(\langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle) \\
&= 2\langle x,y\rangle + 2\langle y,x\rangle \\
&= 4\,\mathrm{Re}\langle x,y\rangle
\end{aligned}$$

Thus $\mathrm{Re}\langle x,y\rangle = (\|x+y\|^2 - \|x-y\|^2)/4$ and in the case $\mathbb{K} = \mathbb{R}$ we have expressed $\langle x,y\rangle$ in terms of vector space operations and the norm. In the case $\mathbb{K} = \mathbb{C}$ we can note

$$\mathrm{Re}\langle x, iy\rangle = \mathrm{Re}(-i\langle x,y\rangle) = \mathrm{Re}(-i(\mathrm{Re}\langle x,y\rangle + i\,\mathrm{Im}\langle x,y\rangle)) = \mathrm{Im}\langle x,y\rangle$$

so that $\mathrm{Im}\langle x,y\rangle = (\|x+iy\|^2 - \|x-iy\|^2)/4$. So again we can express $\langle x,y\rangle = \mathrm{Re}\langle x,y\rangle + i\,\mathrm{Im}\langle x,y\rangle$ in terms of the norm (and vector space operations).

$\Leftarrow$: The idea is that the inner product must be related to the norm by

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

in the case of real scalars $\mathbb{K} = \mathbb{R}$, or in the case $\mathbb{K} = \mathbb{C}$ by

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

What remains to be done is to check that (assuming that the parallelogram identity is true for the norm) these formulae do define inner products on E.

For example, in the case $\mathbb{K} = \mathbb{R}$ we see easily that $\langle x, y\rangle = \langle y, x\rangle$. Then we have

$$\begin{aligned}
4\langle x, y+z\rangle &= \|x+y+z\|^2 - \|x-y-z\|^2 \\
&= 2(\|x+y\|^2 + \|z\|^2) - \|x+y-z\|^2 \\
&\quad - \big(2(\|x-y\|^2 + \|z\|^2) - \|x-y+z\|^2\big) \\
&= 2(\|x+y\|^2 - \|x-y\|^2) + \|x-y+z\|^2 - \|x+y-z\|^2 \\
&= 8\langle x, y\rangle + 2(\|x+z\|^2 + \|y\|^2) - \|x+z+y\|^2 \\
&\quad -2(\|x-z\|^2 + \|y\|^2) + \|x-z-y\|^2 \\
&= 8\langle x, y\rangle + 2(\|x+z\|^2 - \|x-z\|^2) - \|x+y+z\|^2 + \|x-y-z\|^2 \\
&= 8\langle x, y\rangle + 8\langle x, z\rangle - 4\langle x, y+z\rangle
\end{aligned}$$

It follows that $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$.

By symmetry of the inner product, $\langle y+z, x\rangle = \langle y, x\rangle + \langle z, x\rangle$.

It follows that $\langle nx, y\rangle = n\langle x, y\rangle = \langle x, ny\rangle$. (By induction on n it follows easily for $n \in \mathbb{N}$ and it also follows for $n = 0$ and $n \in \mathbb{Z}$ by simple algebraic manipulations. For $n \neq 0$ we deduce $n\langle \frac{1}{n}x, y\rangle = \langle x, y\rangle = n\langle x, \frac{1}{n}y\rangle$ and so $\langle \frac{1}{n}x, y\rangle = \frac{1}{n}\langle x, y\rangle = \langle x, \frac{1}{n}y\rangle$. It follows that for $r = p/q \in \mathbb{Q}$ rational we have $r\langle x, y\rangle = \langle rx, y\rangle = \langle x, ry\rangle$. By continuity of the inner product it follows that $\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle x, \lambda y\rangle$ for all $\lambda \in \mathbb{R}$.

The case $\mathbb{K} = \mathbb{C}$ is similar but longer. $\square$

1.4 Spaces of operators or functionals

1.4.1 Remark. A way to construct new (and interesting) Banach spaces is to consider spaces of linear operators with the operator norm.

We will give the general form and then concentrate on the Hilbert space context in this chapter.

1.4.2 Proposition. *Let E and F be normed spaces (over $\mathbb{K}$) and let $\mathcal{B}(E, F)$ denote the set of all bounded linear operators from E to F. Then we can make $\mathcal{B}(E, F)$ into a normed space using vector space operations defined by*

- *$T+S\colon E \to F$ is defined by $(T+S)(x) = T(x)+S(x)$ (for $T, S \in \mathcal{B}(E, F)$ and $x \in E$);*
- *$\lambda T\colon E \to F$ is defined by $(\lambda T)(x) = \lambda(Tx)$ (for $T \in \mathcal{B}(E, F)$, $\lambda \in \mathbb{K}$ and $x \in E$),*

and the operator norm.

So $(\mathcal{B}(E, F), \|\cdot\|_{\mathrm{op}})$ becomes a normed space over $\mathbb{K}$.

Proof. We must first show that we have a vector space, which involves checking many properties that are quite straightforward.

We do need to know that $T+S, \lambda T \in \mathcal{B}(E, F)$ if $T, S \in \mathcal{B}(E, F)$ and $\lambda \in \mathbb{K}$. That means that addition and multiplication by scalars are properly defined as operations on $\mathcal{B}(E, F)$. To check this notice that $T + S$ is linear (not hard to verify) and that

$$\|(T+S)(x)\| = \|T(x)+S(x)\| \leq \|T\|_{\mathrm{op}}\|x\|+\|S\|_{\mathrm{op}}\|x\| = (\|T\|_{\mathrm{op}}+\|S\|_{\mathrm{op}})\|x\|.$$

This shows that $T + S$ is bounded (and in fact that $\|T + S\|_{\mathrm{op}} \leq \|T\|_{\mathrm{op}} + \|S\|_{\mathrm{op}}$ which we need to show that $\|\cdot\|_{\mathrm{op}}$ satisfies the triangle inequality).

Similarly check that λT is linear and note that

$$\|(\lambda T)(x)\| = \|\lambda Tx\| = |\lambda|\|Tx\| \leq |\lambda|\|T\|_{\mathrm{op}}\|x\|$$

and this shows that λT is bounded. (It also shows that $\|\lambda T\|_{\mathrm{op}} \leq |\lambda|\|T\|_{\mathrm{op}}$. To show that $\|\cdot\|_{\mathrm{op}}$ is a norm we need that $\|\lambda T\|_{\mathrm{op}} = |\lambda|\|T\|_{\mathrm{op}}$. For $\lambda = 0$ this follows from what we have. For $\lambda \neq 0$, apply what we know with T replaced by λT and λ by $1/\lambda$ to get

$$\|T\|_{\mathrm{op}} = \left\|\frac{1}{\lambda}\lambda T\right\|_{\mathrm{op}} \leq \frac{1}{|\lambda|}\|\lambda T\|_{\mathrm{op}} \Rightarrow |\lambda|\|T\|_{\mathrm{op}} \leq \|\lambda T\|_{\mathrm{op}}.$$

Then we conclude $\|\lambda T\|_{\mathrm{op}} = |\lambda|\|T\|_{\mathrm{op}}$.)

The rest of the proof involves showing that there is a zero element for the vector space $\mathcal{B}(E, F)$ [which is the zero linear transformation $Tx = 0$ for all $x \in E$] and that the various rules about associativity, additive inverses and the distributive laws hold. All are quite easy to verify.

We have show above that $\|\cdot\|_{\mathrm{op}}$ has the right properties to be a norm on $\mathcal{B}(E, F)$. □

1.4.3 Theorem. *If E is a normed space and F is a Banach space, then*

$$(\mathcal{B}(E,F), \|\cdot\|_{\mathrm{op}})$$

is a Banach space.

Proof. Let $(T_n)_{n=1}^\infty$ be a Cauchy sequence in $(\mathcal{B}(E,F), \|\cdot\|_{\mathrm{op}})$. Recall that means that given $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ so that

$$n, m \geq N_\varepsilon \Rightarrow \|T_n - T_m\|_{\mathrm{op}} < \varepsilon.$$

For any fixed $x \in E$ $\{T_n(x)\}_{n=1}^\infty$ is a Cauchy sequence in F because

$$\|T_n(x) - T_m(x)\|_F \leq \|T_n - T_m\|_{\mathrm{op}} \|x\|_E$$

is small if n and m are both large. (To be more precise, given $\varepsilon > 0$ take $N = N_{\varepsilon/(\|x\|+1)}$ and then we have

$$\begin{aligned} n, m \geq N \quad &\Rightarrow \quad \|T_n - T_m\|_{\mathrm{op}} < \frac{\varepsilon}{\|x\| + 1} \\ &\Rightarrow \quad \|T_n(x) - T_m(x)\|_F \leq \|T_n - T_m\|_{\mathrm{op}} \|x\| \\ &\qquad \leq \frac{\varepsilon}{\|x\| + 1} \|x\| < \varepsilon.) \end{aligned}$$

Since F is complete, it follows that

$$\lim_{n\to\infty} T_n(x)$$

exists in F (for each $x \in E$). This allows us to define a map $T \colon E \to F$ by

$$T(x) = \lim_{n\to\infty} T_n(x).$$

We will be finished if we show that $T \in \mathcal{B}(E,F)$ and $\lim_{n\to\infty} \|T_n - T\|_{\mathrm{op}} = 0$ (so that $T_n \to T$ in the norm of $\mathcal{B}(E,F)$).

By the Cauchy condition (with $\varepsilon = 1$), we know that we can find N so that $\|T_n - T_m\|_{\mathrm{op}} < 1$ for all $n, m \geq N$. Now take $x \in E$, $\|x\|_E \leq 1$. Then for $n, m \geq N$ we have

$$\|T_n(x) - T_m(x)\|_F = \|(T_n - T_m)(x)\|_F \leq \|T_n - T_m\|_{\mathrm{op}} < 1.$$

Fix $n = N$ and let $m \to \infty$ and use continuity of the norm on F to conclude

$$\|T_N(x) - T(x)\|_F \leq 1.$$

This is true for all $x \in E$ of norm $\|x\|_E \leq 1$ and we can use that to conclude that

$$\begin{aligned} \sup_{\substack{x \in E \\ \|x\| \leq 1}} \|T(x)\|_F &\leq \sup_{\substack{x \in E \\ \|x\| \leq 1}} \|T_N(x)\|_F + \sup_{\substack{x \in E \\ \|x\| \leq 1}} \|T(x) - T_N(x)\|_F \\ &\leq \|T_N\|_{\text{op}} + 1. \end{aligned}$$

We see now that T is bounded. So $T \in \mathcal{B}(E, F)$.

We now repeat the last few steps with $\varepsilon/2$ ($\varepsilon > 0$ arbitrary) where we had 1 before. By the Cauchy condition we can find N so that $\|T_n - T_m\|_{\text{op}} < \varepsilon/2$ for all $n, m \geq N$. Now take $x \in E$, $\|x\|_E \leq 1$. As before we get

$$\|T_n(x) - T_m(x)\|_F = \|(T_n - T_m)(x)\|_F \leq \|T_n - T_m\|_{\text{op}} < \varepsilon/2$$

as long as $n, m \geq N$, $\|x\| \leq 1$. Fix any $n \geq N$ for the moment and let $m \to \infty$ to get

$$\|T_n(x) - T(x)\|_F \leq \frac{\varepsilon}{2}$$

We have this for all x of norm at most 1 and all $n \geq N$. So

$$\|T_n - T\| = \sup_{x \in E, \|x\| \leq 1} \|T(x) - T_n(x)\|_F \leq \frac{\varepsilon}{2} < \varepsilon$$

as long as $n \geq N$. This shows that $T_n \to T$ in $\mathcal{B}(E, F)$.

Thus every Cauchy sequence in the space converges in the space and so $\mathcal{B}(E, F)$ is a Banach space. $\square$

1.4.4 Definition. If E is a normed space, the *dual space* of E is

$$E^* = \mathcal{B}(E, \mathbb{K}) = \{T \colon E \to \mathbb{K} : T \text{ continuous and linear}\}.$$

(We take the usual absolute value norm on $\mathbb{K}$ and the operator norm $\|\cdot_{\text{op}}$ on E^*.)

Elements of E^* are called (continuous) *linear functionals* on E.

The notation E' is sometimes used for E^*.

1.4.5 Corollary. *If E is a normed space, then E^* is a Banach space.*

Proof. Apply Theorem 1.4.3 with $F = \mathbb{K}$ (usual absolute value norm). $\square$

1.4.6 Remark. From the theorem above, E^* is always a Banach space in the operator norm. It will (usually) be a new space, not the same as E. So also is the space $(\mathcal{B}(E, F), \|\cdot\|_{\text{op}})$ a 'new' space we can construct from normed spaces E and F.

1.5 Dual of a Hilbert space

The dual space of a Hilbert space can be identified. First we look at the finite dimensional case (because it is simple, but also because the ideas for ℓ^2 will be quite similar, just a little more elaborate).

1.5.1 Proposition. *Let H be a finite dimensional Hilbert space and $\alpha \in H^*$. Then there exists $y \in H$ such that*

$$\alpha(x) = \langle x, y \rangle \qquad \textit{for all } x \in H.$$

Conversely, given $y \in H$, $\alpha_y(x) = \langle x, y \rangle$ defines an element $\alpha_y \in H^*$ and moreover $\|\alpha_y\| = \|y\|$.

Proof. We know that H has an orthonormal basis. Let us consider one such basis $\phi_1, \phi_1, \ldots, \phi_n$ (where $n = \dim H$).

If $\alpha \colon H \to \mathbb{K}$ is a linear transformation (automatically continuous since $\dim H < \infty$ now), then put $y_j = \overline{\alpha(\phi_j)}$ for $1 \leq j \leq n$ and $y = \sum_{j=1}^n y_j \phi_j$. We can check easily that

$$\alpha(\phi_k) = \langle \phi_k, y \rangle = \sum_{j=1}^n \bar{y}_j \langle \phi_k, \phi_j \rangle = \bar{y}_k = \alpha(\phi_k)$$

and then (using linearity that) for an $x = \sum_{j=1}^n x_j \phi_j \in H$ we have $\alpha(x) = \langle x, y \rangle$.

For the converse, starting with $y \in H$ define $\alpha_y \colon H \to \mathbb{K}$ by $\alpha_y(x) = \langle x, y \rangle$ and it is easy to see that α_y is linear (and bounded too because that is automatic here).

From Cauchy-Schwarz

$$|\alpha_y(x)| = |\langle x, y \rangle| \leq \|x\| \, \|y\|$$

and so $\|\alpha_y\| \leq \|y\|$. So if $y = 0$ then certainly $\|\alpha_y\| = 0$ and if $y \neq 0$ take $x = y/\|y\|$ to get $\alpha(x) = \langle y, y \rangle / \|y\| = \|y\|^2/\|y\| = \|y\|$, which shows $\|\alpha_y\| \geq \|y\|$. So $\|\alpha_y\| = \|y\|$. $\square$

1.5.2 Proposition. *Let $H = \ell^2$ and $\alpha \in H^*$. Then there exists $y \in H$ such that*

$$\alpha(x) = \langle x, y \rangle \qquad \textit{for all } x \in H.$$

Conversely, given $y \in H$, $\alpha_y(x) = \langle x, y \rangle$ defines an element $\alpha_y \in H^*$ and moreover $\|\alpha_y\| = \|y\|$.

Proof. Consider the standard orthonormal basis $e_1, e_2, \ldots$ for ℓ^2.

If $\alpha\colon H \to \mathbb{K}$ is a bounded linear transformation, then put $y_j = \overline{\alpha(e_j)}$ for $j \in \mathbb{N}$.

Modelling on the finite case we would like to put $y = \sum_{j=1}^{\infty} y_j e_j = (y_1, y_2, \ldots)$ but we need to know this is in ℓ^2.

To do this, fix N finite and define $x_j = \bar{y}_j$ for $1 \leq j \leq N$ and $x_j = 0$ for $j > N$. Then $x = \sum_{j=1}^{N} x_j e_j = (x_1, x_2, \ldots, x_N, 0, 0, \ldots) \in \ell^2$ and

$$\alpha(x) = \sum_{j=1}^{N} x_j \alpha(e_j) = \sum_{j=1}^{N} \bar{y}_j y_j = \sum_{j=1}^{N} |y_j|^2$$

Also, since α is bounded

$$|\alpha(x)\| \leq \|\alpha\| \|x\|_2 = \|\alpha\| \sqrt{\sum_{j=1}^{N} |x_j|^2} = \|\alpha\| \sqrt{\sum_{j=1}^{N} |y_j|^2}$$

So $\sum_{j=1}^{N} |y_j|^2 \leq \|\alpha\| \sqrt{\sum_{j=1}^{N} |y_j|^2}$ and it follows that $\sum_{j=1}^{N} |y_j|^2 \leq \|\alpha\|^2$. This it is actually justified to define $y \in \ell^2$ by $y = \sum_{j=1}^{\infty} y_j e_j = (y_1, y_2, \ldots)$.

We can check easily that

$$\alpha(e_k) = \langle e_k, y\rangle = \sum_{j=1}^{n} \bar{y}_j \langle e_k, e_j\rangle = \bar{y}_k = \alpha(e_k).$$

Then for a finite linear combination $x = \sum_{j=1}^{n} x_j e_j \in H$ we have $\alpha(x) = \langle x, y\rangle$. Since these finite linear combinations are dense in ℓ^2 and both $x \mapsto \alpha(x)$ and $x \mapsto \langle x, y\rangle$ are continuous, it follows that $\alpha(x) = \langle x, y\rangle$ for all $x \in \ell^2 = H$.

For the converse, starting with $y \in H$ define $\alpha_y\colon H \to \mathbb{K}$ by $\alpha_y(x) = \langle x, y\rangle$ and it is easy to see that α_y is linear and bounded too by Cauchy-Schwarz:

$$|\alpha_y(x)| = |\langle x, y\rangle| \leq \|x\| \, \|y\|.$$

So we have $\|\alpha_y\| \leq \|y\|$. If $y = 0$ then certainly $\|\alpha_y\| = 0$ and if $y \neq 0$ take $x = y/\|y\|$ to get $\alpha(x) = \langle y, y\rangle/\|y\| = \|y\|^2/\|y\| = \|y\|$, which shows $\|\alpha_y\| \geq \|y\|$. So $\|\alpha_y\| = \|y\|$. □

1.5.3 Theorem (Riesz representation theorem)**.** *Let H be a (separable) Hilbert space and $\alpha \in H^*$. Then there exists $y \in H$ such that*

$$\alpha(x) = \langle x, y\rangle \qquad \text{for all } x \in H.$$

Conversely, given $y \in H$, $\alpha_y(x) = \langle x, y\rangle$ defines an element of H^* and moreover $\|\alpha_y\| = \|y\|$.

Proof. We proved this for $\dim H < \infty$ in Proposition 1.5.1 and it remains to establish it for the separable infinite dimensional case. But in that case Theorem 1.2.14 allows us to reduce to the case of ℓ^2 (which we have just established in Proposition 1.5.2).

Indeed if $T\colon H \to \ell^2$ is a linear isometric isomorphism, then $\alpha \in H^*$ implies $\alpha \circ T^{-1} \in (\ell^2)^*$ (easy to check) and so by Proposition 1.5.2) there is $z \in \ell^2$ so that $(\alpha \circ T^{-1})(x) = \langle x, z\rangle$ for $x \in \ell^2$. But then for $\phi \in H$, take $x = T(\phi)$ to get

$$\alpha(\phi) = (\alpha \circ T^{-1})(x) = \langle x, z\rangle = \langle Tx, Tz\rangle_H = \langle \phi, Tz\rangle_H$$

With $y = Tz$ we get α of the right form.

The rest of the proof is no different from the earlier cases. $\square$

1.5.4 Remark. In fact the Riesz representation theorem holds for all Hilbert spaces, separable or not, but we would need the machinery of uncountable orthonormal bases to prove the general statement.

It cannot hold for incomplete inner product spaces (because the dual space is always complete).

It is usual to state this theorem as $H^* = H$ for H Hilbert, but that is not quite accurate. We need to fix a basis to make that work out right in the case $\mathbb{K} = \mathbb{C}$ (or at least have a way to do some sort of complex conjugation of vectors in H).

Here is a more precise statement.

1.5.5 Corollary. *If H is a Hilbert space (over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) then there is an $\mathbb{R}$-linear isometric identification $T\colon H \to H^*$ given by*

$$T(y)(x) = \langle x, y\rangle.$$

In the case $\mathbb{K} = \mathbb{C}$, we also have that T is conjugate-linear, that is $T(\lambda y) = \bar{\lambda} T(y)$.

Proof. By the Riesz representation theorem, T is a bijection. (In the proofs we used α_y rather than $T(y)$.)

It is easy to see that T is $\mathbb{R}$-linear and by the Riesz representation theorem, we also know that $\|T(y)\| = \|y\|$. The fact that T is conjugate linear is also easy to check. $\square$

1.6 Algebra of operators

1.6.1 Notation. A commonly studied object related to a Hilbert space H is the space of all bounded operators $T\colon H \to H$. We denote this space by $\mathcal{B}(H)$. (The notation $\mathcal{L}(H)$ is also used frequently).

This is the same as the space $\mathcal{B}(H, H)$ in the notation of Theorem 1.4.3.

To avoid annoyingly trivial cases we will assume that $H \neq \{0\}$.

By Theorem 1.4.3, for H a Hilbert space $\mathcal{B}(H)$ is a Banach space (in the norm $\|\cdot\|_{\mathrm{op}}$).

1.6.2 Example. If $y, z \in H$ we can define $T\colon H \to H$ by $T(x) = \langle x, y\rangle z$. To show it is bounded, note that

$$\|T(x)\| = |\langle x, y\rangle| \|z\| \leq \|x\|\|y\|\|z\|$$

(by Cauchy-Schwarz) so that $\|T\| \leq \|y\|\|z\|$.

Since the range of T is $\mathbb{C}z$ (at least if $y \neq 0$), a one dimensional space (if $z \neq 0$), T is called a rank one operator (unless $T = 0$).

We could at least make more examples of operators by adding up finitely make rank one operators (which would then have finite dimensional range). A more complicated process would be to take an infinite sum

$$T(x) = \sum_{n=1}^{\infty} \langle x, y_n\rangle z_n$$

where we choose the $y_n, z_n \in H$ so that $\sum_{n=1}^{\infty} \|x_n\|\|y_n\| < \infty$.

On $H = L^2[0,1]$ we can instead consider multiplication operators $M_g(f) = gf$ where $g \in L^\infty[0,1]$.

Simpler examples are multiplications on $H = \ell^2$ by $g \in \ell^\infty$. There

$$M_g(x) = M_g(x_1, x_2, \ldots) = (g_1 x_1, g_2 x_2, \ldots).$$

In both cases it is in fact the case that $\|M_g\| = \|g\|_\infty$.

A finite dimensional example of this type would be $H = \mathbb{K}^n$ and

$$M_g(x) = M_g(x_1, x_2, \ldots, x_n) = (g_1 x_1, g_2 x_2, \ldots, g_n x_n).$$

If we write our vectors are column matrices we have

$$M_g \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} g_1 & 0 & \cdots & 0 \\ 0 & g_2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & g_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

So multiplication operators M_g correspond to diagonal matrices in this case, and of course $\mathcal{B}(\mathbb{K}^n)$ can be viewed as the $n \times n$ matrices with entries over $\mathbb{K}$. The diagonal (multiplication) operators (in the standard basis of $\mathbb{K}^n$) are then particularly nice operators.

One can represent bounded operators T on ℓ^2 as infinite matrices, by writing $T(e_j)$ down column j. For H infinite dimensional separable (or finite dimensional) we can fix an orthonormal basis for H and write a matrix for T. But this can make T complicated unless the orthonormal basis is well chosen for the particular T.

$\mathcal{B}(H)$ also has an algebra structure, where we define multiplication of two operators via composition. We recall the definition.

1.6.3 Definition. An *algebra* A over a field $\mathbb{K}$ is a vector space over $\mathbb{K}$ that comes with a multiplication (product operation) $A \times A \to A$ ($(a,b) \mapsto ab$) with the properties

i) $(a_1+\lambda a_2)b = a_1 b+\lambda a_2 b$ and $a(b_1+\lambda b_2) = ab_1+\lambda ab_2$ (whenever $a, b, a_1, b_1, a_1, b_2 \in A$, $\lambda \in \mathbb{K}$)

   That is bilinearity of the product.

ii) $(ab)c = a(bc)$ for $a, b, c \in A$ (associativity)

If there is an element $1_A \in A$ such that $1_A a = a = a 1_A$ holds for all $a \in A$, then we say A is a *unital algebra*.

1.6.4 Lemma. *If H is a Hilbert space then $\mathcal{B}(H)$ with the product of $S, T \in \mathcal{B}(H)$, $ST\colon H \to H$, defined by $(ST)(x) = S(T(x))$ for $x \in H$ is an algebra.*

Proof. If $S, T \in \mathcal{B}(H)$, then $ST\colon H \to H$ is defined by $(ST)(x) = S(T(x))$ for $x \in H$. ST is continuous as it is the composition of two continuous maps. We can easily check the algebra properties: associativity of the product $S(TU) = (ST)U$, $\lambda(ST) = (\lambda S)T = S(\lambda T)$ and the distributive laws. As we know from finite dimensions (where composition of linear transformations on $\mathbb{K}^n$ corresponds to matrix multiplication of $n \times n$ matrices) the algebra $\mathcal{B}(H)$ is not usually commutative. The identity operator on H is a multiplicative identity for this algebra. $\square$

In fact $\mathcal{B}(H)$ is a Banach space as well as an algebra and these are called Banach algebras when the norm interacts nicely with the algebra product.

1.6.5 Definition. A *normed algebra* A is an algebra over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ which has a norm that satisfies $\|ab\| \leq \|a\|\|b\|$ for all $a, b \in A$.

We say a normed algebra is a *Banach algebra* if it is also complete.

A *unital normed algebra* is a normed algebra that has a unit (identity element) 1_A such that $\|1_A\| = 1$ and a *unital Banach algebra* is a Banach algebra that is also a unital normed algebra.

1.6.6 Lemma. *If H is a nonzero Hilbert space, then $\mathcal{B}(H)$ with the product given by composition and with the operator norm, is a unital Banach algebra.*

Proof. We can estimate the norm of the product

$$\|ST\|_{\text{op}} = \sup_{x \in H, \|x\|_H \leq 1} \|S(T(x))\|_H \leq \sup_{y \in H, \|y\|_H \leq \|T\|} \|Sy\|_H \leq \|S\|_{\text{op}} \|T\|_{\text{op}}.$$

This establishes the inequality $\|ST\| \leq \|S\| \, \|T\|$. The identity operator $\mathrm{id}_H \colon H \to H$, given by $\mathrm{id}_H(x) = x$ for $x \in H$, is a unit of norm 1 for $\mathcal{B}(H)$. $\square$

There is one further piece of structure on $\mathcal{B}(H)$. Every $T \in \mathcal{B}(H)$ has an adjoint operator $T^* \in \mathcal{B}(H)$ which is uniquely determined by the property

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for $x, y \in H$. To prove that such a T^* exists and to prove that $\|T^*\|_{\text{op}} = \|T\|_{\text{op}}$ we use the Riesz representation theorem and the following lemma.

1.6.7 Lemma. *If H is a non-zero inner product space and $x \in H$, then*

$$\|x\| = \sup_{y \in H, \|y\| = 1} |\langle x, y \rangle|$$

Proof. Fix $x \in H$. By Cauchy-Schwarz, $|\langle x, y \rangle| \leq \|x\|\|y\| = \|x\|$ for each $yinH$ with $\|y\| = 1$. So $\sup_{y \in H, \|y\| = 1} |\langle x, y \rangle| \leq \|x\|$. If $x = 0$ we must have equality and if $x \neq 0$ take $y = x/\|x\|$ to get $\|y\| = 1$ and $\langle x, y \rangle = \|x\|$ so that the supremum is also $\geq \|x\|$. $\square$

1.6.8 Proposition. *If H is a nonzero Hilbert space and $T \in \mathcal{B}(H)$, then*

$$\|T\| = \sup_{x, y \in H, \|x\| = \|y\| = 1} |\langle T(x), y \rangle|$$

Proof. We know $\|T\| = \sup_{x \in H, \|x\|=1} \|T(x)\|$ (using $H \neq \{0\}$). Using Lemma 1.6.7,

$$\begin{aligned}\sup_{x \in H, \|x\|=1} \|T(x)\| &= \sup_{x \in H, \|x\|=1} \sup_{y \in H, \|y\|=1} |\langle T(x), y\rangle| \\ &= \sup_{x,y \in H, \|x\|=\|y\|=1} |\langle T(x), y\rangle|\end{aligned}$$

□

1.6.9 Corollary. *If H is a nonzero Hilbert space and $T \colon H \to H$ is a linear transformation with*

$$\sup_{x,y \in H, \|x\|=\|y\|=1} |\langle T(x), y\rangle| < \infty$$

then $T \in \mathcal{B}(H)$.

Proof. This is a corollary of the previous proof, rather that a corollary of the statement (where we assumed $T \in \mathcal{B}(H)$). □

1.6.10 Proposition. *If H is a Hilbert space and $T \in \mathcal{B}(H)$, then there exists $T^* \in \mathcal{B}(H)$ which is uniquely determined by the property*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for $x, y \in H$. Moreover $\|T^*\|_{\mathrm{op}} = \|T\|_{\mathrm{op}}$.

Proof. We can suppose H is nonzero as if $H = \{0\}$ then $T = 0$ and T^* has to be 0 also.

To define T^*y, fix $y \in H$ and consider the map $\alpha \colon H \to \mathbb{K}$ given by $\alpha(x) = \langle Tx, y\rangle$. This is a bounded (same as continuous) linear map and so (by Corollary 1.5.5) there is some unique $w \in H$ with $\alpha(x) = \langle x, w\rangle$. Define $T^*(y) = w$ (this is the only possible choice), and then we have

$$\langle Tx, y\rangle = \alpha(x) = \langle x, w\rangle = \langle x, T^*(y)\rangle$$

(for all $x \in H$). Since $y \in H$ was arbitrary, we also have this for all $y \in H$. Now it is quite easy to check that T^* is a linear transformation from H to H.

By Proposition 1.6.8

$$\begin{aligned}\|T\| &= \sup_{x,y \in H, \|x\|=\|y\|=1} |\langle T(x), y\rangle| \\ &= \sup_{x,y \in H, \|x\|=\|y\|=1} |\langle x, T^*(y)\rangle| \\ &= \sup_{x,y \in H, \|x\|=\|y\|=1} |\langle T^*(y), x\rangle| \\ &= \|T^*\|\end{aligned}$$

(using Corollary 1.6.9 and Proposition 1.6.8). □

1.6.11 Remark. This adjoint operation makes $\mathcal{B}(H)$ a *Banach *-algebra*: we have $(T^*)^* = T$, $(\lambda T)^* = \bar{\lambda} T^*$, $(ST)^* = T^* S^*$ and $(S+T)^* = S^* + T^*$ for $S, T \in \mathcal{B}(H)$ and $\lambda \in \mathbb{K}$. (The case $\mathbb{K} = \mathbb{C}$ is the one usually studied.)

Moreover the important property

1.6.12 Proposition (C^* identity)**.** *For H a Hilbert space and $T \in \mathcal{B}(H)$, we have $\|T\|^2 = \|T^*T\|$.*

Proof.

$$\begin{aligned}
\|T\|^2 &= \sup_{x \in H, \|x\| \leq 1} \|Tx\|^2 \\
&= \sup_{x \in H, \|x\| \leq 1} \langle Tx, Tx \rangle \\
&= \sup_{x \in H, \|x\| \leq 1} \langle x, T^*Tx \rangle \\
&\leq \sup_{x \in H, \|x\| \leq 1} \|x\| \|T^*Tx\| \\
&= \|TT^*\| \leq \|T\| \|T^*\| = \|T\|^2
\end{aligned}$$

□

1.6.13 Remark. Closed *-subalgebras of $\mathcal{B}(H)$ are known as *C^*-algebras*. (A *-subalgebra is a subalgebra that contains T^* whenever it contains T. By closed we mean closed with respect to the norm topology, or contains limits of convergent sequences with all terms in the subalgebra. All C^*-algebras are then Banach *-algebras, but they also satisfy the property $\|T\|^2 = \|T^*T\|$.)

C^*-algebras have been studied intensively for almost a century, partly as a mathematical basis for quantum theory.

A substantial theorem called the Gelfand-Naimark theorem characterises C^*-algebras abstractly (up to isometric isomorphism of Banach *-algebras) as those Banach algebras with a $*$ operation which satisfy the C^*-identity.

A Gram-Schmidt

A.0.1 Theorem (Gram-Schmidt Orthonormalisation)**.** *Suppose that either*

$$\psi_1, \psi_2, \ldots, \psi_N$$

is a finite linearly independent sequence or that $\psi_1, \psi_2, \ldots$ is an infinite linearly independent sequence in an inner product space H. (By this we mean linearly independent in the usual algebraic sense, or the infinite sequence case that $\psi_1, \psi_1, \ldots, \psi_n$ is linearly independent for each n.)

Then the Gram-Schmidt process,

$$\begin{aligned}
\phi_1 &= \frac{\psi_1}{\|\psi_1\|} \\
\phi_2 &= \frac{\psi_2 - \langle \psi_2, \phi_1 \rangle \phi_1}{\|\psi_2 - \langle \psi_2, \phi_1 \rangle \phi_1\|} \\
\phi_3 &= \frac{\psi_3 - \langle \psi_3, \phi_2 \rangle \phi_2 - \langle \psi_3, \phi_1 \rangle \phi_1}{\|\psi_3 - \langle \psi_3, \phi_2 \rangle \phi_2 - \langle \psi_3, \phi_1 \rangle \phi_1\|} \\
\phi_n &= \frac{\psi_n - \sum_{j=1}^{n-1} \langle \psi_n, \phi_j \rangle \phi_j}{\left\|\psi_n - \sum_{j=1}^{n-1} \langle \psi_n, \phi_j \rangle \phi_j\right\|}
\end{aligned}$$

produces an orthonormal sequence $\phi_1, \phi_2, \ldots$ in H such that

$$\operatorname{span}\{\psi_1, \psi_2, \ldots, \psi_n\} = \operatorname{span}\{\phi_1, \phi_2, \ldots, \phi_n\}$$

(for $1 \leq n \leq N$ in the finite case, all $n \geq 1$ in the infinite case).

Proof. From the finite dimensional version of Gram-Schmidt, we know that for each n, the finite linear combinations $b_1\phi_1 + b_2\phi_2 + \cdots + b_n\phi_n$ are the same as the linear combinations $a_1\psi_1 + a_2\psi_2 + \cdots + a_n\psi_n$. Also $\phi_1, \phi_2, \ldots, \phi_n$ are orthonormal for each n. □

Amendments: Jan 23, 2017: Swop notations ϕ and ψ in Theorem A.0.1 (so that the vectors $\phi_1, \phi_2, \ldots$ are the orthonormal ones). Fix ϕ_{n+1} to ψ_{n+1} in proof of Theorem 1.2.11.

Jan 30, 2017: Add some extra clarification to Theorem 1.3.1.

Feb 5, 2017: Fix Remark 1.6.11 so that the field can be $\mathbb{R}$ or $\mathbb{C}$.

Richard M. Timoney (February 5, 2017)