XII. Hilbert and Banach Spaces, 570-602

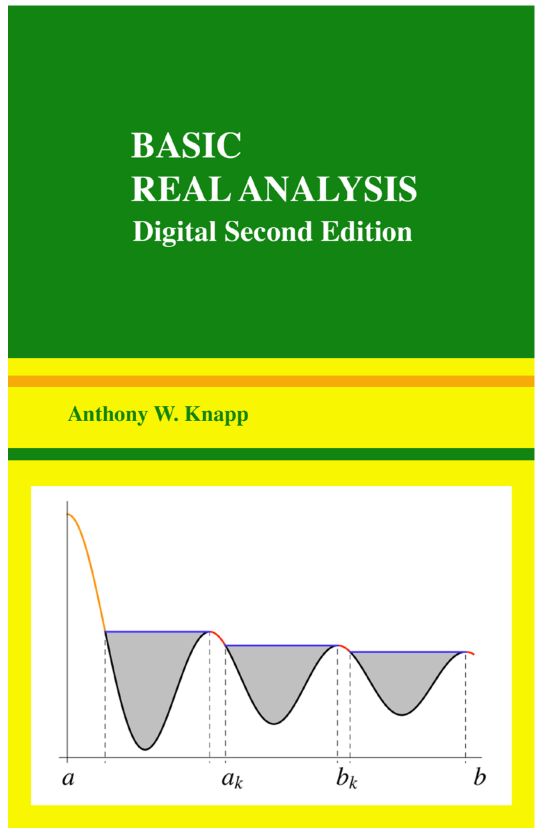
DOI: 10.3792/euclid/9781429799997-12

from

Basic Real Analysis
Digital Second Edition

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799997
ISBN: 978-1-4297-9999-7



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: `aknapp@math.stonybrook.edu`
Homepage: `www.math.stonybrook.edu/~aknapp`

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"
Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28–01, 26–01, 42–01, 54–01, 34–01, 30–01, 32–01.

First Edition, ISBN-13 978-0-8176-3250-2

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the author

CHAPTER XII

Hilbert and Banach Spaces

Abstract. This chapter develops the beginnings of abstract functional analysis, a subject designed to study properties of functions by treating the functions as the members of a space and formulating the properties as properties of the space.

Section 1 defines Banach spaces as complete normed linear spaces and gives a number of examples of these. The space of bounded linear operators from one normed linear space to another is a normed linear space, and it is a Banach space if the range is a Banach space.

Sections 2–3 concern Hilbert spaces. These are Banach spaces whose norms are induced by inner products. Section 2 shows that closed vector subspaces of such a space have orthogonal complements, and it shows the role of orthonormal bases for such a space. Section 3 concentrates on bounded linear operators from a Hilbert space to itself and constructs the adjoint of each such operator.

Sections 4–6 prove the three main abstract theorems about the norm topology of general normed linear spaces—the Hahn–Banach Theorem, the Uniform Boundedness or Banach–Steinhaus Theorem, and the Interior Mapping Principle. A number of consequences of these theorems are given. The second and third of the theorems require some hypothesis of completeness.

The topic of Hilbert and Banach spaces continues in Chapter IV of the companion volume, *Advanced Real Analysis*.

1. Definitions and Examples

Functional analysis puts into practice an idea from the early twentieth century, that sometimes properties of functions become clearer when the functions are regarded as the members of a space and the properties are formulated as properties of the space. We encountered some simple examples of this situation already in Chapter II in the examples of metric spaces. Uniform convergence was encoded in the metric on spaces of functions, and other kinds of convergence were captured by other metrics. In Chapter V we introduced the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ of functions (or really equivalence classes of functions), all of which were proved to be complete. The property of completeness was a useful property of the space as a whole that led, for one thing, to the Riesz–Fischer Theorem in Chapter VI. More complicated properties led us to various kinds of differentiability of integrals in $\mathbb{R}^n$ in Chapters VI and IX and to boundedness of the Hilbert transform in Chapter IX. The development of measure theory on locally compact Hausdorff spaces in Chapter XI rested on an analysis of positive linear functionals on the space of continuous functions of compact support.

The different spaces—of functions, measures, and whatever else—that arise in this way have some properties in common, and we study them in this chapter in a setting that emphasizes these common properties. We shall work with normed linear spaces, which were defined in Section V.9. With such spaces the field of scalars $\mathbb{F}$ can be either $\mathbb{R}$ or $\mathbb{C}$. Recall then that a **normed linear space** X is a vector space over $\mathbb{F}$ with a **norm**, i.e., a function $\|\cdot\|$ from X to $[0, +\infty)$ such that $\|x\| \geq 0$ with equality if and only if $x = 0$, $\|cx\| = |c|\|x\|$ if c is a scalar, and $\|x + y\| \leq \|x\| + \|y\|$. The norm yields a metric $d(x, y) = \|x - y\|$, and we can then speak of the norm topology on X. Proposition 5.55 showed that addition and scalar multiplication are continuous, that the closure of any vector subspace of X is a vector subspace, and that the set of all finite linear combinations of members of a subset S of X is dense in the smallest closed subspace containing S.

Completeness plays an increasingly important role as one studies such spaces, and it is customary to introduce a definition to incorporate this notion: a normed linear space X is a **Banach space** if X is complete as a metric space. The metric-space completion of a normed linear space is automatically a normed linear space that is complete, hence is a Banach space.

Let us consider some examples of normed linear spaces, some old and some new. Except as indicated, they will all be Banach spaces.

EXAMPLES.

(1) Euclidean space $\mathbb{R}^n$ and complex Euclidean space $\mathbb{C}^n$, written briefly as $\mathbb{F}^n$. The space consists of n-tuples of scalars $a = (a_1, \dots, a_n)$ with $\|a\|$ equal to the Euclidean norm $|a|$ of Section II.1, namely $\|a\| = \big(\sum_{k=1}^n |a_k|\big)^{1/2}$. It was remarked in Section II.7 that these spaces are complete, hence are Banach spaces.

(2) Finite-dimensional normed linear spaces. It can be shown that each finite-dimensional normed linear space X is complete.[1] In fact, any linear map carrying a vector-space basis of X to a vector-space basis of some $\mathbb{F}^n$, normed as in the previous example, can be shown to be uniformly continuous with a uniformly continuous inverse, and the completeness of X follows.

(3) $B(S)$, the space of bounded scalar-valued functions on a nonempty set S with the supremum norm, defined in Section II.1. Proposition 2.44 establishes the completeness.

(4) $C(S)$, the space of bounded continuous scalar-valued functions on a metric space or topological space S, defined in Section II.4 in the metric case and Section X.5 in general. The norm is the supremum norm. Corollary 2.45 and Proposition 10.30 establish the completeness of $C(S)$. When S is locally compact Hausdorff,

[1] Section IV.1 of the companion volume, *Advanced Real Analysis*, proves a more general result.

we defined $C_0(S)$ in Section XI.4 to be the subspace of $C(S)$ of all members vanishing at infinity. This is complete. However, the subspace $C_{\text{com}}(S)$ of continuous scalar-valued functions of compact support is usually not complete.

(5) $L^p(S, \mathcal{A}, \mu)$, the space of equivalence classes of p^{th}-power integrable functions on a measure space $(S, \mathcal{A}, \mu)$. This is a normed linear space for $1 \le p < \infty$ with norm $\|f\|_p = \big(\int_S |f(s)|^p \, d\mu(s)\big)^{1/p}$. These spaces were introduced in Section V.9 for $p = 1$ and $p = 2$ and in Section IX.1 for general p. Theorem 5.59 established the completeness for $p = 1$ and $p = 2$, and Theorem 9.6 established the completeness for general p.

(6) $L^\infty(S, \mathcal{A}, \mu)$, the space of equivalence classes of essentially bounded functions on a measure space $(S, \mathcal{A}, \mu)$. This is a normed linear space with norm the essential supremum norm. This space was introduced in Section V.9 and was proved to be complete in Theorem 5.59.

(7) Sequence spaces c, c_0, and ℓ_n^p and ℓ^p for $1 \le p \le \infty$. These are special cases of various examples above. The space ℓ_n^p is $L^p(S, \mathcal{A}, \mu)$ when $S = \{1, 2, \dots, n\}$, $\mathcal{A}$ is the set of all subsets, and μ is counting measure, the norm being $\|(a_1, \dots, a_n)\| = \big(\sum_{k=1}^n |a_k|^p\big)^{1/p}$ if $p < \infty$ and being $\|(a_1, \dots, a_n)\| = \max_{1 \le k \le n} |a_k|$ if $p = \infty$. The space ℓ_n^p specializes to $\mathbb{F}^n$ when $p = 2$. The space ℓ^p is the version of ℓ_n^p when S is the set of positive integers; the members of this space are thus all sequences for which the norm is finite. The sequence spaces c and c_0 can be regarded as subspaces of $C(S)$ when S is the set of positive integers. The space c consists of all convergent sequences, and c_0 is the space of sequences vanishing at infinity; in both cases the norm is the supremum norm. All these examples are Banach spaces. They tend to be useful in testing guesses about properties of normed linear spaces. We shall not need them explicitly, and this traditional notation for them will not recur after the end of this section.

(8) $M(S)$, S being a compact Hausdorff space. This is the space of regular Borel signed or complex measures on S, introduced as $M(S, \mathbb{R})$ or $M(S, \mathbb{C})$ in Section XI.4. The norm is the total-variation norm. Theorems 11.26 and 11.28 identify these spaces with duals of spaces of continuous functions, and Proposition 12.1 below will show that they are complete as a consequence.

(9) $C^N([a, b])$, the space of scalar-valued functions on a bounded interval $[a, b]$ with N bounded derivatives, the norm being

$$\|f\| = \sum_{j=1}^{N} \sup_{a \le s \le b} |f^{(j)}(s)|.$$

It is shown in Problem 2 at the end of the chapter that this space is complete. This space is an indication of how normed linear spaces can carry information about derivatives. Indeed, normed linear spaces carrying information about derivatives play a significant role in the subject of partial differential equations.[2]

(10) $H^\infty(D)$, the space of bounded functions in the open unit disk $D = \{|z| < 1\}$ in $\mathbb{C}$ such that the function is given by a convergent power series. The norm is the supremum norm. It is shown in Problem 3 at the end of the chapter that this space is complete.

(11) $A(D)$, the space of bounded continuous functions on the closed unit disk whose restriction to the open unit disk is given by a convergent power series. The norm is the supremum norm. It is shown in Problem 3 at the end of the chapter that this space is complete.

Two further kinds of normed linear spaces are worth mentioning now. One is that any real or complex inner-product space X in the sense of Section II.1 gives an example of a normed linear space. Recall that an **inner product** on X is a function $(\cdot\,,\,\cdot)$ from $X \times X$ to $\mathbb{F}$ that is linear in the first variable, is conjugate linear in the second variable, is symmetric if $\mathbb{F} = \mathbb{R}$ or Hermitian symmetric if $\mathbb{F} = \mathbb{C}$, and has $(x, x) \geq 0$ for all x with equality if and only if $x = 0$. Such an inner product satisfies the Schwarz inequality $|(x, y)| \leq (x, x)^{1/2}(y, y)^{1/2}$, according to Lemma 2.2, and then the definition $\|x\| = (x, x)^{1/2}$ makes X into a normed linear space, according to Proposition 2.3.

As a normed linear space, an inner-product space may or may not be complete. Any space $L^2(S, \mathcal{A}, \mu)$, with $(f, g) = \int_S f\bar{g}\,d\mu$, is an example in which the associated normed linear space is complete. An inner-product space whose associated normed linear space is complete is called a **Hilbert space**.

The other kind of normed linear space worth mentioning now involves bounded linear operators. Recall from Section V.9 that a linear function $L : X \to Y$ between two normed linear spaces with respective norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ is often called a **linear operator**. Proposition 5.57 showed that a linear operator L is continuous at a point if and only if it is continuous everywhere, if and only if it is uniformly continuous, if and only if it is **bounded** in the sense that $\|L(x)\|_Y \leq M\|x\|_X$ for some constant M and all x in X. The least such constant M is called the **operator norm** of L, written $\|L\|$. We can define addition and scalar multiplication on bounded linear operators from X to Y by addition and scalar multiplication of their values:

$$(L_1 + L_2)(x) = L_1(x) + L_2(x) \qquad \text{and} \qquad (cL)(x) = cL(x).$$

Then $L_1 + L_2$ and cL are linear operators by the elementary theory of vector

[2]This is one of the themes of the companion volume, *Advanced Real Analysis*.

spaces, and the inequalities

$$\begin{aligned}\|(L_1+L_2)(x)\|_Y = \|L_1(x)+L_2(x)\|_Y &\le \|L_1(x)\|_Y + \|L_2(x)\|_Y \\ &\le \|L_1\|\|x\|_X + \|L_2\|\|x\|_X = (\|L_1\|+\|L_2\|)\|x\|_X\end{aligned}$$

and $$\|(cL)(x)\|_Y = \|cL(x)\|_Y = |c|\|L(x)\|_Y \le |c|\|L\|\|x\|_X$$

show that L_1+L_2 and cL are bounded with $\|L_1+L_2\| \le \|L_1\|+\|L_2\|$ and $\|cL\| \le |c|\|L\|$. Applying the latter conclusion to c^{-1} when $c \ne 0$ gives $\|L\| = \|c^{-1}(cL)\| \le |c|^{-1}\|cL\| \le |c|^{-1}|c|\|L\| = \|L\|$, and we conclude that $\|cL\| = |c|\|L\|$. Since it is plain that $\|L\| \ge 0$ with equality if and only if $L=0$, the set of bounded linear operators from X to Y, with the operator norm, is a normed linear space. We denote this normed linear space by $\mathcal{B}(X, Y)$.

Proposition 12.1. If X and Y are normed linear spaces and if Y is complete, then the normed linear space $\mathcal{B}(X, Y)$ is a Banach space.

REMARKS. In the special case in which Y is the set $\mathbb{F}$ of scalars, the linear operators are called **linear functionals**, in terminology we have used repeatedly. The normed linear space $\mathbb{F} = \mathbb{F}^1$ is complete, and therefore the normed linear space of bounded linear functionals on X is a Banach space. The space of bounded linear functionals is called the **dual space** of X and is denoted by X^*. More explicitly the norm of an element x^* of X^* is[3]

$$\|x^*\| = \sup_{\|x\|\le 1} |x^*(x)|.$$

Proposition 12.1 is implicitly saying that X^* is always complete.

PROOF. Let $\{L_n\}$ be a Cauchy sequence in $\mathcal{B}(X, Y)$. Since in any metric space the members of a Cauchy sequence are at a bounded distance from any particular element, the sequence $\{\|L_n\|\}$ is bounded. Let $C = \sup_n \|L_n\|$.

If x is in X, then $\{L_n(x)\}$ is a Cauchy sequence since $\|L_m(x) - L_n(x)\|_Y \le \|L_m - L_n\|\|x\|_X$. By completeness of Y, $L(x) = \lim_n L_n(x)$ exists. Continuity of addition and scalar multiplication in X implies that $L(x+x') = \lim_n L_n(x+x') = \lim_n(L_n(x) + L_n(x')) = \lim_n L_n(x) + \lim_n L_n(x') = L(x) + L(x')$ and that $L(cx) = \lim_n L_n(cx) = \lim_n(cL_n(x)) = c\lim_n L_n(x) = cL(x)$. Therefore L is a linear operator.

[3] A superscript * has also been used in this book to indicate a one-point compactification, but there need never be any confusion about this notation. One-point compactifications arise in practice only for locally compact Hausdorff spaces, and one can show that a normed linear space is locally compact only if it is finite dimensional, For finite-dimensional normed linear spaces it is always clear from the context whether * refers to the dual space or to the one-point compactification.

For boundedness of L, we have $\|L_n(x)\|_Y \le \|L_n\|\|x\|_X \le C\|x\|_X$ for all n. Hence continuity of the norm function implies that $\|L(x)\|_Y = \|\lim L_n(x)\|_Y \le \liminf_n \|L_n(x)\|_Y \le C\|x\|_X$, and L is bounded with $\|L\| \le C$.

To complete the proof, we show that $\|L_n - L\| \to 0$. Assuming the contrary, we can pass to a subsequence and then change notation so that $\|L_n - L\| \ge \epsilon$ for some $\epsilon > 0$ for all n. Then for each n, we can find x_n in X with $\|x_n\|_X = 1$ such that $\|L_n(x_n) - L(x_n)\|_Y \ge \epsilon/2$. Choose and fix N so that $m \ge N$ implies $\|L_N - L_m\| \le \epsilon/4$. Whenever $m \ge N$, the triangle inequality gives

$$\begin{aligned}\|L_m(x_N) - L(x_N)\|_Y &\ge \|L_N(x_N) - L(x_N)\|_Y - \|L_N(x_N) - L_m(x_N)\|_Y \\ &\ge \tfrac{\epsilon}{2} - \|L_N - L_m\|\|x_N\|_X = \tfrac{\epsilon}{2} - \|L_N - L_m\| \ge \tfrac{\epsilon}{4},\end{aligned}$$

in contradiction to the fact that $\lim_m L_m(x_N) = L(x_N)$. □

EXAMPLES OF DUAL SPACES.

(1) $L^p(S, \mathcal{A}, \mu)^* \cong L^{p'}(S, \mathcal{A}, \mu)$ if $1 \le p < \infty$, μ is σ-finite, and p' is the **dual index** with $\frac{1}{p} + \frac{1}{p'} = 1$, according to the Riesz Representation Theorem (Theorem 9.19). Specifically to each x^* in $L^p(S, \mathcal{A}, \mu)^*$ corresponds a unique g in $L^{p'}(S, \mathcal{A}, \mu)$ with $x^*(f) = \int_S fg\,d\mu$ for all f in $L^p(S, \mathcal{A}, \mu)$, and this g has $\|x^*\| = \|g\|_{p'}$. It can be shown that the hypothesis of σ-finiteness of μ can be dropped if $1 < p < \infty$, but Problem 4 at the end of Chapter IX shows that the hypothesis cannot be completely dropped for $p = 1$.

(2) $(\ell_n^p)^* \cong \ell_n^{p'}$ and $(\ell^p)^* \cong \ell^{p'}$ for $1 \le p < \infty$ if p' is the dual index. This is a special case of Example 1. In particular, the first of these duality results for $p = 2$ says that $(\mathbb{R}^n)^* \cong \mathbb{R}^n$ and $(\mathbb{C}^n)^* \cong \mathbb{C}^n$.

(3) $C(S)^* \cong M(S)$ if S is a compact Hausdorff space, according to Theorems 11.26 and 11.28. Specifically to each x^* in $C(S)^*$ corresponds a unique ρ in $M(S)$ with $x^*(f) = \int_S f\,d\rho$ for all f in $C(S)$, and this ρ has $\|x^*\| = \|\rho\|$. Since $M(S)$ is in this way identified as the dual space of some normed linear space, it follows from Proposition 12.1 that $M(S)$ is a Banach space.

(4) $(\ell_n^\infty)^* \cong \ell_n^1$ and $(c_0)^* \cong \ell^1$. The isomorphism $(\ell_n^\infty)^* \cong \ell_n^1$ is the special case of Example 3 in which $S = \{1, \dots, n\}$. To see the isomorphism $(c_0)^* \cong \ell^1$, we take S to be the set of positive integers and form the one-point compactification S^*. The continuous scalar-valued functions on S^*, with their supremum norm, can be identified with the normed linear space c of convergent sequences. Thus Example 3 in this setting says that $c^* \cong M(S^*)$. The members of c_0 are the members of c that vanish at ∞, and any point mass at ∞ in a member of $M(S^*)$ has no effect on the subspace c_0. It readily follows that the dual of c_0 consists of the members of $M(S^*)$ with no point mass at ∞, and these elements, with their norm, may be identified with ℓ^1.

From one point of view, Hilbert spaces are particularly simple Banach spaces, and we shall study them first. The geometry of Hilbert space will be the topic of the next section, and the section after that will give a brief introduction to bounded linear operators from a Hilbert space to itself.

2. Geometry of Hilbert Space

Hilbert spaces were defined in Section 1 as complete normed linear spaces whose norms arise from an inner product. Euclidean space $\mathbb{R}^n$ and complex Euclidean space $\mathbb{C}^n$ are examples, and every space $L^2(S, \mathcal{A}, \mu)$ with $(f, g) = \int_S f\bar{g}\,d\mu$ is a Hilbert space. We shall see in this section that every Hilbert space shares many geometric facts in common with the finite-dimensional examples $\mathbb{R}^n$ and $\mathbb{C}^n$. The expansion of square integrable functions on $[-\pi, \pi]$ in Fourier series will be seen to be an example of expansion of all members of a Hilbert space in terms of an "orthonormal basis."

Let H be a real or complex Hilbert space with inner product $(\cdot\,,\,\cdot)$ and with norm $\|\cdot\|$ given by $\|u\| = (u, u)^{1/2}$. Lemma 2.2 shows that H satisfies the **Schwarz inequality**

$$|(u, v)| \le \|u\|\|v\| \qquad \text{for all } u \text{ and } v \text{ in } H.$$

The Schwarz inequality implies the estimate

$$|(u, v)-(u_0, v_0)| \le |(u-u_0, v)|+|(u_0, v-v_0)| \le \|u-u_0\|\|v\|+\|u_0\|\|v-v_0\|,$$

from which it follows that the inner product is a continuous function of two variables.

We shall make frequent use of the formula

$$\|u + v\|^2 = \|u\|^2 + 2\,\mathrm{Re}(u, v) + \|v\|^2,$$

which is what one combines with the Schwarz inequality to prove the triangle inequality for the norm. With the additional hypothesis that $(u, v) = 0$, this formula reduces to the **Pythagorean Theorem**

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Direct expansion of the norms squared in terms of the inner product shows that H satisfies the **parallelogram law**

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 \qquad \text{for all } u \text{ and } v \text{ in } H.$$

Actually, there is a converse to this formula, due to Jordan and von Neumann, whose details are left to Problems 19–24 at the end of the chapter: a Banach space is a Hilbert space if its norm satisfies the parallelogram law. The idea is that the inner product in a Hilbert space can be computed from the identity

$$(u, v) = \frac{1}{4} \sum_k i^k \, \|u + i^k v\|^2,$$

where the sum extends for $k \in \{0, 2\}$ if the scalars are real and extends for $k \in \{0, 1, 2, 3\}$ if the scalars are complex. This identity goes under the name **polarization**. For the result of Jordan and von Neumann, one *defines* (u, v) by this formula, shows that the result is an inner product, and proves that $\|u\|^2 = (u, u)$.

The following lemma, which makes use of the completeness, is the key to all the geometry.

Lemma 12.2. If M is a closed vector subspace of the Hilbert space H and if u is in H, then there is a vector v in M with

$$\|u - v\| = \inf_{w \in M} \|u - w\|.$$

REMARK. Examination of the proof will show that we do not make full use of the assumption that M is closed under addition and scalar multiplication, only that M is closed under passage to convex combinations, i.e., that x and y in M imply that $tx + (1 - t)y$ is in M for all t with $0 \le t \le 1$. Thus it is enough to assume that M is a closed convex set, not necessarily a closed vector subspace.

PROOF. Let $d = \inf_{w \in M} \|u - w\|$, and choose a sequence $\{w_n\}$ in M with $\|u - w_n\| \to d$. By the parallelogram law,

$$\|2u - (w_n + w_m)\|^2 + \|w_n - w_m\|^2 = 2(\|u - w_m\|^2 + \|u - w_n\|^2) \longrightarrow 4d^2.$$

Since $\frac{1}{2}(w_n + w_m)$ is in M,

$$\|2u - (w_n + w_m)\|^2 = 4\|u - \tfrac{1}{2}(w_n + w_m)\|^2 \ge 4d^2.$$

We conclude that $\|w_n - w_m\|^2 \to 0$, and $\{w_n\}$ is Cauchy. By completeness of H, $\{w_n\}$ is convergent. If $v = \lim w_n$, then v is in M since M is topologically closed. Since $\|u - w_n\| \to d$, continuity of the norm gives $\|u - v\| = d$. □

Two vectors u and v in H are said to be **orthogonal** if $(u, v) = 0$. The set of all vectors orthogonal to a subset M of H is denoted by $M^\perp$. In symbols,

$$M^\perp = \{u \in H \mid (u, v) = 0 \text{ for all } v \in M\}.$$

We see by inspection that $M^\perp$ is a closed vector subspace. Moreover, $M \cap M^\perp = 0$ since any u in $M \cap M^\perp$ must have $(u, u) = 0$. The subspace $M^\perp$ will be of greatest interest when M is a closed vector subspace, as a consequence of the following proposition.

Proposition 12.3 (Projection Theorem). If M is a closed vector subspace of the Hilbert space H, then every u in H decomposes uniquely as $u = v + w$ with v in M and w in $M^\perp$.

REMARKS. One writes $H = M \oplus M^\perp$ to express this unique decomposition of vector spaces. Because of this proposition, $M^\perp$ is often called the **orthogonal complement** of the closed vector subspace M. It is essential that M be closed in this proposition. In fact, consider the vector subspace M of polynomials in $L^2([0, 1])$. This is dense as a consequence of the Weierstrass Approximation Theorem, and consequently no L^2 function other than 0 can be in $M^\perp$. Thus not every member of L^2 is the sum of a member of M and a member of $M^\perp$.

PROOF. Uniqueness follows from the fact that $M \cap M^\perp = 0$. For existence let u be in H, and choose v in M by Lemma 12.2 with $\|u - v\| = \inf_{w\in M} \|u - w\|$. If m is any member of M with $\|m\| = 1$, then the vector $v + (u - v, m)m$ is in M and the formula $\|x - y\|^2 = \|x\|^2 - 2\,\mathrm{Re}(x, y) + \|y\|^2$ gives

$$\begin{aligned}\|u - v\|^2 &\le \|u - v - (u - v, m)m\|^2\\ &= \|u - v\|^2 - 2|(u - v, m)|^2 + |(u - v, m)|^2\\ &= \|u - v\|^2 - |(u - v, m)|^2.\end{aligned}$$

Hence $(u - v, m) = 0$. Since every nonzero member of M is a scalar multiple of a member with $\|m\| = 1$, $u - v$ is in $M^\perp$. □

Corollary 12.4. If M is a closed vector subspace of the Hilbert space H, then $M^{\perp\perp} = M$.

PROOF. From the definition we see that $M \subseteq M^{\perp\perp}$. If u is in $M^{\perp\perp}$, write $u = m + m^\perp$ with $m \in M$ and $m^\perp \in M^\perp$ by Proposition 12.3. Then $0 = m^\perp + (m - u)$ with $m^\perp \in M^\perp$ and $m - u \in M^{\perp\perp}$. By the uniqueness in the decomposition $H = M^\perp \oplus M^{\perp\perp}$ of Proposition 12.3, $m^\perp = 0$ and $m - u = 0$. Therefore $u = m$ is in M, and $M^{\perp\perp} = M$. □

Theorem 12.5 (Riesz Representation Theorem). If ℓ is a continuous linear functional on the Hilbert space H, then there exists a unique v in H with $\ell(u) = (u, v)$ for all v in H. This vector v has the property that $\|\ell\| = \|v\|$.

REMARKS. It is instructive to compare this result with the version of the Riesz Representation Theorem in Theorem 9.19, which applies to $L^p(S, \mathcal{A}, \mu)$ for $1 \le p < \infty$ and in particular to $L^2(S, \mathcal{A}, \mu)$. That theorem associates to a continuous linear functional ℓ on this L^2 space a member g of the space such that $\ell(f) = \int_S fg\,d\mu$ for all f in the space. The present theorem, applied with $H = L^2(S, \mathcal{A}, \mu)$, instead yields a member v of the space such that $\ell(f) = \int_S f\bar{v}\,d\mu$

for all f in the space. The connection, of course, is that the function g is $\bar{v}$. The space $L^2(S, \mathcal{A}, \mu)$ has a canonically defined notion of complex conjugation, but an abstract Hilbert space does not. Because of the existence of this canonical conjugation, Theorem 9.19 gives us a canonical linear isometry of $L^2(S, \mathcal{A}, \mu)^*$ onto $L^2(S, \mathcal{A}, \mu)$, whereas Theorem 12.5 gives us a canonical isometry that is merely conjugate linear.

PROOF. Uniqueness is immediate since if $(u, v) = 0$ for all u, then $(u, v) = 0$ for $u = v$, and hence $v = 0$. Let us prove existence. If $\ell = 0$, take $v = 0$. Otherwise let $M = \{u \mid \ell(u) = 0\}$. This is a vector subspace since ℓ is linear, and it is closed since ℓ is continuous. By Proposition 12.3 and the fact that M is not all of H, $M^\perp$ contains a nonzero vector w. This vector w must have $\ell(w) \neq 0$ since $M \cap M^\perp = 0$, and we let v be the member of $M^\perp$ given by

$$v = \frac{\overline{\ell(w)}}{\|w\|^2}\, w.$$

For any u in H, we have $\ell\big(u - \frac{\ell(u)}{\ell(w)}\, w\big) = 0$, and hence $u - \frac{\ell(u)}{\ell(w)}\, w$ is in M. Since v is in $M^\perp$, $u - \frac{\ell(u)}{\ell(w)}\, w$ is orthogonal to v. Thus

$$(u, v) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ v\Big) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ \frac{\overline{\ell(w)}}{\|w\|^2}\, w\Big) = \ell(u)\, \frac{\ell(w)}{\ell(w)}\, \frac{\|w\|^2}{\|w\|^2} = \ell(u).$$

This proves existence.

For the norm equality every u in H has $|\ell(u)| = |(u, v)| \leq \|u\|\|v\|$ by the Schwarz inequality. Taking the supremum over all u with $\|u\| \leq 1$ gives $\|\ell\| \leq \|v\|$. On the other hand, $|(u, v)| = |\ell(u)| \leq \|\ell\|\|u\|$; putting $u = v$ gives $\|v\| \leq \|\ell\|$. Thus $\|\ell\| = \|v\|$. $\square$

A subset S of H is **orthonormal** if each vector in S has norm 1 and if each pair of distinct vectors in S is orthogonal. For example, relative to the inner product $(f, g) = \frac{1}{2\pi}\int_{\pi}^{\pi} f\bar{g}\, dx$, the functions $x \mapsto e^{inx}$ are orthonormal as n varies through the integers. An orthonormal set S is linearly independent; in fact, if $v_1, \dots, v_n$ are members of S with $\sum_i c_i v_i = 0$, then the computation $0 = \big(v_j, \sum_i c_i v_i\big) = \sum_i c_i (v_j, v_i) = c_j \|v_j\|^2 = c_j$ shows that $c_j = 0$ for all j.

We encountered other examples of orthogonal sets, beyond the functions e^{inx}, in Chapter IV in connection with solving certain ordinary differential equations. Such an orthogonal set becomes orthonormal when each member is scaled by the reciprocal of its norm. One example was the system of Legendre polynomials $P_n(x)$, which were introduced in Section IV.8: the differential equation $(1 - t^2)y'' - 2ty' + n(n + 1)y = 0$ has polynomial solutions $y(t)$ that are unique up to a scalar, and $P_n(t)$ is a suitably normalized polynomial solution, necessarily of degree n. These can be shown to be orthogonal[4] in $L^2([-1, 1], dt)$.

Another example was constructed from the Bessel function

$$J_0(t) = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{2^{2n}(n!)^2},$$

which was defined in Section IV.8. There are infinitely many distinct positive real numbers k_n such that $J_0(k_n) = 0$, and it can be shown that the functions $x \mapsto J_0(k_n x)$ are orthogonal[5] in $L^2([0, 1], x\,dx)$.

If an ordered set of n linearly independent vectors in H is given, the **Gram–Schmidt orthogonalization process**, which appears in Problem 6 at the end of the present chapter, gives an algorithm for replacing the set with an orthonormal set having the same linear span.

Let M be a closed vector subspace of H, so that $H = M \oplus M^{\perp}$ by Proposition 12.3. The linear projection operator E of H on M along $M^{\perp}$, given by the identity on M and the 0 operator on $M^{\perp}$, is called the **orthogonal projection** of H on M. The linear operator E is bounded with $\|E\| \leq 1$ because if $u \in H$ decomposes as $u = m + m^{\perp}$, the Pythagorean Theorem gives

$$\|E(u)\|^2 = \|E(m + m^{\perp})\|^2 = \|m\|^2 \leq \|m\|^2 + \|m^{\perp}\|^2 = \|u\|^2.$$

We are going to derive a formula for E in terms of orthonormal sets.

Lemma 12.6. If $\{u_j\}$ is an orthonormal sequence in the Hilbert space H and if $\{c_j\}$ is a sequence of scalars, then $\sum_{j=1}^{\infty} c_j u_j$ converges if and only if $\sum_{j=1}^{\infty} |c_j|^2 < \infty$, and in this case

$$\Big\| \sum_{j=1}^{\infty} c_j u_j \Big\| = \Big(\sum_{j=1}^{\infty} |c_j|^2 \Big)^{1/2}.$$

When the series converges, the sum $\sum_{j=1}^{\infty} c_j u_j$ is independent of the order of the terms.

PROOF. For $m \geq n$, we have

$$\big\| \sum_{j=n}^{m} c_j u_j \big\|^2 = \big(\sum_{i=n}^{m} c_i u_i,\ \sum_{j=n}^{m} c_j u_j \big) = \sum_{i,j} c_i \bar{c}_j (u_i, u_j) = \sum_{j=n}^{m} |c_j|^2.$$

This shows that the sequence $\big\{ \sum_{j=1}^{p} c_j u_j \big\}$ is Cauchy in H if and only if $\sum_{j=1}^{\infty} |c_j|^2$ is convergent, and the first conclusion follows since H is complete. When

[4] The verification appears in the problems in the companion volume, *Advanced Real Analysis*.

[5] Again the verification appears in the problems in the companion volume, *Advanced Real Analysis*.

$\{\sum_{j=1}^{p} c_j u_j\}$ is convergent, we denote its limit by $\sum_{j=1}^{\infty} c_j u_j$, and continuity of the norm yields $\|\sum_{j=1}^{\infty} c_j u_j\| = \lim_p \|\sum_{j=1}^{p} c_j u_j\|$. Since we have seen that $\|\sum_{j=1}^{p} c_j u_j\| = \big(\sum_{j=1}^{p} |c_j|^2\big)^{1/2}$, the second conclusion of the lemma follows.

Let $u = \sum_j c_j u_j$, and let $\sum_k c_{j_k} u_{j_k}$ be a rearrangement, necessarily convergent by what has already been proved. Suppose that the rearrangement has sum u'. The equality just proved shows that $\|u\|^2 = \sum_{i=1}^{\infty} |c_i|^2 = \|u'\|^2$ since rearrangements of series of nonnegative reals have the same sums. Continuity of the inner product, together with the same computation as made above, gives

$$(u, u') = \lim_{p,q} \big(\textstyle\sum_{i=1}^{p} c_i u_i,\ \sum_{k=1}^{q} c_{j_k} u_{j_k}\big) = \lim_{p,q} \displaystyle\sum_{\substack{1 \le i \le p,\\ i = j_k \text{ with } k \le q}} |c_i|^2.$$

The limit on the right is $\sum_{i=1}^{\infty} |c_i|^2$ since $\sum_k |c_{j_k}|^2$ is a rearrangement of $\sum_i |c_i|^2$, and hence $(u, u') = \sum_{i=1}^{\infty} |c_i|^2 = \|u\|^2 = \|u'\|^2$. Therefore $\|u - u'\|^2 = (u, u) - 2\,\mathrm{Re}(u, u') + (u', u') = \|u\|^2 - 2\|u\|^2 + \|u\|^2 = 0$, and $u' = u$. $\square$

Proposition 12.7. Let S be an orthonormal set in the Hilbert space H, and let M be the smallest closed vector subspace of H containing S. For each u in H, there are at most countably many members v_α of S such that $(u, v_\alpha) \ne 0$, and thus the series

$$E(u) = \sum_{v_\alpha \in S} (u, v_\alpha) v_\alpha$$

has only countably many nonzero terms. The series converges independently of the order of the nonzero terms, E is the orthogonal projection of H on M, and E satisfies

$$\|E(u)\|^2 = \sum_{v_\alpha \in S} |(u, v_\alpha)|^2 \le \|u\|^2.$$

REMARK. The final inequality of the proposition is **Bessel's inequality**.

PROOF. Let $v_{\alpha_1}, \dots, v_{\alpha_n}$ be a finite subset of S, and form the vector $u' = \sum_{j=1}^{n} (u, v_{\alpha_j}) v_{\alpha_j}$. Taking the inner product of both sides with u gives

$$(u', u) = \sum_{j=1}^{n} (u, v_{\alpha_j})(v_{\alpha_j}, u) = \sum_{j=1}^{n} |(u, v_{\alpha_j})|^2,$$

and Lemma 12.6 gives

$$\|u'\|^2 = \sum_{j=1}^{n} |(u, v_{\alpha_j})|^2.$$

Therefore $0 \le \|u-u'\|^2 = \|u\|^2 - 2\,\mathrm{Re}(u, u') + \|u'\|^2 = \|u\|^2 - 2\|u'\|^2 + \|u'\|^2 = \|u\|^2 - \|u'\|^2$, and we obtain

$$\|u'\|^2 \le \|u\|^2. \tag{$*$}$$

In other words,

$$\sum_{j=1}^{n} |(u, v_{\alpha_j})|^2 \le \|u\|^2, \tag{$**$}$$

no matter what finite subset $v_{\alpha_1}, \dots, v_{\alpha_n}$ of S we use.

The sum of uncountably many positive real numbers is infinite, since otherwise there could be only finitely many greater than $1/n$ for each n. Since $\|u\|^2 < \infty$, $(**)$ implies that there can be only countably many α's with $|(u, v_\alpha)|^2$ nonzero. This proves the first conclusion. If we enumerate those α's and apply Lemma 12.6, we obtain the convergence of $\sum_{v_\alpha \in S}(u, v_\alpha)v_\alpha$ to a sum independent of the order of the terms.

It is evident from the formula that E is linear and that $E(u) = 0$ if u is in $M^\perp$. Inequality $(**)$ shows that the partial sums u' of $E(u)$ have $\|u'\| \le \|u\|$, and the continuity of the norm therefore implies that $\|E(u)\| \le \|u\|$ for all u. Hence E is continuous. Since $E(v_\alpha) = v_\alpha$ for all α, E is the identity on all finite linear combinations of members of S. The continuity of E thus implies that E is the identity on all of M. Hence E is the orthogonal projection as asserted. The final assertion of the proposition follows from Lemma 12.6 and the inequality $\|E(u)\| \le \|u\|$, which we have already proved. $\square$

Corollary 12.8. If S is an orthonormal set in the Hilbert space H, then the following are equivalent:

(a) S is maximal among orthonormal subsets of H,
(b) $u = \sum_{v_\alpha \in S} (u, v_\alpha)v_\alpha \quad$ for all u in H,
(c) $\|u\|^2 = \sum_{v_\alpha \in S} |(u, v_\alpha)|^2 \quad$ for all u in H,
(d) $(u, v) = \sum_{v_\alpha \in S} (u, v_\alpha)\overline{(v, v_\alpha)} \quad$ for all u and v in H.

REMARKS. Condition (b) is summarized by saying that the orthonormal set S is an **orthonormal basis** of H. If H is infinite-dimensional, an orthonormal basis is not a basis in the ordinary linear-algebra sense; a passage to the limit is usually needed to expand vectors in terms of the basis. Condition (c), or sometimes condition (d), is called **Parseval's equality**. Thus the corollary says that the orthonormal set S is maximal if and only if it is an orthonormal basis, if and only if Parseval's equality holds.

PROOF. Let M be the smallest closed vector subspace of H containing S. Then S is maximal if and only if $M^\perp = 0$, and we replace (a) by this condition. If $M^\perp = 0$, then E is the identity operator in Proposition 12.7, and the proposition shows that (b) holds. If (b) holds, Proposition 12.7 says that (c) holds. On the other hand, if (c) holds, then Proposition 12.7 says that $\|u\| = \|E(u)\|$ for all u. For a vector u in $M^\perp$, which must have $E(u) = 0$, this says that $\|u\| = 0$. Thus $M^\perp = 0$, and (a) holds. Hence (a), (b), and (c) are equivalent. Finally (c) and (d) are equivalent by polarization. $\square$

In the context of Fourier series, Parseval's equality ((c) in Corollary 12.8) was proved as Theorem 6.49, and that theorem showed also that any member of $L^2\big([-\pi, \pi], \frac{1}{2\pi}\,dx\big)$ is the sum of its Fourier series in the sense of convergence in L^2. This conclusion was (b) in the corollary. The corollary is showing that the equivalence of (b) and (c) is just a result in abstract Hilbert-space theory. The extra content of Theorem 6.49 is that these conditions are actually satisfied by the system of exponential functions.

One can show that the other two examples we gave in this section of orthogonal sets give orthonormal bases when normalized—the Legendre polynomials $P_n(t)$ on $[-1, 1]$ with respect to dt and the functions $J_0(k_n t)$ on $[0, 1]$ with respect to $t\,dt$.

Proposition 12.9. Let (X, μ) and (Y, ν) be σ-finite measure spaces, and suppose that $L^2(X, \mu)$ has a countable orthonormal basis $\{u_i\}$ and $L^2(Y, \nu)$ has a countable orthonormal basis $\{v_j\}$. Then $\{(x, y) \mapsto u_i(x)v_j(y)\}$ is an orthonormal basis of $L^2(X \times Y, \mu \times \nu)$.

PROOF. The functions $u_i(x)v_j(y)$ are orthonormal, and Corollary 12.8 shows that it is enough to prove that this orthonormal set is maximal. Suppose that $w(x, y)$ is an L^2 function on $X \times Y$ orthogonal to all of them. Then

$$0 = \int_X \int_Y w(x, y)\,\overline{u_i(x)}\,\overline{v_j(y)}\,d\nu(y)\,d\mu(x) = \int_X (w(x,\,\cdot\,), v_j)\,\overline{u_i(x)}\,d\mu(x)$$

for all i and j. Since $\{u_i\}$ is an orthonormal basis of $L^2(X, \mu)$, $x \mapsto (w(x,\,\cdot\,), v_j)$ is the 0 function in $L^2(X, \mu)$ for each j. In other words, $(w(x,\,\cdot\,), v_j) = 0$ for a.e. x $[d\mu]$ for that j. Since the number of j's is countable, $(w(x,\,\cdot\,), v_j) = 0$ for all j for a.e. x $[d\mu]$. Any such x has $0 = \sum_j |(w(x,\,\cdot\,), v_j)|^2 = \int_Y |w(x, y)|^2\,d\nu(y)$. Integrating in x, we see that w is the 0 function in $L^2(X \times Y, \mu \times \nu)$. $\square$

Proposition 12.10. Any orthonormal set in a closed vector subspace M of a Hilbert space H can be extended to an orthonormal basis of M. In particular any closed vector subspace M of H has an orthonormal basis.

PROOF. As a closed subset of a complete space, M is complete, and therefore M is a Hilbert space in its own right. Order by inclusion all orthonormal subsets of M containing the given set. The given set is one such, and the union of the members of a chain is an orthonormal set forming an upper bound for the chain. By Zorn's Lemma we can find a maximal orthonormal set S in M containing the given one. This satisfies (a) in Corollary 12.8 and hence is an orthonormal basis. This proves the first conclusion, and the second conclusion follows from the first by taking the given orthonormal set in M to be empty. $\square$

Proposition 12.11. Any two orthonormal bases of a Hilbert space have the same cardinality.

REMARKS. Cardinality is discussed in Section A10 of Appendix A. The "same cardinality" whose existence is proved in the proposition is called the **Hilbert space dimension** of the Hilbert space. Problem 7 at the end of the chapter shows that two Hilbert spaces are isomorphic as Hilbert spaces if and only if they have the same Hilbert space dimension. Despite the apparent definitive sound of this result, one must not attach too much significance to the proposition. Hilbert spaces that arise in practice tend to have some additional structure, and an isomorphism of this kind need not preserve the additional structure.

PROOF. Fix two orthonormal bases $U = \{u_\alpha\}$ and $V = \{v_\beta\}$ of a Hilbert space H. We define two members u_α and $u_{\alpha'}$ of U to be equivalent if there exists a sequence

$$u_{\alpha_1}, v_{\beta_1}, u_{\alpha_2}, v_{\beta_2}, \ldots, u_{\alpha_{n-1}}, v_{\beta_{n-1}}, u_{\alpha_n} \tag{$*$}$$

with $u_{\alpha_1} = u_\alpha$ and $u_{\alpha_n} = u_{\alpha'}$, with each u_{α_j} in U and each v_{β_j} in V, and with each consecutive pair having nonzero inner product. Define an equivalence relation in V similarly.

Each equivalence class is countable. In fact, consider the class of u_{α_1}, and consider sequences of a fixed length. Proposition 12.7 shows that only countably many members of V can have nonzero inner product with u_{α_1}, only countably many members of U can have nonzero inner product with that, and so on. Thus there are only countably many sequences of any particular length. The countable union of these countable sets is countable, and thus there are only finitely many sequences connecting u_{α_1} to anything. Hence u_{α_1} can be equivalent to only countably many members of U.

Let U_1 and V_1 be equivalence classes in U and V, respectively, and suppose that u_{α_0} and v_{β_0} are members of U_1 and V_1 with nonzero inner product. Expand u_{α_0} in terms of V as $u_{\alpha_0} = \sum_\beta (u_{\alpha_0}, v_\beta)v_\beta$, retaining only the terms with $(u_{\alpha_0}, v_\beta) \neq 0$. One of the terms making a contribution is the one with $v_\beta = v_{\beta_0}$, and it follows that any other term with $(u_{\alpha_0}, v_{\beta'}) \neq 0$ has $v_{\beta'}$ equivalent to v_β. Hence we have

$$u_{\alpha_0} = \sum_{v_\beta \in V_1} (u_{\alpha_0}, v_\beta) v_\beta \qquad \text{and similarly} \qquad v_{\beta_0} = \sum_{u_\alpha \in U_1} (v_{\beta_0}, u_\alpha) u_\alpha.$$

If $u_{\alpha_0'}$ is another member of U_1 and we expand it in terms of V, retaining only the nonzero terms, then the v_β's that occur have to be equivalent to one another. So we have $u_{\alpha_0'} = \sum_{v_\beta \in V_2} (u_{\alpha_0}, v_\beta) v_\beta$ for some equivalence class V_2 within V. If we form a sequence $(*)$ connecting u_{α_0} and $u_{\alpha_0'}$, we see that at least one member of V_2 is connected to at least one member of V_1. Thus $V_1 = V_2$. Consequently every member of U_1 lies in the smallest closed vector subspace containing V_1, and every member of V_1 lies in the smallest closed subspace containing U_1. In other words, U_1 and V_1 are orthonormal bases for the same closed vector subspace of H.

If U_1 is finite, then linear algebra shows that V_1 is finite and has the same number of elements. Since U_1 and V_1 are countable, the only way that either can be infinite is if both are countably infinite. In any event, U_1 and V_1 have the same cardinality. Thus we have a one-one function carrying U_1 onto V_1. Repeating this process for each equivalence class within U, we obtain a one-one function carrying U onto V. □

3. Bounded Linear Operators on Hilbert Spaces

In this section we briefly study bounded linear operators from a Hilbert space H to itself. In the finite-dimensional case we often make a correspondence between matrices and linear operators by using the standard basis of the space of column vectors. If $\{e_i\}_{i=1}^n$ is this basis, then the correspondence between a matrix $A = [A_{ij}]$ and a linear operator L is given by $A_{ij} = (L(e_j), e_i)$. If $u = \sum_j u_j e_j$ and $v = \sum_i v_i e_i$ are column vectors, then $L(u) = \sum_j u_j L(e_j)$ and hence $(L(u), v) = \sum_{i,j} u_j \bar{v}_i (L(e_j), e_i) = \sum_{ij} \bar{v}_i A_{ij} u_j$.

We could extend these formulas to the case of a general Hilbert space, not necessarily finite-dimensional, by using a particular orthonormal basis as the generalization of $\{e_i\}$. But no particular such basis recommends itself, and we work without any choice of basis as much as possible, except for purposes of motivation. Instead, we may think of the function $(u, v) \mapsto (L(u), v)$ as a more appropriate—and canonical—analog of the matrix of L. Just as the operator norm of L is given by a formula that views L as an operator, namely

$$\|L\| = \sup_{\|u\| \leq 1} \|L(u)\|,$$

so there is a formula for computing the norm in terms of the function of two variables, namely

$$\|L\| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L(u), v)|.$$

To verify this formula, fix u and let v have norm ≤ 1. Application of the Schwarz inequality gives $|(L(u), v)| \le \|L(u)\|\|v\| \le \|L(u)\|$. On the other hand, if $L(u) \neq 0$, we take $v = \|L(u)\|^{-1}L(u)$; this v has $\|v\| = 1$, and we obtain $|(L(u), v)| = \|L(u)\|^{-1}(L(u), L(u)) = \|L(u)\|$. Hence $\sup_{\|v\|\le 1} |(L(u), v)\| = \|L(u)\|$. Taking the supremum over $\|u\| \le 1$ shows that the two expressions for $\|L\|$ are equal.

We shall work with the "adjoint" L^* of a bounded linear operator L. In terms of matrices in the finite-dimensional case, the matrix of L^* is to be the conjugate transpose of the matrix of L. In other words, the $(i, j)^{\text{th}}$ entry $(L^*(e_j), e_i))$ of the matrix for L^* is to be $\overline{(L(e_i), e_j)} = (e_j, L(e_i))$. Passing to our functions of two variables, we want to arrange that $(L^*(u), v) = (u, L(v))$ for all u and v. Let us prove existence and uniqueness of such a bounded linear operator.

Proposition 12.12. Let $L : H \to H$ be a bounded linear operator on the Hilbert space H. For each u in H, there exists a unique vector $L^*(u)$ in H such that

$$(L^*(u), v) = (u, L(v)) \qquad \text{for all } v \text{ in } H.$$

As u varies, this formula defines L^* as a bounded linear operator on H, and $\|L^*\| = \|L\|$.

PROOF. The function $v \mapsto (L(v), u)$ is a linear functional on H satisfying $|(L(v), u)| \le \|L\|\|u\|\|v\|$, hence having norm $\le \|L\|\|u\|$. Being bounded, the linear functional is given by $(L(v), u) = (v, w)$ for some unique w in H, according to Theorem 12.5. We define $L^*(u) = w$, and then we have $(L^*(u), v) = (u, L(v))$. This formula shows that L^* is a linear operator, and the computation

$$\|L^*\| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L^*(u), v)| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(u, L(v))| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L(v), u)| = \|L\|$$

shows that $\|L^*\| = \|L\|$. $\square$

The bounded linear operator L^* in the proposition is called the **adjoint** of L. The mapping $L \mapsto L^*$ is conjugate linear. We shall be especially interested in the case that $L^* = L$, in which case we say that L is **self adjoint**.

An example of a self-adjoint operator is the orthogonal projection E on a closed vector subspace M as defined before Lemma 3.6. In fact, if u in H decomposes according to $H = M \oplus M^\perp$ as $u = u' + u''$, then the computation $(1 - E)(u) =$

$u - u' = u''$ shows that $1 - E$ is the orthogonal projection on $M^\perp$. Hence $(E(u), (1-E)(v)) = 0$ for all u and v in H, and also $((1-E)(u), E(v)) = 0$. The first of these says that $(E(u), v) = (E(u), E(v))$, and the second says that $(E(u), E(v)) = (u, E(v))$. Combining these, we obtain $(E(u), v) = (u, E(v))$. Comparison of this formula with the formula in Proposition 12.12 shows that $E = E^*$.

The Banach space $\mathcal{B}(H, H)$ is closed under composition. In fact, if L and M are in $\mathcal{B}(H, H)$, then linear algebra shows LM to be linear, and the computation $\|(LM)(u)\| = \|L(M(u))\| \le \|L\|\|M(u)\| \le \|L\|\|M\|\|u\|$ shows that

$$\|LM\| \le \|L\|\|M\|.$$

Hence LM is in $\mathcal{B}(H, H)$ if L and M are. Within $\mathcal{B}(H, H)$, we have $(LM)^* = M^*L^*$.

The structure of abstract bounded linear operators on Hilbert spaces is one of the topics in Chapter IV of the companion volume, *Advanced Real Analysis*.

4. Hahn–Banach Theorem

We return now to the setting of general normed linear spaces or Banach spaces. There are three main theorems concerning the norm topology of such spaces—the Hahn–Banach Theorem, the Uniform Boundedness Theorem, and the Interior Mapping Principle. These three theorems are the main subject matter of the remainder of this chapter. Further properties of normed linear spaces and Banach spaces are established in Chapter IV of the companion volume, *Advanced Real Analysis*.

We shall often use symbols $x, y, \ldots$ for members of a normed linear space and symbols $x^*, y^*, \ldots$ for linear functionals. This notation has the advantage of allowing us to use symbols like x^{**} for linear functionals on a space of linear functionals, an important notion as we shall see.

We begin with the Hahn–Banach Theorem, which ensures the existence of many continuous linear functionals on a normed linear space. The theorem has applications even in situations in which one has a concrete realization of the dual space, because it shows that any closed vector subspace is characterized by the continuous linear functionals that vanish on the subspace.

Theorem 12.13 (Hahn–Banach Theorem). If Y is a vector subspace of a normed linear space X and if y^* is a continuous linear functional on Y, then there exists a continuous linear functional x^* on X with $\|x^*\| = \|y^*\|$ such that

$$x^*(y) = y^*(y) \qquad \text{for all } y \in Y.$$

The theorem as stated is derived from the following lemma, which itself goes under the name "Hahn–Banach Theorem" and has other applications quite distinct from Theorem 12.13 that are beyond the scope of this book.

Lemma 12.14. Let X be a real vector space, and let p be a real-valued function on X with

$$p(x+x') \le p(x) + p(x') \qquad \text{and} \qquad p(tx) = tp(x)$$

for all x and x' in X and all real $t \ge 0$. If f is a linear functional on a vector subspace Y of X with $f(y) \le p(y)$ for all y in Y, then there exists a linear functional F on X with $F(y) = f(y)$ for all $y \in Y$ and $F(x) \le p(x)$ for all $x \in X$. $\square$

PROOF. Form the collection of all linear functionals on vector subspaces of X that extend f and that are dominated by p, and partially order the collection by saying that one is $\le$ another if the second is an extension of the first. If we have a chain of such extensions, then we can obtain an upper bound for the chain by taking the union of the domains and using the common value of the linear functionals on an element of this domain as the value of the linear functional forming the upper bound. The result is linear because any two members of the domain must lie in the domain of a single member of the chain. By Zorn's Lemma let f_0, with domain Y_0, be a maximal extension. We shall prove that $Y_0 = X$.

In fact, suppose that y_1 is a vector in X but not Y_0. Every vector in the vector subspace Y_1 spanned by y_1 and Y_0 has a unique representation as $y + cy_1$, where y is in Y_0 and c is in $\mathbb{R}$. Define f_1 on Y_1 by

$$f_1(y + cy_1) = f_0(y) + ck, \tag{$*$}$$

where k is a real number to be specified. For a suitable choice of k, f_1 will be bounded by p and will contradict the maximality of (f_0, Y_0).

Let y and y' be in Y_0. Then

$$f_0(y') - f_0(y) = f_0(y'-y) \le p(y'-y) \le p(y'+y_1) + p(-y_1-y),$$

and hence $\qquad -p(-y_1-y) - f_0(y) \le p(y'+y_1) - f_0(y').$

Take the supremum of the left side over y and the infimum of the right side over y', let k be any real number in between, and define f_1 on Y_1 by $(*)$.

To complete the proof, we are to check that $f_1(x) \le p(x)$ for all x in Y_1. Thus suppose that $x = y + cy_1$ is arbitrary in Y_1. If $c = 0$, then $f_1(x) \le p(x)$ by the assumption on Y_0. If $c > 0$, then

$$f_1(x) = f_0(y)+ck \le f_0(y)+c[p(c^{-1}y+y_1)-f_0(c^{-1}y)] = p(y+cy_1) = p(x).$$

If $c < 0$, then

$$f_1(x) = f_0(y)+ck \le f_0(y)+c[-p(-y_1-c^{-1}y)-f_0(c^{-1}y)] = p(y+cy_1) = p(x).$$

In any case, $f_1(x) \le p(x)$. $\square$

PROOF OF THEOREM 12.13. If the field of scalars is $\mathbb{R}$, then Theorem 12.13 follows immediately from Lemma 12.14 with $p(x) = \|y^*\|\|x\|$ and $f = y^*$.

If the field of scalars is $\mathbb{C}$, if y^* is given, and if, as we may, we regard X as a real normed linear space, then $\operatorname{Re} y^*$ defined by $(\operatorname{Re} y^*)(y) = \operatorname{Re}(y^*(y))$ is a real linear functional on Y with

$$|(\operatorname{Re} y^*)(y)| \le |y^*(y)| \le \|y^*\|\|y\| \qquad \text{for all } y \in Y.$$

By what has already been proved, we can extend $\operatorname{Re} y^*$ without an increase in norm to a real linear functional F defined on all of X. Define

$$x^*(x) = F(x) - iF(ix).$$

We show that x^* has the required properties. Certainly $x^*(x+x') = x^*(x)+x^*(x')$ and $x^*(cx) = cx^*(x)$ for c real. Furthermore

$$x^*(ix) = F(ix) - iF(i^2x) = i[F(x) - iF(ix)] = ix^*(x).$$

Thus x^* is complex linear. On Y, we have

$$(\operatorname{Re} y^*)(iy) + i(\operatorname{Im} y^*)(iy) = y^*(iy) = iy^*(y) = -(\operatorname{Im} y^*)(y) + i(\operatorname{Re} y^*)(y),$$

and thus $(\operatorname{Re} y^*)(iy) = -(\operatorname{Im} y^*)(y)$. Substituting this identity into the definition of x^*, we obtain

$$x^*(y) = (\operatorname{Re} y^*)(y) - i(\operatorname{Re} y^*)(iy) = (\operatorname{Re} y^*)(y) + i(\operatorname{Im} y^*)(y) = y^*(y)$$

for y in Y. Thus x^* is an extension of y^*. Finally if $x^*(x) = re^{i\theta}$ for r and θ real and $r \ge 0$, then

$$|x^*(x)| = x^*(e^{-i\theta}x) = F(e^{-i\theta}x) \le \|y^*\|\|e^{-i\theta}x\| = \|y^*\|\|x\|,$$

since the nonnegative number $x^*(e^{-i\theta}x)$ has 0 imaginary part. Thus $\|x^*\| \le \|y^*\|$. The reverse inequality follows because x^* is an extension of y^*, and the proof is complete. $\square$

Corollary 12.15. If Y is a closed vector subspace of a normed linear space X and if x_0 is a vector of X not in Y, then there exists an x^* in the dual X^* with

$$x^*(y) = 0 \qquad \text{for all } y \in Y$$

and

$$x^*(x_0) = 1.$$

The norm of x^* can be taken to be the reciprocal of the distance from x_0 to Y.

PROOF. Let $d > 0$ be the distance from x_0 to Y, and let Z be the linear span of x_0 and Y. Every x in Z has a unique expansion as $x = y + cx_0$ for some scalar c and some y in Y. For such an x, let $z^*(x) = c$. Let us see that the linear function z^* on Z satisfies

$$\|z^*\| = d^{-1}. \tag{$*$}$$

First we check that $|z^*(x)| \le d^{-1}\|x\|$: if $c \ne 0$, then

$$\|x\| = \|y + cx_0\| = |c|\|c^{-1}y + x_0\| \ge |c|d = d|z^*(x)|,$$

while if $c = 0$, then $z^*(x) = 0$. Thus $|z^*(x)| \le d^{-1}\|x\|$ for all x, and we obtain $\|z^*\| \le d^{-1}$. For the reverse inequality, let $\{y_n\}$ be a sequence in Y, not necessarily convergent, with $\lim_n \|x_0 - y_n\| = d$. Then

$$1 = z^*(x_0 - y_n) \le \|z^*\|\|x_0 - y_n\| \longrightarrow d\|z^*\|,$$

and hence $\|z^*\| \ge d^{-1}$. This proves $(*)$. Applying Theorem 12.13 to z^*, we obtain the corollary. □

EXAMPLE. To illustrate Corollary 12.15, we re-prove the result of Proposition 11.21a that $C(S)$ is dense in $L^p(S, \mu)$ if S is a compact Hausdorff space, μ is a regular Borel measure on S, and p satisfies $1 \le p < \infty$. For definiteness let us suppose that the underlying scalars are real. If $C(S)$ were not dense, then the corollary would produce a continuous linear functional ℓ on $L^p(S, \mu)$ that vanishes on $C(S)$ but is not identically 0 on $L^p(S, \mu)$. Theorem 9.19 says that ℓ has to be given by integration with some member g of $L^{p'}(S, \mu)$, where p' is the dual index: $\ell(f) = \int_S fg\,d\mu$ for all f in $L^p(S, \mu)$. Since ℓ vanishes on $C(S)$, we have $\int_S fg\,d\mu = 0$ for all $f \in C(S)$. Thus $\int_S fg^+\,d\mu = \int_S fg^-\,d\mu$ for all $f \in C(S)$. Here $g^+\,d\mu$ and $g^-\,d\mu$ are Borel measures on S, regular by Proposition 11.20, and they yield the same positive linear functional on $C(S)$. Applying the uniqueness in the Riesz Representation Theorem (Theorem 11.1), we obtain $g^+\,d\mu = g^-\,d\mu$ and therefore $g^+ = g^-$ almost everywhere. Since g^+ and g^- are nowhere both nonzero, $g^+ = g^- = 0$ almost everywhere. Hence g is the 0 function, and $\ell = 0$, contradiction.

Corollary 12.16. If X is a normed linear space and if $x_0 \ne 0$ is a vector in X, then there is an x^* in X^* with

$$\|x^*\| = 1 \qquad \text{and} \qquad x^*(x_0) = \|x_0\|.$$

PROOF. Apply Corollary 12.15 with $Y = 0$ and multiply by $\|x_0\|$ the linear functional that is produced by that corollary. □

Corollary 12.16, when applied to $x_0 = x - x'$, shows that there are enough continuous linear functionals on a normed linear space X to separate points. Also, it implies that the only vector x_0 in X with $x^*(x_0) = 0$ for all x^* in X^* is $x_0 = 0$. The third corollary we have already seen for L^p spaces with $1 \le p < \infty$ in Proposition 9.8, at least when the measure space is σ-finite.

Corollary 12.17. If X is a normed linear space and x_0 is in X, then

$$\|x_0\| = \sup_{\|x^*\| \le 1} |x^*(x_0)|.$$

PROOF. If $\|x^*\| \le 1$, then $|x^*(x_0)| \le \|x^*\|\|x_0\| \le \|x_0\|$, and therefore $\sup_{\|x^*\| \le 1} |x^*(x_0)| \le \|x_0\|$. The linear functional of Corollary 12.16 shows that equality holds. □

We have seen for σ-finite measure spaces that the dual X^* of $X = L^1(S, \mu)$ may be identified with $L^\infty(S, \mu)$ via integration. In turn every member of $L^1(S, \mu)$ then acts as a continuous linear functional on $L^\infty(S, \mu)$ via integration. This change of point of view amounts to the implementation of a certain canonically defined linear mapping of X into X^{**}, which we now define for general normed linear spaces.

Let X be a normed linear space, and let X^{**} be the dual of X^*. We define a linear operator $\iota : X \to X^{**}$ by

$$(\iota(x))(x^*) = x^*(x) \qquad \text{for all } x^* \in X^*,$$

and we call ι the **canonical map** of X into X^{**}.

Corollary 12.18. If X is a normed linear space, then the canonical map $\iota : X \to X^{**}$ has $\|\iota(x)\| = \|x\|$ for all x and in particular is one-one. Consequently if X is complete, then $\iota(X)$ is a closed vector subspace of X^{**}.

PROOF. We have

$$\|\iota(x)\| = \sup_{\|x^*\| \le 1} |(\iota(x))(x^*)| = \sup_{\|x^*\| \le 1} |x^*(x)| = \|x\|,$$

the last step holding by Corollary 12.17. This proves the first conclusion. Because ι preserves norms, X complete implies that $\iota(X)$ is a complete subset of the complete space X^{**} and is therefore closed, by Corollary 2.43. □

A Banach space X is said to be **reflexive** if the canonical map carries X onto X^{**}. *Warning:* This is a more restrictive condition than to say that there is some norm-preserving linear mapping of X onto X^{**}.

Finite-dimensional normed linear spaces are reflexive since linear functionals in this case are automatically continuous and since the vector-space dual of a finite-dimensional vector space has the same dimension as the space itself. Hilbert spaces are reflexive as a consequence of the Riesz Representation Theorem in its form in Theorem 12.5. The spaces $L^p(S, \mu)$ for a σ-finite measure space, when $1 < p < \infty$, are reflexive as a consequence of the Riesz Representation Theorem[6] in its form in Theorem 9.19. However, $L^1(S, \mu)$ and $L^\infty(S, \mu)$ are often not reflexive, as is shown below in Proposition 12.19 and Corollary 12.21.

Proposition 12.19. If (S, μ) is a σ-finite measure space with infinitely many disjoint sets of positive measure, then $L^1(S, \mu)$ is not reflexive.

PROOF. Theorem 9.19 shows that the Banach space $X = L^1(S, \mu)$ has $X^* \cong L^\infty(S, \mu)$, the isomorphism being given by integration. Therefore it is enough to produce a continuous linear functional on $L^\infty(S, \mu)$ that is not given by integration with an L^1 function.

Thus let $\{E_n\}$ be a sequence of disjoint sets of positive measure, and let Y be the vector subspace of functions in $L^\infty(S, \mu)$ that are constant on each E_n and have values on the E_n's tending to a finite limit as n tends to infinity. Let y^* of such a function be the limit. Then y^* is a linear functional on Y of norm 1. By the Hahn–Banach Theorem (Theorem 12.13), there exists a linear functional x^* defined on all of $L^\infty(S, \mu)$, having norm 1, and restricting to y^* on Y. Suppose that there is some g in $L^1(S, \mu)$ with $x^*(f) = \int_S fg\,d\mu$ for all f in Y, quite apart from all f in $L^\infty(S, \mu)$. If f is 1 on E_n and is 0 elsewhere, then $x^*(f) = 0$, and hence $\int_{E_n} g\,d\mu = 0$. In other words, $\int_{E_n} g\,d\mu = 0$ for every n. If we next take f to be 1 on $\bigcup_{n=1}^\infty E_n$ and to be 0 elsewhere, then $x^*(f) = 1$. On the other hand, this f has

$$x^*(f) = \int_S fg\,d\mu = \int_{\bigcup_n E_n} g\,d\mu = \sum_{n=1}^\infty \int_{E_n} g\,d\mu = 0,$$

and we have a contradiction. $\square$

Proposition 12.20. If X is a Banach space and its dual X^* is reflexive, then X is reflexive.

PROOF. Let $\iota : X \to X^{**}$ and $\iota^* : X^* \to X^{***}$ be the canonical maps. Arguing by contradiction, suppose that X is not reflexive. Since $\iota(X)$ is a closed proper vector subspace of X^{**}, Corollary 12.15 produces a nonzero member

[6] Actually, the σ-finiteness is not needed for $1 < p < \infty$.

x^{***} of X^{***} such that $x^{***}(\iota(X)) = 0$. Since X^* is reflexive by assumption, there exists x^* in X^* with $x^{***} = \iota^*(x^*)$. If x is in X, then we have $0 = x^{***}(\iota(x)) = (\iota^*(x^*))(\iota(x)) = (\iota(x))(x^*) = x^*(x)$, and hence $x^* = 0$. But then $x^{***} = \iota^*(x^*) = 0$, and we have a contradiction. □

Corollary 12.21. If (S, μ) is a σ-finite measure space with infinitely many disjoint sets of positive measure, then $L^\infty(S, \mu)$ is not reflexive.

PROOF. Theorem 9.19 shows that the Banach space $X = L^1(S, \mu)$ has $X^* \cong L^\infty(S, \mu)$, the isomorphism being given by integration. If X^* were reflexive, then X would have to be reflexive by Proposition 12.20, in contradiction to Proposition 12.19. □

5. Uniform Boundedness Theorem

The second main theorem about the norm topology of normed linear spaces is the Uniform Boundedness Theorem, also known as the Banach–Steinhaus Theorem. This result involves a parametrized family of linear operators from one normed linear space into another, and it is assumed that the domain is complete. Two kinds of boundedness as a function of one variable are assumed—boundedness of each linear operator as a function on (the unit ball of) the domain and boundedness in the parameter for each fixed member of the domain. The conclusion is boundedness in the two variables jointly.

Theorem 12.22 (Uniform Boundedness Theorem). If $\{L_\alpha\}$ is a set of bounded linear operators from a Banach space X into a normed linear space Y such that

$$\|L_\alpha(x)\| \le C_x \qquad \text{for all } \alpha,$$

then there is a constant C independent of x such that $\|L_\alpha\| \le C$ for all α.

PROOF. For each positive integer n, the set

$$F_n = \big\{x \in X \;\big|\; \|L_\alpha(x)\| \le n \text{ for all } \alpha\big\}$$

is closed in X, being the intersection of inverse images of closed sets in Y under continuous functions, and $\bigcup_{n=1}^\infty F_n = X$ by assumption. By the Baire Category Theorem (Theorem 2.53b), one of the sets, say F_N, contains a nonempty open subset B of X. Then $\|L_\alpha(x)\| \le N$ for all α and for all x in B. If B contains the open ball in X of radius $2r > 0$ and center b, then $\|x\| \le r$ implies that $x + b$ is in B and that

$$\|L_\alpha(x)\| = \|L_\alpha(x+b) - L_\alpha(b)\| \le \|L_\alpha(x+b)\| + \|L_\alpha(b)\| \le N + C_b,$$

independently of α. Hence $\|x\| \le 1$ implies

$$\|L_\alpha(x)\| = r^{-1}\|L_\alpha(rx)\| \le r^{-1}(N + C_b).$$

In other words, $\|L_\alpha\| \le r^{-1}(N + C_b)$. □

EXAMPLE. Let us use the theorem to give a proof that the Fourier series of a continuous periodic function need not converge at some point. Consider the Banach space X of all continuous periodic functions f on $[-\pi, \pi]$ with the supremum norm. Let D_n be the Dirichlet kernel as in Section I.10, given by

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = \frac{\sin((n+\frac{1}{2})t)}{\sin \frac{1}{2}t}.$$

The n^{th} partial sum of the Fourier series of f is

$$s_n(f; x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_n(t)\, dt.$$

Define linear functionals ℓ_n on X by

$$\ell_n(f) = s_n(f; 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(-t) D_n(t)\, dt.$$

Each of these is bounded; specifically $\|\ell_n\| \leq 2n+1$ because $\|D_n\|_{\sup} \leq 2n+1$. If the Fourier series of each continuous function f were to converge at 0, then $\lim_n \ell_n(f)$ would exist for each f, and hence we would have $|\ell_n(f)| \leq C_f$ for a constant C_f independent of n. The Uniform Boundedness Theorem would say that $\|\ell_n\| \leq C$ for some constant C independent of n. The norm equality of Theorem 11.26 or 11.28 would then allow us to conclude that $\int_{-\pi}^{\pi} |D_n(t)|\, dt$ is bounded. In fact, the numbers $\int_{-\pi}^{\pi} |D_n(t)|\, dt$ are unbounded, according to the following proposition, and thus there exists a continuous periodic function whose Fourier series diverges at $x = 0$.

Proposition 12.23. The numbers

$$L_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\, dt$$

have the property that

$$L_n = 4\pi^{-2} \log n + O(1),$$

where $O(1)$ denotes an expression bounded as a function of n. Hence L_n is unbounded with n.

REMARK. The numbers L_n are sometimes called **Lebesgue constants**.

PROOF. By writing $\sin((n+\frac{1}{2})t) = \sin nt \cos \frac{1}{2}t + \cos nt \sin \frac{1}{2}t$, we see that

$$D_n(t) = \sin nt \cot \tfrac{1}{2}t + \cos nt = 2t^{-1} \sin nt + h_n(t),$$

where $h_n(t)$ is bounded in the pair (n,t) for $|t| \leq \pi$. If we let $O(1)$ denote an expression bounded as a function of n, then

$$\begin{aligned}
L_n &= \frac{2}{2\pi} \int_{-\pi}^{\pi} \frac{|\sin nt|}{|t|}\, dt + O(1) \\
&= \frac{2}{\pi} \int_0^{\pi} \frac{|\sin nt|}{t}\, dt + O(1) \\
&= \frac{2}{\pi} \sum_{k=0}^{n-1} \int_{k\pi/n}^{(k+1)\pi/n} \frac{|\sin nt|}{t}\, dt + O(1) \\
&= \frac{2}{\pi} \int_0^{\pi/n} \frac{\sin nt}{t}\, dt + \frac{2}{\pi} \int_0^{\pi/n} (\sin nt) \Big[\sum_{k=1}^{n-1} \frac{1}{t + k\pi/n} \Big]\, dt + O(1).
\end{aligned}$$

The first term on the right side is bounded, and the sum in brackets lies between

$$\pi^{-1} n(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n-1}) \qquad \text{and} \qquad \pi^{-1} n(\tfrac{1}{2} + \cdots + \tfrac{1}{n}),$$

which are upper and lower Riemann sums for $\pi^{-1} n \int_1^n t^{-1}\, dt$ and have difference $\pi^{-1} n(1 - \frac{1}{n})$. Thus the sum in brackets is equal to $\pi^{-1} n(\log n + O(1))$. The integral of $\sin nt$ over $[0, \pi/n]$ is $2/n$, and the result follows. □

6. Interior Mapping Principle

The third main theorem about the norm topology of normed linear spaces is the Interior Mapping Principle. This result involves a single bounded linear operator from one normed linear space into another, and it is assumed that the domain and the range are both complete. The theorem is that if the operator is onto the range, then it carries open sets to open sets.

Theorem 12.24 (Interior Mapping Principle). If L is a continuous linear operator from a Banach space X *onto* a Banach space Y, then L carries open subsets of X to open subsets of Y.

PROOF. Let B_r be the closed ball in X with center 0 and radius r, and let U_s be the open ball in Y with center 0 and radius s. The proof is in three steps.

The first step is to show that $(L(B_1))^{\mathrm{cl}}$ contains an open neighborhood of 0 in Y. To do so, we use the fact that L is onto Y to write

$$Y = L(X) = L\big(\bigcup_{n=1}^{\infty} B_n\big) = \bigcup_{n=1}^{\infty} L(B_n).$$

Thus $Y = \bigcup_{n=1}^{\infty}(L(B_n))^{\mathrm{cl}}$, and the Baire Category Theorem (Theorem 2.53b) shows that one of the sets $(L(B_n))^{\mathrm{cl}}$ contains a nonempty open set. Since L is linear and since multiplication by $2n$ is a homeomorphism of Y, $(L(B_n))^{\mathrm{cl}} = (L(2nB_{1/2}))^{\mathrm{cl}} = (2nL(B_{1/2}))^{\mathrm{cl}} = (2n)(L(B_{1/2}))^{\mathrm{cl}}$, and we see that $(L(B_{1/2}))^{\mathrm{cl}}$ contains some nonempty open subset V of Y. If v and v' are in V, they are in $(L(B_{1/2}))^{\mathrm{cl}}$ and there exist sequences $\{v_n\}$ and $\{v_n'\}$ in $L(B_{1/2})$ with $v_n \to v$ and $v_n' \to v'$. By linearity, $v_n - v_n'$ is in $L(B_1)$, and passage to the limit shows that $v - v'$ is in $L(B_1)^{\mathrm{cl}}$. The set $V - V$ of such differences $v - v'$ is the union over $v' \in V$ of $V - v'$, hence is the union of open sets and is open. Since 0 is in $V - V$, the set $V - V$ is an open neighborhood of 0 lying in $L(B_1)^{\mathrm{cl}}$.

The second step is to show that the image of any neighborhood of 0 in X is a neighborhood of 0 in Y. The previous step shows that $(L(B_1))^{\mathrm{cl}} \supseteq U_s$ for some $s > 0$, and we show for every $c > 0$ that $L(B_c) \supseteq U_{sc/2}$. For $t > 0$, multiplication of the inclusion $(L(B_1))^{\mathrm{cl}} \supseteq U_s$ by t shows that

$$(L(B_t))^{\mathrm{cl}} \supseteq U_{st} \tag{$*$}$$

since multiplication by t is a homeomorphism of Y and L is linear. If y is in $U_{sc/2}$, we are to produce x in B_c with $L(x) = y$, and we do so by successive approximations. Specifically we construct inductively the terms x_n of a convergent series in X with sum x, as follows: Condition $(*)$ with $t = c/2$ allows us to choose a member x_1 of $B_{c/2}$ with $\|y - L(x_1)\| < 2^{-2}sc$. If $x_1, \dots, x_{n-1}$ have been constructed with each x_j in $B_{2^{-j}c}$ and with

$$\|y - L(x_1 + \cdots + x_{n-1})\| < 2^{-n}sc,$$

then $y - L(x_1 + \cdots + x_{n-1})$ is in $U_{2^{-n}sc}$. Condition $(*)$ with $t = 2^{-n}c$ shows that we can find x_n in $B_{2^{-n}c}$ with

$$\|y - L(x_1 + \cdots + x_{n-1}) - L(x_n)\| < 2^{-(n+1)}sc.$$

We now have

$$\|y - L(x_1 + \cdots + x_{n-1} + x_n)\| < 2^{-(n+1)}sc.$$

This completes the inductive construction of the x_n's, and we shall prove that the series $\sum x_n$ is convergent in X. Since X is complete, it is enough to show that the partial sums of $\sum x_n$ are Cauchy. If $q \geq p$, then

$$\big\| \textstyle\sum_{n=1}^q x_n - \sum_{n=1}^p x_n \big\| = \big\| \sum_{n=p+1}^q x_n \big\| \leq \sum_{n=p+1}^q \|x_n\| \leq \sum_{n=p+1}^q 2^{-n}c.$$

The right side is $\leq 2^{-p}c$, and the partial sums of $\sum x_n$ are indeed Cauchy. Let $x = \sum_{n=1}^\infty x_n$. Taking $p = 0$ and using the continuity of the norm, we see that $\|x\| \leq c$. By continuity of L, we have $y = \lim_n L(x_1 + \cdots + x_n) = L(x)$. Consequently the member y of $U_{sc/2}$ is of the form $L(x)$ for some x in B_c, as was asserted.

The third step is to show that each open set of X is mapped to an open set of Y by L. Let U be open in X, let x be in U, and let N be an open neighborhood of 0 in X such that $x + N \subseteq U$. The previous step shows that there is some open neighborhood V of 0 in Y such that $V \subseteq L(N)$. Then $L(x) + V$ is an open neighborhood in Y of $L(x)$ with

$$L(x) + V \subseteq L(x) + L(N) = L(x + N) \subseteq L(U).$$

Therefore $L(U)$ contains a neighborhood about each of its points and must be open. $\square$

Corollary 12.25. A one-one continuous linear operator L of a Banach space X onto a Banach space Y has a continuous linear inverse.

PROOF. Since L is one-one onto, L^{-1} exists. For L^{-1} to be continuous, the inverse image under L^{-1} of each open set is to be open. In other words, the direct image under L of any open set is to be open. But this is just the conclusion of Theorem 12.24. $\square$

EXAMPLE. Let $\mathcal{F}$ be the Fourier coefficient mapping, which carries functions in $L^1\big(\frac{1}{2\pi}\,dx\big)$ to doubly infinite sequences $\{c_n\}$ vanishing at infinity. The linear operator $\mathcal{F}$ has norm 1 when the space of doubly infinite sequences is given the supremum norm $\|\{c_n\}\|_{\text{sup}} = \sup_n |c_n|$. Corollary 6.50 shows that $\mathcal{F}$ is one-one. Let us see that there is some doubly infinite sequence vanishing at infinity that is not the sequence of Fourier coefficients of some L^1 function. If this were not so, then Corollary 12.25 would say that $\mathcal{F}^{-1}$ is bounded. We can obtain a contradiction if we produce a sequence $\{f_n\}$ of L^1 functions with $\|f_n\|_1 = 1$ for all n and with $\lim_n \|\mathcal{F}(f_n)\|_{\text{sup}} = 0$. Form the Dirichlet kernel D_n as defined in Section I.10 and reproduced in the previous section. Its Fourier coefficients c_k are 1 for $|k| \leq n$ and are 0 for $|k| > n$, and thus $\|\mathcal{F}(D_n)\|_{\text{sup}} = 1$. Put $f_n = D_n\big/\|D_n\|_1$. Then $\|f_n\|_1 = 1$ for all n, and $\|\mathcal{F}(f_n)\|_{\text{sup}} = 1\big/\|D_n\|_1$. Proposition 12.23 shows that in fact $\lim_n 1/\|D_n\|_1 = 0$, and we obtain the desired contradiction. The conclusion is that the image of $\mathcal{F}$ on L^1 fails to include some doubly infinite sequence $\{c_n\}$ vanishing at infinity.

If $f : X \to Y$ is a function between Hausdorff spaces, the graph of f is the subset $G = \{(x, f(x)) \mid x \in X\}$ of $X \times Y$. If f is continuous, then G is a closed set, as we see immediately by using nets. The converse fails because $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ and $f(x) = 1/x$ for $x > 0$ is a discontinuous function with closed graph.

We shall be interested in the converse under the additional condition that our function f is linear. Our spaces being metric spaces, the condition that the graph be closed is that whenever $\{(x_n, f(x_n))\}$ converges to some (x, y), then x is in the domain of f and $f(x) = y$.

Linearity by itself is not enough to get an affirmative result. In fact, let $X = C([0, 1])$, let X_0 be the vector subspace of functions with a continuous derivative, and let $L : X_0 \to X$ be the derivative operator $F \mapsto F'$. If $\lim_n F_n = F$ in X and $\lim_n F_n' = H$, then Theorem 1.23 shows that F' exists and equals H. Hence the linear operator $L : X_0 \to X$ has closed graph. However, L is unbounded since the function $x \mapsto x^n$ has norm 1 and its derivative has norm n.

Corollary 12.26 (Closed Graph Theorem). If $L : X \to Y$ is a linear operator from a Banach space X into a Banach space Y such that the graph of L is a closed subset of $X \times Y$, then L is a bounded linear operator.

PROOF. Make $X \oplus Y$ into a Banach space by defining $\|(x, y)\| = \|x\|_X + \|y\|_Y$. The graph $G = \{(x, L(x)) \mid x \in X\}$ of L is a vector subspace of $X \oplus Y$ since L is linear, and it is closed by hypothesis. Thus G is a Banach space. The linear operator $P : G \to X$ given by $P((x, L(x)) = x$ is one-one and onto, and Corollary 12.25 shows that the linear operator $P^{-1} : X \to G$ given by $P^{-1}(x) = (x, L(x))$ is continuous. If E denotes the projection of $X \oplus Y$ to the Y coordinate, then E is bounded with norm ≤ 1, and hence the restriction $E\big|_G : G \to Y$ is bounded with norm ≤ 1. Therefore the composition $(E\big|_G) \circ P^{-1} : X \to Y$ is bounded. But $(E\big|_G)(P^{-1}(x)) = E(x, L(x)) = L(x)$, and thus L is bounded. $\square$

EXAMPLE. Suppose that a Banach space X is the vector-space direct sum of two closed vector subspaces: $X = Y \oplus Y'$. Let $E : X \to Y$ be the projection of X on Y given by $E(y + y') = y$. Corollary 12.26 implies that E is bounded. In fact, let $x_n = y_n + y_n'$ define a sequence in X, so that (x_n, y_n) defines a sequence in the graph of E. Suppose that $\lim_n (x_n, y_n) = (x_0, y_0)$ in $X \times X$, i.e., that $\lim_n x_n = x_0$ and $\lim_n y_n = y_0$. Here x_0 is in X, and y_0 is in Y since Y is closed. Then $y_0' = \lim_n y_n' = \lim_n x_n - \lim_n y_n = x_0 - y_0$, and this is in Y' since Y' is closed. The equality $x_0 = y_0 + y_0'$ shows that $E(x_0) = y_0$, and therefore (x_0, y_0) is in the graph of E. In other words, the graph of E is closed. We conclude from Corollary 12.26 that E is bounded.

7. Problems

1. Let X be a normed linear space.
   (a) Prove that the closure of the open ball of radius r and center x_0 is the closed ball of radius r and center x_0.
   (b) If X is complete, prove that any decreasing sequence of closed balls has nonempty intersection.

2. The normed linear space $C^{(N)}([a, b])$ was defined in Section 1. Prove that it is complete.

3. The normed linear space $H^\infty(D)$ and its vector subspace $A(D)$ were defined in Section 1. Prove that $H^\infty(D)$ is complete and that $A(D)$ is a closed subspace, hence complete.

4. Let X be a Banach space, let Y be a closed vector subspace, and define $\|x+Y\| = \inf_{y\in Y} \|x + y\|$ for $x + Y$ in the quotient vector space X/Y.
   (a) Show that $\|\cdot + Y\|$ is a norm for X/Y.
   (b) By replacing a Cauchy sequence $\{x_n + Y\}$ in X/Y by a subsequence such that $\|x_{n_k} - x_{n_{k+1}} + Y\| \le 2^{-k}$, show that the subsequence can be lifted to a Cauchy sequence in X and deduce that X/Y is a Banach space.

5. Let $v_1, \dots, v_n$ be vectors in an inner-product space. Their **Gram matrix** is the Hermitian matrix of inner products given by $G(v_1, \dots, v_n) = [(v_i, v_j)]$, and $\det G(v_1, \dots, v_n)$ is called their **Gram determinant**.
   (a) If $c_1, \dots, c_n$ are in $\mathbb{C}$, let $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Prove that $c^{\mathrm{tr}} G(v_1, \dots, v_n)\bar{c} = \|c_1v_1 + \cdots + c_nv_n\|^2$.
   (b) Making use of the finite-dimensional Spectral Theorem, prove that there exists a unitary matrix u such that the matrix $u^{-1}G(v_1, \dots, v_n)u$ is diagonal with diagonal entries ≥ 0.
   (c) Prove that $\det G(v_1, \dots, v_n) \ge 0$ with equality if and only if $v_1, \dots, v_n$ are linearly dependent. (This generalizes the Schwarz inequality.)

6. **(Gram–Schmidt orthogonalization process)** Let $(u_1, \dots, u_n)$ be a linearly independent ordered set in an inner-product space, and inductively define $v_1' = u_1$, $v_1 = \|v_1'\|^{-1}v_1'$, $v_k' = u_k - \sum_{j=1}^{k-1} (u, v_j)v_j$, and $v_k = \|v_k'\|^{-1}v_k'$. Prove that the vectors $v_1, \dots, v_n$ are well defined, that $v_1, \dots, v_n$ are orthonormal, and that for each k with $1 \le k \le n$, $\operatorname{span}\{v_1, \dots, v_k\} = \operatorname{span}\{u_1, \dots, u_k\}$.

7. Let H_1 and H_2 be Hilbert spaces with respective orthonormal bases $\{u_\alpha\}$ and $\{v_\beta\}$. If there is a one-one function carrying the one orthonormal basis onto the other, prove that there is a bounded linear operator $F : H_1 \to H_2$ carrying H_1 onto H_2 and preserving distances. Deduce that H_1 and H_2 are isomorphic as Hilbert spaces if and only if they have the same Hilbert space dimension.

8. Let (S, μ) be a σ-finite measure space, and let f be in $L^\infty(S, \mu)$.
   (a) Show that multiplication by f is a bounded linear operator on $L^2(S, \mu)$, and find the norm of this operator.
   (b) Find the adjoint of the operator in (a).

9. Suppose that X is a normed linear space and that its dual X^* is separable in its norm topology, with $\{x_n^*\}$ as a countable dense set. For each n, choose x_n in X with $\|x_n\| \leq 1$ and $|x_n^*(x_n)| \geq \frac{1}{2}\|x_n^*\|$. Prove that the linear span of $\{x_n\}$ is dense in X, and conclude that X^* separable implies X separable.

10. By considering the discontinuous indicator function $I_{\{s_0\}}$, where s_0 is a limit point of S, prove that the Banach space $C(S)$ is not reflexive if S is compact Hausdorff and infinite.

11. Without using the Baire Category Theorem, prove that the Uniform Boundedness Theorem for linear functionals implies the same theorem for linear operators.

12. Suppose for each n that $L_n : X \to X'$ is a bounded linear operator from a normed linear space X to a Banach space X' such that $\|L_n\| \leq C$ with C independent of n. Suppose in addition that $\{L_n(y)\}$ converges for each y in a dense subset Y of X. Prove that $L(x) = \lim_n L_n(x)$ exists for all x in X and that the resulting function $L : X \to X'$ is a bounded linear operator with $\|L\| \leq C$.

13. Let X be a normed linear space, and let $\{x_\alpha\}$ be a subset of X. If $\sup_\alpha |x^*(x_\alpha)| < \infty$ for each x^* in X^*, prove that $\sup_\alpha \|x_\alpha\| < \infty$.

14. Let X be a Banach space. A subset E of X is **convex** if it contains all points $(1-t)x + ty$ with $0 \leq t \leq 1$ whenever it contains x and y.
    (a) Show that any closed ball $\{y \mid |y - x| \leq r\}$ is convex.
    (b) Give an example of a decreasing sequence of nonempty bounded closed convex sets in a Banach space with empty intersection.

15. Let X and Y be Banach spaces, and let L be a bounded linear operator from X onto Y. Suppose that $\{y_n\}$ is a convergent sequence in Y with limit y_0. Prove that there exists a constant M and a sequence $\{x_n\}$ in X such that $\|x_n\| \leq M\|y_n\|$ for all n, $L(x_n) = y_n$ for all n, and $\{x_n\}$ is convergent.

Problems 16–18 introduce "Banach limits," a kind of universal summability method. Let X be the real Banach space of real-valued bounded sequences $s = \{s_n\}_{n=1}^\infty$ with the supremum norm.

16. Let X_0 be the smallest closed vector subspace of X containing all sequences with terms $s_1, s_2 - s_2, s_3 - s_2, \ldots$ such that $\{s_n\}$ is in X. Prove that the sequence e with all terms 1 is not in X_0.

17. A **Banach limit** is defined to be any member x^* of X^* with $\|x^*\| = 1, x^*(e) = 1$, and $x^*(x_0) = 0$ for all x_0 in X_0. Prove that a Banach limit exists.

18. Let $\text{LIM}_{n\to\infty} s_n$ denote the value of a Banach limit when applied to the member $\{s_n\}$ of X. Prove that this satisfies
    (a) $\text{LIM}_{n\to\infty} s_n \geq 0$ if $s_n \geq 0$ for all n.
    (b) $\text{LIM}_{n\to\infty} s_{n+1} = \text{LIM}_{n\to\infty} s_n$ for every $\{s_n\}$ in X.
    (c) $\text{LIM}_{n\to\infty} s_n = 0$ if all terms s_n are 0 for n sufficiently large.
    (d) $\liminf_n s_n \leq \text{LIM}_{n\to\infty} s_n \leq \limsup_n s_n$ for all $\{s_n\}$ in X.
    (e) $\text{LIM}_{n\to\infty} s_n = c$ if $\{s_n\}$ is convergent with limit c.

Problems 19–24 establish the Jordan and von Neumann Theorem that a normed linear space satisfying the parallelogram law acquires its norm from an inner product, the definition of the inner product being $(x, y) = \sum_k \frac{i^k}{4}\|x + i^k y\|^2$, where the sum extends for $k \in \{0, 2\}$ if the scalars are real and extends for $k \in \{0, 1, 2, 3\}$ if the scalars are complex. The norm is recovered from the inner product by the usual formula $(x, x) = \|x\|^2$. Thus let X be a normed linear space with norm $\|\cdot\|$ such that the parallelogram law holds.

19. Check from the definition of (x, y) that $(x, x) = \|x\|^2$, that $(x, x) \geq 0$ with equality if and only if $x = 0$, and that $(x, y) = \overline{(y, x)}$.

20. Prove the identity
$$\|x + y + z\|^2 = \|x + y\|^2 + \|x + z\|^2 + \|y + z\|^2 - \|x\|^2 - \|y\|^2 - \|z\|^2$$
for all x, y, z in X.

21. Derive the formula $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ from the identity in the previous problem.

22. Let D be the set of rationals if the scalars are real, or the set of all $a + bi$ with a and b rational if the scalars are complex. Using the definition of (x, y) and the result of the previous problem, prove that $(rx, y) = r(x, y)$ if r is in D.

23. By considering $\|x - ry\|^2$ for r in D with r tending to $(x, y)/\|y\|^2$, prove that $(\cdot\,, \cdot)$ satisfies the Schwarz inequality.

24. By estimating $|r(x, y) - (cx, y)|$ with the Schwarz inequality when c is a scalar and r is a member of D tending to c, prove that $c(x, y) = (cx, y)$, thereby completing the proof that $(\cdot\,, \cdot)$ is an inner product.

Problems 25–27 establish some properties of the Banach space $\mathcal{B}(X, Y)$, where X and Y are Banach spaces.

25. Prove that the function from $\mathcal{B}(X, Y) \times X$ to Y given by $(L, x) \mapsto L(x)$ is continuous.

26. Prove that if L and L_n are members of $L(X, Y)$ such that $\lim_n L_n(x) = L(x)$ for all $x \in X$, then $\|L\| \leq \sup_n \|L_n\| < \infty$. Give an example where $\|L\| < \sup_n \|L_n\|$.

27. Prove that if L and L_n are members of $L(X, Y)$ such that $\lim_n L_n(x) = L(x)$ for all $x \in X$ and if $\{u_n\}$ is a sequence in X with $\lim u_n = u$, then $\lim_n L_n(u_n) = L(u)$.