321 2008–09

Chapter 4: Hilbert Spaces

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

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

(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.

4.2 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).

4.3 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.)

4.4 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.

4.5 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)

$$\sum_{\phi \in S} |\langle x, \phi\rangle|^2 \leq \|x\|^2 \qquad \text{(Bessel's inequality)}. \tag{4.1}$$

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

$$\sum_n |\langle x, \phi_n\rangle|^2 \leq \|x\|^2. \tag{4.2}$$

The sum is independent of the order in which the $\phi_1, \phi_2, \ldots$ are listed.

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

$$\sum_{k=1}^{n} |a_k|^2 = \sum_{k=1}^{n} |\langle x, \phi_k \rangle|^2 \leq \langle x, x \rangle = \|x\|^2. \tag{4.3}$$

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 (4.3)

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

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 (4.3) to get Bessel's inequality in the form (4.2). $\square$

4.6 Corollary. *Let H be a Hilbert space (completeness is essential now). Let S be any maximal orthonormal subset of H (such subsets always exist by Zorn's lemma). Let $x \in H$. Then*

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

This sum is to be interpreted as follows. We can list all the $\phi \in S$ for which $\langle x, \phi \rangle \neq 0$ as $\phi_1, \phi_2, \ldots$. The list could be finite or infinite. If it is finite then there is no real problem interpreting the sum. If it is infinite, we mean

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

as a limit in the norm of H. Moreover the order of the list $\phi_1, \phi_2, \ldots$ is immaterial.

Maximal orthonormal subsets of a Hilbert space are called *orthonormal bases* because of this result. They are also sometimes known as *complete orthonormal systems*.

Note the difference between this kind of orthonormal basis and the finite kind encountered in finite dimensional inner product spaces, where no infinite summations are required.

The simplest example of this kind of orthonormal basis, apart from the finite dimensional ones, is the standard basis of ℓ^2. We'll spell that out now, but the verification of the example is quite straightforward. It does not use Zorn's lemma.

4.7 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

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

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

Let

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

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 $x = (x_j)_{j=1}^\infty = (x_1, x_2, \ldots) \in \ell^2$, 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

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

and

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

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 $x = \sum_{j=1}^\infty \langle x, e_j \rangle e_j$.

Proof (of Corollary 4.6). The fact that such maximal orthonormal subsets exist is easy to verify with Zorn's lemma. (Consider the collection of all orthonormal subsets of H ordered by inclusion. The empty set is one option and so the collection is nonempty. For any linearly ordered sub-collection, the union is orthonormal and is an upper bound for the sub-collection.)

Fix one such S and let $S_x = \{\phi \in S : \langle x, \phi \rangle \neq 0\} = \{\phi_1, \phi_2, \ldots\}$. The case where S_x is finite is simpler than the infinite case. In the infinite case consider the series

$$\sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$$

in H. We show it is Cauchy. By a calculation with inner products we can show

$$\left\| \sum_{j=n}^{m} \langle x, \phi_j \rangle \phi_j \right\|^2 = \sum_{j=n}^{m} |\langle x, \phi_j \rangle|^2.$$

By Bessel's inequality (4.1) the series of positive scalars $\sum_{n=1}^{\infty} |\langle x, \phi_n \rangle|^2$ converges, hence it is Cauchy and so our series in H also satisfies the Cauchy criterion.

Since H is complete, $y = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$ exists in H. By continuity of the inner product, we find

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

Hence

$$\langle x - y, \phi_k \rangle = 0$$

for each k. Also, repeating the above argument with ϕ_k replaced by one of the $\phi \in S$ with $\langle x, \phi \rangle = 0$ shows that $\langle x - y, \phi \rangle = 0$ for all $\phi \in S$.

If $x \neq y$, we could get a strictly larger orthonormal subset of H than S by taking $S \cup \left\{ \frac{x-y}{\|x-y\|} \right\}$. That would contradict the maximality of S. Therefore $x = y = \sum_{n=1}^{\infty} \langle x, \phi_n \rangle \phi_n$. □

4.8 Corollary. *If H is a Hilbert space and $S \subset H$ is an orthonormal basis for H, then for each $x, y \in H$ we have*

$$\langle x, y \rangle = \sum_{\phi \in S} \langle x, \phi \rangle \overline{\langle y, \phi \rangle}$$

and

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

Proof. From Proposition 4.5 (i) we see that $S_{x,y} = \{\phi \in S : \langle x, \phi\rangle \neq 0$ or $\langle y, \phi\rangle \neq 0\} = S_x \cup S_y$ is countable. If we list it as $S_{x,y} = \{\phi_1, \phi_2, \ldots\}$ then 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. □

4.9 Theorem. *A Hilbert space H is separable (that is, has a countable dense subset) if and only if it has one countable orthonormal basis if and only if every orthonormal basis for H is countable.*

Proof. $\Rightarrow$: Suppose H is separable and consider any orthonormal basis $S \subset H$. Then if $\phi_1, \phi_2 \in S$ and $\phi_1 \neq \phi_2$ we conclude $\|\phi_1 - \phi_2\| = \sqrt{\langle \phi_1 - \phi_2, \phi_1 - \phi_2\rangle} = \sqrt{2}$. Thus $S \subset H$ has the discrete topology.

Now H a separable metric space $\Rightarrow$ S is a separable metric space $\Rightarrow$ S has a countable dense subset. But S is the only subset of itself that is dense in S (since S has the discrete topology) and so S must be countable.

$\Leftarrow$: Suppose for the converse that there is one countable orthonormal basis $S = \{\phi_1, \phi_2, \ldots\}$ 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 Corollary 4.6. 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). □

4.10 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

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

By Corollary 4.8, 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 $\sum_{n=1}^{\infty} a_n \phi_n$ is a Cauchy sequence in H 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. □

4.11 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 implies that $L^2[0, 2\pi]$ is a separable Hilbert space (since it has a countable orthonormal basis) and 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 S is a maximal orthonormal subset of $L^2[0, 2\pi]$. That means that no non-zero $f \in L^2[0, 2\pi]$ can be orthogonal to all the functions $\phi_n(t)$. If such an f did exist it would be orthogonal to all the finite linear combinations

$$p(t) = \sum_{j=-n}^{n} a_j \phi_j(t) = \sum_{j=-n}^{n} \frac{a_j}{\sqrt{2\pi}} e^{ijt}.$$

Such $p(t)$ are known as *trigonometric polynomials* and what we need to do this proof is the fact that the trigonometric polynomials are dense in $L^2[0, 2\pi]$. The proof of this depends on

- Lusin's theorem — a theorem in measure theory that allows us to show that continuous functions $g\colon [0, 2\pi] \to \mathbb{K}$ which satisfy the 2π-periodicity condition $g(0) = g(2\pi)$ are dense in $L^2[0, 2\pi]$.
- A theorem of Weierstrass that shows that each such continuous function can be approximated uniformly on $[0, 2\pi]$ by trigonometric polynomials.
- The fact that $\|g - p\|_2 \le \sqrt{2\pi}\|g - p\|_\infty$ which shows then that these continuous functions can be approximated by trigonometric polynomials in the norm of $L^2[0, 2\pi]$.
- The final step is then to observe that if $f \in L^2[0, 2\pi]$ was orthogonal to all trigonometric polynomials, then $\langle f, f\rangle = 0$ because $f = \lim_{n\to\infty} p_n$ for some sequence of trigonometric polynomials p_n (by density of the polynomials) and so by continuity of the inner product $\langle f, f\rangle = \lim_{n\to\infty}\langle f, p_n\rangle = 0$. Thus $f = 0$.

As Lusin's theorem would take us into measure theory and the proof of the Weierstrass theorem is quite lengthy, we skip this proof.

4.12 Theorem (Gram-Schmidt Orthonormalisation)**.** *Suppose $\phi_1, \phi_2, \ldots$ is a linearly independent set in a Hilbert space H. (By this we mean linearly independent in the usual algebraic sense.) Suppose that the finite linear combinations*

$$a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n$$

are dense in H.

Then the Gram-Schmidt process,

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

produces an orthonormal basis $\psi_1, \psi_2, \ldots$ for H.

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

Hence the linear combinations $\sum_{j=1}^n b_j\psi_j$ are dense in H and $\{\psi_j : j \in \mathbb{N}\}$ is an orthonormal set.

It follows that $\{\psi_j : j \in \mathbb{N}\}$ is a maximal orthonormal set because if $\psi \in H$ is orthogonal to that set then it is orthogonal to the linear combinations $x = \sum_{j=1}^n b_j\psi_j$. There is a sequence x_k of such linear combinations such that $x_k \to \psi$ in H. It follows that $\langle x_n, \psi\rangle = 0 \to \langle \psi, \psi\rangle$ and so that $\|\psi\|^2 = 0$, or $\psi = 0$. $\square$

4.13 Example. By similar reasoning to that outlined for the (omitted) proof of Example 4.11, it can be shown 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 $L^2[-1, 1]$. (Another version of the Weierstrass theorem mentioned earlier says that continuous functions on $[-1, 1]$ can be approximated uniformly by polynomials, hence approximated in $L^2[-1, 1]$-norm by polynomials. Lusin's theorem is used to show that arbitrary elements of $L^2[-1, 1]$ can be approximated by polynomials and this shows the density.)

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 that

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

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

The dual space of a Hilbert space can be identified.

4.14 Theorem (Riesz representation theorem)**.** *Let H be a 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 $y \in H$, $\alpha(x) = \langle x, y \rangle$ defines an element of H^* and moreover $\|\alpha\| = \|y\|$.

Proof. We give most of the details, but not quite all of them.

For $H = \ell^2$, we have nearly done this already because

$$(\ell^2)^* = \ell^2$$

by Examples 3.4 (i). It is a short step from this to the result for all separable Hilbert spaces (using Theorem 4.10).

For a non-separable Hilbert space H, we know there is an orthonormal basis $\{\phi_i : i \in I\}$ for H (with I uncountable). Let $\alpha \in H^*$. We claim that $S = \{i \in I : \alpha(\phi_i) \neq 0\}$ is countable. Let $S_n = \{i \in I : |\alpha(\phi_i)| \geq 1/n\}$. If $i_1, i_2, \ldots, i_k \in S_n$ then we can take

$$x = \sum_{j=1}^{k} \frac{\overline{\alpha(\phi_i)}}{|\alpha(\phi_i)|\sqrt{k}} \phi_i$$

and compute

$$\alpha(x) = \sum_{j=1}^{k} \frac{\overline{\alpha(\phi_i)}}{|\alpha(\phi_i)|\sqrt{k}} \alpha(\phi_i) = \sum_{j=1}^{k} \frac{|\alpha(\phi_i)|}{\sqrt{k}} \geq \frac{\sqrt{k}}{n}$$

But also

$$|\alpha(x)| \leq \|\alpha\| \|x\| = \|\alpha\| \sqrt{\sum_{j=1}^{k} \frac{|\alpha(\phi_i)|^2}{|\alpha(\phi_i)|^2 k}} = \|\alpha\|$$

and so we conclude $\|\alpha\| \geq \frac{\sqrt{k}}{n}$, or $k \leq n\|\alpha\|^2$. This shows that S_n is finite and so $S = \bigcup_n S_n$ must be countable.

As every $x \in I$ can be written

$$x = \sum_{i \in I} \langle x, \phi_i \rangle \phi_i = \sum_{i \in S} \langle x, \phi_i \rangle \phi_i + \sum_{i \in S \setminus I} \langle x, \phi_i \rangle \phi_i$$

(using Bessels inequality to show that the [countably nonzero] sums converge), it follows that $H = H_1 \oplus_2 H_2$ where H_1 is the closed linear span of $\{\phi_i : i \in S\}$ and H_2 is the closed linear span of $\{\phi_i : i \in I \setminus S\}$. It takes just a little checking to see this.

> Some details filled in here.
>
> One can check that H_1 and H_2 are Hilbert spaces (complete since closed, and linear subspaces of H). Also $\{\phi_i : i \in S\}$ is an orthonormal basis for H_1 and $\{\phi_i : i \in I \setminus S\}$ is an orthonormal basis for H_2.
>
> If y is a finite linear combination of elements of $\{\phi_i : i \in S\}$ and z is a finite linear combination of elements of $\{\phi_i : i \in I \setminus S\}$, then it is easy to see that $\langle x, y\rangle = 0$. Taking limits, one sees that $\langle x, y\rangle = 0$ for $y \in H_1$ and $z \in H_2$.
>
> Now it follows that $H_1 \cap H_2 = \{0\}$ and $\|y+z\| = \sqrt{\langle y+z, y+z\rangle} = \sqrt{\|y\|^2 + \|z\|^2}$ (for $y \in H_1$, $z \in H_2$).

As $\alpha(\phi_i) = 0$ for $i \in I \setminus S$, it follows that $\alpha(z) = 0$ for finite linear combinations of $\{\phi_i : i \in I \setminus S\}$. Taking limits (using continuity of α) it follows that $\alpha(z) = 0$ for all $z \in H_2$.

So for $x \in H$, we have $x = y + z$ for (unique) $y \in H_1$, $z \in H_2$, and $\alpha(x) = \alpha(y)$. Applying the separable case to the restriction of α to H_1, there is $y_1 \in H - 1$ with $\alpha(y) = \langle y, y_1\rangle$ for all $y \in H_1$. It then follows that $\alpha(x) = \alpha(y) = \langle y, y_1\rangle = \langle x, y_1\rangle$ for all $x \in H$. □

It is usual to state this theorem as $H^* = H$ for H Hilbert, but that is not quite accurate. Here is a more precise statement

4.15 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. 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. □

4.16 Corollary. *Hilbert spaces are reflexive Banach spaces.*

Proof. This follows from the Riesz representation theorem. □

4.17 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 $x, y \in E$.

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

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

$\Leftarrow$: 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. $\square$

4.18 Remark. 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 3.3.

By Theorem 3.3, for H a Hilbert space $\mathcal{B}(H)$ is a Banach space (in the norm $\|\cdot\|_{op}$).

It also has an algebra structure, where we define multiplication of two operators via composition. 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.

We can estimate the norm of the product

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

This inequality $\|ST\| \leq \|S\|\,\|T\|$ together with the fact that the identity operator has norm 1 makes $\mathcal{B}(H)$ a *Banach algebra*.

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 $\|T^*\|_{op} = \|T\|_{op}$ we use the Riesz representation theorem and the Hahn-Banach theorem.

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 there is some unique $w \in H$ with $\alpha(x) = \langle x, w\rangle$. Define $T^*(y) = w$, 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.

We know $\|T\| = \sup_{x\in H, \|x\|=1} \|T(x)\|$. From the Hahn-Banach theorem (recall Corollary 3.7) $\|z\| = \sup_{\alpha \in H^*, \|\alpha\|=1} |\alpha(z)|$ holds for $z \in H$. By the Riesz-representation theorem, every $\alpha \in H^*$ with $\|\alpha\| = 1$ has the form $\alpha(z) = \langle z, y\rangle$ for $y \in H$ with $\|y\| = 1$ (and conversely every α which can be represented in this way has $\|\alpha\| = 1$. So

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

By reversing the steps we used to write $\|T\|$ as $\sup_{x,y\in H, \|x\|=\|y\|=1} |\langle T(x), y\rangle|$, this last expression if the same as $\|T^*\|$.

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 $\lambda \in \mathbb{C}$. Moreover the important property $\|T\|^2 = \|T^*T\|$ follows easily:

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

Closed *-subalgebras of $\mathcal{B}(H)$ are known as *C*-algebras*. (A *-subalgebras 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 *-algebra, but they also satisfy the property $\|T\|^2 = \|T^*T\|$.)