5. Hilbert spaces

Definition 5.1. Let H be a (complex) vector space. A *scalar product* (or *inner product*) is a map $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ with the following properties:
(1) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0 \iff x = 0$;
(2) $\langle x, y\rangle = \overline{\langle y, x\rangle}$;
(3) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$;
(4) $\langle x, cy\rangle = c\langle x, y\rangle$.

(3), (4) say that a scalar product is linear in the second argument, and by combining this with (2), we see that it is *antilinear* in the first argument, that is $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, as usual, but $\langle cx, y\rangle = \overline{c}\,\langle x, y\rangle$.

Example 5.1. It is easy to check that

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$$

defines a scalar product on $H = \ell^2$. Indeed, the series converges by Hölder's inequality with $p = q = 2$, and once we know that, it is clear that (1)–(4) from above hold.

In fact, this works for arbitrary index sets I: there is a similarly defined scalar product on $\ell^2(I)$. I mention this fact here because we will actually make brief use of this space later in this chapter.

Similarly,

$$\langle f, g\rangle = \int_X \overline{f(x)} g(x)\, d\mu(x)$$

defines a scalar product on $L^2(X, \mu)$.

Theorem 5.2. *Let H be a space with a scalar product. Then:*
(a) $\|x\| := \sqrt{\langle x, x\rangle}$ defines a norm on H;
(b) The Cauchy-Schwarz inequality *holds:*

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

(c) We have equality in (b) if and only if x, y are linearly dependent.

Proof. We first discuss parts (b) and (c). Let $x, y \in H$. Property (1) from Definition 5.1 shows that

$$0 \leq \langle cx + y, cx + y\rangle = |c|^2\|x\|^2 + \|y\|^2 + c\langle y, x\rangle + \overline{c}\langle x, y\rangle, \tag{5.1}$$

for arbitrary $c \in \mathbb{C}$. If $x \neq 0$, we can take $c = -\langle x, y\rangle/\|x\|^2$ here (note that (1) makes sure that $\|x\| = \sqrt{\langle x, x\rangle} > 0$, even though we don't know yet that this really is a norm). Then (5.1) says that

$$0 \leq \|y\|^2 - \frac{|\langle x, y\rangle|^2}{\|x\|^2},$$

and this implies the Cauchy-Schwarz inequality. Moreover, we can get equality in (5.1) only if $cx + y = 0$, so x, y are linearly dependent in this case. Conversely, if $y = cx$ or $x = cy$, then it is easy to check that we do get equality in (b).

We can now prove (a). Property (1) from Definition 5.1 immediately implies condition (1) from Definition 2.1. Moreover, $\|cx\| = \sqrt{\langle cx, cx\rangle} = \sqrt{|c|^2\langle x, x\rangle} = |c|\,\|x\|$, and the triangle inequality follows from the Cauchy-Schwarz inequality, as follows:

$$\begin{aligned} \|x+y\|^2 = \langle x+y, x+y\rangle &= \|x\|^2 + \|y\|^2 + 2 \text{ Re } \langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Notice that we recover the usual norms on ℓ^2 and L^2, respectively, if we use the scalar products introduced in Example 5.1. It now seems natural to ask if every norm is of the form $\|x\| = \sqrt{\langle x, x\rangle}$ for some scalar product $\langle\cdot,\cdot\rangle$. This question admits a neat, satisfactory answer (although it must be admitted that this result doesn't seem to have meaningful applications):

Exercise 5.1. Let H be a vector space with a scalar product, and introduce a norm $\|\cdot\|$ on H as in Theorem 5.2(a). Then $\|\cdot\|$ satisfies the *parallelogram identity:*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{5.2}$$

One can now show that (5.2) is also a sufficient condition for a norm to come from a scalar product (the *Jordan-von Neumann Theorem*). This converse is much harder to prove; we don't want to discuss it in detail here. However, I will mention how to get this proof started. The perhaps somewhat surprising fact is that the norm already completely determines its scalar product (assuming now that the norm does come from a scalar product). In fact, we can be totally explicit, as Proposition 5.3 below will show. A slightly more general version is often useful; to state this, we need an additional definition: A *sesquilinear form* is a map $s : H\times H \to \mathbb{C}$ that is linear in the second argument and antilinear in the first ("sesquilinear" = one and a half linear):

$$s(x, cy + dz) = cs(x, y) + ds(x, z)$$
$$s(cx + dy, z) = \overline{c}s(x, z) + \overline{d}s(y, z)$$

A scalar product has these properties, but this new notion is more general.

Proposition 5.3 (The polarization identity)**.** *Let s be a sesquilinear form, and let $q(x) = s(x, x)$. Then*

$$s(x, y) = \frac{1}{4}\left[q(x + y) - q(x - y) + iq(x - iy) - iq(x + iy)\right].$$

Exercise 5.2. Prove Proposition 5.3, by a calculation.

This is an extremely useful tool and has many applications. The polarization identity suggest the principle "it is often enough to know what happens on the diagonal."

In the context of the Jordan-von Neumann Theorem, it implies that the scalar product can be recovered from its norm, as already mentioned above. This is in fact immediate now because if $s(x, y) = \langle x, y\rangle$, then $q(x) = \|x\|^2$, so the polarization identity gives us $\langle x, y\rangle$ in terms of the norms of $x \pm y$, $x \pm iy$.

Exercise 5.3. Use the result from Exercise 5.1 to prove that the norms $\|\cdot\|_p$ on ℓ^p are not generated by a scalar product for $p \neq 2$.

Given a scalar product on a space H, we always equip H with the norm from Theorem 5.2(a) also. So all constructions and results on normed spaces can be used in this setting, and, in particular, we have a topology on H. The following observation is similar to the result from Exercise 2.2(b).

Corollary 5.4. *The scalar product is continuous: if $x_n \to x$, $y_n \to y$, then also $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Exercise 5.4. Deduce this from the Cauchy-Schwarz inequality.

As usual, complete spaces are particularly important, so they again get a special name:

Definition 5.5. A complete space with scalar product is called a *Hilbert space.*

Or we could say a Hilbert space is a Banach space whose norm comes from a scalar product. By Example 5.1, ℓ^2 and L^2 are Hilbert spaces (we of course know that these are Banach spaces, so there's nothing new to check here). On the other hand, Exercise 5.3 says that ℓ^p cannot be given a Hilbert space structure (that leaves the norm intact) if $p \neq 2$. Hilbert spaces are very special Banach spaces. Roughly speaking, the scalar product allows us to introduce angles between vectors, and this additional structure makes things much more pleasant. There is no such notion on a general Banach space.

In particular, a scalar product leads to a natural notion of orthogonality, and this can be used to great effect. In the sequel, H will always assumed to be a Hilbert space. We say that $x, y \in H$ are *orthogonal* if $\langle x, y\rangle = 0$. In this case, we also write $x \perp y$. If $M \subseteq H$ is an arbitrary subset of H, we define its *orthogonal complement* by

$$M^{\perp} = \{x \in H : \langle x, m\rangle = 0 \text{ for all } m \in M\}.$$

Exercise 5.5. Prove the following formula, which is reminiscent of the Pythagorean theorem: If $x \perp y$, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2. \tag{5.3}$$

Theorem 5.6. *(a) $M^{\perp}$ is a closed subspace of H.*
(b) $M^{\perp} = L(M)^{\perp} = \overline{L(M)}^{\perp}$

Here, $L(M)$ denotes the linear span of M, that is, $L(M)$ is the smallest subspace containing M. A more explicit description is also possible: $L(M) = \{\sum_{j=1}^{n} c_j m_j : c_j \in \mathbb{C}, m_j \in M, n \in \mathbb{N}\}$

Proof. (a) To show that $M^{\perp}$ is a subspace, let $x, y \in M^{\perp}$. Then, for arbitrary $m \in M$, $\langle x + y, m\rangle = \langle x, m\rangle + \langle y, m\rangle = 0$, so $x + y \in M^{\perp}$ also. A similar argument works for multiples of vectors from $M^{\perp}$.

If $x_n \in M^{\perp}$, $x \in H$, $x_n \to x$ and $m \in M$ is again arbitrary, then, by the continuity of the scalar product (Corollary 5.4),

$$\langle x, m\rangle = \lim_{n\to\infty} \langle x_n, m\rangle = 0,$$

so $x \in M^{\perp}$ also and $M^{\perp}$ turns out to be closed, as claimed.

(b) From the definition of $A^{\perp}$, it is clear that $A^{\perp} \supseteq B^{\perp}$ if $A \subseteq B$. Since obviously $M \subseteq L(M) \subseteq \overline{L(M)}$, we see that $\overline{L(M)}^{\perp} \subseteq L(M)^{\perp} \subseteq M^{\perp}$. On the other hand, if $x \in M^{\perp}$, then $\langle x, m\rangle = 0$ for all $m \in M$. Since the scalar product is linear in the second argument, this implies that $\langle x, y\rangle = 0$ for all $y \in L(M)$. Since the scalar product is also continuous, it now follows that in fact $\langle x, z\rangle = 0$ for all $z \in \overline{L(M)}$, that is, $x \in \overline{L(M)}^{\perp}$. □

Exercise 5.6. (a) Show that the closure of a subspace is a subspace again. (This shows that $\overline{L(M)}$ can be described as the smallest closed subspace containing M.)
(b) Show that $L(\overline{M}) \subseteq \overline{L(M)}$.
(c) Show that it can happen that $L(\overline{M}) \neq \overline{L(M)}$.
Suggestion: Consider $M = \{e_n : n \geq 1\} \subseteq \ell^2$.

Theorem 5.7. *Let $M \subseteq H$ be a closed subspace of H, and let $x \in H$. Then there exists a unique best approximation to x in M, that is, there exists a unique $y \in M$ such that*

$$\|x - y\| = \inf_{m \in M} \|x - m\|.$$

Proof. Write $d = \inf_{m \in M} \|x - m\|$ and pick a sequence $y_n \in M$ with $\|x - y_n\| \to d$. The parallelogram identity (5.2) implies that

$$\begin{aligned}\|y_m - y_n\|^2 &= \|(y_m - x) - (y_n - x)\|^2 \\ &= 2\|y_m - x\|^2 + 2\|y_n - x\|^2 - \|y_m + y_n - 2x\|^2 \\ &= 2\|y_m - x\|^2 + 2\|y_n - x\|^2 - 4\|(1/2)(y_m + y_n) - x\|^2.\end{aligned}$$

Now if $m, n \to \infty$, then the first two terms in this final expression both converge to $2d^2$, by the choice of y_n. Since $(1/2)(y_m + y_n) \in M$, we have $\|(1/2)(y_m + y_n) - x\| \geq d$ for all m, n. It follows that $\|y_m - y_n\| \to 0$ as $m, n \to \infty$, so y_n is a Cauchy sequence. Let $y = \lim_{n\to\infty} y_n$. Since M is closed, $y \in M$, and by the continuity of the norm, $\|x - y\| = \lim \|x - y_n\| = d$, so y is a best approximation.

To prove the uniqueness of y, assume that $y' \in M$ also satisfies $\|x - y'\| = d$. Then, by the above calculation, with y_m, y_n replaced by y, y', we have

$$\begin{aligned}\|y - y'\|^2 &= 2\|y - x\|^2 + 2\|y' - x\|^2 - 4\|(1/2)(y + y') - x\|^2 \\ &= 4d^2 - 4\|(1/2)(y + y') - x\|^2.\end{aligned}$$

Again, since $(1/2)(y + y') \in M$, this last norm is $\geq d$, so the whole expression is ≤ 0 and we must have $y = y'$, as desired. □

These best approximations can be used to project orthogonally onto closed subspaces of a Hilbert space. More precisely, we have the following:

Theorem 5.8. *Let $M \subseteq H$ be a closed subspace. Then every $x \in H$ has a unique representation of the form $x = y + z$, with $y \in M$, $z \in M^\perp$.*

Proof. Use Theorem 5.7 to define $y \in M$ as the best approximation to x from M, that is, $\|x - y\| \leq \|x - m\|$ for all $m \in M$. Let $z = x - y$. We want to show that $z \in M^\perp$. If $w \in M$, $w \neq 0$, and $c \in \mathbb{C}$, then

$$\|z\|^2 \leq \|x - (y + cw)\|^2 = \|z - cw\|^2 = \|z\|^2 + |c|^2\,\|w\|^2 - 2\,\mathrm{Re}\,c\langle z, w\rangle.$$

In particular, with $c = \frac{\langle w,z\rangle}{\|w\|^2}$, this shows that $|\langle w, z\rangle|^2 \leq 0$, so $\langle w, z\rangle = 0$, and since this holds for every $w \in M$, we see that $z \in M^\perp$, as desired.

To show that the decomposition from Theorem 5.8 is unique, suppose that $x = y + z = y' + z'$, with $y, y' \in M$, $z, z' \in M^\perp$. Then $y - y' = z' - z \in M \cap M^\perp = \{0\}$, so $y = y'$ and $z = z'$. □

Corollary 5.9. *For an arbitrary subset $A \subseteq H$, we have $A^{\perp\perp} = \overline{L(A)}$.*

Proof. From the definition of $(\dots)^\perp$, we see that $B \subseteq B^{\perp\perp}$, so Theorem 5.6(b) implies that $\overline{L(A)} \subseteq A^{\perp\perp}$.

On the other hand, if $x \in A^{\perp\perp}$, we can use Theorem 5.8 to write $x = y + z$ with $y \in \overline{L(A)}$, $z \in \overline{L(A)}^\perp = A^\perp$. The last equality again follows from Theorem 5.6(b). By what we just showed, we then also have $y \in A^{\perp\perp}$ and thus $z = x - y \in A^\perp \cap A^{\perp\perp} = \{0\}$, so $x = y \in \overline{L(A)}$. □

We now introduce a linear operator that produces the decomposition from Theorem 5.8. Let $M \subseteq H$ be a closed subspace. We then define $P_M : H \to H$, $P_M x = y$, where $y \in M$ is as in Theorem 5.8; P_M is called the *(orthogonal) projection* onto M.

Proposition 5.10. *$P_M \in B(H)$, $P_M^2 = P_M$, and if $M \neq \{0\}$, then $\|P_M\| = 1$.*

Proof. We will only compute the operator norm of P_M here.

Exercise 5.7. Prove that P_M is linear and $P_M^2 = P_M$.

Write $x = P_M x + z$. Then $P_M x \in M$, $z \in M^\perp$, so, by the Pythagorean formula (5.3),

$$\|x\|^2 = \|P_M x\|^2 + \|z\|^2 \geq \|P_M x\|^2.$$

Thus $P_M \in B(H)$ and $\|P_M\| \leq 1$. On the other hand, if $x \in M$, then $P_M x = x$, so $\|P_M\| = 1$ if $M \neq \{0\}$. □

We saw in Chapter 4 that $(\ell^2)^* = \ell^2$, $(L^2)^* = L^2$. This is no coincidence.

Theorem 5.11 (Riesz Representation Theorem)**.** *Every $F \in H^*$ has the form $F(x) = \langle y, x \rangle$, for some $y = y_F \in H$. Moreover, $\|F\| = \|y_F\|$.*

We can say more: conversely, every $y \in H$ generates a bounded, linear functional $F = F_y$ via $F_y(x) = \langle y, x \rangle$. So we can define a map $I : H \to H^*$, $y \mapsto F_y$. This map is injective (why?), and, by the Riesz representation theorem, I is also surjective and isometric, so we obtain an identification of H with H^*. This is a convenient way to summarize the Riesz representation theorem, but note that I narrowly fails to be a (Banach space) isomorphism: it is *antilinear*, that is, $F_{y+z} = F_y + F_z$, as usual, but $F_{cy} = \overline{c} F_y$.

Exercise 5.8. Deduce from this that Hilbert spaces are reflexive. If we ignore the identification maps and just pretend that $H = H^*$ and proceed formally, then this becomes obvious: $H^{**} = (H^*)^* = H^* = H$.

Please give a careful argument. Recall that you really need to show that $j(H) = H^{**}$, where j was defined in Chapter 4. (This is surprisingly awkward to write down; perhaps you want to use the fact that $F : X \to \mathbb{C}$ is antilinear precisely if $\overline{F}$ is linear.)

Exercise 5.9. Let X be a (complex) vector space and let $F : X \to \mathbb{C}$ be a linear functional, $F \neq 0$.
(a) Show that codim $N(F) = 1$, that is, show that there exists a one-dimensional subspace $M \subseteq X$, $M \cap N(F) = \{0\}$, $M + N(F) = X$. (This is an immediate consequence of linear algebra facts, but you can also do it by hand.)
(b) Let F, G be linear functionals with $N(F) = N(G)$. Then $F = cG$ for some $c \in \mathbb{C}$, $c \neq 0$.

Proof of Theorem 5.11. This is surprisingly easy; Exercise 5.9 provides the motivation for the following argument and also explains why this procedure (take an arbitrary vector from $N(F)^{\perp}$) works.

If $F = 0$, we can of course just take $y = 0$. If $F \neq 0$, then $N(F) \neq H$, and $N(F)$ is a closed subspace because F is continuous. Therefore, Corollary 5.9 shows that $N(F)^{\perp} \neq \{0\}$. Pick a vector $z \in N(F)^{\perp}$, $z \neq 0$. Then, for arbitrary $x \in H$, we have $F(z)x - F(x)z \in N(F)$, so

$$0 = \langle z, F(z)x - F(x)z \rangle = F(z)\langle z, x \rangle - F(x)\|z\|^2.$$

Rearranging, we obtain $F(x) = \langle y, x \rangle$, with $y = \frac{\overline{F(z)}}{\|z\|^2} z$.

Since $|\langle y, x \rangle| \leq \|y\| \, \|x\|$, it is clear that $\|F\| \leq \|y\|$. On the other hand, $F(y) = \|y\|^2$, so $\|F\| = \|y\|$. □

Exercise 5.10. Corollary 4.2(b), when combined with the Riesz Representation Theorem, implies that

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

Give a quick direct proof of this fact.

Exercise 5.11. We don't need the Hahn-Banach Theorem on Hilbert spaces because the Riesz Representation Theorem gives a much more explicit description of the dual space. Show that it in fact implies the following stronger version of Hahn-Banach: If $F_0 : H_0 \to \mathbb{C}$ is a bounded linear functional on a subspace H_0, then there exists a *unique* bounded linear extension $F : H \to \mathbb{C}$ with $\|F\| = \|F_0\|$.

Remark: If you want to avoid using the Hahn-Banach Theorem here, you could as a first step extend F_0 to $\overline{H_0}$, by using Exercise 2.26.

The Riesz Representation Theorem also shows that on a Hilbert space, $x_n \xrightarrow{w} x$ if and only if $\langle y, x_n \rangle \to \langle y, x \rangle$; compare Exercise 4.7.

Exercise 5.12. Assume that $x_n \xrightarrow{w} x$.
(a) Show that $\|x\| \le \liminf_{n\to\infty} \|x_n\|$.
(b) Show that it can happen that $\|x\| < \liminf \|x_n\|$.
(c) On the other hand, if $\|x_n\| \to \|x\|$, then $\|x_n - x\| \to 0$ (prove this).

As our final topic in this chapter, we discuss orthonormal bases in Hilbert spaces.

Definition 5.12. A subset $\{x_\alpha : \alpha \in I\}$ is called an *orthogonal system* if $\langle x_\alpha, x_\beta\rangle = 0$ for all $\alpha \neq \beta$. If, in addition, all x_α are normalized (so $\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta}$), we call $\{x_\alpha\}$ an *orthonormal system* (ONS). A maximal ONS is called an *orthonormal basis* (ONB).

Theorem 5.13. *Every Hilbert space has an ONB. Moreover, any ONS can be extended to an ONB.*

This looks very plausible (if an ONS isn't maximal yet, just keep adding vectors). The formal proof depends on Zorn's Lemma; we don't want to do it here.

Theorem 5.14. *Let $\{x_\alpha\}$ be an ONB. Then, for every $y \in H$, we have the expansions*

$$y = \sum_{\alpha\in I} \langle x_\alpha, y\rangle x_\alpha,$$
$$\|y\|^2 = \sum_{\alpha\in I} |\langle x_\alpha, y\rangle|^2 \qquad \textit{(Parseval's identity).}$$

If, conversely, $c_\alpha \in \mathbb{C}$, $\sum_{\alpha\in I} |c_\alpha|^2 < \infty$, then the series $\sum_{\alpha\in I} c_\alpha x_\alpha$ converges to an element $y \in H$.

To make this precise, we need to define sums over arbitrary index sets. We encountered this problem before, in Chapter 2, when defining the spaces $\ell^p(I)$, and we will use the same procedure here: $\sum_{\alpha\in I} w_\alpha = z$ means that $w_\alpha \neq 0$ for at most countably many $\alpha \in I$ and if $\{\alpha_n\}$ is an arbitrary enumeration of these α's, then $\lim_{N\to\infty} \sum_{n=1}^N w_{\alpha_n} = z$. In this definition, we can have $w_\alpha, z \in H$ or $\in \mathbb{C}$. In this latter case, we can also again use counting measure on I to obtain a more elegant formulation.

Theorem 5.14 can now be rephrased in a more abstract way. Consider the map

$$U : H \to \ell^2(I), \qquad (Uy)_\alpha = \langle x_\alpha, y\rangle.$$

Theorem 5.14 says that this is well defined, bijective, and isometric. Moreover, U is also obviously linear. So, summing up, we have a bijection $U \in B(H, \ell^2)$ that also preserves the scalar product: $\langle Ux, Uy\rangle = \langle x, y\rangle$.

Exercise 5.13. Prove this last statement. *Hint:* Polarization!

Such maps are called *unitary;* they preserve the complete Hilbert space structure. In other words, we can now say that Theorem 5.14 shows that $H \cong \ell^2(I)$ for an arbitrary Hilbert space H; more specifically, I can be taken as the index set of an ONB. So we have a one-size-fits-all model space, namely $\ell^2(I)$; there is no such universal model for Banach spaces.

There is a version of Theorem 5.14 for ONS; actually, we will prove the two results together.

Theorem 5.15. *Let $\{x_\alpha\}$ be an ONS. Then, for every $y \in H$, we have*

$$P_{\overline{L(x_\alpha)}}\, y = \sum_{\alpha \in I} \langle x_\alpha, y\rangle x_\alpha,$$

$$\|y\|^2 \geq \sum_{\alpha \in I} |\langle x_\alpha, y\rangle|^2 \qquad \text{(Bessel's inequality).}$$

Proof of Theorems 5.14, 5.15. We start by establishing Bessel's inequality for finite ONS $\{x_1, \ldots, x_N\}$. Let $y \in H$ and write

$$y = \sum_{n=1}^N \langle x_n, y\rangle x_n + \left(y - \sum_{n=1}^N \langle x_n, y\rangle x_n \right).$$

A calculation shows that the two terms on the right-hand side are orthogonal, so

$$\begin{aligned} \|y\|^2 &= \left\| \sum_{n=1}^N \langle x_n, y\rangle x_n \right\|^2 + \left\| y - \sum_{n=1}^N \langle x_n, y\rangle x_n \right\|^2 \\ &= \sum_{n=1}^N |\langle x_n, y\rangle|^2 + \left\| y - \sum_{n=1}^N \langle x_n, y\rangle x_n \right\|^2 \geq \sum_{n=1}^N |\langle x_n, y\rangle|^2. \end{aligned}$$

This is Bessel's inequality for finite ONS. It now follows that the sets $\{\alpha \in I : |\langle x_\alpha, y\rangle| \geq 1/n\}$ are finite, so $\{\alpha : \langle x_\alpha, y\rangle \neq 0\}$ is countable. Let $\{\alpha_1, \alpha_2, \ldots\}$ be an enumeration. Then, by Bessel's inequality (we're still referring to the version for finite ONS), $\lim_{N\to\infty} \sum_{n=1}^N |\langle x_{\alpha_n}, y\rangle|^2$ exists, and, since we have absolute convergence here, the limit is independent of the enumeration. If we recall how $\sum_{\alpha\in I} \ldots$ was defined, we see that we have proved the general version of Bessel's inequality.

As the next step, define $y_n = \sum_{j=1}^n \langle x_{\alpha_j}, y\rangle x_{\alpha_j}$. If $n \geq m$ (say), then

$$\|y_m - y_n\|^2 = \left\| \sum_{j=m+1}^n \langle x_{\alpha_j}, y\rangle x_{\alpha_j} \right\|^2 = \sum_{j=m+1}^n \left| \langle x_{\alpha_j}, y\rangle \right|^2.$$

This shows that y_n is a Cauchy sequence. Let $y' = \lim y_n = \sum_{j=1}^{\infty} \langle x_{\alpha_j}, y\rangle x_{\alpha_j}$. By the continuity of the scalar product,

$$\langle x_{\alpha_k}, y - y'\rangle = \langle x_{\alpha_k}, y\rangle - \sum_{j=1}^{\infty} \langle x_{\alpha_j}, y\rangle \delta_{jk} = 0$$

for all $k \in \mathbb{N}$, and if $\alpha \in I \setminus \{\alpha_j\}$, then it also follows that

$$\langle x_\alpha, y - y'\rangle = -\langle x_\alpha, y'\rangle = -\sum_{j=1}^{\infty} \langle x_{\alpha_j}, y\rangle \langle x_\alpha, x_{\alpha_j}\rangle = 0.$$

So $y - y' \in \{x_\alpha\}^\perp = \overline{L(x_\alpha)}^\perp$, and, by its construction, $y' \in \overline{L(x_\alpha)}$. Thus $y' = P_{\overline{L(x_\alpha)}} y$, as claimed in Theorem 5.15. It now also follows that $\sum_{\alpha \in I} \langle x_\alpha, y\rangle x_\alpha$ exists because we always obtain the same limit $y' = P_{\overline{L(x_\alpha)}} y$, no matter how the α_j are arranged.

To obtain Theorem 5.14, we observe that if $\{x_\alpha\}$ is an ONB, then $\overline{L(x_\alpha)} = H$. Indeed, if this were not true, then the closed subspace $\overline{L(x_\alpha)}$ would have to have a non-zero orthogonal complement, by Corollary 5.9, and we could pass to a bigger ONS by adding a normalized vector from this orthogonal complement. So $\overline{L(x_\alpha)} = H$ if $\{x_\alpha\}$ is an ONB, but then also $y' = y$, and Parseval's identity now follows from the continuity of the norm:

$$\|y\|^2 = \lim_{N\to\infty} \left\| \sum_{i=1}^{N} \langle x_{\alpha_i}, y\rangle x_{\alpha_i} \right\|^2 = \lim_{N\to\infty} \sum_{i=1}^{N} |\langle x_{\alpha_i}, y\rangle|^2 = \sum_{\alpha \in I} |\langle x_\alpha, y\rangle|^2$$

Finally, similar arguments show that $\sum c_\alpha x_\alpha$ exists for $c \in \ell^2(I)$ (consider the partial sums and check that these form a Cauchy sequence). □

We can try to summarize this as follows: once an ONB is fixed, we may use the coefficients with respect to this ONB to manipulate vectors; in particular, there is an easy formula (Parseval's identity) that will give the norm in terms of these coefficients. The situation is quite similar to linear algebra: coefficients with respect to a fixed basis is all we need to know about vectors. Note, however, that ONB's are not bases in the sense of linear algebra: we use *infinite* linear combinations (properly defined as limits) to represent vectors. *Algebraic* bases on infinite-dimensional Hilbert spaces exist, too, but they are almost entirely useless (for example, they can never be countable).

Exercise 5.14. Show that $\{e_n : n \in \mathbb{N}\}$ is an ONB of $\ell^2 = \ell^2(\mathbb{N})$.

Exercise 5.15. Show that $\{e^{inx} : n \in \mathbb{Z}\}$ is an ONB of $L^2\left((-\pi,\pi); \frac{dx}{2\pi}\right)$.

Suggestion: You should not try to prove the maximality directly, but rather refer to suitable facts from Analysis (continuous functions are dense in L^2, and they can be uniformly approximated, on compact sets, by trigonometric polynomials).

Exercise 5.16. (a) For $f \in L^2(-\pi,\pi)$, define the nth *Fourier coefficient* as

$$\widehat{f}_n = \int_{-\pi}^{\pi} f(x)e^{-inx}\, dx,$$

and use the result from Exercise 5.15 to establish the following formula, which is also known as *Parseval's identity:*

$$\sum_{n=-\infty}^{\infty} \left|\widehat{f}_n\right|^2 = 2\pi \int_{-\pi}^{\pi} |f(x)|^2\, dx$$

(b) Prove that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Suggestion: Part (a) with $f(x) = x$.

Exercise 5.17. The *Rademacher functions* $R_0(x) = 1$,

$$R_n(x) = \begin{cases} 1 & x \in \bigcup_{k=0}^{2^{n-1}-1} [k2^{1-n}, (2k+1)2^{-n}) \\ -1 & \text{else} \end{cases}$$

form an ONS, but not an ONB in $L^2(0,1)$. (Please plot the first few functions to get your bearings here.)

Exercise 5.18. (a) Let $U : H_1 \to H_2$ be a unitary map between Hilbert spaces, and let $\{x_\alpha\}$ be an ONB of H_1. Show that $\{Ux_\alpha\}$ is an ONB of H_2.
(b) Conversely, let $U : H_1 \to H_2$ be a linear map that maps an ONB to an ONB again. Show that U is unitary.

Exercise 5.19. Show that a Hilbert space is separable precisely if it has a countable ONB.