II: Hilbert Spaces

Gentlemen: there's lots of room left in Hilbert space — *S. MacLane*

II.1 The geometry of Hilbert space

Finite-dimensional vector spaces have three kinds of properties whose generalizations we will study in the next four chapters: linear properties, metric properties, and geometric properties. In this chapter we study vector spaces that have an inner product, a generalization of the usual dot product on finite dimensional vector spaces. The geometric properties of these spaces follow from the notion of angle which is implicit in the definition of inner product.

Definition A complex vector space V is called an **inner product space** if there is a complex-valued function $(\cdot, \cdot)$ on $V \times V$ that satisfies the following four conditions for all $x, y, z \in V$ and $\alpha \in \mathbb{C}$:

(i) $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$
(ii) $(x, y + z) = (x, y) + (x, z)$
(iii) $(x, \alpha y) = \alpha(x, y)$
(iv) $(x, y) = \overline{(y, x)}$

The function $(\cdot, \cdot)$ is called an inner product.

We note that (ii), (iii), and (iv) imply that $(x, \alpha y + \beta z) = \alpha(x, y) + \beta(x, z)$ and that $(\alpha x, y) = \bar{\alpha}(x, y)$. The reader should be aware that some texts use a convention different from the one introduced in (iii); they take the inner product to be linear in the *first* vector and conjugate-linear in the *second.*

Example 1 ($\mathbb{C}^n$) Let $\mathbb{C}^n$ denote the set of all n-tuples of complex numbers. For $x = \langle x_1, \ldots, x_n \rangle$ and $y = \langle y_1, \ldots, y_n \rangle$ in $\mathbb{C}^n$ define

$$(x, y) = \sum_{j=1}^{n} \bar{x}_j y_j$$

Example 2 Let $C[a, b]$ denote the complex-valued continuous functions on the interval $[a, b]$. For $f(x), g(x) \in C[a, b]$ define

$$(f, g) = \int_a^b \overline{f(x)} g(x)\, dx$$

We now develop those geometrical notions that extend to arbitrary inner product spaces.

Definition Two vectors, x and y, in an inner product space V are said to be **orthogonal** if $(x, y) = 0$. A collection $\{x_i\}$ of vectors in V is called an **orthonormal set** if $(x_i, x_i) = 1$ for all i, and $(x_i, x_j) = 0$ if $i \neq j$.

We introduce the shorthand $\|x\| = \sqrt{(x, x)}$. We will shortly see that $\|\cdot\|$ is in fact a norm.

Theorem II.1 (Pythagorean theorem) Let $\{x_n\}_{n=1}^{N}$ be an orthonormal set in an inner product space V. Then for all $x \in V$,

$$\|x\|^2 = \sum_{n=1}^{N} |(x, x_n)|^2 + \left\| x - \sum_{n=1}^{N} (x_n, x)x_n \right\|^2$$

Proof We write x as

$$x = \sum_{n=1}^{N} (x_n, x)x_n + \left(x - \sum_{n=1}^{N} (x_n, x)x_n \right)$$

A short computation using the properties of inner products shows that

$$\sum_{n=1}^{N} (x_n, x)x_n \qquad \text{and} \qquad x - \sum_{n=1}^{N} (x_n, x)x_n$$

are orthogonal. Thus,

$$(x, x) = \left\| \sum_{n=1}^{N} (x_n, x)x_n \right\|^2 + \left\| x - \sum_{n=1}^{N} (x_n, x)x_n \right\|^2$$
$$= \sum_{n=1}^{N} |(x_n, x)|^2 + \left\| x - \sum_{n=1}^{N} (x_n, x)x_n \right\|^2 \quad \blacksquare$$

Corollary (Bessel's inequality) Let $\{x_n\}_{n=1}^{N}$ be an orthonormal set in an inner product space, V. Then for all $x \in V$,

$$\|x\|^2 \geq \sum_{n=1}^{N} |(x, x_n)|^2$$

Corollary (the Schwarz inequality) If x and y are vectors in an inner product space V, then

$$|(x, y)| \leq \|x\| \, \|y\|$$

Proof The case $y = 0$ is trivial, so suppose $y \neq 0$. The vector $y/\|y\|$ by itself forms an orthonormal set, so applying Bessel's inequality to any $x \in V$ we get

$$\|x\|^2 \geq |(x, y/\|y\|)|^2 = \frac{|(x, y)|^2}{\|y\|^2}$$

from which $|(x, y)| \leq \|x\| \, \|y\|$ follows. $\blacksquare$

Another useful geometric equality is the parallelogram law (Problem 4):

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

In Section I.2 we defined normed linear spaces and observed that every normed linear space is a metric space. The following theorem shows that every inner product space is a normed linear space.

Theorem II.2 Every inner product space V is a normed linear space with the norm $\|x\| = (x, x)^{1/2}$

Proof Since V is a vector space, we need only verify that $\|\cdot\|$ has all the properties of a norm. All of these properties, except the triangle inequality, follow immediately from the properties (i)–(iv) of inner products. Suppose $x, y \in V$. Then

$$\begin{aligned}\|x + y\|^2 &= (x, x) + (x, y) + (y, x) + (y, y) \\ &= (x, x) + 2\,\mathrm{Re}(x, y) + (y, y) \\ &\leq (x, x) + 2|(x, y)| + (y, y) \\ &\leq (x, x) + 2(x, x)^{1/2}(y, y)^{1/2} + (y, y)\end{aligned}$$

by the Schwarz inequality. Thus

$$\|x + y\|^2 \le (\|x\| + \|y\|)^2$$

which proves the triangle inequality. ∎

This theorem shows that we have a natural metric,

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

in V. We thus have the notions of convergence, completeness, and density defined for metric spaces in Section I.2. In particular, we can always complete V to a normed linear space $\tilde{V}$ in which V is isometrically embedded as a dense subset. In fact, $\tilde{V}$ is also an inner product space since the inner product can be extended from V to $\tilde{V}$ by continuity (Problem 1).

Definition A complete inner product space is called a **Hilbert space.** Inner product spaces are sometimes called pre-Hilbert spaces.

Definition Two Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ are said to be **isomorphic** if there is a linear operator U from $\mathscr{H}_1$ *onto* $\mathscr{H}_2$ such that $(Ux, Uy)_{\mathscr{H}_2} = (x, y)_{\mathscr{H}_1}$ for all $x, y \in \mathscr{H}_1$. Such an operator is called **unitary**.

We elaborate these ideas and show the reader what types of Hilbert spaces he is likely to meet by a series of examples.

Example 2 (revisited) Define $L^2[a, b]$ to be the set of complex-valued measurable functions on $[a, b]$, a finite interval, that satisfy $\int_a^b |f(x)|^2\, dx < \infty$. We define an inner product by

$$(f, g) = \int_a^b \overline{f(x)}g(x)\, dx$$

Observe that the inner makes sense since

$$|\overline{f(x)}g(x)| \le \tfrac{1}{2}|f(x)|^2 + \tfrac{1}{2}|g(x)|^2$$

so that $\overline{f(x)}g(x)$ is in $L^1[a, b]$. A proof similar to the Riesz–Fisher theorem (Theorem I.12) shows that $L^2[a, b]$ is complete and is therefore a Hilbert space. It is not too difficult to show (Problem 2) that $L^2[a, b]$ is the completion of $C[a, b]$ in the norm

$$\|f\| = \left(\int_a^b |f(x)|^2\, dx\right)^{1/2}$$

Example 3 (ℓ_2) Define ℓ_2 to be the set of sequences $\{x_n\}_{n=1}^{\infty}$ of complex numbers which satisfy $\sum_{n=1}^{\infty} |x_n|^2 < \infty$ with the inner product

$$(\{x_n\}_{n=1}^{\infty}, \{y_n\}_{n=1}^{\infty}) = \sum_{n=1}^{\infty} \bar{x}_n y_n$$

In Section II.3 we will see that any Hilbert space that has a countable dense set and is not finite dimensional is isomorphic to ℓ_2. In this sense, ℓ_2 is the canonical example of a Hilbert space.

Example 4 ($L^2(\mathbb{R}^n, d\mu)$) Let μ be a Borel measure on $\mathbb{R}^n$. $L^2(\mathbb{R}^n, d\mu)$ is the set of complex-valued measurable functions on $\mathbb{R}^n$ which satisfy $\int_{\mathbb{R}^n} |f(x)|^2 \, d\mu < \infty$. $L^2(\mathbb{R}^n, d\mu)$ is a Hilbert space under the inner product

$$(f, g) = \int_{\mathbb{R}^n} \overline{f(x)} g(x) \, d\mu$$

Example 5 (direct sum) Suppose that $\mathscr{H}_1$ and $\mathscr{H}_2$ are Hilbert spaces. Then the set of pairs $\langle x, y \rangle$ with $x \in \mathscr{H}_1$, $y \in \mathscr{H}_2$ is a Hilbert space with inner product

$$(\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle) = (x_1, x_2)_{\mathscr{H}_1} + (y_1, y_2)_{\mathscr{H}_2}$$

This space is called the **direct sum** of the spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ and is denoted by $\mathscr{H}_1 \oplus \mathscr{H}_2$. If μ_1 and μ_2 are mutually singular Borel measures on $\mathbb{R}$ and $\mu = \mu_1 + \mu_2$, then $L^2(\mathbb{R}, d\mu)$ is isomorphic in a natural way to $L^2(\mathbb{R}, d\mu_1) \oplus L^2(\mathbb{R}, d\mu_2)$ (Problem 3). We can also construct countable direct sums as follows. Suppose $\{\mathscr{H}_n\}_{n=1}^{\infty}$ is a sequence of Hilbert spaces. Let $\mathscr{H}$ denote the set of sequences $\{x_n\}_{n=1}^{\infty}$, with $x_n \in \mathscr{H}_n$, which satisfy

$$\sum_{n=1}^{\infty} \|x_n\|_{\mathscr{H}_n}^2 < \infty$$

$\mathscr{H}$ is a Hilbert space under the natural inner product and is denoted by

$$\mathscr{H} = \bigoplus_{n=1}^{\infty} \mathscr{H}_n$$

Example 6 (vector-valued functions) Suppose $\langle X, \mu \rangle$ is a measure space and $\mathscr{H}'$ is a Hilbert space. Let $L^2(X, d\mu; \mathscr{H}')$ be the set of measurable functions on X with values in $\mathscr{H}'$ which satisfy

$$\int \|f(x)\|_{\mathscr{H}'}^2 \, d\mu(x) < \infty$$

This set is a Hilbert space with the inner product

$$(f, g) = \int_X (f(x), g(x))_{\mathscr{H}'} \, d\mu(x)$$

Of course, we have not said what it means for a vector-valued function to be measurable. For this definition and related matters see Problem 12 and the appendix to Section IV.5.

II.2 The Riesz lemma †

In the examples in Section II.1 we showed several ways of constructing new Hilbert spaces from old ones. Another way to do this is to restrict attention to a closed subspace $\mathscr{M}$ of the given Hilbert space $\mathscr{H}$. Under the natural inner product that it inherits as a subspace of $\mathscr{H}$, $\mathscr{M}$ is a Hilbert space. We denote by $\mathscr{M}^\perp$ the set of vectors in $\mathscr{H}$ which are orthogonal to $\mathscr{M}$; $\mathscr{M}^\perp$ is called the **orthogonal complement** of $\mathscr{M}$. It follows from the linearity of the inner product that $\mathscr{M}^\perp$ is a linear subspace of $\mathscr{H}$ and an elementary argument (Problem 6) shows that $\mathscr{M}^\perp$ is closed. Thus $\mathscr{M}^\perp$ is also a Hilbert space. $\mathscr{M}$ and $\mathscr{M}^\perp$ have only the zero element in common. The following theorem shows that there are vectors perpendicular to any closed proper subspace, indeed there are enough of them so that

$$\mathscr{H} = \mathscr{M} + \mathscr{M}^\perp = \{x + y \mid x \in \mathscr{M}, y \in \mathscr{M}^\perp\}$$

This important geometric property is one of the main reasons that Hilbert spaces are easier to handle than Banach spaces (Chapter III). In the following lemma and theorem, the reader should keep the finite-dimensional case in mind (see Figure II.1).

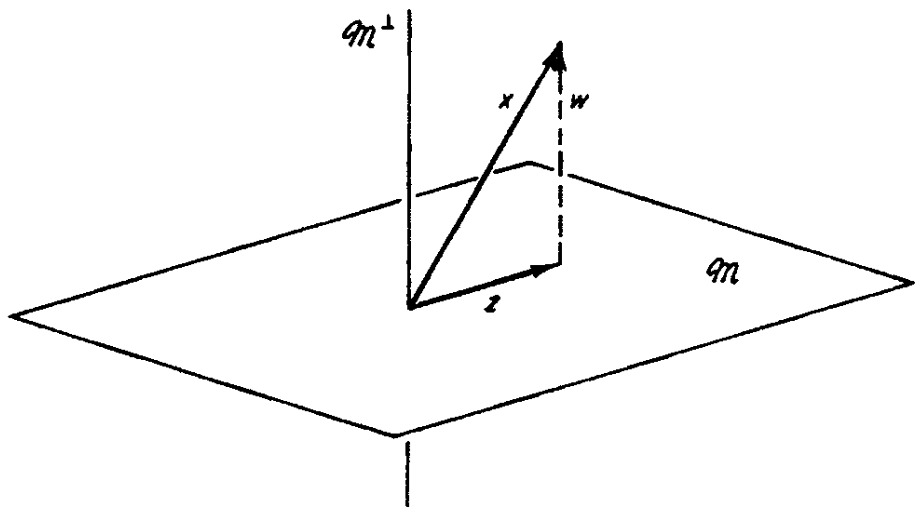

FIGURE II.1 The projection of x on $\mathscr{M}$.

† A supplement to this section begins on p. 344.

Lemma Let $\mathscr{H}$ be a Hilbert space, $\mathscr{M}$ a closed subspace of $\mathscr{H}$, and suppose $x \in \mathscr{H}$. Then there exists in $\mathscr{M}$ a unique element z closest to x.

Proof Let $d = \inf_{y \in \mathscr{M}} \|x - y\|$. Choose a sequence $\{y_n\}$, $y_n \in \mathscr{M}$, so that

$$\|x - y_n\| \to d$$

Then

$$\begin{aligned}
\|y_n - y_m\|^2 &= \|(y_n - x) - (y_m - x)\|^2 \\
&= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - \|-2x + y_n + y_m\|^2 \\
&= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\|x - \tfrac{1}{2}(y_n + y_m)\|^2 \\
&\le 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2 \\
&\underset{\substack{m\to\infty \\ n\to\infty}}{\longrightarrow} 2d^2 + 2d^2 - 4d^2 = 0
\end{aligned}$$

The second equality follows from the parallelogram law; the inequality follows from the fact that $\frac{1}{2}(y_n + y_m) \in \mathscr{M}$. Thus $\{y_n\}$ is Cauchy and since $\mathscr{M}$ is closed, $\{y_n\}$ converges to an element z of $\mathscr{M}$. It follows easily that $\|x - z\| = d$. Uniqueness is left as an exercise. ∎

Theorem II.3 (the projection theorem) Let $\mathscr{H}$ be a Hilbert space, $\mathscr{M}$ a closed subspace. Then every $x \in \mathscr{H}$ can be uniquely written $x = z + w$ where $z \in \mathscr{M}$ and $w \in \mathscr{M}^\perp$.

Proof Let x be in $\mathscr{H}$. Then by the lemma there is a unique element $z \in \mathscr{M}$ closest to x. Define $w = x - z$, then we clearly have $x = z + w$. Let $y \in \mathscr{M}$ and $t \in \mathbb{R}$. If $d = \|x - z\|$, then

$$\begin{aligned}
d^2 \le \|x - (z + ty)\|^2 &= \|w - ty\|^2 \\
&= d^2 - 2t \operatorname{Re}(w, y) + t^2\|y\|^2
\end{aligned}$$

Thus, $-2t \operatorname{Re}(w, y) + t^2\|y\|^2 \ge 0$ for all t, which implies $\operatorname{Re}(w, y) = 0$. A similar argument using ti instead of t shows that $\operatorname{Im}(w, y) = 0$. Thus, $w \in \mathscr{M}^\perp$. Uniqueness is left as an exercise. ∎

The projection theorem sets up a natural isomorphism between $\mathscr{M} \oplus \mathscr{M}^\perp$ and $\mathscr{H}$ given by

$$\langle z, w \rangle \mapsto z + w$$

We will often suppress the isomorphism and simply write $\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^\perp$.

We have already defined in Section I.2 what we mean by a bounded linear transformation from one Hilbert space $\mathscr{H}$ to another $\mathscr{H}'$. We will denote by

$\mathscr{L}(\mathscr{H}, \mathscr{H}')$ the set of such transformations. $\mathscr{L}(\mathscr{H}, \mathscr{H}')$ is clearly a vector space and it becomes a Banach space under the norm

$$\|T\| = \sup_{\|x\|_{\mathscr{H}}=1} \|Tx\|_{\mathscr{H}'}.$$

The proof of this fact, though not difficult, is postponed until Chapter III where it is proven in greater generality. For the time being we are interested in the special case where $\mathscr{H}' = \mathbb{C}$:

Definition The space $\mathscr{L}(\mathscr{H}, \mathbb{C})$ is called the **dual space** of $\mathscr{H}$ and is denoted by $\mathscr{H}^*$. The elements of $\mathscr{H}^*$ are called **continuous linear functionals.**

The following important theorem which characterizes $\mathscr{H}^*$ is due to F. Riesz and M. Fréchet.

Theorem II.4 (the Riesz lemma) For each $T \in \mathscr{H}^*$, there is a unique $y_T \in \mathscr{H}$ such that $T(x) = (y_T, x)$ for all $x \in \mathscr{H}$. In addition $\|y_T\|_{\mathscr{H}} = \|T\|_{\mathscr{H}^*}$.

Proof Let $\mathscr{N}$ be the set of $x \in \mathscr{H}$ such that $T(x) = 0$. By the continuity of T, $\mathscr{N}$ is a closed subspace. If $\mathscr{N} = \mathscr{H}$, then $T(x) = 0 = (0, x)$ for all x and we are finished; so assume $\mathscr{N}$ is not all of $\mathscr{H}$. Then by the projection theorem there is a nonzero vector x_0 in $\mathscr{N}^\perp$. Define $y_T = \overline{T(x_0)}\|x_0\|^{-2}x_0$. We will verify that the vector y_T has the right properties. First, if $x \in \mathscr{N}$, then $T(x) = 0 = (y_T, x)$. Further, if $x = \alpha x_0$, then

$$T(x) = T(\alpha x_0) = \alpha T(x_0) = (\overline{T(x_0)}\|x_0\|^{-2}x_0, \alpha x_0) = (y_T, \alpha x_0)$$

Since the functions $T(\cdot)$ and $(y_T, \cdot)$ are linear and agree on $\mathscr{N}$ and x_0, they must agree on the space spanned by $\mathscr{N}$ and x_0. But $\mathscr{N}$ and x_0 span $\mathscr{H}$ since every element $y \in \mathscr{H}$ can be written

$$y = \left(y - \frac{T(y)}{T(x_0)}x_0\right) + \frac{T(y)}{T(x_0)}x_0$$

Thus $T(x) = (y_T, x)$ for all $x \in \mathscr{H}$. If $T(x) = (y', x)$ also, then $\|y' - y_T\|^2 = T(y' - y_T) - T(y' - y_T) = 0$ so $y' = y_T$, proving uniqueness.

To prove that $\|T\|_{\mathscr{H}^*} = \|y_T\|_{\mathscr{H}}$ we observe that

$$\|T\| = \sup_{\|x\|\le 1} |T(x)| = \sup_{\|x\|\le 1} |(y_T, x)| \le \sup_{\|x\|\le 1} \|y_T\|\,\|x\| = \|y_T\|$$

and

$$\|T\| = \sup_{\|x\|\le 1} |T(x)| \ge \left|T\left(\frac{y_T}{\|y_T\|}\right)\right| = \left(y_T, \frac{y_T}{\|y_T\|}\right) = \|y_T\| \quad \blacksquare$$

We note that the Schwarz inequality shows that the converse of the Riesz lemma is true. Namely, each $y \in \mathscr{H}$ defines a continuous linear functional T_y on $\mathscr{H}$ by $T_y(x) = (y, x)$. The Riesz lemma has the following corollary which is very important in applications.

Corollary Let $B(\cdot, \cdot)$ be a function from $\mathscr{H} \times \mathscr{H}$ to $\mathbb{C}$ which satisfies:

(i) $B(x, \alpha y + \beta z) = \alpha B(x, y) + \beta B(x, z)$
(ii) $B(\alpha x + \beta y, z) = \bar{\alpha} B(x, z) + \bar{\beta} B(y, z)$
(iii) $|B(x, y)| \le C\|x\| \, \|y\|$

for all $x, y, z \in \mathscr{H}$, $\alpha, \beta \in \mathbb{C}$. Then there is a unique bounded linear transformation A, from $\mathscr{H}$ to $\mathscr{H}$ so that

$$B(x, y) = (Ax, y) \qquad \text{for all} \quad x, y \in \mathscr{H}$$

The norm of A is the smallest constant C such that (iii) holds.

Proof Fix x, then (i) and (iii) show that $B(x, \cdot)$ is a continuous linear functional on $\mathscr{H}$. Thus by the Riesz lemma there is an $x' \in \mathscr{H}$ so that

$$B(x, y) = (x', y) \qquad \text{for all} \quad y \in \mathscr{H}$$

Define $Ax = x'$. It is not difficult to show that A is a continuous linear operator with the right properties (Problem 8). ∎

A bilinear function on $\mathscr{H}$ obeying (i) and (ii) is called a **sesquilinear form**.

II.3 Orthonormal bases

We have already defined what it means for a set of vectors to be orthonormal. In this section we develop this idea further; in particular we want to extend the idea of a "basis," so useful for finite-dimensional vector spaces, to complete inner product spaces. If S is an orthonormal set in a Hilbert space $\mathscr{H}$ and no other orthonormal set contains S as a proper subset, then S is called an **orthonormal basis** (or a **complete orthonormal system**) for $\mathscr{H}$.

Theorem II.5 Every Hilbert space $\mathscr{H}$ has an orthonormal basis.

Proof Consider the collection $\mathscr{C}$ of orthonormal sets in V. We order $\mathscr{C}$ by inclusion; that is, we say $S_1 \prec S_2$ if $S_1 \subset S_2$. With this definition of $\prec$, $\mathscr{C}$ is partially ordered; it is also nonempty since if v is any element of V, the set consisting only of $v/\|v\|$ is an orthonormal set. Now let $\{S_\alpha\}_{\alpha\in A}$ be any linearly ordered subset of $\mathscr{C}$. Then $\bigcup_{\alpha\in A} S_\alpha$ is an orthonormal set which contains each S_α and is thus an upper bound for $\{S_\alpha\}_{\alpha\in A}$. Since every linearly ordered subset of $\mathscr{C}$ has an upper bound, we can apply Zorn's lemma (Theorem I.2) and conclude that $\mathscr{C}$ has a maximal element; that is, an orthonormal system not properly contained in any other orthonormal system. ▮

The following theorem shows that as in the finite-dimensional case every element of a Hilbert space can be expressed as a linear combination (possibly infinite) of basis elements.

Theorem II.6 Let $\mathscr{H}$ be a Hilbert space and $S = \{x_\alpha\}_{\alpha\in A}$ an orthonormal basis. Then for each $y \in \mathscr{H}$,

$$y = \sum_{\alpha\in A} (x_\alpha, y)x_\alpha \tag{II.1}$$

and

$$\|y\|^2 = \sum_{\alpha\in A} |(x_\alpha, y)|^2 \tag{II.2}$$

The equality in (II.1) means that the sum on the right-hand side converges (independent of order) to y in $\mathscr{H}$. Conversely, if $\sum_{\alpha\in A} |c_\alpha|^2 < \infty$, $c_\alpha \in \mathbb{C}$, then $\sum_{\alpha\in A} c_\alpha x_\alpha$ converges to an element of $\mathscr{H}$.

Proof We have already shown in Section II.1 (Bessel's inequality) that for any finite subset $A' \subset A$, $\sum_{\alpha\in A'} |(x_\alpha, y)|^2 \le \|y\|^2$. Thus $(x_\alpha, y) \neq 0$ for at most a countable number of α's in A which we order in some way $\alpha_1, \alpha_2, \alpha_3, \ldots$. Furthermore, since $\sum_{j=1}^N |(x_{\alpha_j}, y)|^2$ is monotone increasing and bounded, it converges to a finite limit as $N \to \infty$. Let $y_n = \sum_{j=1}^n (x_{\alpha_j}, y)x_{\alpha_j}$. Then for $n > m$,

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

Therefore $\{y_n\}$ is a Cauchy sequence and converges to an element y' of $\mathscr{H}$. Observe that

$$\begin{aligned}(y - y', x_{\alpha_\ell}) &= \lim_{n\to\infty} \left(y - \sum_{j=1}^n (x_{\alpha_j}, y)x_{\alpha_j}, x_{\alpha_\ell} \right) \\ &= (y, x_{\alpha_\ell}) - (y, x_{\alpha_\ell}) = 0\end{aligned}$$

And if $\alpha \neq \alpha_\ell$ for some ℓ we have

$$(y - y', x_\alpha) = \lim_{n\to\infty} \left(y - \sum_{j=1}^n (x_{\alpha_j}, y)x_{\alpha_j}, x_\alpha \right) = 0$$

Therefore $y - y'$ is orthogonal to all the x_α in S. Since S is a complete orthonormal system we must have $y - y' = 0$. Thus

$$y = \lim_{n\to\infty} \sum_{j=1}^{n} (x_{\alpha_j}, y) x_{\alpha_j}$$

and (II.1) holds. Furthermore,

$$\begin{aligned} 0 &= \lim_{n\to\infty} \left\| y - \sum_{j=1}^{n} (x_{\alpha_j}, y) x_{\alpha_j} \right\|^2 \\ &= \lim_{n\to\infty} \left(\|y\|^2 - \sum_{j=1}^{n} |(x_{\alpha_j}, y)|^2 \right) \\ &= \|y\|^2 - \sum_{\alpha\in A} |(x_\alpha, y)|^2 \end{aligned}$$

so that (II.2) holds also. We omit the easy proof of the converse statement. ∎

We note that (II.2) is called Parseval's relation. The coefficients (x_α, y) are often called the **Fourier coefficients** of y with respect to the basis $\{x_\alpha\}$. The reason for this terminology will become apparent shortly.

We now describe a useful procedure, called Gram–Schmidt orthogonalization, for constructing an orthonormal set from an arbitrary sequence of independent vectors. Suppose the independent vectors $u_1, u_2, \ldots$ are given and define

$$\begin{aligned} w_1 &= u_1, & v_1 &= w_1/\|w_1\| \\ w_2 &= u_2 - (v_1, u_1)v_k, & v_2 &= w_2/\|w_2\| \\ &\vdots & &\vdots \\ w_n &= u_n - \sum_{k=1}^{n-1} (v_k, u_n)v_k, & v_n &= w_n/\|w_n\| \\ &\vdots & &\vdots \end{aligned}$$

The family $\{v_j\}$ is an orthonormal set and has the property that for each m, $\{u_j\}_{j=1}^m$ and $\{v_j\}_{j=1}^m$ span the same vector space. In particular, the set of finite linear combinations of all the v's is the same as the finite linear combinations of the u's (see Figure II.2).

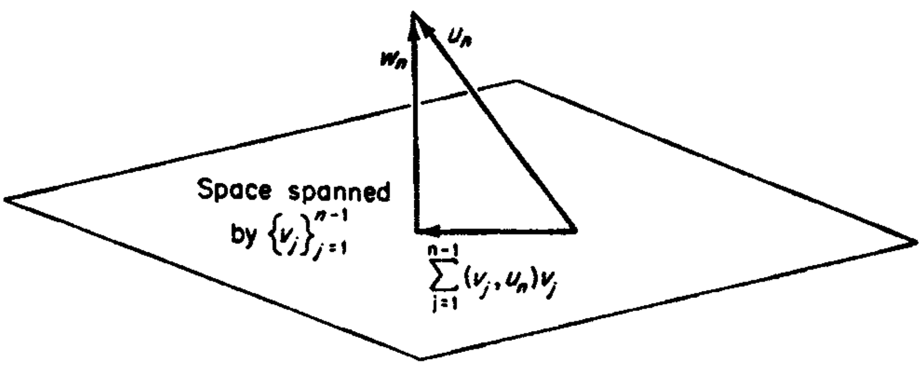

FIGURE II.2 Gram–Schmidt orthogonalization.

We remark that the Legendre polynomials (up to constant multiples) are obtained by applying the Gram–Schmidt process to the functions 1, x, x^2, $x^3, \ldots$, on the interval $[-1, 1]$ with the usual L^2 inner product.

Definition A metric space which has a countable dense subset is said to be **separable.**

Most Hilbert spaces that arise in practice are separable. The following theorem characterizes them up to isomorphism.

Theorem II.7 A Hilbert space $\mathscr{H}$ is separable if and only if it has a countable orthonormal basis S. If there are $N < \infty$ elements in S, then $\mathscr{H}$ is isomorphic to $\mathbb{C}^N$. If there are countably many elements in S, then $\mathscr{H}$ is isomorphic to ℓ_2 (Example 3, Section II.1).

Proof Suppose $\mathscr{H}$ is separable and let $\{x_n\}$ be a countable dense set. By throwing out some of the x_n's we can get a subcollection of independent vectors whose span (finite linear combinations) is the same as the $\{x_n\}$ and is thus dense. Applying the Gram–Schmidt procedure to this subcollection we obtain a countable complete orthonormal system. Conversely, if $\{y_n\}$ is a complete orthonormal system for a Hilbert space $\mathscr{H}$ then it follows from Theorem II.6 that the set of finite linear combinations of the y_n with rational coefficients is dense in $\mathscr{H}$. Since this set is countable, $\mathscr{H}$ is separable.

Suppose $\mathscr{H}$ is separable and $\{y_n\}_{n=1}^{\infty}$ is a complete orthonormal system. We define a map $\mathscr{U}: \mathscr{H} \to \ell_2$ by

$$\mathscr{U}: x \to \{(y_n, x)\}_{n=1}^{\infty}$$

Theorem II.6 shows that this map is well defined and onto. It is easy to show it is unitary. The proof that $\mathscr{H}$ is isomorphic to $\mathbb{C}^N$ if S has N elements is similar. ∎

II.4 Tensor products of Hilbert spaces

We described in Sections II.1 and II.2 several ways of making new Hilbert spaces from old ones. In this section we describe the tensor product $\mathscr{H}_1 \otimes \mathscr{H}_2$ of two Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$. The construction of the tensor product which we use is not the most elegant, but is very direct. The reader can easily extend our proofs to construct the tensor product $\mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \cdots \otimes \mathscr{H}_n$ of finitely many Hilbert spaces.

Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be Hilbert spaces. For each $\varphi_1 \in \mathscr{H}_1$, $\varphi_2 \in \mathscr{H}_2$, let $\varphi_1 \otimes \varphi_2$ denote the conjugate bilinear form which acts on $\mathscr{H}_1 \times \mathscr{H}_2$ by

$$(\varphi_1 \otimes \varphi_2)\langle \psi_1, \psi_2 \rangle = (\psi_1, \varphi_1)(\psi_2, \varphi_2)$$

Let $\mathscr{E}$ be the set of finite linear combinations of such conjugate linear forms; we define an inner product $(\cdot, \cdot)$ on $\mathscr{E}$ by defining

$$(\varphi \otimes \psi, \eta \otimes \mu) = (\varphi, \eta)(\psi, \mu)$$

and extending by linearity to $\mathscr{E}$.

Proposition 1 $(\cdot, \cdot)$ is well defined and positive definite.

Proof To show that $(\cdot, \cdot)$ is well defined, we must show that (λ, λ') does not depend on which finite linear combinations are used to express λ and λ'. To do this it is sufficient to show that if μ is a finite sum which is the zero form, then $(\eta, \mu) = 0$ for all $\eta \in \mathscr{E}$. To see that this is true, let $\eta = \sum_{i=1}^{N} c_i(\varphi_i \otimes \psi_i)$, then

$$\begin{aligned}(\eta, \mu) &= \left(\sum_{i=1}^{N} c_i(\varphi_i \otimes \psi_i), \mu\right)\\ &= \sum_{i=1}^{N} c_i \mu\langle \varphi_i, \psi_i\rangle\\ &= 0\end{aligned}$$

since μ is the zero form. Thus, $(\cdot, \cdot)$ is well defined.

Now, suppose $\lambda = \sum_{k=1}^{M} d_k(\eta_k \otimes \mu_k)$. Then $\{\eta_k\}_{k=1}^{M}$ and $\{\mu_k\}_{k=1}^{M}$ span subspaces $M_1 \subset \mathscr{H}_1$ and $M_2 \subset \mathscr{H}_2$ respectively. If we let $\{\varphi_j\}_{j=1}^{N_1}$ and $\{\psi_\ell\}_{\ell=1}^{N_2}$ be orthonormal bases for M_1 and M_2, we can express each η_k in terms of the φ_j's and each μ_k in terms of the ψ_ℓ's obtaining

$$\lambda = \sum_{\substack{j=1\\ \ell=1}}^{M_1, M_2} c_{j\ell}(\varphi_j \otimes \psi_\ell)$$

But,

$$\begin{aligned}(\lambda, \lambda) &= \left(\sum c_{j\ell}(\varphi_j \otimes \psi_\ell), \sum c_{im}(\varphi_i \otimes \psi_m)\right)\\ &= \sum \overline{c_{j\ell}}\, c_{im}(\varphi_j, \varphi_i)(\psi_\ell, \psi_m)\\ &= \sum_{j\ell} |c_{j\ell}|^2\end{aligned}$$

so if $(\lambda, \lambda) = 0$, then all the $c_{j\ell} = 0$ and λ is the zero form. Thus $(\cdot, \cdot)$ is positive definite. ▌

Definition We define $\mathscr{H}_1 \otimes \mathscr{H}_2$ to be the completion of $\mathscr{E}$ under the inner product $(\cdot, \cdot)$ defined above. $\mathscr{H}_1 \otimes \mathscr{H}_2$ is called the **tensor product** of $\mathscr{H}_1$ and $\mathscr{H}_2$.

Proposition 2 If $\{\varphi_k\}$ and $\{\psi_\ell\}$ are orthonormal bases for $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively, then $\{\varphi_k \otimes \psi_\ell\}$ is an orthonormal basis for $\mathscr{H}_1 \otimes \mathscr{H}_2$.

Proof To simplify notation, we consider the case in which both $\mathscr{H}_1$ and $\mathscr{H}_2$ are infinite dimensional and separable. The other cases are similar. The set $\{\varphi_k \otimes \psi_\ell\}$ is clearly orthonormal and therefore we need only show that $\mathscr{E}$ is contained in the closed space S spanned by $\{\varphi_k \otimes \psi_\ell\}$. Let $\varphi \otimes \psi \in \mathscr{E}$. Since $\{\varphi_k\}$ and $\{\psi_\ell\}$ are bases, $\varphi = \sum c_k \varphi_k$ and $\psi = \sum d_\ell \psi_\ell$ where $\sum |c_k|^2 < \infty$ and $\sum |d_\ell|^2 < \infty$. Thus $\sum_{k,\ell} |c_k d_\ell|^2 < \infty$. Therefore by Theorem II.6, there is a vector $\mu = \sum_{k,\ell} c_k d_\ell \varphi_k \otimes \psi_\ell$ in S. By direct computation

$$\|\varphi \otimes \psi - \sum_{\substack{k<M \\ \ell<N}} c_k d_\ell \varphi_k \otimes \psi_\ell\| \to 0$$

as $M, N \to \infty$. ∎

To show how the tensor product arises naturally, we will show how it is related to Hilbert spaces with which the reader is already familiar. First, let $\langle M_1, \mu_1 \rangle$ and $\langle M_2, \mu_2 \rangle$ be measure spaces. We suppose that $L^2(M_1, d\mu_1)$ and $L^2(M_2, d\mu_2)$ are separable (see Problems 24 and 25 of this chapter and Problem 43 of Chapter IV). Let $\{\varphi_k(x)\}$ and $\{\psi_\ell(y)\}$ be bases for $L^2(M_1, d\mu_1)$ and $L^2(M_2, d\mu_2)$ respectively. Then $\{\varphi_k(x)\psi_\ell(y)\}$ is certainly an orthonormal set in $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$. The fact that $\{\varphi_k(x)\psi_\ell(y)\}$ is actually a basis can be seen as follows. Suppose that $f(x, y) \in L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$, and

$$\iint_{M_1 \times M_2} \overline{f(x, y)}\varphi_k(x)\psi_\ell(y)\, d\mu_1(x)\, d\mu_2(y) = 0$$

for all k and ℓ. By Fubini's theorem this can be rewritten

$$\int_{M_2} \left(\int_{M_1} \overline{f(x, y)}\varphi_k(x)\, d\mu_1(x) \right) \psi_\ell(y)\, d\mu_2(y) = 0$$

Since $\{\psi_\ell\}$ is a basis for $L^2(M_2, \mu_2)$, this implies that

$$\textstyle\int_{M_1} \overline{f(x, y)}\varphi_k(x)\, d\mu_1(x) = 0$$

except on a set $S_k \subset M_2$ with $\mu_2(S_k) = 0$. Thus, for $y \notin \bigcup S_k$, $\int_{M_1} f(x, y) \times \varphi_k(x)\, d\mu_1(x) = 0$ for all k, which implies that $f(x, y) = 0$, a.e. $[\mu_1]$. Thus, $f(x, y) = 0$ a.e. $[\mu_1 \otimes \mu_2]$. So, $\{\varphi_k(x)\psi_\ell(y)\}$ is a basis for

$$L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$$

Now, let

$$U: \varphi_k \otimes \psi_\ell \to \varphi_k(x)\psi_\ell(y)$$

Then U takes an orthonormal basis for $L^2(M_1, d\mu_1) \otimes L^2(M_2, d\mu_2)$ onto an orthonormal basis for $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ and extends uniquely to a unitary mapping of

$$L^2(M_1, d\mu_1) \otimes L^2(M_2, d\mu) \quad \text{onto} \quad L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2).$$

Notice that if $f \in L^2(M_1, d\mu_1)$, $g \in L^2(M_2, d\mu_2)$, then

$$\begin{aligned} U(f \otimes g) &= U(\textstyle\sum c_k \varphi_k \otimes \sum d_\ell \psi_\ell) \\ &= U\Big(\sum_{k,\ell} c_k d_\ell \varphi_k \otimes \psi_\ell\Big) \\ &= \sum_{k,\ell} c_k d_\ell \varphi_k(x)\psi_\ell(y) \\ &= f(x)g(y) \end{aligned}$$

Because of this property, we often say that $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ and $L^2(M_1, d\mu_1) \otimes L^2(M_2, d\mu_2)$ are "naturally" isomorphic. Let $M_i = \mathbb{R}$ and $\mu_i =$ Lebesgue measure, then we have shown that $L^2(\mathbb{R}^2)$ is naturally isomorphic to $L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$.

Let us return for a moment to Example 6 of Section II.1: $\langle M, \mu \rangle$ is a measure space and $\mathscr{H}'$ a separable Hilbert space with basis $\{\varphi_k\}$. In Problem 12, the reader is asked to show that each $g \in L^2(M, d\mu; \mathscr{H}')$ is a limit

$$g(x) = \lim_{N \to \infty} \sum_{k=1}^{N} (\varphi_k, g(x))_{\mathscr{H}'} \varphi_k$$

of finite linear combinations of vectors of the form $f_k(x)\varphi_k$, $f_k(x) \in L^2(M, d\mu)$. We now define

$$U\colon \sum_{k=1}^{N} f_k(x) \otimes \varphi_k \to \sum_{k=1}^{N} f_k(x)\varphi_k$$

Then U is a well-defined map from a dense set in $L^2(M, d\mu) \otimes \mathscr{H}'$ onto a dense set in $L^2(M, d\mu; \mathscr{H}')$ which preserves norms, so U extends uniquely to a unitary operator from $L^2(M, d\mu) \otimes \mathscr{H}'$ to $L^2(M, d\mu; \mathscr{H}')$. Notice that under this map, $U(f(x) \otimes \varphi) = f(x)\varphi$ for all $\varphi \in \mathscr{H}'$. In this sense, U is called the natural isomorphism between $L^2(M, d\mu) \otimes \mathscr{H}'$ and $L^2(M, d\mu; \mathscr{H}')$. We summarize this discussion in a theorem:

Theorem II.10 Let $\langle M_1, \mu_1 \rangle$ and $\langle M_2, \mu_2 \rangle$ be measure spaces so that $L^2(M_1, d\mu_1)$ and $L^2(M_2, d\mu_2)$ are separable. Then

(a) There is a unique isomorphism from $L^2(M_1, d\mu_1) \otimes L^2(M_2, d\mu_2)$ to $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ so that $f \otimes g \mapsto fg$.

(b) If $\mathscr{H}'$ is a separable Hilbert space, then there is a unique isomorphism from $L^2(M_1, d\mu_1) \otimes \mathscr{H}'$ to $L^2(M_1, d\mu_1; \mathscr{H}')$ so that $f(x) \otimes \varphi \mapsto f(x)\varphi$.

(c) There is a unique isomorphism from $L^2(M_1 \times M_2, d\mu_1 \otimes d\mu_2)$ to $L^2(M_1, d\mu_1; L^2(M_2, d\mu_2))$ such that $f(x, y)$ is taken into the function $x \mapsto f(x, \cdot)$.

1.5. The C^* algebra of bounded linear operators

We start by introducing a conjugation for operators on a Hilbert space $\mathfrak{H}$. Let $A \in \mathfrak{L}(\mathfrak{H})$. Then the **adjoint operator** is defined via

$$\langle \varphi, A^* \psi \rangle = \langle A\varphi, \psi \rangle \tag{1.45}$$

(compare Corollary 1.9).

Example. If $\mathfrak{H} = \mathbb{C}^n$ and $A = (a_{jk})_{1 \le j,k \le n}$, then $A^* = (a_{kj}^*)_{1 \le j,k \le n}$. ⋄

Lemma 1.11. *Let $A, B \in \mathfrak{L}(\mathfrak{H})$. Then*

(i) $(A + B)^* = A^* + B^*$, $\qquad (\alpha A)^* = \alpha^* A^*$,

(ii) $A^{**} = A$,

(iii) $(AB)^* = B^* A^*$,

(iv) $\|A\| = \|A^*\|$ *and* $\|A\|^2 = \|A^* A\| = \|A A^*\|$.

Proof. (i) and (ii) are obvious. (iii) follows from $\langle \varphi, (AB)\psi \rangle = \langle A^* \varphi, B\psi \rangle = \langle B^* A^* \varphi, \psi \rangle$. (iv) follows from

$$\|A^*\| = \sup_{\|\varphi\| = \|\psi\| = 1} |\langle \psi, A^* \varphi \rangle| = \sup_{\|\varphi\| = \|\psi\| = 1} |\langle A\psi, \varphi \rangle| = \|A\|$$

and

$$\begin{aligned}\|A^*A\| &= \sup_{\|\varphi\|=\|\psi\|=1} |\langle\varphi, A^*A\psi\rangle| = \sup_{\|\varphi\|=\|\psi\|=1} |\langle A\varphi, A\psi\rangle| \\ &= \sup_{\|\varphi\|=1} \|A\varphi\|^2 = \|A\|^2,\end{aligned}$$

where we have used $\|\varphi\| = \sup_{\|\psi\|=1} |\langle\psi, \varphi\rangle|$. □

As a consequence of $\|A^*\| = \|A\|$ observe that taking the adjoint is continuous.

In general, a Banach algebra $\mathcal{A}$ together with an **involution**

$$(a+b)^* = a^* + b^*, \quad (\alpha a)^* = \alpha^* a^*, \quad a^{**} = a, \quad (ab)^* = b^* a^* \tag{1.46}$$

satisfying

$$\|a\|^2 = \|a^* a\| \tag{1.47}$$

is called a C^* **algebra**. The element a^* is called the adjoint of a. Note that $\|a^*\| = \|a\|$ follows from (1.47) and $\|aa^*\| \le \|a\|\|a^*\|$.

Any subalgebra which is also closed under involution is called a $*$-subalgebra. An **ideal** is a subspace $\mathcal{I} \subseteq \mathcal{A}$ such that $a \in \mathcal{I}$, $b \in \mathcal{A}$ imply $ab \in \mathcal{I}$ and $ba \in \mathcal{I}$. If it is closed under the adjoint map, it is called a $*$-ideal. Note that if there is an identity e, we have $e^* = e$ and hence $(a^{-1})^* = (a^*)^{-1}$ (show this).

Example. The continuous functions $C(I)$ together with complex conjugation form a commutative C^* algebra. ⋄

An element $a \in \mathcal{A}$ is called **normal** if $aa^* = a^*a$, **self-adjoint** if $a = a^*$, **unitary** if $aa^* = a^*a = \mathbb{I}$, an (orthogonal) **projection** if $a = a^* = a^2$, and **positive** if $a = bb^*$ for some $b \in \mathcal{A}$. Clearly both self-adjoint and unitary elements are normal.

Problem 1.14. *Let $A \in \mathfrak{L}(\mathfrak{H})$. Show that A is normal if and only if*

$$\|A\psi\| = \|A^*\psi\|, \quad \forall \psi \in \mathfrak{H}.$$

(Hint: Problem 0.14.)

Problem 1.15. *Show that $U : \mathfrak{H} \to \mathfrak{H}$ is unitary if and only if $U^{-1} = U^*$.*

Problem 1.16. *Compute the adjoint of*

$$S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N}), \qquad (a_1, a_2, a_3, \dots) \mapsto (0, a_1, a_2, \dots).$$

Recall that the interior of a set is the largest open subset (that is, the union of all open subsets). A set is called **nowhere dense** if its closure has empty interior. The key to several important theorems about Banach spaces is the observation that a Banach space cannot be the countable union of nowhere dense sets.

Theorem 0.38 (Baire category theorem)**.** *Let X be a complete metric space. Then X cannot be the countable union of nowhere dense sets.*

Proof. Suppose $X = \bigcup_{n=1}^{\infty} X_n$. We can assume that the sets X_n are closed and none of them contains a ball; that is, $X \backslash X_n$ is open and nonempty for every n. We will construct a Cauchy sequence x_n which stays away from all X_n.

Since $X \backslash X_1$ is open and nonempty, there is a closed ball $B_{r_1}(x_1) \subseteq X \backslash X_1$. Reducing r_1 a little, we can even assume $\overline{B_{r_1}(x_1)} \subseteq X \backslash X_1$. Moreover, since X_2 cannot contain $B_{r_1}(x_1)$, there is some $x_2 \in B_{r_1}(x_1)$ that is not in X_2. Since $B_{r_1}(x_1) \cap (X \backslash X_2)$ is open, there is a closed ball $\overline{B_{r_2}(x_2)} \subseteq B_{r_1}(x_1) \cap (X \backslash X_2)$. Proceeding by induction, we obtain a sequence of balls such that

$$\overline{B_{r_n}(x_n)} \subseteq B_{r_{n-1}}(x_{n-1}) \cap (X \backslash X_n).$$

Now observe that in every step we can choose r_n as small as we please; hence without loss of generality $r_n \to 0$. Since by construction $x_n \in \overline{B_{r_N}(x_N)}$ for $n \geq N$, we conclude that x_n is Cauchy and converges to some point $x \in X$. But $x \in \overline{B_{r_n}(x_n)} \subseteq X \backslash X_n$ for every n, contradicting our assumption that the X_n cover X. □

(Sets which can be written as the countable union of nowhere dense sets are said to be of first category. All other sets are second category. Hence we have the name category theorem.)

In other words, if $X_n \subseteq X$ is a sequence of closed subsets which cover X, at least one X_n contains a ball of radius $\varepsilon > 0$.

Now we come to the first important consequence, the **uniform boundedness principle**.

Theorem 0.39 (Banach–Steinhaus). *Let X be a Banach space and Y some normed linear space. Let $\{A_\alpha\} \subseteq \mathfrak{L}(X,Y)$ be a family of bounded operators. Suppose $\|A_\alpha x\| \le C(x)$ is bounded for fixed $x \in X$. Then $\|A_\alpha\| \le C$ is uniformly bounded.*

Proof. Let

$$X_n = \{x |\ \|A_\alpha x\| \le n \text{ for all } \alpha\} = \bigcap_\alpha \{x |\ \|A_\alpha x\| \le n\}.$$

Then $\bigcup_n X_n = X$ by assumption. Moreover, by continuity of A_α and the norm, each X_n is an intersection of closed sets and hence closed. By Baire's theorem at least one contains a ball of positive radius: $B_\varepsilon(x_0) \subset X_n$. Now observe

$$\|A_\alpha y\| \le \|A_\alpha(y + x_0)\| + \|A_\alpha x_0\| \le n + \|A_\alpha x_0\|$$

for $\|y\| < \varepsilon$. Setting $y = \varepsilon \frac{x}{\|x\|}$, we obtain

$$\|A_\alpha x\| \le \frac{n + C(x_0)}{\varepsilon} \|x\|$$

for any x. □

1.6. Weak and strong convergence

Sometimes a weaker notion of convergence is useful: We say that ψ_n **converges weakly** to ψ and write

$$\operatorname*{w-lim}_{n\to\infty} \psi_n = \psi \quad \text{or} \quad \psi_n \rightharpoonup \psi \tag{1.48}$$

if $\langle\varphi, \psi_n\rangle \to \langle\varphi, \psi\rangle$ for every $\varphi \in \mathfrak{H}$ (show that a weak limit is unique).

Example. Let φ_n be an (infinite) orthonormal set. Then $\langle\psi, \varphi_n\rangle \to 0$ for every ψ since these are just the expansion coefficients of ψ. (φ_n does not converge to 0, since $\|\varphi_n\| = 1$.) ◇

Clearly $\psi_n \to \psi$ implies $\psi_n \rightharpoonup \psi$ and hence this notion of convergence is indeed weaker. Moreover, the weak limit is unique, since $\langle\varphi, \psi_n\rangle \to \langle\varphi, \psi\rangle$ and $\langle\varphi, \psi_n\rangle \to \langle\varphi, \tilde{\psi}\rangle$ imply $\langle\varphi, (\psi - \tilde{\psi})\rangle = 0$. A sequence ψ_n is called a **weak Cauchy sequence** if $\langle\varphi, \psi_n\rangle$ is Cauchy for every $\varphi \in \mathfrak{H}$.

Lemma 1.12. *Let $\mathfrak{H}$ be a Hilbert space.*

(i) $\psi_n \rightharpoonup \psi$ *implies* $\|\psi\| \le \liminf \|\psi_n\|$.

(ii) *Every weak Cauchy sequence ψ_n is bounded:* $\|\psi_n\| \le C$.

(iii) *Every weak Cauchy sequence converges weakly.*

(iv) *For a weakly convergent sequence $\psi_n \rightharpoonup \psi$ we have $\psi_n \to \psi$ if and only if* $\limsup \|\psi_n\| \le \|\psi\|$.

Proof. (i) Observe

$$\|\psi\|^2 = \langle\psi, \psi\rangle = \liminf \langle\psi, \psi_n\rangle \le \|\psi\| \liminf \|\psi_n\|.$$

(ii) For every φ we have that $|\langle\varphi, \psi_n\rangle| \le C(\varphi)$ is bounded. Hence by the uniform boundedness principle we have $\|\psi_n\| = \|\langle\psi_n, .\rangle\| \le C$.
(iii) Let φ_m be an orthonormal basis and define $c_m = \lim_{n\to\infty}\langle\varphi_m, \psi_n\rangle$. Then $\psi = \sum_m c_m\varphi_m$ is the desired limit.
(iv) By (i) we have $\lim \|\psi_n\| = \|\psi\|$ and hence

$$\|\psi - \psi_n\|^2 = \|\psi\|^2 - 2\operatorname{Re}(\langle\psi, \psi_n\rangle) + \|\psi_n\|^2 \to 0.$$

The converse is straightforward. □

Clearly an orthonormal basis does not have a norm convergent subsequence. Hence the unit ball in an infinite dimensional Hilbert space is never compact. However, we can at least extract weakly convergent subsequences:

Lemma 1.13. *Let $\mathfrak{H}$ be a Hilbert space. Every bounded sequence ψ_n has a weakly convergent subsequence.*

The second argument which we refer to as the "diagonal sequence trick" is illustrated in:

Theorem I.24 Let $f_n(m)$ be a sequence of functions on the positive integers which is uniformly bounded, i.e. $|f_n(m)| \leq C$ for all n, m. Then there is a subsequence $\{f_{\hat{n}(i)}(m)\}_{i=1}^{\infty}$ so that for each fixed m, $f_{\hat{n}(i)}(m)$ converges as $i \to \infty$.

Proof Consider the sequence $f_n(1)$. It is a bounded set of numbers, so we can find a subsequence $f_{n_1(i)}$ so $f_{n_1(i)}(1) \to f_\infty(1)$, for some number $f_\infty(1)$. Now consider the sequence $f_{n_1(i)}(2)$. We can find a subsequence $f_{n_2(i)}(2) \to f_\infty(2)$ as $i \to \infty$. Proceeding inductively, we find successive subsequences, $f_{n_k(i)}$ so that (a) $f_{n_{k+1}(i)}$ is a subsequence of $f_{n_k(i)}$ and (b) $f_{n_k(i)}(k) \to f_\infty(k)$ as $i \to \infty$. Thus, in particular, $f_{n_k(i)}(j) \to f_\infty(j)$ as $i \to \infty$ for $j = 1, 2, \ldots, k$. To get a subsequence $f_{\hat{n}(i)}$ converging for each j, one is tempted to try to take the limit of the horizontal sequence (see Figure I.8a) but that won't work! (for it may happen $n_k(1) \to \infty$). The simple way out is to take the diagonal sequence $\hat{n}(k) = n_k(k)$. Then $f_{\hat{n}(k)}, f_{\hat{n}(k+1)}, \ldots$ is a subsequence of $f_{n_k(i)}$ so $f_{\hat{n}(i)}(k) \to f_\infty(k)$ as $i \to \infty$ for any k. ∎

Proof. Let φ_k be an orthonormal basis. Then by the usual diagonal sequence argument we can find a subsequence ψ_{n_m} such that $\langle\varphi_k, \psi_{n_m}\rangle$ converges for all k. Since ψ_n is bounded, $\langle\varphi, \psi_{n_m}\rangle$ converges for every $\varphi \in \mathfrak{H}$ and hence ψ_{n_m} is a weak Cauchy sequence. □

Finally, let me remark that similar concepts can be introduced for operators. This is of particular importance for the case of unbounded operators, where convergence in the operator norm makes no sense at all.

A sequence of operators A_n is said to **converge strongly** to A,

$$\operatorname*{s-lim}_{n\to\infty} A_n = A \quad :\Leftrightarrow \quad A_n\psi \to A\psi \quad \forall x \in \mathfrak{D}(A) \subseteq \mathfrak{D}(A_n). \tag{1.49}$$

It is said to **converge weakly** to A,

$$\operatorname*{w-lim}_{n\to\infty} A_n = A \quad :\Leftrightarrow \quad A_n\psi \rightharpoonup A\psi \quad \forall \psi \in \mathfrak{D}(A) \subseteq \mathfrak{D}(A_n). \tag{1.50}$$

Clearly norm convergence implies strong convergence and strong convergence implies weak convergence.

Example. Consider the operator $S_n \in \mathfrak{L}(\ell^2(\mathbb{N}))$ which shifts a sequence n places to the left, that is,

$$S_n\,(x_1, x_2, \dots) = (x_{n+1}, x_{n+2}, \dots), \tag{1.51}$$

and the operator $S_n^* \in \mathfrak{L}(\ell^2(\mathbb{N}))$ which shifts a sequence n places to the right and fills up the first n places with zeros, that is,

$$S_n^*\,(x_1, x_2, \dots) = (\underbrace{0, \dots, 0}_{n \text{ places}}, x_1, x_2, \dots). \tag{1.52}$$

Then S_n converges to zero strongly but not in norm (since $\|S_n\| = 1$) and S_n^* converges weakly to zero (since $\langle\varphi, S_n^*\psi\rangle = \langle S_n\varphi, \psi\rangle$) but not strongly (since $\|S_n^*\psi\| = \|\psi\|$) . ◇

Note that this example also shows that taking adjoints is not continuous with respect to strong convergence! If $A_n \xrightarrow{s} A$, we only have

$$\langle\varphi, A_n^*\psi\rangle = \langle A_n\varphi, \psi\rangle \to \langle A\varphi, \psi\rangle = \langle\varphi, A^*\psi\rangle \tag{1.53}$$

and hence $A_n^* \rightharpoonup A^*$ in general. However, if A_n and A are normal, we have

$$\|(A_n - A)^*\psi\| = \|(A_n - A)\psi\| \tag{1.54}$$

and hence $A_n^* \xrightarrow{s} A^*$ in this case. Thus at least for normal operators taking adjoints is continuous with respect to strong convergence.

Lemma 1.14. *Suppose A_n is a sequence of bounded operators.*

(i) $\operatorname{s-lim}_{n\to\infty} A_n = A$ *implies* $\|A\| \le \liminf_{n\to\infty} \|A_n\|$.

(ii) *Every strong Cauchy sequence A_n is bounded:* $\|A_n\| \le C$.

(iii) *If $A_n\psi \to A\psi$ for ψ in some dense set and $\|A_n\| \le C$, then* $\text{s-lim}_{n\to\infty} A_n = A$.

The same result holds if strong convergence is replaced by weak convergence.

Proof. (i) follows from

$$\|A\psi\| = \lim_{n\to\infty} \|A_n\psi\| \le \liminf_{n\to\infty} \|A_n\|$$

for every ψ with $\|\psi\| = 1$.
(ii) follows as in Lemma 1.12 (i).
(iii) Just use

$$\begin{aligned}\|A_n\psi - A\psi\| &\le \|A_n\psi - A_n\varphi\| + \|A_n\varphi - A\varphi\| + \|A\varphi - A\psi\| \\ &\le 2C\|\psi - \varphi\| + \|A_n\varphi - A\varphi\|\end{aligned}$$

and choose φ in the dense subspace such that $\|\psi - \varphi\| \le \frac{\varepsilon}{4C}$ and n large such that $\|A_n\varphi - A\varphi\| \le \frac{\varepsilon}{2}$.

The case of weak convergence is left as an exercise. (Hint: (2.14).) □

Problem 1.17. *Suppose $\psi_n \to \psi$ and $\varphi_n \rightharpoonup \varphi$. Then $\langle\psi_n, \varphi_n\rangle \to \langle\psi, \varphi\rangle$.*

Problem 1.18. *Let $\{\varphi_j\}_{j=1}^\infty$ be some orthonormal basis. Show that $\psi_n \rightharpoonup \psi$ if and only if ψ_n is bounded and $\langle\varphi_j, \psi_n\rangle \to \langle\varphi_j, \psi\rangle$ for every j. Show that this is wrong without the boundedness assumption.*

Problem 1.19. *A subspace $M \subseteq \mathfrak{H}$ is closed if and only if every weak Cauchy sequence in M has a limit in M. (Hint: $\overline{M} = M^{\perp\perp}$.)*