Chapter

8

Hilbert Spaces

Normed vector spaces and Banach spaces, which were introduced in Chapter 6, capture the notion of distance. In this chapter we introduce inner product spaces, which capture the notion of angle. As we will see, the concept of orthogonality, which corresponds to right angles in the familiar context of $\mathbf{R}^2$ or $\mathbf{R}^3$, plays a particularly important role in inner product spaces.

Just as a Banach space is defined to be a normed vector space in which every Cauchy sequence converges, a Hilbert space is defined to be an inner product space that is a Banach space. Hilbert spaces are named in honor of German mathematician David Hilbert (1862–1943), who helped develop parts of the theory in the early twentieth century.

In this chapter, we will see a clean description of the bounded linear functionals on a Hilbert space. We will also see that every Hilbert space has an orthonormal basis, which make Hilbert spaces look much like standard Euclidean spaces but with infinite sums replacing finite sums.

The Mathematical Institute at the University of Göttingen. This building was opened in 1930, when Hilbert was near the end of his career at the University of Göttingen. Other prominent mathematicians who had taught at the University of Göttingen and made major contributions to mathematics include Courant, Dedekind, Dirichlet, Gauss, Minkowski, Noether, and Riemann.

8A Inner Product Spaces

Inner Products

If $p = 2$, then the dual exponent p' also equals 2. In this special case Hölder's Inequality (7.9) implies that if μ is a measure, then

$$\left|\int fg\,d\mu\right| \leq \|f\|_2\|g\|_2$$

for all $f, g \in \mathcal{L}^2(\mu)$. Thus we can associate with each pair of functions $f, g \in \mathcal{L}^2(\mu)$ a number $\int fg\,d\mu$. An inner product is almost a generalization of this pairing.

If $g = f$ and $\mathbf{F} = \mathbf{R}$, then the left side of the inequality above equals $\|f\|_2^2$. The next definition will help provide a similar connection with $\|\cdot\|_2$ in the case of complex scalars.

8.1 Definition ***complex conjugate;*** $\bar{z}$

Suppose $z \in \mathbf{C}$. The *complex conjugate* of $z \in \mathbf{C}$, denoted $\bar{z}$ (pronounced z-bar), is defined by

$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$

For example, if $z = 5 + 7i$ then $\bar{z} = 5 - 7i$. Note that a complex number z is a real number if and only if $z = \bar{z}$.

The next result gives basic properties of the complex conjugate.

8.2 ***Properties of complex conjugates***

Suppose $w, z \in \mathbf{C}$. Then

product of z and $\bar{z}$
$z\bar{z} = |z|^2$;

sum of z and $\bar{z}$
$z + \bar{z} = 2\operatorname{Re} z$;

difference of z and $\bar{z}$
$z - \bar{z} = 2(\operatorname{Im} z)i$;

additivity and multiplicativity of complex conjugate
$\overline{w + z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\,\bar{z}$;

complex conjugate of complex conjugate
$\bar{\bar{z}} = z$;

absolute value of complex conjugate
$|\bar{z}| = |z|$;

integral of complex conjugate of a function
$\int \bar{f}\,d\mu = \overline{\int f\,d\mu}$ for every measure μ and every $f \in \mathcal{L}^1(\mu)$.

Proof We have

$$z\bar{z} = (\operatorname{Re} z + i \operatorname{Im} z)(\operatorname{Re} z - i \operatorname{Im} z) = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 = |z|^2,$$

which proves the first item.

To prove the last item, suppose μ is a measure and $f \in \mathcal{L}^1(\mu)$. Then

$$\begin{aligned}
\int \overline{f} \, d\mu &= \int (\operatorname{Re} f - i \operatorname{Im} f) \, d\mu \\
&= \int \operatorname{Re} f \, d\mu - i \int \operatorname{Im} f \, d\mu \\
&= \overline{\int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu} \\
&= \overline{\int f \, d\mu},
\end{aligned}$$

as desired.

The straightforward proofs of the remaining items are left to the reader. ■

Now we are ready to define inner products, which abstract the key properties of the pairing $f, g \mapsto \int f \, \overline{g} \, d\mu$ on $L^2(\mu)$, where μ is a measure.

8.3 Definition ***inner product; inner product space***

An *inner product* on a vector space V is a function that takes each ordered pair f, g of elements of V to a number $\langle f, g \rangle \in \mathbf{F}$ and has the following properties:

positivity

$\langle f, f \rangle \in [0, \infty)$ for all $f \in V$;

definiteness

$\langle f, f \rangle = 0$ if and only if $f = 0$;

linearity in first slot

$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$ and $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$ for all $f, g, h \in V$ and all $\alpha \in \mathbf{F}$;

conjugate symmetry

$\langle f, g \rangle = \overline{\langle g, f \rangle}$ for all $f, g \in V$.

A vector space with an inner product on it is called an *inner product space*. The terminology *real inner product space* indicates that $\mathbf{F} = \mathbf{R}$; the terminology *complex inner product space* indicates that $\mathbf{F} = \mathbf{C}$.

If $\mathbf{F} = \mathbf{R}$, then the conjugate symmetry property above can be rewritten as $\langle f, g \rangle = \langle g, f \rangle$ for all $f, g \in V$.

Although most mathematicians define an inner product as above, many physicists use a definition that requires linearity in the second slot instead of the first slot.

8.4 Example *inner product spaces*

- For $n \in \mathbf{Z}^+$, define an inner product on $\mathbf{F}^n$ by

$$\langle (a_1, \dots, a_n), (b_1, \dots, b_n) \rangle = a_1 \overline{b_1} + \dots + a_n \overline{b_n}$$

for $(a_1, \dots, a_n), (b_1, \dots, b_n) \in \mathbf{F}^n$. When thinking of $\mathbf{F}^n$ as an inner product space, we will always mean this inner product unless the context indicates some other inner product.

- Define an inner product on ℓ^2 by

$$\langle (a_1, a_2, \dots), (b_1, b_2, \dots) \rangle = \sum_{k=1}^{\infty} a_k \overline{b_k}$$

for $(a_1, a_2, \dots), (b_1, b_2, \dots) \in \ell^2$. Hölder's Inequality (7.9), as applied to counting measure on $\mathbf{Z}^+$ and taking $p = 2$, implies that the infinite sum above converges absolutely and hence converges to an element of $\mathbf{F}$. When thinking of ℓ^2 as an inner product space, we will always mean this inner product unless the context indicates some other inner product.

- Define an inner product on $C([0,1])$, which is the vector space of continuous functions from $[0,1]$ to $\mathbf{F}$, by

$$\langle f, g \rangle = \int_0^1 f \overline{g}$$

for $f, g \in C([0,1])$. The definiteness requirement for an inner product is satisfied because if $f \colon [0,1] \to \mathbf{F}$ is a continuous function such that $\int_0^1 |f|^2 = 0$, then the function f is identically 0.

- Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Define an inner product on $L^2(\mu)$ by

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

for $f, g \in L^2(\mu)$. Hölder's Inequality (7.9) with $p = 2$ implies that the integral above makes sense as an element of $\mathbf{F}$. When thinking of $L^2(\mu)$ as an inner product space, we will always mean this inner product unless the context indicates some other inner product.

Here we use $L^2(\mu)$ rather than $\mathcal{L}^2(\mu)$ because the definiteness requirement fails on $\mathcal{L}^2(\mu)$ if there exist nonempty sets $E \in \mathcal{S}$ such that $\mu(E) = 0$ (consider $\langle \chi_E, \chi_E \rangle$ to see the problem).

The first two bullet points in this example are special cases of $L^2(\mu)$, taking μ to be counting measure on either $\{1, \dots, n\}$ or $\mathbf{Z}^+$.

As we will see, even though the main examples of inner product spaces are $L^2(\mu)$ spaces, working with the inner product structure is often cleaner and simpler than working with measures and integrals.

8.5 ***Basic properties of an inner product***

Suppose V is an inner product space. Then

(a) $\langle 0, g\rangle = \langle g, 0\rangle = 0$ for every $g \in V$;

(b) $\langle f, g+h\rangle = \langle f, g\rangle + \langle f, h\rangle$ for all $f, g, h \in V$;

(c) $\langle f, \alpha g\rangle = \overline{\alpha}\langle f, g\rangle$ for all $\alpha \in \mathbf{F}$ and $f, g \in V$.

Proof

(a) For $g \in V$, the function $f \mapsto \langle f, g\rangle$ is a linear map from V to $\mathbf{F}$. Because every linear map takes 0 to 0, we have $\langle 0, g\rangle = 0$. Now the conjugate symmetry property of an inner product implies that

$$\langle g, 0\rangle = \overline{\langle 0, g\rangle} = \overline{0} = 0.$$

(b) Suppose $f, g, h \in V$. Then

$$\langle f, g+h\rangle = \overline{\langle g+h, f\rangle} = \overline{\langle g, f\rangle + \langle h, f\rangle} = \overline{\langle g, f\rangle} + \overline{\langle h, f\rangle} = \langle f, g\rangle + \langle f, h\rangle.$$

(c) Suppose $\alpha \in \mathbf{F}$ and $f, g \in V$. Then

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

as desired. ■

If $\mathbf{F} = \mathbf{R}$, then parts (b) and (c) of 8.5 imply that for $f \in V$, the function $g \mapsto \langle f, g\rangle$ is a linear map from V to $\mathbf{R}$. However, if $\mathbf{F} = \mathbf{C}$ and $f \neq 0$, then the function $g \mapsto \langle f, g\rangle$ is not a linear map from V to $\mathbf{C}$ because of the complex conjugate in part (c) of 8.5.

Cauchy–Schwarz Inequality and Triangle Inequality

Now we can define the norm associated with each inner product. We use the word *norm* (which will turn out to be correct) even though it is not yet clear that all the properties required of a norm are satisfied.

8.6 Definition ***norm associated with an inner product;*** $\|\cdot\|$

Suppose V is an inner product space. For $f \in V$, define the *norm* of f, denoted $\|f\|$, by

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

8.7 Example *norms on inner product spaces*

In each of the following examples, the inner product is the standard inner product as defined in Example 8.4.

- If $n \in \mathbf{Z}^+$ and $(a_1, \dots, a_n) \in \mathbf{F}^n$, then
$$\|(a_1, \dots, a_n)\| = \sqrt{|a_1|^2 + \cdots + |a_n|^2}.$$
Thus the norm on $\mathbf{F}^n$ associated with the standard inner product is the usual Euclidean norm.

- If $(a_1, a_2, \dots) \in \ell^2$, then
$$\|(a_1, a_2, \dots)\| = \Bigl(\sum_{k=1}^{\infty} |a_k|^2\Bigr)^{1/2}.$$
Thus the norm associated with the inner product on ℓ^2 is just the standard norm $\|\cdot\|_2$ on ℓ^2 as defined in Example 7.2.

- If μ is a measure and $f \in L^2(\mu)$, then
$$\|f\| = \Bigl(\int |f|^2 \, d\mu\Bigr)^{1/2}.$$
Thus the norm associated with the inner product on $L^2(\mu)$ is just the standard norm $\|\cdot\|_2$ on $L^2(\mu)$ as defined in 7.17.

The definition of an inner product (8.3) implies that if V is an inner product space and $f \in V$, then

- $\|f\| \geq 0$;
- $\|f\| = 0$ if and only if $f = 0$.

The proof of the next result illustrates a frequently used property of the norm on an inner product space: working with the square of the norm is often easier than working directly with the norm.

8.8 *Homogeneity of the norm*

Suppose V is an inner product space, $f \in V$, and $\alpha \in \mathbf{F}$. Then
$$\|\alpha f\| = |\alpha| \, \|f\|.$$

Proof We have
$$\|\alpha f\|^2 = \langle \alpha f, \alpha f \rangle = \alpha \langle f, \alpha f \rangle = \alpha \overline{\alpha} \langle f, f \rangle = |\alpha|^2 \|f\|^2.$$
Taking square roots now gives the desired equality. ■

The next definition plays a crucial role in the study of inner product spaces.

8.9 Definition *orthogonal*

Two elements of an inner product space are called *orthogonal* if their inner product equals 0.

In the definition above, the order of the two elements of the inner product space does not matter because $\langle f, g \rangle = 0$ if and only if $\langle g, f \rangle = 0$. Instead of saying that f and g are orthogonal, sometimes we say that f is orthogonal to g.

8.10 Example *orthogonal elements of an inner product space*

- In $\mathbf{C}^3$, $(2, 3, 5i)$ and $(6, 1, -3i)$ are orthogonal because

$$\langle (2, 3, 5i), (6, 1, -3i) \rangle = 2 \cdot 6 + 3 \cdot 1 + 5i \cdot (3i) = 12 + 3 - 15 = 0.$$

- The elements of $L^2([0, 2\pi])$ represented by $\sin(3t)$ and $\cos(8t)$ are orthogonal because

$$\int_0^{2\pi} \sin(3t) \cos(8t) \, dt = \left[\frac{\cos(5t)}{10} - \frac{\cos(11t)}{22} \right]_{t=0}^{t=2\pi} = 0,$$

where dt denotes integration with respect to Lebesgue measure on $[0, 2\pi]$.

Exercise 7 asks you to prove that if a and b are nonzero elements in $\mathbf{R}^2$, then

$$\langle a, b \rangle = \|a\| \, \|b\| \cos \theta,$$

where θ is the angle between a and b (thinking of a as the vector whose initial point is the origin and whose end point is a, and similarly for b). Thus two elements of $\mathbf{R}^2$ are orthogonal if and only if the cosine of the angle between them is 0, which happens if and only if the vectors are perpendicular in the usual sense of plane geometry. Thus you can think of the word *orthogonal* as a fancy word meaning *perpendicular*.

Law professor Richard Friedman presenting a case before the U.S. Supreme Court in 2010:

Mr. Friedman: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging—
Chief Justice Roberts: I'm sorry. Entirely what?
Mr. Friedman: Orthogonal. Right angle. Unrelated. Irrelevant.
Chief Justice Roberts: Oh.
Justice Scalia: What was that adjective? I liked that.
Mr. Friedman: Orthogonal.
Chief Justice Roberts: Orthogonal.
Mr. Friedman: Right, right.
Justice Scalia: Orthogonal, ooh. (Laughter.)
Justice Kennedy: I knew this case presented us a problem. (Laughter.)

The next theorem is over 2500 years old, although it was not originally stated in the context of inner product spaces.

8.11 *Pythagorean Theorem*

Suppose f and g are orthogonal elements of an inner product space. Then

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2.$$

Proof We have

$$\begin{aligned}\|f+g\|^2 &= \langle f+g, f+g\rangle \\ &= \langle f,f\rangle + \langle f,g\rangle + \langle g,f\rangle + \langle g,g\rangle \\ &= \|f\|^2 + \|g\|^2,\end{aligned}$$

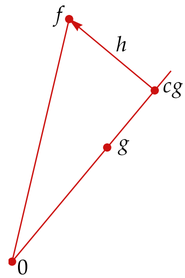
as desired. ■

Exercise 2 shows that whether or not the converse of the Pythagorean Theorem holds depends upon whether $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$.

Suppose f and g are elements of an inner product space V, with $g \neq 0$. Frequently it is useful to write f as some number c times g plus an element h of V that is orthogonal to g. The figure here suggests that such a decomposition should be possible. To find the appropriate choice for c, note that if $f = cg + h$ for some $c \in \mathbf{F}$ and some $h \in V$ with $\langle h, g\rangle = 0$, then we must have

$$\langle f, g\rangle = \langle cg + h, g\rangle = c\|g\|^2,$$

which implies that $c = \dfrac{\langle f,g\rangle}{\|g\|^2}$, which then implies that $h = f - \dfrac{\langle f,g\rangle}{\|g\|^2}g$. Hence we are led to the following result.



Here $f = cg + h$, where h is orthogonal to g.

8.12 *Orthogonal decomposition*

Suppose f and g are elements of an inner product space, with $g \neq 0$. Set $h = f - \dfrac{\langle f,g\rangle}{\|g\|^2}g$. Then

$$\langle h, g\rangle = 0 \quad \text{and} \quad f = \frac{\langle f,g\rangle}{\|g\|^2}g + h.$$

Proof We have

$$\langle h, g\rangle = \Big\langle f - \frac{\langle f,g\rangle}{\|g\|^2}g, g\Big\rangle = \langle f, g\rangle - \frac{\langle f,g\rangle}{\|g\|^2}\langle g, g\rangle = 0,$$

giving the first equation in the conclusion. The second equation in the conclusion follows immediately from the definition of h. ■

The orthogonal decomposition 8.12 will be used in our proof of the next result, which is one of the most important inequalities in mathematics.

8.13 *Cauchy–Schwarz Inequality*

Suppose f and g are elements of an inner product space. Then

$$|\langle f, g\rangle| \leq \|f\| \, \|g\|,$$

with equality if and only if one of f, g is a scalar multiple of the other.

Proof If $g = 0$, then both sides of the desired inequality equal 0. Thus we can assume $g \neq 0$. Consider the orthogonal decomposition

$$f = \frac{\langle f, g\rangle}{\|g\|^2} g + h$$

given by 8.12, where h is orthogonal to g. The Pythagorean Theorem (8.11) implies

$$\begin{aligned} \|f\|^2 &= \left\| \frac{\langle f, g\rangle}{\|g\|^2} g \right\|^2 + \|h\|^2 \\ &= \frac{|\langle f, g\rangle|^2}{\|g\|^2} + \|h\|^2 \end{aligned}$$

8.14
$$\geq \frac{|\langle f, g\rangle|^2}{\|g\|^2}.$$

Multiplying both sides of this inequality by $\|g\|^2$ and then taking square roots gives the desired inequality.

The proof above shows that the Cauchy–Schwarz Inequality is an equality if and only if 8.14 is an equality. This happens if and only if $h = 0$. But $h = 0$ if and only if f is a multiple of g (see 8.12). Thus the Cauchy–Schwarz Inequality is an equality if and only if f is a scalar multiple of g or g is a scalar multiple of f (or both; the phrasing has been chosen to cover cases in which either f or g equals 0). ■

8.15 Example *Cauchy–Schwarz Inequality for* $\mathbf{F}^n$

Applying the Cauchy–Schwarz Inequality with the standard inner product on $\mathbf{F}^n$ to $(|a_1|, \ldots, |a_n|)$ and $(|b_1|, \ldots, |b_n|)$ gives the inequality

$$|a_1 b_1| + \cdots + |a_n b_n| \leq \sqrt{|a_1|^2 + \cdots + |a_n|^2} \sqrt{|b_1|^2 + \cdots + |b_n|^2}$$

for all $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbf{F}^n$.

The inequality in this example was first proved by Cauchy in 1821.

Thus we have a new and clean proof of Hölder's Inequality (7.9) for the special case where μ is counting measure on $\{1, \ldots, n\}$ and $p = p' = 2$.

8.16 Example *Cauchy–Schwarz Inequality for $L^2(\mu)$*

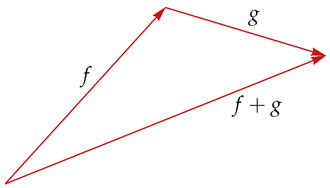
Suppose μ is a measure and $f, g \in L^2(\mu)$. Applying the Cauchy–Schwarz Inequality with the standard inner product on $L^2(\mu)$ to $|f|$ and $|g|$ gives the inequality

$$\int |fg| \, d\mu \leq \Big(\int |f|^2 \, d\mu\Big)^{1/2} \Big(\int |g|^2 \, d\mu\Big)^{1/2}.$$

The inequality above is equivalent to Hölder's Inequality (7.9) for the special case where $p = p' = 2$. However, the proof of the inequality above via the Cauchy–Schwarz Inequality still depends upon Hölder's Inequality to show that the definition of the standard inner product on $L^2(\mu)$ makes sense. See Exercise 16 in this section for a derivation of the inequality above that is truly independent of Hölder's Inequality.

In 1859 Russian mathematician Viktor Bunyakovsky (1804–1889), who had been Cauchy's student in Paris, first proved integral inequalities like the one above. Similar discoveries by German mathematician Hermann Schwarz (1843–1921) in 1885 attracted more attention and led to the name of this inequality.

If we think of the norm determined by an inner product as a length, then the Triangle Inequality has the geometric interpretation that the length of each side of a triangle is less than the sum of the lengths of the other two sides.



8.17 *Triangle Inequality*

Suppose f and g are elements of an inner product space. Then

$$\|f + g\| \leq \|f\| + \|g\|,$$

with equality if and only if one of f, g is a nonnegative multiple of the other.

Proof We have

$$\begin{aligned}
\|f+g\|^2 &= \langle f+g, f+g \rangle \\
&= \langle f, f \rangle + \langle g, g \rangle + \langle f, g \rangle + \langle g, f \rangle \\
&= \langle f, f \rangle + \langle g, g \rangle + \langle f, g \rangle + \overline{\langle f, g \rangle} \\
&= \|f\|^2 + \|g\|^2 + 2\,\mathrm{Re}\langle f, g \rangle \\
&\leq \|f\|^2 + \|g\|^2 + 2|\langle f, g \rangle| && (8.18) \\
&\leq \|f\|^2 + \|g\|^2 + 2\|f\|\,\|g\| && (8.19) \\
&= (\|f\| + \|g\|)^2,
\end{aligned}$$

where 8.19 follows from the Cauchy–Schwarz Inequality (8.13). Taking square roots of both sides of the inequality above gives the desired inequality.

The proof above shows that the Triangle Inequality is an equality if and only if we have equality in 8.18 and 8.19. Thus we have equality in the Triangle Inequality if and only if

8.20 $$\langle f, g\rangle = \|f\|\,\|g\|.$$

If one of f, g is a nonnegative multiple of the other, then 8.20 holds, as you should verify. Conversely, suppose 8.20 holds. Then the condition for equality in the Cauchy–Schwarz Inequality (8.13) implies that one of f, g is a scalar multiple of the other. Clearly 8.20 forces the scalar in question to be nonnegative, as desired. ∎

Applying the previous result to the inner product space $L^2(\mu)$, where μ is a measure, gives a new proof of Minkowski's Inequality (7.14) for the case $p = 2$.

Now we can prove that what we have been calling a norm on an inner product space is indeed a norm.

8.21 $\|\cdot\|$ *is a norm*

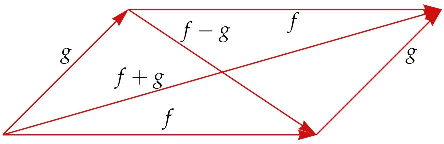
Suppose V is an inner product space and $\|f\|$ is defined as usual by

$$\|f\| = \sqrt{\langle f, f\rangle}$$

for $f \in V$. Then $\|\cdot\|$ is a norm on V.

Proof The definition of an inner product implies that $\|\cdot\|$ satisfies the positive definite requirement for a norm. The homogeneity and triangle inequality requirements for a norm are satisfied because of 8.8 and 8.17. ∎

The next result has the geometric interpretation that the sum of the squares of the lengths of the diagonals of a parallelogram equals the sum of the squares of the lengths of the four sides.



8.22 *Parallelogram Equality*

Suppose f and g are elements of an inner product space. Then

$$\|f+g\|^2 + \|f-g\|^2 = 2\|f\|^2 + 2\|g\|^2.$$

Proof We have

$$\begin{aligned}\|f+g\|^2 + \|f-g\|^2 &= \langle f+g, f+g\rangle + \langle f-g, f-g\rangle \\ &= \|f\|^2 + \|g\|^2 + \langle f, g\rangle + \langle g, f\rangle \\ &\qquad + \|f\|^2 + \|g\|^2 - \langle f, g\rangle - \langle g, f\rangle \\ &= 2\|f\|^2 + 2\|g\|^2,\end{aligned}$$

as desired. ∎

EXERCISES 8A

1 Let V denote the vector space of bounded continuous functions from $\mathbf{R}$ to $\mathbf{F}$. Let $r_1, r_2, \dots$ be a list of the rational numbers. For $f, g \in V$, define

$$\langle f, g \rangle = \sum_{k=1}^{\infty} \frac{f(r_k)\overline{g(r_k)}}{2^k}.$$

Show that $\langle \cdot, \cdot \rangle$ is an inner product on V.

2 Suppose f and g are elements of an inner product space and

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

(a) Prove that if $\mathbf{F} = \mathbf{R}$, then f and g are orthogonal.

(b) Give an example to show that if $\mathbf{F} = \mathbf{C}$, then f and g can satisfy the equation above without being orthogonal.

3 Find $a, b \in \mathbf{R}^3$ such that a is a scalar multiple of $(1, 6, 3)$, b is orthogonal to $(1, 6, 3)$, and $(5, 4, -2) = a + b$.

4 Prove that

$$16 \leq (a + b + c + d)\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)$$

for all positive numbers a, b, c, d, with equality if and only if $a = b = c = d$.

5 Prove that the square of the average of each finite list of real numbers containing at least two distinct real numbers is less than the average of the squares of the numbers in that list.

6 Suppose f and g are elements of an inner product space and $\|f\| \leq 1$ and $\|g\| \leq 1$. Prove that

$$\sqrt{1 - \|f\|^2}\sqrt{1 - \|g\|^2} \leq 1 - |\langle f, g \rangle|.$$

7 Suppose a and b are nonzero elements of $\mathbf{R}^2$. Prove that

$$\langle a, b \rangle = \|a\| \, \|b\| \cos\theta,$$

where θ is the angle between a and b (thinking of a as the vector whose initial point is the origin and whose end point is a, and similarly for b).

Hint: Draw the triangle formed by a, b, and $a - b$; then use the law of cosines.

8 The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $a, b \in \mathbf{R}^n$ is defined to be

$$\arccos \frac{\langle a, b \rangle}{\|a\| \, \|b\|},$$

where the motivation for this definition comes from the previous exercise. Explain why the Cauchy–Schwarz Inequality is needed to show that this definition makes sense.

9 (a) Suppose f and g are elements of a real inner product space. Prove that f and g have the same norm if and only if $f+g$ is orthogonal to $f-g$.

(b) Use part (a) to show that the diagonals of a parallelogram are perpendicular to each other if and only if the parallelogram is a rhombus.

10 Suppose f and g are elements of an inner product space. Prove that $\|f\| = \|g\|$ if and only if $\|sf + tg\| = \|tf + sg\|$ for all $s, t \in \mathbf{R}$.

11 Suppose f and g are elements of an inner product space and $\|f\| = \|g\| = 1$ and $\langle f, g\rangle = 1$. Prove that $f = g$.

12 Suppose f and g are elements of a real inner product space. Prove that

$$\langle f, g\rangle = \frac{\|f+g\|^2 - \|f-g\|^2}{4}.$$

13 Suppose f and g are elements of a complex inner product space. Prove that

$$\langle f, g\rangle = \frac{\|f+g\|^2 - \|f-g\|^2 + \|f+ig\|^2 i - \|f-ig\|^2 i}{4}.$$

14 Suppose f, g, h are elements of an inner product space. Prove that

$$\|h - \tfrac{1}{2}(f+g)\|^2 = \frac{\|h-f\|^2 + \|h-v\|^2}{2} - \frac{\|f-g\|^2}{4}.$$

15 Prove that a norm satisfying the parallelogram equality comes from an inner product. In other words, show that if V is a normed vector space whose norm $\|\cdot\|$ satisfies the parallelogram equality, then there is an inner product $\langle\cdot,\cdot\rangle$ on V such that $\|f\| = \langle f, f\rangle^{1/2}$ for all $f \in V$.

16 Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Let V be the subspace of $L^2(\mu)$ defined by

$$V = \{f \in L^2(\mu) : \|f\|_\infty < \infty \text{ and } \mu(\{x \in X : f(x) \neq 0\}) < \infty\}.$$

For $f, g \in V$, define $\langle f, g\rangle$ by

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

The integral above makes sense without the use of Hölder's Inequality because of the definition of V.

(a) Show that the Cauchy–Schwarz Inequality implies that

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2$$

for all $f, g \in V$ (again, without using Hölder's Inequality).

(b) Now suppose $f, g \in L^2(\mu)$. Let $f_1, f_2, \dots$ and $g_1, g_2, \dots$ be the increasing sequences of simple functions that approximate $|f|$ and $|g|$ as constructed in 2.82. Show that each f_k and each g_k is in V. Then apply part (a) to f_k and g_k and use an appropriate limit theorem to conclude (without using Hölder's Inequality) that $\|fg\|_1 \leq \|f\|_2 \|g\|_2$.

17 Suppose $V_1, \dots, V_m$ are inner product spaces. Show that the equation

$$\langle (f_1, \dots, f_m), (g_1, \dots, g_m) \rangle = \langle f_1, g_1 \rangle + \cdots + \langle f_m, g_m \rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.
[*Each of the inner product spaces $V_1, \dots, V_m$ may have a different inner product, even though the same notation is used for the inner product on all these inner product spaces.*]

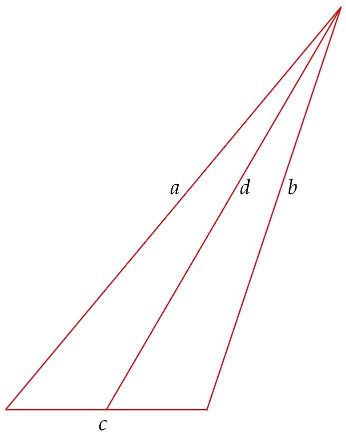
18 Suppose V is an inner product space. Make $V \times V$ an inner product space as in the exercise above. Prove that the function that takes an ordered pair $(f, g) \in V \times V$ to the inner product $\langle f, g \rangle \in \mathbf{F}$ is a continuous function from $V \times V$ to $\mathbf{F}$.

19 Suppose $p \in [1, \infty]$. Show that the usual norm on ℓ^p comes from an inner product if and only if $p = 2$.

20 Suppose $p \in [1, \infty]$. Show that the usual norm on $L^p(\mathbf{R})$ comes from an inner product if and only if $p = 2$.

21 Use inner products to prove Apollonius's Identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$



8B Orthogonality

Orthogonal Projections

The previous section developed inner product spaces following the pattern used in the author's textbook *Linear Algebra Done Right* and other linear algebra books. Linear algebra focuses mainly on finite-dimensional vector spaces. Many interesting results about infinite-dimensional inner product spaces require an additional hypothesis, which we now introduce.

8.23 Definition ***Hilbert space***

A *Hilbert space* is an inner product space that is a Banach space with the norm determined by the inner product.

8.24 Example *Hilbert spaces*

- Suppose μ is a measure. Then $L^2(\mu)$ with its usual inner product is a Hilbert space (by 7.23).
- As a special case of the first bullet point, if $n \in \mathbf{Z}^+$ then taking μ to be counting measure on $\{1, \ldots, n\}$ shows that $\mathbf{F}^n$ with its usual inner product is a Hilbert space.
- As another special case of the first bullet point, taking μ to be counting measure on $\mathbf{Z}^+$ shows that ℓ^2 with its usual inner product is a Hilbert space.
- Every closed subspace of a Hilbert space is a Hilbert space (by Exercise 13 in Section 6B).

8.25 Example *not Hilbert spaces*

- The inner product space ℓ^1, where $\langle (a_1, a_2, \ldots), (b_1, b_2, \ldots) \rangle = \sum_{k=1}^{\infty} a_k \overline{b_k}$, is not a Hilbert space.
- The inner product space $C([0,1])$, where $\langle f, g \rangle = \int_0^1 f\overline{g}$, is not a Hilbert space.

The next definition makes sense in the context of normed vector spaces.

8.26 Definition ***distance from a point to a set***

Suppose U is a nonempty subset of a normed vector space V and $f \in V$. The *distance* from f to U, denoted $\text{distance}(f, U)$, is defined by

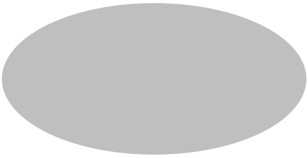
$$\text{distance}(f, U) = \inf\{\|f - g\| : g \in U\}.$$

Notice that $\text{distance}(f, U) = 0$ if and only if $f \in \overline{U}$.

8.27 Definition *convex set*

- A subset of a vector space is called *convex* if the subset contains the line segment connecting each pair of points in it.
- More precisely, suppose V is a vector space and $U \subset V$. Then U is called *convex* if

$$tf + (1-t)g \in U \text{ for all } t \in [0,1] \text{ and all } f, g \in U.$$



Convex subset of $\mathbf{R}^2$.

Nonconvex subset of $\mathbf{R}^2$.

8.28 Example *convex sets*

- Every subspace of a vector space is convex, as you should verify.
- If V is a normed vector space, $f \in V$, and $r > 0$, then the open ball centered at f with radius r is convex, as you should verify.

The next example shows that the distance from an element of a Banach space to a closed subspace is not necessarily attained by some element of the closed subspace. After this example, we will prove that this behavior cannot happen in a Hilbert space.

8.29 Example *no closest element to a closed subspace of a Banach space*

In the Banch space $C([0,1])$ with norm $\|g\| = \sup_{x\in[0,1]}|g(x)|$, let

$$U = \{g \in C([0,1]) : \int_0^1 g = 0 \text{ and } g(1) = 0\}.$$

Then U is a closed subspace of $C([0,1])$.

Let $f \in C([0,1])$ be defined by $f(x) = 1 - x$. For $k \in \mathbf{Z}^+$, let

$$g_k(x) = \frac{1}{2} - x + \frac{x^k}{2} + \frac{x-1}{k+1}.$$

Then $g_k \in U$ and $\lim_{k\to\infty}\|f - g_k\| = \frac{1}{2}$, which implies that distance$(f, U) \leq \frac{1}{2}$.

If $g \in U$, then $\int_0^1 (f - g) = \frac{1}{2}$ and $(f - g)(1) = 0$. These conditions imply that $\|f - g\| > \frac{1}{2}$.

Thus distance$(f, U) = \frac{1}{2}$ but there does not exist $g \in U$ such that $\|f - g\| = \frac{1}{2}$.

In the next result, we use for the first time the hypothesis that V is a Hilbert space. This result is not true for inner product spaces (see Exercise 11).

8.30 *Distance to a closed convex set is attained in a Hilbert space*

- The distance from an element of a Hilbert space to a nonempty closed convex set is attained by a unique element of the nonempty closed convex set.
- More specifically, suppose V is a Hilbert space, $f \in V$, and U is a nonempty closed convex subset of V. Then there exists a unique $h \in U$ such that

$$\|f - h\| = \text{distance}(f, U).$$

Proof First we prove the existence of an element of U that attains the distance to f. To do this, suppose $g_1, g_2, \dots$ is a sequence of elements of U such that

8.31
$$\lim_{k\to\infty} \|f - g_k\| = \text{distance}(f, U).$$

Then for $j, k \in \mathbf{Z}^+$ we have

$$\begin{aligned}\|g_j - g_k\|^2 &= \|(f - g_k) - (f - g_j)\|^2 \\ &= 2\|f - g_k\|^2 + 2\|f - g_j\|^2 - \|2f - (g_k + g_j)\|^2 \\ &= 2\|f - g_k\|^2 + 2\|f - g_j\|^2 - 4\Big\|f - \frac{g_k + g_j}{2}\Big\|^2\end{aligned}$$

8.32
$$\le 2\|f - g_k\|^2 + 2\|f - g_j\|^2 - 4\big(\text{distance}(f, U)\big)^2,$$

where the second equality comes from the Parallelogram Equality (8.22) and the last line holds because the convexity of U implies that $(g_k + g_j)/2 \in U$. Now the inequality above and 8.31 imply that $g_1, g_2, \dots$ is a Cauchy sequence. Thus there exists $h \in V$ such that

8.33
$$\lim_{k\to\infty} \|g_k - h\| = 0.$$

Because U is a closed subset of V and each $g_k \in U$, we know that $h \in U$. Now 8.31 and 8.33 imply that

$$\|f - h\| = \text{distance}(f, U),$$

which completes the existence proof of the existence part of this result.

To prove the uniqueness part of this result, suppose h and $\tilde{h}$ are elements of U such that

8.34
$$\|f - h\| = \|f - \tilde{h}\| = \text{distance}(f, U).$$

Then

$$\|h - \tilde{h}\|^2 \le 2\|f - h\|^2 + 2\|f - \tilde{h}\|^2 - 4\big(\text{distance}(f, U)\big)^2$$

8.35
$$= 0,$$

where the first line above follows from 8.32 (with g_m replaced by h and g_n replaced by $\tilde{h}$) and the last line above follows from 8.34. Now 8.35 implies that $h = \tilde{h}$, completing the proof of uniqueness. ■

Example 8.29 showed that the existence part of the previous result can fail in a Banach space. Exercise 12 shows that the uniqueness part can also fail in a Banach space. These observations highlight the advantanges of working in a Hilbert space.

8.36 Definition ***orthogonal projection;*** P_U

Suppose U is a nonempty closed convex subset of a Hilbert space V. The *orthogonal projection* of V onto U is the function $P_U \colon V \to V$ defined by setting $P_U(f)$ equal to the unique element of U that is closest to f.

The definition above makes sense because of 8.30. We will often use the notation $P_U f$ instead of $P_U(f)$. To test your understanding of the definition above, make sure that you can show that if U is a nonempty closed convex subset of a Hilbert space V, then

- $P_U f = f$ if and only if $f \in U$;
- $P_U \circ P_U = P_U$.

8.37 Example *orthogonal projection onto closed unit ball*

Suppose U is the closed unit ball $\{g \in V : \|g\| \leq 1\}$ in a Hilbert space V. Then

$$P_U f = \begin{cases} f & \text{if } \|f\| \leq 1, \\ \dfrac{f}{\|f\|} & \text{if } \|f\| > 1, \end{cases}$$

as you should verify.

8.38 Example *orthogonal projection onto a closed subspace*

Suppose U is the closed subspace of ℓ^2 consisting of the elements of ℓ^2 whose even coordinates all equal 0:

$$U = \Big\{(a_1, 0, a_3, 0, a_5, 0, \dots) : \text{each } a_k \in \mathbf{F} \text{ and } \sum_{k=1}^{\infty} |a_{2k-1}|^2 < \infty\Big\}.$$

Then for $b = (b_1, b_2, b_3, b_4, b_5, b_6, \dots) \in \ell^2$, we have

$$P_U b = (b_1, 0, b_3, 0, b_5, 0, \dots),$$

as you should verify.

Note that in this example the function P_U is a linear map from ℓ^2 to ℓ^2 (unlike the behavior in Example 8.37).

Also, notice that $b - P_U b = (0, b_2, 0, b_4, 0, b_6, \dots)$ and thus $b - P_U b$ is orthogonal to every element of U.

The next result shows that the properties stated in the last two paragraphs of the example above hold whenever U is a closed subspace of a Hilbert space.

8.39 *Orthogonal projection onto a closed subspace*

Suppose U is a closed subspace of a Hilbert space V and $f \in V$. Then

(a) $f - P_U f$ is orthogonal to g for every $g \in U$;

(b) if $h \in U$ and $f - h$ is orthogonal to g for every $g \in U$, then $h = P_U f$;

(c) $P_U \colon V \to V$ is a linear map;

(d) $\|P_U f\| \le \|f\|$, with equality if and only if $f \in U$.

Proof The figure below illustrates (a). To prove (a), suppose $g \in U$. Then for all $\alpha \in \mathbf{F}$ we have

$$\begin{aligned}\|f - P_U f\|^2 &\le \|f - P_U f + \alpha g\|^2\\ &= \langle f - P_U f + \alpha g, f - P_U f + \alpha g\rangle\\ &= \|f - P_U f\|^2 + |\alpha|^2\|g\|^2 + 2\,\mathrm{Re}\,\overline{\alpha}\langle f - P_U f, g\rangle.\end{aligned}$$

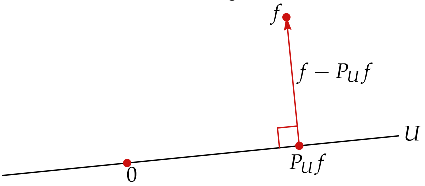
Let $\alpha = -t\langle f - P_U f, g\rangle$ for $t > 0$. A tiny bit of algebra applied to the inequality above implies

$$2|\langle f - P_U f, g\rangle|^2 \le t|\langle f - P_U f, g\rangle|^2\|g\|^2$$

for all $t > 0$. Thus $\langle f - P_U f, g\rangle = 0$, completing the proof of (a).

To prove (b), suppose $h \in U$ and $f - h$ is orthogonal to g for every $g \in U$. If $g \in U$, then $h - g \in U$ and hence $f - h$ is orthogonal to $h - g$. Thus

$$\begin{aligned}\|f - h\|^2 &\le \|f - h\|^2 + \|h - g\|^2\\ &= \|(f - h) + (h - g)\|^2\\ &= \|f - g\|^2,\end{aligned}$$



$f - P_U f$ is orthogonal to each element of U.

where the first equality above follows from the Pythagorean Theorem (8.11). Thus

$$\|f - h\| \le \|f - g\|$$

for all $g \in U$. Hence h is the element of U that minimizes the distance to f, which implies that $h = P_U f$, completing the proof of (b).

To prove (c), suppose $f_1, f_2 \in V$. If $g \in U$, then (a) implies that $\langle f_1 - P_U f_1, g\rangle = \langle f_2 - P_U f_2, g\rangle = 0$, and thus

$$\langle (f_1 + f_2) - (P_U f_1 + P_U f_2), g\rangle = 0.$$

The equation above and (b) now imply that

$$P_U(f_1 + f_2) = P_U f_1 + P_U f_2.$$

The equation above and the equation $P_U(\alpha f) = \alpha P_U f$ for $\alpha \in \mathbf{F}$ (whose proof is left to the reader) show that P_U is a linear map, proving (c).

The proof of (d) is left as an exercise for the reader. ∎

Orthogonal Complements

8.40 Definition *orthogonal complement;* $U^\perp$

Suppose U is a subset of an inner product space V. The *orthogonal complement* of U is denoted by $U^\perp$ and is defined by

$$U^\perp = \{h \in V : \langle g, h\rangle = 0 \text{ for all } g \in U\}.$$

In other words, the orthogonal complement of a subset U of an inner product space V is the set of elements of V that are orthogonal to every element of U.

8.41 Example *orthogonal complement*

Suppose U is the set of elements of ℓ^2 whose even coordinates all equal 0:

$$U = \{(a_1, 0, a_3, 0, a_5, 0, \dots) : \text{each } a_k \in \mathbf{F} \text{ and } \sum_{k=1}^{\infty} |a_{2k-1}|^2 < \infty\}.$$

Then $U^\perp$ equals the set of elements of ℓ^2 whose odd coordinates all equal 0:

$$U^\perp = \{0, a_2, 0, a_4, 0, a_6, \dots) : \text{each } a_k \in \mathbf{F} \text{ and } \sum_{k=1}^{\infty} |a_{2k}|^2 < \infty\},$$

as you should verify.

8.42 ***Properties of orthogonal complement***

Suppose U is a subset of an inner product space V. Then

(a) $U^\perp$ is a closed subspace of V;

(b) $U \cap U^\perp \subset \{0\}$;

(c) if $W \subset U$, then $U^\perp \subset W^\perp$;

(d) $(\overline{U})^\perp = U^\perp$;

(e) $U \subset (U^\perp)^\perp$.

Proof To prove (a), suppose $h_1, h_2, \dots$ is a sequence in $U^\perp$ that converges to some $h \in V$. If $g \in U$, then $\langle g, h\rangle = \lim_{n\to\infty}\langle g, h_n\rangle = 0$, and thus $h \in U^\perp$. Thus $U^\perp$ is a closed subset of V. The proof of (a) is completed by showing that $U^\perp$ is a subspace of V, which is left to the reader.

To prove (b), suppose $g \in U \cap U^\perp$. Then $\langle g, g\rangle = 0$, which implies that $g = 0$, proving (b).

To prove (e), suppose $g \in U$. Thus $\langle g, h\rangle = 0$ for all $h \in U^\perp$, which implies that $g \in (U^\perp)^\perp$. Hence $U \subset (U^\perp)^\perp$, proving (e).

The proofs of (c) and (d) are left to the reader. ■

The results in the rest of this subsection have as a hypothesis that V is a Hilbert space. These results do not hold when V is only an inner product space.

8.43 *Orthogonal complement of the orthogonal complement*

Suppose U is a subspace of a Hilbert space V. Then

$$\overline{U} = (U^{\perp})^{\perp}.$$

Proof Applying 8.42(a) to $U^{\perp}$, we see that $(U^{\perp})^{\perp}$ is a closed subspace of V. Now taking closures of both sides of the inclusion $U \subset (U^{\perp})^{\perp}$ [8.42(e)] shows that $\overline{U} \subset (U^{\perp})^{\perp}$.

To prove the inclusion in the other direction, suppose $f \in (U^{\perp})^{\perp}$. Because $f \in (U^{\perp})^{\perp}$ and $P_{\overline{U}} f \in \overline{U} \subset (U^{\perp})^{\perp}$ (by the previous paragraph), we see that

$$f - P_{\overline{U}} f \in (U^{\perp})^{\perp}.$$

Also,

$$f - P_{\overline{U}} f \in U^{\perp}$$

by 8.39(a). Hence $f - P_{\overline{U}} f \in U^{\perp} \cap (U^{\perp})^{\perp}$. Now 8.42(b) (applied to $U^{\perp}$ in place of U) implies that $f - P_{\overline{U}} f = 0$, which implies that $f \in \overline{U}$. Thus $(U^{\perp})^{\perp} \subset \overline{U}$, completing the proof. ∎

As a special case, the result above implies that if U is a closed subspace of a Hilbert space V, then $U = (U^{\perp})^{\perp}$.

Another special case of the result above is sufficiently useful to deserve stating separately, as we do in the next result.

8.44 *Necessary and sufficient condition for a subspace to be dense*

Suppose U is a subspace of a Hilbert space V. Then

$$\overline{U} = V \text{ if and only if } U^{\perp} = \{0\}.$$

Proof First suppose $\overline{U} = V$. Then using 8.42(d), we have

$$U^{\perp} = \overline{U}^{\perp} = V^{\perp} = \{0\}.$$

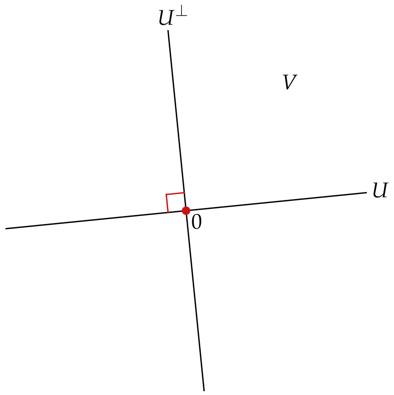
To prove the other direction, now suppose $U^{\perp} = \{0\}$. Then 8.43 implies that

$$\overline{U} = (U^{\perp})^{\perp} = \{0\}^{\perp} = V,$$

completing the proof. ∎

The next result states that if U is a closed subspace of a Hilbert space V, then V is the *direct sum* of U and $U^\perp$, often written $V = U \oplus U^\perp$, although we do not need to use this terminology or notation further.

The key point to keep in mind is that the next result shows that the picture here represents what happens in general for a closed subspace U of a Hilbert space V: every element of V can be uniquely written as an element of U plus an element of $U^\perp$.



8.45 *Orthogonal decomposition*

Suppose U is a closed subspace of a Hilbert space V. Then every element $f \in V$ can be uniquely written in the form

$$f = g + h,$$

where $g \in U$ and $h \in U^\perp$. Furthermore, $g = P_U f$ and $h = f - P_U f$.

Proof Suppose $f \in V$. Then

$$f = P_U f + (f - P_U f),$$

where $P_U f \in U$ [by definition of $P_U f$ as the element of U that is closest to f] and $f - P_U f \in U^\perp$ [by 8.39(a)]. Thus we have the desired decomposition of f as the sum of an element of U and an element of $U^\perp$.

To prove the uniqueness of this decomposition, suppose

$$f = g_1 + h_1 = g_2 + h_2,$$

where $g_1, g_2 \in U$ and $h_1, h_2 \in U^\perp$. Then $g_1 - g_2 = h_2 - h_1 \in U \cap U^\perp$, which implies that $g_1 = g_2$ and $h_1 = h_2$, as desired. ∎

In the next definition, the function I depends upon the vector space V. Thus a notation such as I_V might be more precise. However, the domain of I should always be clear from the context.

8.46 Definition ***identity map; I***

Suppose V is a vector space. The *identity map* I is the linear map from V to V defined by $If = f$ for $f \in V$.

The next result highlights the close relationship between orthogonal projections and orthogonal complements.

8.47 *range and null space of orthogonal projections*

Suppose U is a closed subspace of a Hilbert space V. Then

(a) $\operatorname{range} P_U = U$ and $\operatorname{null} P_U = U^\perp$;

(b) $\operatorname{range} P_{U^\perp} = U^\perp$ and $\operatorname{null} P_{U^\perp} = U$;

(c) $P_{U^\perp} = I - P_U$.

Proof The definition of $P_U f$ as the closest point in U to f implies $\operatorname{range} P_U \subset U$. Because $P_U g = g$ for all $g \in U$, we also have $U \subset \operatorname{range} P_U$. Thus $\operatorname{range} P_U = U$.

If $f \in \operatorname{null} P_U$, then $f \in U^\perp$ [by 8.39(a)]. Thus $\operatorname{null} P_U \subset U^\perp$. Conversely, if $f \in U^\perp$, then 8.39(b) (with $h = 0$) implies that $P_U f = 0$; hence $U^\perp \subset \operatorname{null} P_U$. Thus $\operatorname{null} P_U = U^\perp$, completing the proof of (a).

Replace U by $U^\perp$ in (a), getting $\operatorname{range} P_{U^\perp} = U^\perp$ and $\operatorname{null} P_{U^\perp} = (U^\perp)^\perp = U$ (where the last equality comes from 8.43), completing the proof of (b).

Finally, if $f \in U$, then

$$P_{U^\perp} f = 0 = f - P_U f = (I - P_U) f,$$

where the first equality above holds because $\operatorname{null} P_{U^\perp} = U$ [by (b)].

If $f \in U^\perp$, then

$$P_{U^\perp} f = f = f - P_U f = (I - P_U) f,$$

where the second equality above holds because $\operatorname{null} P_U = U^\perp$ [by (a)].

The last two displayed equations show that $P_{U^\perp}$ and $I - P_U$ agree on U and agree on $U^\perp$. Because $P_{U^\perp}$ and $I - P_U$ are both linear maps and because each element of V equals some element of U plus some element of $U^\perp$ (by 8.45), this implies that $P_{U^\perp} = I - P_U$, completing the proof of (c). ∎

8.48 Example $P_{U^\perp} = I - P_U$

Suppose U is the closed subspace of $L^2(\mathbf{R})$ defined by

$$U = \{f \in L^2(\mathbf{R}) : f(x) = 0 \text{ for almost all } x < 0\}.$$

Then, as you should verify,

$$U^\perp = \{f \in L^2(\mathbf{R}) : f(x) = 0 \text{ for almost all } x \geq 0\}.$$

Furthermore, you should also verify that if $f \in L^2(\mathbf{R})$, then

$$P_U f = f\chi_{[0,\infty)} \quad \text{and} \quad P_{U^\perp} f = f\chi_{(-\infty,0)}.$$

Thus $P_{U^\perp} f = f(1 - \chi_{[0,\infty)}) = (I - P_U) f$ and hence $P_{U^\perp} = I - P_U$, as asserted in 8.47(c).

Riesz Representation Theorem

Suppose h is an element of a Hilbert space V. Define $\varphi\colon V \to \mathbf{F}$ by $\varphi(f) = \langle f, h \rangle$ for $f \in V$. The properties of an inner product imply that φ is a linear functional. The Cauchy–Schwarz Inequality (8.13) implies that $|\varphi(f)| \leq \|f\|\,\|h\|$ for all $f \in V$, which implies that φ is a bounded linear functional on V. The next result states that every bounded linear functional on V arises in this fashion.

To motivate the proof of the next result, note that if φ is as in the paragraph above, then $\operatorname{null}\varphi = \{h\}^{\perp}$. Thus $h \in (\operatorname{null}\varphi)^{\perp}$ [by 8.42(e)]. Hence in the proof of the next result, to find h we start with an element of $(\operatorname{null}\varphi)^{\perp}$ and then multiply it by a scalar to make everything come out right.

8.49 *Riesz Representation Theorem*

Suppose φ is a bounded linear functional on a Hilbert space V. Then there exists a unique $h \in V$ such that

$$\varphi(f) = \langle f, h \rangle$$

for all $f \in V$. Furthermore, $\|\varphi\| = \|h\|$.

Proof If $\varphi = 0$, take $h = 0$. Thus we can assume $\varphi \neq 0$. Hence $\operatorname{null}\varphi$ is a closed subspace of V that is not equal to V (see 6.45). The subspace $(\operatorname{null}\varphi)^{\perp}$ is not equal to $\{0\}$ (by 8.44). Thus there exists $g \in (\operatorname{null}\varphi)^{\perp}$ with $\|g\| = 1$. Let

$$h = \overline{\varphi(g)}g.$$

Then

$$\varphi(h) = |\varphi(g)|^2 = \|h\|^2.$$

Now suppose $f \in V$. Then

$$\langle f, h \rangle = \left\langle f - \frac{\varphi(f)}{\|h\|^2}h, h \right\rangle + \left\langle \frac{\varphi(f)}{\|h\|^2}h, h \right\rangle$$

8.50 $$= \left\langle \frac{\varphi(f)}{\|h\|^2}h, h \right\rangle$$

$$= \varphi(f),$$

where 8.50 holds because $f - \frac{\varphi(f)}{\|h\|^2}h \in \operatorname{null}\varphi$ and h is orthogonal to all elements of $\operatorname{null}\varphi$.

We have now proved the existence of $h \in V$ such that $\varphi(f) = \langle f, h \rangle$ for all $f \in V$. To prove uniqueness, suppose $\tilde{h} \in V$ has the same property. Then

$$\langle h - \tilde{h}, h - \tilde{h} \rangle = \langle h - \tilde{h}, h \rangle - \langle h - \tilde{h}, \tilde{h} \rangle = \varphi(h - \tilde{h}) - \varphi(h - \tilde{h}) = 0,$$

which implies that $h = \tilde{h}$, which proves uniqueness.

The Cauchy–Schwarz Inequality implies that $|\varphi(f)| = |\langle f, h \rangle| \leq \|f\|\,\|h\|$ for all $f \in V$, which implies that $\|\varphi\| \leq \|h\|$. Because $\varphi(h) = \langle h, h \rangle = \|h\|^2$, we also have $\|\varphi\| \geq \|h\|$. Thus $\|\varphi\| = \|h\|$, completing the proof. ■

Hungarian mathematician Frigyes Riesz (1880–1956) proved 8.49 in 1907.

Suppose that μ is a measure and $p \in (1, \infty]$. In 7.25 we considered the natural map of $L^{p'}(\mu)$ into $\big(L^p(\mu)\big)'$, and we showed that this maps preserves norms. In the special case where $p = p' = 2$, the Riesz Representation Theorem (8.49) shows that this map is surjective. In other words, if φ is a bounded linear functional on $L^2(\mu)$, then there exists $h \in L^2(\mu)$ such that

$$\varphi(f) = \int fh \, d\mu$$

for all $f \in L^2(\mu)$ (take h to be the complex conjugate of the function given by 8.49). Hence we can identify the dual of $L^2(\mu)$ with $L^2(\mu)$. In 9.19 we will deal with other values of p.

EXERCISES 8B

1 Show that each of the inner product spaces in Example 8.25 is not a Hilbert space.

2 Prove or disprove: The inner product space in Exercise 1 in Section 8A is a Hilbert space.

3 Suppose $V_1, V_2, \ldots$ are Hilbert spaces. Let

$$V = \Big\{(f_1, f_2, \ldots) \in V_1 \times V_2 \times \cdots : \sum_{k=1}^{\infty} \|f_k\|^2 < \infty.\Big)$$

Show that the equation

$$\langle (f_1, f_2, \ldots), (g_1, g_2, \ldots) \rangle = \sum_{k=1}^{\infty} \langle f_k, g_k \rangle$$

defines an inner product on V that makes V a Hilbert space.
[*Each of the Hilbert spaces $V_1, V_2, \ldots$ may have a different inner product, even though the same notation is used for the norm and inner product on all these Hilbert spaces.*]

4 Prove that if V is a normed vector space, $f \in V$, and $r > 0$, then the open ball $B(f, r)$ centered at f with radius r is convex.

5 Suppose V is a normed vector space and U is a closed subset of V. Prove that U is convex if and only if

$$\frac{f+g}{2} \in U \text{ for all } f, g \in U.$$

6 Prove that if U is a convex subset of a normed vector space, then $\overline{U}$ is also convex.

7 Prove that if U is a convex subset of a normed vector space, then the interior of U is also convex.
[*The interior of U is the set* $\{f \in U : B(f,r) \subset U$ *for some* $r > 0\}$.]

8 Suppose V is a Hilbert space, U is a nonempty closed convex subset of V, and $h \in U$ is the unique element of U with smallest norm (obtained by taking $f = 0$ in 8.30). Prove that

$$\operatorname{Re}\langle g, h\rangle \geq \|h\|^2$$

for all $g \in U$.

9 Suppose V is a Hilbert space. A *closed half-space* of V is a set of the form

$$\{g \in V : \operatorname{Re}\langle g, h\rangle \geq c\}$$

for some $h \in V$ and some $c \in \mathbf{R}$. Prove that every closed convex subset of V is the intersection of all the closed half-spaces that contain it.

10 Give an example of a nonempty closed subset U of the Hilbert space ℓ^2 and $a \in \ell^2$ such that there does not exist $b \in U$ with $\|a - b\| = \operatorname{distance}(a, U)$.
[*By 8.30, U cannot be a convex subset of ℓ^2.*]

11 Give an example of an inner product space V, a closed subspace U of V, and $f \in V$ such that there does not exist $h \in U$ with $\|f - h\| = \operatorname{distance}(f, U)$.
[*By 8.30, V cannot be a Hilbert space.*]

12 In the real Banach space $\mathbf{R}^2$ with norm defined by $\|(x,y)\|_\infty = \max\{|x|,|y|\}$, give an example of a closed convex set $U \subset \mathbf{R}^2$ and $z \in \mathbf{R}^2$ such that there exist infinitly many choices of $w \in U$ with $\|z - w\| = \operatorname{distance}(z, U)$.

13 Suppose f and g are elements of an inner product space. Prove that $\langle f, g\rangle = 0$ if and only if

$$\|f\| \leq \|f + \alpha g\|$$

for all $\alpha \in \mathbf{F}$.

14 Suppose U is a closed subspace of a Hilbert space V and $T\colon V \to V$ is linear. Prove that $Tf \in U$ for every $f \in U$ if and only if $TP_U = P_U T P_U$.
[*If $Tf \in U$ for every $f \in U$, then U is called an* invariant subspace *for T.*]

15 Suppose U is a closed subspace of a Hilbert space V and $f \in V$. Prove that $\|P_U f\| \leq \|f\|$, with equality if and only if $f \in U$.
[*This exercise asks you to prove 8.39(d).*]

16 Suppose U is a closed subspace of a Hilbert space V. Suppose also that W is a normed vector space and $S\colon U \to W$ is a bounded linear map. Prove that there exists a bounded linear map $T\colon V \to W$ such that $T|_U = S$ and $\|T\| = \|S\|$.
[*The result in this exercise is not true if the hypothesis that V is a Hilbert space is replaced by the hypothesis that V is a Banach space.*]

17 Give an example of an inner product space V and a closed subspace U of V such that $(U^\perp)^\perp \neq U$.

18 Suppose U and W are subspaces of a Hilbert space V. Prove that $\overline{U} = \overline{W}$ if and only if $U^{\perp} = W^{\perp}$.

19 Suppose U and W are closed subspaces of a Hilbert space. Prove that $P_U P_W = 0$ if and only if $\langle f, g \rangle = 0$ for all $f \in U$ and all $g \in W$.

20 Verify the assertions in Example 8.48.

21 Show that every inner product space is a subspace of some Hilbert space.

Hint: See Exercise 21 in Section 6B.

22 Give an example of an inner product space V and a bounded linear functional φ on V such that there does not exist $h \in V$ with $\varphi(f) = \langle f, h \rangle$ for all $f \in V$.

23 (a) Give an example of a Banach space V and a bounded linear functional φ on V such that

$$|\varphi(f)| < \|\varphi\| \, \|f\|$$

for all $f \in V \setminus \{0\}$.

(b) Show that there does not exist an example as in part (a) where V is a Hilbert space.

24 (a) Suppose φ and ψ are bounded linear functionals on a Hilbert space V such that $\|\varphi + \psi\| = \|\varphi\| + \|\psi\|$. Prove that one of φ, ψ is a multiple of the other.

(b) Give an example to show that the result in part (a) fails if the hypothesis that V is a Hilbert space is replaced by the hypothesis that V is a Banach space.

25 Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Show that each bounded linear functional on $L^2(\mu)$ lives on a set of σ-finite measure. More precisely, show that if φ is a bounded linear functional on $L^2(\mu)$, then there exists a sequence $E_1, E_2, \ldots$ in $\mathcal{S}$ such that $\mu(E_k) < \infty$ for each $k \in \mathbf{Z}^+$ and

$$\varphi(f) = \varphi\big(f \chi_{\bigcup_{k=1}^{\infty} E_k}\big).$$

for all $f \in L^2(\mu)$.

26 Give an example of an inner product space V and a convex subset U of V such that $\overline{U} = V$ and such that the complement $V \setminus U$ is also a convex subset of V with $\overline{V \setminus U} = V$.
[*This exercise should stretch your geometric intuition because this behavior cannot happen in finite dimensions.*]

8C Orthonormal Bases

Bessel's Inequality

To develop the theory of orthonormal bases in Hilbert spaces, we need to generalize the notion of a sequence. A sequence is simply a function whose domain is $\mathbf{Z}^+$, with the value of this function at $n \in \mathbf{Z}^+$ denoted with a subscript n instead of the usual functional notation. To generalize, we allow $\mathbf{Z}^+$ to be replaced by an arbitrary set.

8.51 Definition *family*

A *family* $\{e_k\}_{k \in K}$ in a set V is a function e from a set K to V, with the value of the function e at $k \in K$ denoted by e_k.

Even though a family in V is a function mapping into V and thus is not a subset of V, the set terminology and the bracket notation $\{e_k\}_{k \in K}$ are useful (note that the range of a family in V really is a subset of V).

8.52 Definition *orthonormal family*

A family $\{e_k\}_{k \in K}$ in an inner product space is called an *orthonormal family* if

$$\langle e_j, e_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \end{cases}$$

for all $j, k \in K$.

In other words, a family $\{e_k\}_{k \in K}$ is an orthonormal family if e_j and e_k are orthogonal for all distinct $j, k \in K$ and $\|e_k\| = 1$ for all $k \in K$.

8.53 Example *orthonormal families*

- For $k \in \mathbf{Z}^+$, let e_k be the element of ℓ^2 all of whose coordinates are 0 except for the k^{th} coordinate, which is 1:

  $$e_k = (0, \ldots, 0, 1, 0, \ldots).$$

  Then $\{e_k\}_{k \in \mathbf{Z}^+}$ is an orthonormal family in ℓ^2. In this case, our family is a sequence; thus we can call $\{e_k\}_{k \in \mathbf{Z}^+}$ an *orthonormal sequence*.

- More generally, suppose K is a nonempty set. The Hilbert space $L^2(\mu)$, where μ is counting measure on K, is often denoted by $\ell^2(K)$. For $k \in K$, define a function $e_k \colon K \to \mathbf{F}$ by

  $$e_k(j) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

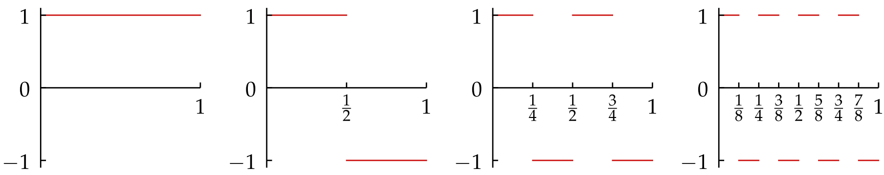
  Then $\{e_k\}_{k \in K}$ is an orthonormal family in $\ell^2(K)$.

- For k a nonnegative integer, define $e_k \colon [0,1) \to \mathbf{F}$ by

$$e_k(x) = \begin{cases} 1 & \text{if } x \in \left[\frac{n-1}{2^k}, \frac{n}{2^k}\right) \text{ for some odd integer } n, \\ -1 & \text{if } x \in \left[\frac{n-1}{2^k}, \frac{n}{2^k}\right) \text{ for some even integer } n. \end{cases}$$

This orthonormal family was invented by Hungarian mathematician Alfréd Haar (1885–1933), who was a student of Hilbert.

The figure below shows the graphs of e_0, e_1, e_2, and e_3. The pattern of these graphs should convince you that $\{e_k\}_{k\in\{0,1,\dots\}}$ is an orthonormal family in $L^2([0,1))$.



The graph of e_0. *The graph of e_1.* *The graph of e_2.* *The graph of e_3.*

- Now we modify the example in the previous bullet point by translating the functions in the previous bullet point by arbitrary integers. Specifically, for k a nonnegative integer and $m \in \mathbf{Z}$, define $e_{k,m} \colon \mathbf{R} \to \mathbf{F}$ by

$$e_{k,m}(x) = \begin{cases} 1 & \text{if } x \in \left[m + \frac{n-1}{2^k}, m + \frac{n}{2^k}\right) \text{ for some odd integer } n \in [1, 2^k], \\ -1 & \text{if } x \in \left[m + \frac{n-1}{2^k}, m + \frac{n}{2^k}\right) \text{ for some even integer } n \in [1, 2^k], \\ 0 & \text{if } x \notin [m, m+1). \end{cases}$$

Then $\{e_{k,m}\}_{(k,m)\in\{0,1,\dots\}\times\mathbf{Z}}$ is an orthonormal family in $L^2(\mathbf{R})$.

This example illustrates the usefulness of considering families that are not sequences. Although $\{0,1,\dots\} \times \mathbf{Z}$ is a countable set and hence we could rewrite $\{e_{k,m}\}_{(k,m)\in\{0,1,\dots\}\times\mathbf{Z}}$ as a sequence, doing so would be awkward and would be less clean than the $e_{k,m}$ notation.

- For $k \in \mathbf{Z}$, define $e_k \colon [0, 2\pi) \to \mathbf{F}$ by

$$e_k(\theta) = \begin{cases} \frac{1}{\sqrt{\pi}} \sin(k\theta) & \text{if } k > 0, \\ \frac{1}{\sqrt{2\pi}} & \text{if } k = 0, \\ \frac{1}{\sqrt{\pi}} \cos(k\theta) & \text{if } k < 0. \end{cases}$$

Then $\{e_k\}_{k\in\mathbf{Z}}$ is an orthonormal family in $L^2([0, 2\pi))$, as you should verify (see Exercise 1 for useful formulas that will help in this verification).

This orthonormal family $\{e_k\}_{k\in\mathbf{Z}}$ leads to the classical theory of Fourier series, as we will see in more depth in Chapter 11.

The next result gives our first indication of why orthonormal families are so useful.

8.54 ***Finite orthonormal families***

Suppose J is a finite set and $\{e_j\}_{j \in J}$ is an orthonormal family in an inner product space. Then

$$\Big\| \sum_{j \in J} \alpha_j e_j \Big\|^2 = \sum_{j \in J} |\alpha_j|^2$$

for every family $\{\alpha_j\}_{j \in J}$ in $\mathbf{F}$.

Proof Suppose $\{\alpha_j\}_{j \in J}$ is a family in $\mathbf{F}$. Standard properties of inner products show that

$$\begin{aligned} \Big\| \sum_{j \in J} \alpha_j e_j \Big\|^2 &= \Big\langle \sum_{j \in J} \alpha_j e_j, \sum_{k \in J} \alpha_k e_k \Big\rangle \\ &= \sum_{j,k \in J} \alpha_j \overline{\alpha_k} \langle e_j, e_k \rangle \\ &= \sum_{j \in J} |\alpha_j|^2, \end{aligned}$$

as desired. ■

Suppose J is a finite set and $\{e_j\}_{j \in J}$ is an orthonormal family in an inner product space. The result above implies that if $\sum_{j \in J} \alpha_j e_j = 0$, then $\alpha_j = 0$ for every $j \in J$.

Linear algebra, and algebra more generally, deals with sums of only finite many terms. However, in analysis we often want to sum infinitely many terms. For example, earlier we defined the infinite sum of a sequence $g_1, g_2, \dots$ in a normed vector space to be the limit as $n \to \infty$ of the partial sums $\sum_{k=1}^{n} g_k$ if that limit exists (see 6.23).

The next definition captures a more powerful method of dealing with infinite sums. The sum defined below is called an *unordered sum* because the set K is not assumed to come with any ordering. A finite unordered sum is defined in the obvious way.

8.55 Definition ***unordered sum; $\sum_{k \in K} f_k$***

Suppose $\{f_k\}_{k \in K}$ is a family in a normed vector space V. The *unordered sum* $\sum_{k \in K} f_k$ is said to *converge* if there exists $g \in V$ such that for every $\varepsilon > 0$, there exists a finite subset J of K such that

$$\Big\| g - \sum_{j \in J'} f_j \Big\| < \varepsilon$$

for all finite sets J' with $J \subset J' \subset K$. If this happens, we set $\sum_{k \in K} f_k = g$. If there is no such $g \in V$, then $\sum_{k \in K} f_k$ is left undefined.

Exercises at the end of this section ask you to develop basic properties of unordered sums, including the following:

- Suppose $\{a_k\}_{k \in K}$ is a family in $\mathbf{R}$ and $a_k \geq 0$ for each $k \in K$. Then the unordered sum $\sum_{k \in K} a_k$ converges if and only if
$$\sup\Big\{\sum_{j \in J} a_j : J \text{ is a finite subset of } K\Big\} < \infty.$$
Furthermore, if $\sum_{k \in K} a_k$ converges then it equals the supremum above. If $\sum_{k \in K} a_k$ does not converge, then the supremum above equals ∞ and we write $\sum_{k \in K} a_k = \infty$ (this notation should be used only when $a_k \geq 0$ for each $k \in K$).

- Suppose $\{a_k\}_{k \in K}$ is a family in $\mathbf{R}$. Then the unordered sum $\sum_{k \in K} a_k$ converges if and only if $\sum_{k \in K} |a_k| < \infty$. Thus convergence of an unordered summation in $\mathbf{R}$ is the same as absolute convergence. As we will soon see, the situation in more general Hilbert spaces is quite different.

Now we can extend 8.54 to infinite sums.

8.56 *Linear combinations of an orthonormal family*

Suppose $\{e_k\}_{k \in K}$ is an orthonormal family in a Hilbert space V. Suppose $\{\alpha_k\}_{k \in K}$ is a family in $\mathbf{F}$. Then

(a) the unordered sum $\displaystyle\sum_{k \in K} \alpha_k e_k$ converges $\iff$ $\displaystyle\sum_{k \in K} |\alpha_k|^2 < \infty$.

Furthermore, if $\sum_{k \in K} \alpha_k e_k$ converges, then

(b) $$\Big\| \sum_{k \in K} \alpha_k e_k \Big\|^2 = \sum_{k \in K} |\alpha_k|^2.$$

Proof First suppose that $\sum_{k \in K} \alpha_k e_k$ converges, with $\sum_{k \in K} \alpha_k e_k = g$. Suppose $\varepsilon > 0$. Then there exists a finite set $J \subset K$ such that
$$\Big\| g - \sum_{j \in J'} \alpha_j e_j \Big\| < \varepsilon$$
for all finite sets J' with $J \subset J' \subset K$. If J' is a finite set with $J \subset J' \subset K$, then the inequality above implies that
$$\|g\| - \varepsilon < \Big\| \sum_{j \in J'} \alpha_j e_j \Big\| < \|g\| + \varepsilon,$$
which (using 8.54) implies that
$$\|g\| - \varepsilon < \Big(\sum_{j \in J'} |\alpha_j|^2 \Big)^{1/2} < \|g\| + \varepsilon.$$
Thus $\|g\| = \big(\sum_{k \in K} |\alpha_k|^2\big)^{1/2}$, completing the proof of one direction of (a) and the proof of (b).

To prove the other direction of (a), now suppose $\sum_{k\in K}|\alpha_k|^2 < \infty$. Thus there exists an increasing sequence $J_1 \subset J_2 \subset \cdots$ of finite subsets of K such that for each $m \in \mathbf{Z}^+$,

8.57
$$\sum_{j\in J'\setminus J_m} |\alpha_j|^2 < \frac{1}{m^2}$$

for every finite set J' such that $J_m \subset J' \subset K$. For each $m \in \mathbf{Z}^+$, let

$$g_m = \sum_{j\in J_m} \alpha_j e_j.$$

If $n > m$, then 8.54 implies that

$$\|g_n - g_m\|^2 = \sum_{j\in J_n\setminus J_m} |\alpha_j|^2 < \frac{1}{m^2}.$$

Thus $g_1, g_2, \dots$ is a Cauchy sequence and hence converges to some element g of V.

Temporarily fixing $m \in \mathbf{Z}^+$ and taking the limit of the equation above as $n \to \infty$, we see that

$$\|g - g_m\| \le \frac{1}{m}.$$

To show that $\sum_{k\in K} \alpha_k e_k = g$, suppose $\varepsilon > 0$. Let $m \in \mathbf{Z}^+$ be such that $\frac{2}{m} < \varepsilon$. Suppose J' is a finite set with $J_m \subset J' \subset K$. Then

$$\begin{aligned}\Big\|g - \sum_{j\in J'} \alpha_j e_j\Big\| &\le \|g - g_m\| + \Big\|g_m - \sum_{j\in J'} \alpha_j e_j\Big\| \\ &\le \frac{1}{m} + \Big\|\sum_{j\in J'\setminus J_m} \alpha_j e_j\Big\| \\ &= \frac{1}{m} + \Big(\sum_{j\in J'\setminus J_m} |\alpha_j|^2\Big)^{1/2} \\ &< \varepsilon,\end{aligned}$$

where the third line comes from 8.54 and the last line comes from 8.57. Thus $\sum_{k\in K} \alpha_k e_k = g$, completing the proof. ■

8.58 Example *a convergent unordered sum need not converge absolutely*

Suppose $\{e_k\}_{k\in\mathbf{Z}^+}$ is the orthogonal family in ℓ^2 defined by setting e_k equal to the sequence that equals 0 everywhere except for a 1 in the k^{th} slot. Then by 8.56, the unordered sum

$$\sum_{k\in\mathbf{Z}^+} \tfrac{1}{k} e_k$$

converges in ℓ^2 (because $\sum_{k\in\mathbf{Z}^+} \frac{1}{k^2} < \infty$) even though $\sum_{k\in\mathbf{Z}^+} \|\frac{1}{k} e_k\| = \infty$. Note that $\sum_{k\in\mathbf{Z}^+} \frac{1}{k} e_k = (1, \frac{1}{2}, \frac{1}{3}, \dots) \in \ell^2$.

Now we prove an important inequality.

8.59 *Bessel's Inequality*

Suppose $\{e_k\}_{k\in K}$ is an orthonormal family in an inner product space V and $f \in V$. Then

$$\sum_{k\in K} |\langle f, e_k\rangle|^2 \leq \|f\|^2.$$

Bessel's Inequality is named in honor of German astronomer Friedrich Bessel (1784–1846), who discovered this inequality in 1828 in the special case of the trigonometric orthonormal family given by the last bullet point in Example 8.53.

Proof Suppose J is a finite subset of K. Then

$$f = \Bigl(\sum_{j\in J} \langle f, e_j\rangle e_j\Bigr) + \Bigl(f - \sum_{j\in J} \langle f, e_j\rangle e_j\Bigr),$$

where the first term in parentheses above is orthogonal to the second term in parentheses above (as you should verify).

Applying the Pythagorean Theorem (8.11) to the equation above gives

$$\begin{aligned}
\|f\|^2 &= \Bigl\|\sum_{j\in J} \langle f, e_j\rangle e_j\Bigr\|^2 + \Bigl\|f - \sum_{j\in J} \langle f, e_j\rangle e_j\Bigr\|^2 \\
&\geq \Bigl\|\sum_{j\in J} \langle f, e_j\rangle e_j\Bigr\|^2 \\
&= \sum_{j\in J} |\langle f, e_j\rangle|^2,
\end{aligned}$$

where the last equality follows from 8.54. Because the inequality above holds for every finite set $J \subset K$, we conclude that $\|f\|^2 \geq \sum_{k\in K} |\langle f, e_k\rangle|^2$, as desired. ■

The following definition is probably familiar to you from linear algebra. Only finite sums appear in this definition, even though we might be working with an infinity family.

8.60 Definition *span*

Suppose $\{e_k\}_{k\in K}$ is a family in a vector space V. Then $\operatorname{span}\{e_k\}_{k\in K}$ is defined to be the set of all sums of the form

$$\sum_{j\in J} \alpha_j e_j,$$

where J is a finite subset of K and $\{\alpha_j\}_{j\in J}$ is a family in $\mathbf{F}$.

Bessel's Inequality now allows us to prove the following beautiful result showing that the closure of the span of an orthonormal family equals a set of infinite sums.

8.61 *closure of the span of an orthonormal family*

Suppose $\{e_k\}_{k\in K}$ is an orthonormal family in a Hilbert space V. Then

(a) $\overline{\operatorname{span}\{e_k\}_{k\in K}} = \Big\{\sum_{k\in K} \alpha_k e_k : \{\alpha_k\}_{k\in K} \text{ is a family in } \mathbf{F} \text{ and } \sum_{k\in K} |\alpha_k|^2 < \infty\Big\}.$

Furthermore,

(b) $$f = \sum_{k\in K} \langle f, e_k\rangle e_k$$

for every $f \in \overline{\operatorname{span}\{e_k\}_{k\in K}}$.

Proof The right side of (a) above makes sense because of 8.56(a). Furthermore, the right side of (a) above is a subspace of V because $\ell^2(K)$ [which equals $\mathcal{L}^2(\mu)$, where μ is counting measure on K] is closed under addition and scalar multiplication by 7.5.

Suppose first $\{\alpha_k\}_{k\in K}$ is a family in $\mathbf{F}$ and $\sum_{k\in K}|\alpha_k|^2 < \infty$. Let $\varepsilon > 0$. Then there is a finite subset J of K such that

$$\sum_{j\in K\setminus J} |\alpha_j|^2 < \varepsilon^2.$$

The inequality above and 8.56(b) imply that

$$\Big\|\sum_{k\in K} \alpha_k e_k - \sum_{j\in J} \alpha_j e_j\Big\| < \varepsilon.$$

The definition of the closure (see 6.30) now implies that $\sum_{k\in K} \alpha_k e_k \in \overline{\operatorname{span}\{e_k\}_{k\in K}}$, showing that the right side of (a) is contained in the left side of (a).

To prove the inclusion in the other direction, now suppose $f \in \overline{\operatorname{span}\{e_k\}_{k\in K}}$. Let

8.62 $$g = \sum_{k\in K} \langle f, e_k\rangle e_k,$$

where the sum above converges by Bessel's Inequality (8.59) and by 8.56(a). The direction of the inclusion that we just proved implies that $g \in \overline{\operatorname{span}\{e_k\}_{k\in K}}$. Thus

8.63 $$g - f \in \overline{\operatorname{span}\{e_k\}_{k\in K}}.$$

Equation 8.62 implies that $\langle g, e_j\rangle = \langle f, e_j\rangle$ for each $j \in K$, as you should verify (which will require using the Caunchy–Schwarz Inequality if done rigorously). Hence

$$\langle g - f, e_k\rangle = 0 \quad \text{for every } k \in K.$$

This implies that

$$g - f \in \big(\operatorname{span}\{e_j\}_{j\in K}\big)^{\perp} = \big(\overline{\operatorname{span}\{e_j\}_{j\in K}}\big)^{\perp},$$

where the equality above comes from 8.42(d). Now 8.63 and the inclusion above imply that $f = g$ [see 8.42(b)], which along with 8.62 implies that f is in the right side of (a), completing the proof of (a).

The equations $f = g$ and 8.62 also imply (b). ∎

Parseval's Identity

Let's review some basic linear algebra concepts, but now in the context of vector spaces that might be infinite-dimensional.

8.64 Definition ***linearly independent; basis***

Suppose $\{e_k\}_{k \in K}$ is a family in a vector space V.

- $\{e_k\}_{k \in K}$ is called *linearly independent* if there does not exist a finite nonempty subset J of K and a family $\{\alpha_j\}_{j \in J}$ in $\mathbf{F} \setminus \{0\}$ such that $\sum_{j \in J} \alpha_j e_j = 0$.
- $\{e_k\}_{k \in K}$ is called a *basis* of V if $\{e_k\}_{k \in K}$ is linearly independent and $\operatorname{span}\{e_k\}_{k \in K} = V$.

Note that 8.54 implies that every orthonormal family in an inner product space is linearly independent.

8.65 Definition ***finite-dimensional; infinite-dimensional***

- A vector space V is called *finite-dimensional* if there exists a finite set K and a family $\{e_k\}_{k \in K}$ in V such that $\operatorname{span}\{e_k\}_{k \in K} = V$.
- A vector space is called *infinite-dimensional* if it is not finite-dimensional.

Linear algebra deals mainly with finite-dimensional vector spaces, but infinite-dimensional vector spaces frequently appear in analysis. The notion of a basis is not so useful when doing analysis with infinite-dimensional vector spaces because the definition of span does not take advantage of the possibility of summing an infinite number of elements.

However, 8.61 tells us that taking the closure of the span of an orthonormal family can capture the sum of infinitely many elements. Thus we make the following definition.

8.66 Definition ***orthonormal basis***

An orthornormal family $\{e_k\}_{k \in K}$ in a Hilbert space V is called an *orthonormal basis* of V if

$$\overline{\operatorname{span}\{e_k\}_{k \in K}} = V.$$

Because every orthonormal family is linearly independent, the definition above differs from the definition of a basis (in the case of an orthonormal family) only by considering the closure of the span rather the span. An important point to keep in mind is that despite the terminology, an orthonormal basis is not necessarily a basis in the sense of 8.64. In fact, if K is an infinite set and $\{e_k\}_{k \in K}$ is an orthonormal basis of V, then $\{e_k\}_{k \in K}$ is not a basis of V (see Exercise 9).

8.67 Example *orthonormal bases*

- For $n \in \mathbf{Z}^+$ and $k \in \{1, \dots, n\}$, let e_k be the element of $\mathbf{F}^n$ all of whose coordinates are 0 except the k^{th} coordinate, which is 1:

$$e_k = (0, \dots, 0, 1, 0, \dots, 0).$$

Then $\{e_k\}_{k \in \{1,\dots,n\}}$ is an orthonormal basis of $\mathbf{F}^n$.

- Let $e_1 = \bigl(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\bigr)$, $e_2 = \bigl(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\bigr)$, and $e_3 = \bigl(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\bigr)$. Then $\{e_k\}_{k \in \{1,2,3\}}$ is an orthonormal basis of $\mathbf{F}^3$, as you should verify.
- All five examples of orthonormal families in Example 8.53 are orthonormal bases. The exercises ask you to verify that we have an orthonormal basis in the first four examples of orthornormal families in Example 8.53; we will deal with the last example in Example 8.53 (the trigonometric functions) in Chapter 11.

The next result shows why orthonormal bases are so useful—a Hilbert space with orthonormal basis $\{e_k\}_{k \in K}$ behaves like $\ell^2(K)$.

8.68 *Parseval's Identity*

Suppose $\{e_k\}_{k \in K}$ is an orthonormal basis of a Hilbert space V and $f, g \in V$. Then

(a) $f = \sum_{k \in K} \langle f, e_k \rangle e_k$;

(b) $\langle f, g \rangle = \sum_{k \in K} \langle f, e_k \rangle \overline{\langle g, e_k \rangle}$;

(c) $\|f\|^2 = \sum_{k \in K} |\langle f, e_k \rangle|^2$.

Proof The equation in (a) follows immediately from 8.61(b) and the definition of an orthonormal basis.

To prove (b), note that

$$\begin{aligned} \langle f, g \rangle &= \Bigl\langle \sum_{k \in K} \langle f, e_k \rangle e_k, g \Bigr\rangle \\ &= \sum_{k \in K} \langle f, e_k \rangle \langle e_k, g \rangle \\ &= \sum_{k \in K} \langle f, e_k \rangle \overline{\langle g, e_k \rangle}, \end{aligned}$$

where the first equation follows from (a) and the second equation follows from the definition of an unordered sum and the Cauchy–Schwarz Inequality.

Equation (c) follows from setting $g = f$ in (b). An alternative proof: equation (c) follows from 8.56(b) and the equation $f = \sum_{k \in K} \langle f, e_k \rangle e_k$ from (a). ■

Equation (c) is called Parseval's Identity *in honor of French mathematician Marc-Antoine Parseval (1755–1836), who discovered a special case in 1799.*

The Gram–Schmidt Process

8.69 Definition *separable*

A normed vector space is called *separable* if it has a countable subset whose closure equals the whole space.

8.70 Example *separable normed vector spaces*

- Suppose $n \in \mathbf{Z}^+$. Then $\mathbf{F}^n$ with the usual Hilbert space norm is separable because the closure of the countable set

$$\{(c_1, \dots, c_n) \in \mathbf{F}^n : \text{each } c_j \text{ is rational}\}$$

  equals $\mathbf{F}^n$ (in case $\mathbf{F} = \mathbf{C}$: to say that a complex number is rational in this context means that both the real and imaginary parts of the complex number are rational numbers in the usual sense).

- The Hilbert space ℓ^2 is separable because the closure of the countable set

$$\bigcup_{n=1}^{\infty} \{(c_1, \dots, c_n, 0, 0, \dots) \in \ell^2 : \text{ each } c_j \text{ is rational}\}$$

  is ℓ^2.

- The Hilbert spaces $L^2([0,1])$ and $L^2(\mathbf{R})$ are separable, as an exercise asks you to verify [hint: consider finite linear combinations with rational coefficients of functions of the form $\chi_{(c,d)}$, where c and d are rational numbers].

A moment's thought about the definition of closure (see 6.30) shows that a normed vector space V is separable if and only if there exists a countable subset C of V such that every open ball in V contains at least one element of C.

8.71 Example *nonseparable normed vector spaces*

- Suppose K is an uncountable set. Then the Hilbert space $\ell^2(K)$ is not separable. To see this, note that $\|\chi_{\{j\}} - \chi_{\{k\}}\| = \sqrt{2}$ for all $j, k \in K$ with $j \neq k$. Hence

$$\Big\{B\big(\chi_{\{k\}}, \tfrac{\sqrt{2}}{2}\big) : k \in K\Big\}$$

  is an uncountable collection of disjoint open balls in $\ell^2(K)$; no countable set can have at least one element in each of these balls.

- The Banach space $L^\infty([0,1])$ is not separable. Here $\|\chi_{[0,s]} - \chi_{[0,t]}\| = 1$ for all $s, t \in [0,1]$ with $s \neq t$. Thus $\big\{B\big(\chi_{[0,t]}, \frac{1}{2}\big) : t \in [0,1]\big\}$ is an uncountable collection of disjoint open balls in $L^\infty([0,1])$.

We will present two proofs about the existence of orthonormal bases of Hilbert spaces. The first proof works only for separable Hilbert spaces, but it gives a useful algorithm, called the *Gram–Schmidt process*, for constructing orthonormal sequences. The second proof works for all Hilbert spaces, but it uses a result that depends upon the Axiom of Choice.

Which proof should you read? In practice, the Hilbert spaces that you will encounter will almost certainly be separable. Thus the first proof suffices, and it has the additional benefit of introducing you to a widely-used algorithm. The second proof uses an entirely different approach and has the advantage of applying to separable and nonseparable Hilbert spaces. For maximum learning, read both proofs!

8.72 *Existence of orthonormal bases for separable Hilbert spaces*

Every separable Hilbert space has an orthonormal basis.

Proof Suppose V is a separable Hilbert space and $\{f_1, f_2, \dots\}$ is a countable subset of V whose closure equals V. We will inductively define an orthonormal sequence $\{e_k\}_{k\in\mathbf{Z}^+}$ such that

8.73
$$\operatorname{span}\{f_1, \dots, f_n\} \subset \operatorname{span}\{e_1, \dots, e_n\}$$

for each $n \in \mathbf{Z}^+$. This will imply that $\overline{\operatorname{span}\{e_k\}_{k\in\mathbf{Z}^+}} = V$, which will mean that $\{e_k\}_{k\in\mathbf{Z}^+}$ is an orthonormal basis of V.

To get started with the induction, set $e_1 = f_1/\|f_1\|$ (we can assume that $f_1 \neq 0$).

Now suppose $n \in \mathbf{Z}^+$ and $e_1, \dots, e_n$ have been chosen so that $\{e_k\}_{k\in\{1,\dots,n\}}$ is an orthonormal family in V and 8.73 holds. If $f_k \in \operatorname{span}\{e_1, \dots, e_n\}$ for every $k \in \mathbf{Z}^+$, then $\{e_k\}_{k\in\{1,\dots,n\}}$ is an orthonormal basis of V (completing the proof) and the process should be stopped. Otherwise, let m be the smallest positive integer such that

8.74
$$f_m \notin \operatorname{span}\{e_1, \dots, e_n\}.$$

Define e_{n+1} by

8.75
$$e_{n+1} = \frac{f_m - \langle f_m, e_1\rangle e_1 - \cdots - \langle f_m, e_n\rangle e_n}{\|f_m - \langle f_m, e_1\rangle e_1 - \cdots - \langle f_m, e_n\rangle e_n\|}.$$

Clearly $\|e_{n+1}\| = 1$ (8.74 guarantees there is no division by 0). If $k \in \{1, \dots, n\}$, then the equation above implies that $\langle e_{n+1}, e_k\rangle = 0$. Thus $\{e_k\}_{k\in\{1,\dots,n+1\}}$ is an orthonormal family in V. Also, 8.73 and the choice of m as the smallest positive integer satisfying 8.74 imply that

Danish mathematician Jørgen Gram (1850–1916) and German mathematician Erhard Schmidt (1876–1959) popularized this process that constructs orthonormal sequences.

$$\operatorname{span}\{f_1, \dots, f_{n+1}\} \subset \operatorname{span}\{e_1, \dots, e_{n+1}\},$$

completing the induction and completing the proof. ■

Before considering nonsepable Hilbert spaces, we take a short detour to illustrate how the Gram–Schmidt process used in the previous proof can be used to find closest elements to subspaces. We begin with a result connecting the orthogonal projection onto a closed subspace with an orthonormal basis of that subspace.

8.76 *orthogonal projection in terms of orthonormal basis*

Suppose that U is a closed subspace of a Hilbert space V and $\{e_k\}_{k\in K}$ is an orthonormal basis of U. Then

$$P_U f = \sum_{k\in K} \langle f, e_k\rangle e_k$$

for all $f \in V$.

Proof Let $f \in V$. If $k \in K$, then

8.77 $$\langle f, e_k\rangle = \langle f - P_U f, e_k\rangle + \langle P_U f, e_k\rangle = \langle P_U f, e_k\rangle,$$

where the last equality follows from 8.39(a). Now

$$P_U f = \sum_{k\in K} \langle P_U f, e_k\rangle e_k = \sum_{k\in K} \langle f, e_k\rangle e_k,$$

where the first equality follows from Parseval's Identity [8.68(a)] as applied to U and its orthonormal basis $\{e_k\}_{k\in K}$, and the second equality follows from 8.77. ∎

8.78 Example *best approximation*

Find the polynomial g of degree at most 10 that minimizes

$$\int_{-1}^{1} \Big|\sqrt{|x|} - g(x)\Big|^2\, dx.$$

Solution We will work in the real Hilbert space $L^2([-1,1])$ with the usual inner product $\langle g, h\rangle = \int_{-1}^{1} gh$. For $k \in \{0, 1, \dots, 10\}$, let $f_k \in L^2([-1,1])$ be defined by $f_k(x) = x^k$. Let U be the subspace of $L^2([-1,1])$ defined by

$$U = \operatorname{span}\{f_k\}_{k\in\{0,\dots,10\}}.$$

Apply the Gram–Schmidt process from the proof of 8.72 to $\{f_k\}_{k\in\{0,\dots,10\}}$, producing an orthonormal basis $\{e_k\}_{k\in\{0,\dots,10\}}$ of U, which is a closed subspace of $L^2([-1,1])$ (see Exercise 8). The point here is that $\{e_k\}_{k\in\{0,\dots,10\}}$ can be computed explicitly and exactly by using 8.75 and evaluating some integrals (using software that can do exact rational arithmetic will make the process easier), getting $e_0(x) = 1/\sqrt{2}$, $e_1(x) = \sqrt{6}x/2, \dots$ up to

$$e_{10}(x) = \frac{\sqrt{42}}{512}(-63 + 3465x^2 - 30030x^4 + 90090x^6 - 109395x^8 + 46189x^{10}).$$

Define $f \in L^2([-1,1])$ by $f(x) = \sqrt{|x|}$. Because U is the subspace of $L^2([-1,1])$ consisting of polynomials of degree at most 10 and $P_U f$ equals the element of U closest to f (see 8.36), the formula in 8.76 tells us that the solution g to our minimization problem is given by the formula

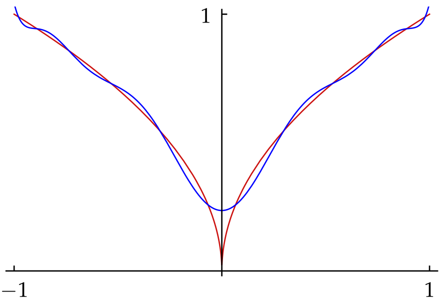
$$g = \sum_{k=0}^{10} \langle f, e_k \rangle e_k.$$

Using the explicit expressions for $e_0, \dots, e_{10}$ and again evaluating some integrals, this gives

$$g(x) = \frac{693 + 15015x^2 - 64350x^4 + 139230x^6 - 138567x^8 + 51051x^{10}}{2944}.$$

The figure here shows the graph of $f(x) = \sqrt{|x|}$ (red) and the graph of its closest polynomial g (blue) of degree at most 10; here *closest* means as measured in the norm of $L^2([-1,1])$.

The approximation of f by g is pretty good, especially considering that f is not differentiable at 0 and thus a Taylor series expansion for f does not make sense.



Existence of Orthonormal Bases

Now we introduce terminology that will be needed in our proof that every Hilbert space (including those Hilbert spaces that are nonseparable) has an orthonormal basis.

8.79 Definition ***chain***

A collection $\mathcal{C}$ of subsets of a set V is called a *chain* if $J, K \in \mathcal{C}$ implies $J \subset K$ or $K \subset J$.

8.80 Example *a chain of orthonormal families*

Suppose $\{e_k\}_{k \in \mathbf{Z}^+}$ is an orthonormal sequence in a Hilbert space V. Let

$$\mathcal{C} = \big\{\{e_1, \dots, e_n\} : n \in \mathbf{Z}^+\big\}.$$

Then the collection $\mathcal{C}$ is a chain because for each pair of elements of $\mathcal{C}$, one of them must be contained in the other.

8.81 Definition ***maximal element***

Suppose $\mathcal{A}$ is a collection of subsets of a set V. A set $K \in \mathcal{A}$ is called a *maximal element* of $\mathcal{A}$ if there does not exist $K' \in \mathcal{A}$ such that $K \subsetneq K'$.

A subset K of a Hilbert space V can be thought of as a family in V by considering $\{e_f\}_{f \in K}$, where $e_f = f$. With this convention, a set $K \subset V$ is an *orthonormal subset* of V if $\|f\| = 1$ for all $f \in K$ and $\langle f, g \rangle = 0$ for all $f, g \in K$ with $f \neq g$.

8.82 *orthonormal bases as maximal elements*

Suppose V is a Hilbert space, $\mathcal{A}$ is the collection of all orthonormal subsets of V, and K is an orthonormal subset of V. Then K is an orthonormal basis of V if and only if K is a maximal element of $\mathcal{A}$.

Proof First suppose K is an orthonormal basis of V. Parseval's Identity [8.68(a)] implies that the only element of V that is orthogonal to every element of K is 0. Thus there does not exist an orthonormal subset of V that strictly contains K. In other words, K is a maximal element of $\mathcal{A}$.

To prove the other direction, suppose now that K is a maximal element of $\mathcal{A}$. Let U denote the span of K. Then $U^\perp = \{0\}$ (because if f is a nonzero element of $U^\perp$, then $K \cup \{f/\|f\|\}$ is an orthonormal subset of V that strictly contains K). Hence $\overline{U} = V$ (by 8.44), which implies that K is an orthonormal basis of V. ■

The next result follows from the Axiom of Choice, although it is not as intuitively believable as the Axiom of Choice. Because the techniques used to prove the next result are so different from techniques used elsewhere in this book, the reader is asked either to accept this result without proof or find one of the good proofs available via the internet. The version of Zorn's Lemma stated here is simpler than the standard more general version, but this version is all that we need now.

8.83 *Zorn's Lemma*

Suppose V is a set and $\mathcal{A}$ is a collection of subsets of V with the property that the union of all the sets in $\mathcal{C}$ is in $\mathcal{A}$ for every chain $\mathcal{C} \subset \mathcal{A}$. Then $\mathcal{A}$ contains a maximal element.

Now we are ready to prove that every Hilbert space has an orthonormal basis.

8.84 *Existence of orthonormal bases for all Hilbert spaces*

Every Hilbert space has an orthonormal basis.

Proof Suppose V is a Hilbert space. Let $\mathcal{A}$ be the collection of all orthonormal subsets of V. Suppose $\mathcal{C} \subset \mathcal{A}$ is a chain. Let L be the union of all the sets in $\mathcal{C}$. If $f \in L$, then $\|f\| = 1$ because f is an element of some orthonormal subset of V that is contained in $\mathcal{C}$.

If $f, g \in L$ with $f \neq g$, then there exist orthonormal subsets J and K in $\mathcal{C}$ such that $f \in J$ and $g \in K$. Because $\mathcal{C}$ is a chain, either $J \subset K$ or $K \subset J$. Either way, there is an orthonormal subset of V that contains both f and g. Thus $\langle f, g \rangle = 0$.

We have shown that L is an orthonormal subset of V; in other words, $L \in \mathcal{A}$. Thus Zorn's Lemma (8.83) implies that $\mathcal{A}$ has a maximal element. Now 8.82 implies that V has an orthonormal basis. ■

Riesz Representation Theorem, Revisited

Now that we know that every Hilbert space has an orthonormal basis, we can give a completely different proof of the Riesz Representation Theorem (8.49) than the proof we gave earlier.

Note that the new proof below of the Riesz Representation Theorem gives the formula 8.86 for h in terms of an orthonormal basis. One interesting feature of this formula is that h is uniquely determined by φ and thus h does not depend upon the choice of an orthonormal basis. Hence despite its appearance, the right side of 8.86 is independent of the choice of an orthonormal basis.

8.85 *Riesz Representation Theorem*

Suppose φ is a bounded linear functional on a Hilbert space V and $\{e_k\}_{k\in K}$ is an orthonormal basis of V. Let

8.86
$$h = \sum_{k\in K} \overline{\varphi(e_k)} e_k.$$

Then

8.87
$$\varphi(f) = \langle f, h\rangle$$

for all $f \in V$. Furthermore, $\|\varphi\| = \bigl(\sum_{k\in K} |\varphi(e_k)|^2\bigr)^{1/2}$.

Proof First we must show that the sum defining h makes sense. To do this, suppose J is a finite subset of K. Then

$$\sum_{j\in J} |\varphi(e_j)|^2 = \varphi\Bigl(\sum_{j\in J} \overline{\varphi(e_j)} e_j\Bigr) \le \|\varphi\| \Bigl\|\sum_{j\in J} \overline{\varphi(e_j)} e_j\Bigr\| = \|\varphi\| \Bigl(\sum_{j\in J} |\varphi(e_j)|^2\Bigr)^{1/2},$$

where the last equality follows from 8.54. Dividing by $\bigl(\sum_{j\in J} |\varphi(e_j)|^2\bigr)^{1/2}$ gives

$$\Bigl(\sum_{j\in J} |\varphi(e_j)|^2\Bigr)^{1/2} \le \|\varphi\|.$$

Because the inequality above holds for every finite subset J of K, we conclude that

$$\sum_{k\in K} |\varphi(e_k)|^2 \le \|\varphi\|^2.$$

Thus the sum defining h makes sense (by 8.56) in equation 8.86.

Now 8.86 shows that $\langle h, e_j\rangle = \overline{\varphi(e_j)}$ for each $j \in K$. Thus if $f \in V$ then

$$\varphi(f) = \varphi\Bigl(\sum_{k\in K} \langle f, e_k\rangle e_k\Bigr) = \sum_{k\in K} \langle f, e_k\rangle \varphi(e_k) = \sum_{k\in K} \langle f, e_k\rangle \overline{\langle h, e_k\rangle} = \langle f, h\rangle,$$

where the first and last equalities follow from 8.68 and the second equality follows from the boundedness/continuity of φ. Thus 8.87 holds.

Finally, the Cauchy–Schwarz Inequality, 8.87, and the equation $\varphi(h) = \langle h, h\rangle$ show that $\|\varphi\| = \|h\| = \bigl(\sum_{k\in K} |\varphi(e_k)|^2\bigr)^{1/2}$. ∎

EXERCISES 8C

1 Verify that the family $\{e_k\}_{k\in\mathbf{Z}}$ as defined in the last bullet point of Example 8.53 is an orthonormal family in $L^2([0,2\pi))$. The following formulas should help:

$$(\sin x)(\cos y) = \frac{\sin(x+y)+\sin(x-y)}{2};$$

$$(\sin x)(\sin y) = \frac{\cos(x-y)-\cos(x+y)}{2};$$

$$(\cos x)(\cos y) = \frac{\cos(x+y)+\cos(x-y)}{2}.$$

2 Suppose $\{a_k\}_{k\in K}$ is a family in $\mathbf{R}$ and $a_k \geq 0$ for each $k \in K$. Prove the unordered sum $\sum_{k\in K} a_k$ converges if and only if

$$\sup\Big\{\sum_{j\in J} a_j : J \text{ is a finite subset of } K\Big\} < \infty.$$

Furthermore, prove that if $\sum_{k\in K} a_k$ converges then it equals the supremum above.

3 Suppose $\{a_k\}_{k\in K}$ is family in $\mathbf{R}$. Prove that the unordered sum $\sum_{k\in K} a_k$ converges if and only if $\sum_{k\in K}|a_k| < \infty$.

4 Suppose $\{e_k\}_{k\in K}$ is an orthonormal family in an inner product space V. Prove that if $f \in V$, then $\{k \in K : \langle f, e_k\rangle \neq 0\}$ is a countable set.

5 Suppose $\{f_k\}_{k\in K}$ and $\{g_k\}_{k\in K}$ are families in a normed vector space such that $\sum_{k\in K} f_k$ and $\sum_{k\in K} g_k$ converge. Prove that $\sum_{k\in K}(f_k+g_k)$ converges and

$$\sum_{k\in K}(f_k+g_k) = \sum_{k\in K} f_k + \sum_{k\in K} g_k.$$

6 Suppose $\{f_k\}_{k\in K}$ is a family in a normed vector space such that $\sum_{k\in K} f_k$ converges. Prove that if $c \in \mathbf{F}$, then $\sum_{k\in K}(cf_k)$ converges and

$$\sum_{k\in K}(cf_k) = c\sum_{k\in K} f_k.$$

7 Suppose $\{f_k\}_{k\in\mathbf{Z}^+}$ is a family in a normed vector space. Prove that the unordered sum $\sum_{k\in\mathbf{Z}^+} f_k$ converges if and only if the usual ordered sum $\sum_{k=1}^{\infty} f_{p(k)}$ converges for every injective function $p\colon \mathbf{Z}^+ \to \mathbf{Z}^+$.

8 Explain why 8.61 implies that if K is a finite set and $\{e_k\}_{k\in K}$ is an orthonormal family in a Hilbert space V, then $\operatorname{span}\{e_k\}_{k\in K}$ is a closed subspace of V.

9 Suppose V is an infinite-dimensional Hilbert space. Prove that there does not exist a basis of V that is an orthonormal family.

10 Show that the orthonormal family given in the first bullet point of Example 8.53 is an orthonormal basis of ℓ^2.

11 Show that the orthonormal family given in the second bullet point of Example 8.53 is an orthonormal basis of $\ell^2(K)$.

12 Show that the orthonormal family given in the third bullet point of Example 8.53 is an orthonormal basis of $L^2\bigl([0,1)\bigr)$.

13 Show that the orthonormal family given in the fourth bullet point of Example 8.53 is an orthonormal basis of $L^2(\mathbf{R})$.

14 Prove the converse of Parseval's Identity. More specifically, prove that if $\{e_k\}_{k\in K}$ is an orthonormal family in a Hilbert space V and

$$\|f\|^2 = \sum_{k\in K} |\langle f, e_k\rangle|^2$$

for every $f \in V$, then $\{e_k\}_{k\in K}$ is an orthonormal basis of V.

15 Show that the Hilbert space $L^2([0,1])$ is separable.

16 Show that the Hilbert space $L^2(\mathbf{R})$ is separable.

17 Show that the Banach space ℓ^∞ is not separable.

18 Prove that every subspace of a separable normed vector space is separable.

19 Suppose H is an infinite-dimensional Hilbert space. Prove that there does not exist a translation invariant measure on the Borel subsets of H that assigns positive but finite measure to each open ball in H.
[*A subset of H is called a* Borel set *if it is in the smallest σ-algebra containing all the open subsets of H. A measure μ on the Borel subsets of H is called* translation invariant *if $\mu(f+E) = \mu(E)$ for every $f \in H$ and every Borel set E of H.*]

20 Find the polynomial g of degree at most 4 that minimizes $\int_0^1 \left|x^5 - g(x)\right|^2 dx$.

21 Prove that each orthonormal family in a Hilbert space can be extended to an orthonormal basis of the Hilbert space. Specifically, suppose $\{e_j\}_{j\in J}$ is an orthonormal family in a Hilbert space V. Prove that there exists a set K containing J and an orthonormal basis $\{f_k\}_{k\in K}$ of V such that $f_j = e_j$ for every $j \in J$.

22 Prove that every vector space has a basis.

23 Suppose that V is an infinite-dimensional normed vector space. Prove that there exists a linear functional $\varphi\colon V \to \mathbf{F}$ that is not continuous.

24 Find the polynomial g of degree at most 4 such that

$$f\left(\tfrac{1}{2}\right) = \int_0^1 fg$$

for every polynomial f of degree at most 4.

Exercises 25–30 are for readers familiar with analytic functions.

25 Suppose Ω is a nonempty open subset of $\mathbf{C}$. The *Bergman space* $L^2_a(\Omega)$ is defined to be the set of analytic functions $f\colon \Omega \to \mathbf{C}$ such that

$$\int_\Omega |f|^2 \, d\lambda_2 < \infty,$$

where λ_2 is the usual Lebesgue measure on $\mathbf{R}^2$, which is identified with $\mathbf{C}$. For $f, g \in L^2_a(\Omega)$, define $\langle f, g\rangle$ to be $\int_\Omega f\,\overline{g}\, d\lambda_2$.

(a) Show that $L^2_a(\Omega)$ is a Hilbert space.

(b) Show that if $w \in \Omega$, then $f \mapsto f(w)$ is a bounded linear functional on $L^2_a(\Omega)$.

26 Let $\mathbf{D}$ denote the open unit disk in $\mathbf{C}$; thus

$$\mathbf{D} = \{z \in \mathbf{C} : |z| < 1\}.$$

(a) Find an orthonormal basis of $L^2_a(\mathbf{D})$.

(b) Suppose $f \in L^2_a(\mathbf{D})$ has Taylor series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

for $z \in \mathbf{D}$. Find a formula for $\|f\|$ in terms of $a_0, a_1, a_2, \dots$.

(c) Suppose $w \in \mathbf{D}$. By the previous exercise and the Riesz Representation Theorem (8.49 and 8.85), there exists $K_w \in L^2_a(\mathbf{D})$ such that

$$f(w) = \langle f, K_w\rangle \text{ for all } f \in L^2_a(\mathbf{D}).$$

Find an explicit formula for K_w.

27 Suppose Ω is the annulus defined by

$$\Omega = \{z \in \mathbf{C} : 1 < |z| < 2\}.$$

(a) Find an orthonormal basis of $L^2_a(\Omega)$.

(b) Suppose $f \in L^2_a(\Omega)$ has Laurent series

$$f(z) = \sum_{k=-\infty}^{\infty} a_k z^k$$

for $z \in \mathbf{D}$. Find a formula for $\|f\|$ in terms of $\dots, a_{-1}, a_0, a_1, \dots$.

28 Prove that if $f \in L^2_a(\mathbf{D} \setminus \{0\})$, then f has a removable singularity at 0 (meaning that f can be extended to a function that is analytic on $\mathbf{D}$).

29 The *Dirichlet space* $\mathcal{D}$ is defined to be the set of analytic functions $f\colon \mathbf{D} \to \mathbf{C}$ such that

$$\int_{\mathbf{D}} |f'|^2 \, d\lambda_2 < \infty.$$

For $f, g \in \mathcal{D}$, define $\langle f, g \rangle$ to be $f(0)\overline{g(0)} + \int_{\mathbf{D}} f' \, \overline{g'} \, d\lambda_2$.

(a) Show that $\mathcal{D}$ is a Hilbert space.

(b) Show that if $w \in \Omega$, then $f \mapsto f(w)$ is a bounded linear functional on $\mathcal{D}$.

(c) Find an orthonormal basis of $\mathcal{D}$.

(d) Suppose $f \in \mathcal{D}$ has Taylor series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

for $z \in \mathbf{D}$. Find a formula for $\|f\|$ in terms of $a_0, a_1, a_2, \dots$.

(e) Suppose $w \in \mathbf{D}$. Find an explicit formula for $K_w \in \mathcal{D}$ such that

$$f(w) = \langle f, K_w \rangle \text{ for all } f \in \mathcal{D}.$$

30 (a) Prove that the Dirichlet space $\mathcal{D}$ is contained in the Bergman space $L^2_a(\mathbf{D})$.

(b) Prove that there exists a function $f \in L^2_a(\mathbf{D})$ such that f is uniformly continuous on $\mathbf{D}$ and $f \notin \mathcal{D}$.