CHAPTER

FOUR

ELEMENTARY HILBERT SPACE THEORY

Inner Products and Linear Functionals

4.1 Definition A complex vector space H is called an *inner product space* (or *unitary space*) if to each ordered pair of vectors x and $y \in H$ there is associated a complex number (x, y), the so-called "inner product" (or "scalar product") of x and y, such that the following rules hold:

(*a*) $(y, x) = \overline{(x, y)}$. (The bar denotes complex conjugation.)
(*b*) $(x + y, z) = (x, z) + (y, z)$ if x, y, and $z \in H$.
(*c*) $(\alpha x, y) = \alpha(x, y)$ if x and $y \in H$ and α is a scalar.
(*d*) $(x, x) \geq 0$ for all $x \in H$.
(*e*) $(x, x) = 0$ only if $x = 0$.

Let us list some immediate consequences of these axioms:

(*c*) implies that $(0, y) = 0$ for all $y \in H$.
(*b*) and (*c*) may be combined into the statement: *For every $y \in H$, the mapping $x \rightarrow (x, y)$ is a linear functional on H.*
(*a*) and (*c*) show that $(x, \alpha y) = \bar{\alpha}(x, y)$.
(*a*) and (*b*) imply the second distributive law:

$$(z, x + y) = (z, x) + (z, y).$$

By (*d*), we may define $\|x\|$, the *norm* of the vector $x \in H$, to be the non-negative square root of (x, x). Thus

(*f*) $$\|x\|^2 = (x, x).$$

4.2 The Schwarz Inequality *The properties* 4.1 (*a*) *to* (*d*) *imply that*

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

for all x and $y \in H$.

PROOF Put $A = \|x\|^2$, $B = |(x, y)|$, and $C = \|y\|^2$. There is a complex number α such that $|\alpha| = 1$ and $\alpha(y, x) = B$. For any real r, we then have

$$(x - r\alpha y, x - r\alpha y) = (x, x) - r\alpha(y, x) - r\bar{\alpha}(x, y) + r^2(y, y). \tag{1}$$

The expression on the left is real and not negative. Hence

$$A - 2Br + Cr^2 \geq 0 \tag{2}$$

for every real r. If $C = 0$, we must have $B = 0$; otherwise (2) is false for large positive r. If $C > 0$, take $r = B/C$ in (2), and obtain $B^2 \leq AC$. ////

4.3 The Triangle Inequality *For x and $y \in H$, we have*

$$\|x + y\| \leq \|x\| + \|y\|.$$

PROOF By the Schwarz inequality,

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

4.4 Definition It follows from the triangle inequality that

$$\|x - z\| \leq \|x - y\| + \|y - z\| \qquad (x, y, z \in H). \tag{1}$$

If we define the distance between x and y to be $\|x - y\|$, all the axioms for a metric space are satisfied; here, for the first time, we use part (*e*) of Definition 4.1.

Thus H is now a metric space. If this metric space is *complete*, i.e., if every Cauchy sequence converges in H, then H is called a *Hilbert space*.

Throughout the rest of this chapter, the letter H will denote a Hilbert space.

4.5 Examples

(*a*) For any fixed n, the set C^n of all n-tuples

$$x = (\xi_1, \ldots, \xi_n),$$

where $\xi_1, \ldots, \xi_n$ are complex numbers, is a Hilbert space if addition and scalar multiplication are defined componentwise, as usual, and if

$$(x, y) = \sum_{j=1}^{n} \xi_j \bar{\eta}_j \qquad (y = (\eta_1, \ldots, \eta_n)).$$

(b) If μ is any positive measure, $L^2(\mu)$ is a Hilbert space, with inner product

$$(f, g) = \int_X f\bar{g}\, d\mu.$$

The integrand on the right is in $L^1(\mu)$, by Theorem 3.8, so that (f, g) is well defined. Note that

$$\|f\| = (f, f)^{1/2} = \left\{\int_X |f|^2\, d\mu\right\}^{1/2} = \|f\|_2.$$

The completeness of $L^2(\mu)$ (Theorem 3.11) shows that $L^2(\mu)$ is indeed a Hilbert space. [We recall that $L^2(\mu)$ should be regarded as a space of *equivalence classes* of functions; compare the discussion in Sec. 3.10.]

For $H = L^2(\mu)$, the inequalities 4.2 and 4.3 turn out to be special cases of the inequalities of Hölder and Minkowski.

Note that Example (a) is a special case of (b). What is the measure in (a)?

(c) The vector space of all continuous complex functions on [0, 1] is an inner product space if

$$(f, g) = \int_0^1 f(t)\overline{g(t)}\, dt$$

but is not a Hilbert space.

4.6 Theorem *For any fixed $y \in H$, the mappings*

$$x \to (x, y), \quad x \to (y, x), \quad x \to \|x\|$$

are continuous functions on H.

PROOF The Schwarz inequality implies that

$$|(x_1, y) - (x_2, y)| = |(x_1 - x_2, y)| \le \|x_1 - x_2\|\,\|y\|,$$

which proves that $x \to (x, y)$ is, in fact, uniformly continuous, and the same is true for $x \to (y, x)$. The triangle inequality $\|x_1\| \le \|x_1 - x_2\| + \|x_2\|$ yields

$$\|x_1\| - \|x_2\| \le \|x_1 - x_2\|,$$

and if we interchange x_1 and x_2 we see that

$$\big|\,\|x_1\| - \|x_2\|\,\big| \le \|x_1 - x_2\|$$

for all x_1 and $x_2 \in H$. Thus $x \to \|x\|$ is also uniformly continuous. ////

4.7 Subspaces A subset M of a vector space V is called a *subspace* of V if M is itself a vector space, relative to the addition and scalar multiplication which are defined in V. A necessary and sufficient condition for a set $M \subset V$ to be a subspace is that $x + y \in M$ and $\alpha x \in M$ whenever x and $y \in M$ and α is a scalar.

In the vector space context, the word "subspace" will always have this meaning. Sometimes, for emphasis, we may use the term "linear subspace" in place of subspace.

For example, if V is the vector space of all complex functions on a set S, the set of all bounded complex functions on S is a subspace of V, but the set of all $f \in V$ with $|f(x)| \le 1$ for all $x \in S$ is not. The real vector space R^3 has the following subspaces, and no others: (a) R^3, (b) all planes through the origin 0, (c) all straight lines through 0, and (d) $\{0\}$.

A *closed subspace* of H is a subspace that is a closed set relative to the topology induced by the metric of H.

Note that *if M is a subspace of H, so is its closure $\bar{M}$.* To see this, pick x and y in $\bar{M}$ and let α be a scalar. There are sequences $\{x_n\}$ and $\{y_n\}$ in M that converge to x and y, respectively. It is then easy to verify that $x_n + y_n$ and αx_n converge to $x + y$ and αx, respectively. Thus $x + y \in \bar{M}$ and $\alpha x \in \bar{M}$.

4.8 Convex Sets A set E in a vector space V is said to be *convex* if it has the following geometric property: Whenever $x \in E$, $y \in E$, and $0 < t < 1$, the point

$$z_t = (1 - t)x + ty$$

also lies in E. As t runs from 0 to 1, one may visualize z_t as describing a straight line segment in V, from x to y. Convexity requires that E contain the segments between any two of its points.

It is clear that every subspace of V is convex.

Also, if E is convex, so is each of its translates

$$E + x = \{y + x : y \in E\}.$$

4.9 Orthogonality If $(x, y) = 0$ for some x and $y \in H$, we say that x is orthogonal to y, and sometimes write $x \perp y$. Since $(x, y) = 0$ implies $(y, x) = 0$, the relation $\perp$ is symmetric.

Let $x^\perp$ denote the set of all $y \in H$ which are orthogonal to x; and if M is a subspace of H, let $M^\perp$ be the set of all $y \in H$ which are orthogonal to every $x \in M$.

Note that $x^\perp$ is a subspace of H, since $x \perp y$ and $x \perp y'$ implies $x \perp (y + y')$ and $x \perp \alpha y$. Also, $x^\perp$ is precisely the set of points where the continuous function $y \to (x, y)$ is 0. Hence $x^\perp$ is a *closed* subspace of H. Since

$$M^\perp = \bigcap_{x \in M} x^\perp,$$

$M^\perp$ is an intersection of closed subspaces, and it follows that *$M^\perp$ is a closed subspace of H.*

4.10 Theorem *Every nonempty, closed, convex set E in a Hilbert space H contains a unique element of smallest norm.*

In other words, there is one and only one $x_0 \in E$ such that $\|x_0\| \le \|x\|$ for every $x \in E$.

PROOF An easy computation, using only the properties listed in Definition 4.1, establishes the identity

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \qquad (x \text{ and } y \in H). \tag{1}$$

This is known as the *parallelogram law*: If we interpret $\|x\|$ to be the length of the vector x, (1) says that the sum of the squares of the diagonals of a parallelogram is equal to the sum of the squares of its sides, a familiar proposition in plane geometry.

Let $\delta = \inf\{\|x\|: x \in E\}$. For any x and $y \in E$, we apply (1) to $\frac{1}{2}x$ and $\frac{1}{2}y$ and obtain

$$\tfrac{1}{4}\|x - y\|^2 = \tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}\|y\|^2 - \left\|\frac{x + y}{2}\right\|^2. \tag{2}$$

Since E is convex, $(x + y)/2 \in E$. Hence

$$\|x - y\|^2 \le 2\|x\|^2 + 2\|y\|^2 - 4\delta^2 \qquad (x \text{ and } y \in E). \tag{3}$$

If also $\|x\| = \|y\| = \delta$, then (3) implies $x = y$, and we have proved the uniqueness assertion of the theorem.

The definition of δ shows that there is a sequence $\{y_n\}$ in E so that $\|y_n\| \to \delta$ as $n \to \infty$. Replace x and y in (3) by y_n and y_m. Then, as $n \to \infty$ and $m \to \infty$, the right side of (3) will tend to 0. This shows that $\{y_n\}$ is a Cauchy sequence. Since H is complete, there exists an $x_0 \in H$ so that $y_n \to x_0$, i.e., $\|y_n - x_0\| \to 0$, as $n \to \infty$. Since $y_n \in E$ and E is closed, $x_0 \in E$. Since the norm is a continuous function on H (Theorem 4.6), it follows that

$$\|x_0\| = \lim_{n \to \infty} \|y_n\| = \delta. \qquad ////$$

4.11 Theorem *Let M be a closed subspace of a Hilbert space H.*

(*a*) *Every $x \in H$ has then a unique decomposition*

$$x = Px + Qx$$

into a sum of $Px \in M$ and $Qx \in M^\perp$.

(*b*) *Px and Qx are the nearest points to x in M and in $M^\perp$, respectively.*

(*c*) *The mappings $P: H \to M$ and $Q: H \to M^\perp$ are linear.*

(*d*) $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$.

Corollary *If $M \ne H$, then there exists $y \in H$, $y \ne 0$, such that $y \perp M$.*

P and Q are called the *orthogonal projections* of H onto M and $M^\perp$.

PROOF As regards the uniqueness in (*a*), suppose that $x' + y' = x'' + y''$ for some vectors x', x'' in M and y', y'' in $M^\perp$. Then

$$x' - x'' = y'' - y'.$$

Since $x' - x'' \in M$, $y'' - y' \in M^\perp$, and $M \cap M^\perp = \{0\}$ [an immediate consequence of the fact that $(x, x) = 0$ implies $x = 0$], we have $x'' = x'$, $y'' = y'$.

To prove the existence of the decomposition, note that the set

$$x + M = \{x + y: y \in M\}$$

is closed and convex. Define Qx to be the element of smallest norm in $x + M$; this exists, by Theorem 4.10. Define $Px = x - Qx$.

Since $Qx \in x + M$, it is clear that $Px \in M$. Thus P maps H into M.

To prove that Q maps H into $M^\perp$ we show that $(Qx, y) = 0$ for all $y \in M$. Assume $\|y\| = 1$, without loss of generality, and put $z = Qx$. The minimizing property of Qx shows that

$$(z, z) = \|z\|^2 \le \|z - \alpha y\|^2 = (z - \alpha y, z - \alpha y)$$

for every scalar α. This simplifies to

$$0 \le -\alpha(y, z) - \bar{\alpha}(z, y) + \alpha\bar{\alpha}.$$

With $\alpha = (z, y)$, this gives $0 \le -|(z, y)|^2$, so that $(z, y) = 0$. Thus $Qx \in M^\perp$.

We have already seen that $Px \in M$. If $y \in M$, it follows that

$$\|x - y\|^2 = \|Qx + (Px - y)\|^2 = \|Qx\|^2 + \|Px - y\|^2$$

which is obviously minimized when $y = Px$.

We have now proved (*a*) and (*b*). If we apply (*a*) to x, to y, and to $\alpha x + \beta y$, we obtain

$$P(\alpha x + \beta y) - \alpha Px - \beta Py = \alpha Qx + \beta Qy - Q(\alpha x + \beta y).$$

The left side is in M, the right side in $M^\perp$. Hence both are 0, so P and Q are linear.

Since $Px \perp Qx$, (*d*) follows from (*a*).

To prove the corollary, take $x \in H$, $x \notin M$, and put $y = Qx$. Since $Px \in M$, $x \ne Px$, hence $y = x - Px \ne 0$. ////

We have already observed that $x \to (x, y)$ is, for each $y \in H$, a continuous linear functional on H. It is a very important fact that *all* continuous linear functionals on H are of this type.

4.12 Theorem *If L is a continuous linear functional on H, then there is a unique $y \in H$ such that*

$$Lx = (x, y) \qquad (x \in H). \tag{1}$$

PROOF If $Lx = 0$ for all x, take $y = 0$. Otherwise, define

$$M = \{x: Lx = 0\}. \tag{2}$$

The linearity of L shows that M is a subspace. The continuity of L shows that M is closed. Since $Lx \neq 0$ for some $x \in H$, Theorem 4.11 shows that $M^\perp$ does not consist of 0 alone.

Hence there exists $z \in M^\perp$, with $\|z\| = 1$. Put

$$u = (Lx)z - (Lz)x. \tag{3}$$

Since $Lu = (Lx)(Lz) - (Lz)(Lx) = 0$, we have $u \in M$. Thus $(u, z) = 0$. This gives

$$Lx = (Lx)(z, z) = (Lz)(x, z). \tag{4}$$

Thus (1) holds with $y = \alpha z$, where $\bar{\alpha} = Lz$.

The uniqueness of y is easily proved, for if $(x, y) = (x, y')$ for all $x \in H$, set $z = y - y'$; then $(x, z) = 0$ for all $x \in H$; in particular, $(z, z) = 0$, hence $z = 0$. ////

Orthonormal Sets

4.13 Definitions If V is a vector space, if $x_1, \ldots, x_k \in V$, and if $c_1, \ldots, c_k$ are scalars, then $c_1x_1 + \cdots + c_kx_k$ is called a *linear combination* of $x_1, \ldots, x_k$. The set $\{x_1, \ldots, x_k\}$ is called *independent* if $c_1x_1 + \cdots + c_kx_k = 0$ implies that $c_1 = \cdots = c_k = 0$. A set $S \subset V$ is independent if every finite subset of S is independent. The set $[S]$ of all linear combinations of all finite subsets of S (also called the set of all *finite linear combinations* of members of S) is clearly a vector space; $[S]$ is the smallest subspace of V which contains S; $[S]$ is called the *span* of S, or the space spanned by S.

A set of vectors u_α in a Hilbert space H, where α runs through some index set A, is called *orthonormal* if it satisfies the orthogonality relations $(u_\alpha, u_\beta) = 0$ for all $\alpha \neq \beta$, $\alpha \in A$, and $\beta \in A$, and if it is normalized so that $\|u_\alpha\| = 1$ for each $\alpha \in A$. In other words, $\{u_\alpha\}$ is orthonormal provided that

$$(u_\alpha, u_\beta) = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta. \end{cases} \tag{1}$$

If $\{u_\alpha: \alpha \in A\}$ is orthonormal, we associate with each $x \in H$ a complex function $\hat{x}$ on the index set A, defined by

$$\hat{x}(\alpha) = (x, u_\alpha) \qquad (\alpha \in A). \tag{2}$$

One sometimes calls the numbers $\hat{x}(\alpha)$ the *Fourier coefficients* of x, relative to the set $\{u_\alpha\}$.

We begin with some simple facts about *finite* orthonormal sets.

4.14 Theorem *Suppose that $\{u_\alpha: \alpha \in A\}$ is an orthonormal set in H and that F is a finite subset of A. Let M_F be the span of $\{u_\alpha: \alpha \in F\}$.*

(a) If φ is a complex function on A that is 0 outside F, then there is a vector $y \in M_F$, namely

$$y = \sum_{\alpha \in F} \varphi(\alpha)u_\alpha \tag{1}$$

that has $\hat{y}(\alpha) = \varphi(\alpha)$ for every $\alpha \in A$. Also,

$$\|y\|^2 = \sum_{\alpha \in F} |\varphi(\alpha)|^2. \tag{2}$$

(b) If $x \in H$ and

$$s_F(x) = \sum_{\alpha \in F} \hat{x}(\alpha)u_\alpha \tag{3}$$

then

$$\|x - s_F(x)\| < \|x - s\| \tag{4}$$

for every $s \in M_F$, except for $s = s_F(x)$, and

$$\sum_{\alpha \in F} |\hat{x}(\alpha)|^2 \leq \|x\|^2. \tag{5}$$

PROOF Part (*a*) is an immediate consequence of the orthogonality relations 4.13(1).

In the proof of (*b*), let us write s_F in place of $s_F(x)$, and note that $\hat{s}_F(\alpha) = \hat{x}(\alpha)$ for all $\alpha \in F$. This says that $(x - s_F) \perp u_\alpha$ if $\alpha \in F$, hence $(x - s_F) \perp (s_F - s)$ for every $s \in M_F$, and therefore

$$\|x - s\|^2 = \|(x - s_F) + (s_F - s)\|^2 = \|x - s_F\|^2 + \|s_F - s\|^2. \tag{6}$$

This gives (4). With $s = 0$, (6) gives $\|s_F\|^2 \leq \|x\|^2$, which is the same as (5), because of (2). ////

The inequality (4) states that the "partial sum" $s_F(x)$ of the "Fourier series" $\sum \hat{x}(\alpha)u_\alpha$ of x is the unique best approximation to x in M_F, relative to the metric defined by the Hilbert space norm.

4.15 We want to drop the finiteness condition that appears in Theorem 4.14 (thus obtaining Theorems 4.17 and 4.18) without even restricting ourselves to sets that are necessarily countable. For this reason it seems advisable to clarify the meaning of the symbol $\sum_{\alpha \in A} \varphi(\alpha)$ when α ranges over an arbitrary set A.

Assume $0 \leq \varphi(\alpha) \leq \infty$ for each $\alpha \in A$. Then

$$\sum_{\alpha \in A} \varphi(\alpha) \tag{1}$$

denotes the supremum of the set of all finite sums $\varphi(\alpha_1) + \cdots + \varphi(\alpha_n)$, where $\alpha_1, \ldots, \alpha_n$ are distinct members of A.

A moment's consideration will show that *the sum* (1) *is thus precisely the Lebesgue integral of* φ *relative to the counting measure* μ *on* A.

In this context one usually writes $\ell^p(A)$ for $L^p(\mu)$. A complex function φ with domain A is thus in $\ell^2(A)$ if and only if

$$\sum_{\alpha \in A} |\varphi(\alpha)|^2 < \infty. \tag{2}$$

Example 4.5(*b*) shows that $\ell^2(A)$ is a Hilbert space, with inner product

$$(\varphi, \psi) = \sum_{\alpha \in A} \varphi(\alpha)\overline{\psi(\alpha)}. \tag{3}$$

Here, again, the sum over A stands for the integral of $\varphi\bar{\psi}$ with respect to the counting measure; note that $\varphi\bar{\psi} \in \ell^1(A)$ because φ and ψ are in $\ell^2(A)$.

Theorem 3.13 shows that *the functions* φ *that are zero except on some finite subset of* A *are dense in* $\ell^2(A)$.

Moreover, if $\varphi \in \ell^2(A)$, then $\{\alpha \in A: \varphi(\alpha) \neq 0\}$ *is at most countable*. For if A_n is the set of all α where $|\varphi(\alpha)| > 1/n$, then the number of elements of A_n is at most

$$\sum_{\alpha \in A_n} |n\varphi(\alpha)|^2 \le n^2 \sum_{\alpha \in A} |\varphi(\alpha)|^2.$$

Each A_n $(n = 1, 2, 3, \ldots)$ is thus a finite set.

The following lemma about complete metric spaces will make it easy to pass from finite orthonormal sets to infinite ones.

4.16 Lemma *Suppose that*

(*a*) X *and* Y *are metric spaces,* X *is complete,*
(*b*) $f: X \to Y$ *is continuous,*
(*c*) X *has a dense subset* X_0 *on which* f *is an isometry, and*
(*d*) $f(X_0)$ *is dense in* Y.
Then f *is an isometry of* X *onto* Y.

The most important part of the conclusion is that f maps X *onto all of* Y.

Recall that an isometry is simply a mapping that preserves distances. Thus, by assumption, the distance between $f(x_1)$ and $f(x_2)$ in Y is equal to that between x_1 and x_2 in X, for all points x_1, x_2 in X_0.

PROOF The fact that f is an isometry on X is an immediate consequence of the continuity of f, since X_0 is dense in X.

Pick $y \in Y$. Since $f(X_0)$ is dense in Y, there is a sequence $\{x_n\}$ in X_0 such that $f(x_n) \to y$ as $n \to \infty$. Thus $\{f(x_n)\}$ is a Cauchy sequence in Y. Since f is an isometry on X_0, it follows that $\{x_n\}$ is also a Cauchy sequence. The completeness of X implies now that $\{x_n\}$ converges to some $x \in X$, and the continuity of f shows that $f(x) = \lim f(x_n) = y$. ////

4.17 Theorem *Let* $\{u_\alpha: \alpha \in A\}$ *be an orthonormal set in* H, *and let* P *be the space of all finite linear combinations of the vectors* u_α.

The inequality

$$\sum_{\alpha \in A} |\hat{x}(\alpha)|^2 \le \|x\|^2 \tag{1}$$

holds then for every $x \in H$, *and* $x \to \hat{x}$ *is a continuous linear mapping of* H *onto* $\ell^2(A)$ *whose restriction to the closure* $\bar{P}$ *of* P *is an isometry of* $\bar{P}$ *onto* $\ell^2(A)$.

PROOF Since the inequality 4.14(5) holds for every finite set $F \subset A$, we have (1), the so-called *Bessel inequality*.

Define f on H by $f(x) = \hat{x}$. Then (1) shows explicitly that f maps H into $\ell^2(A)$. The linearity of f is obvious. If we apply (1) to $x - y$ we see that

$$\|f(y) - f(x)\|_2 = \|\hat{y} - \hat{x}\|_2 \le \|y - x\|.$$

Thus f is continuous. Theorem 4.14(*a*) shows that f is an isometry of P onto the dense subspace of $\ell^2(A)$ consisting of those functions whose support is a finite set $F \subset A$. The theorem follows therefore from Lemma 4.16, applied with $X = \bar{P}$, $X_0 = P$, $Y = \ell^2(A)$; note that $\bar{P}$, being a closed subset of the complete metric space H, is itself complete. ////

The fact that the mapping $x \to \hat{x}$ carries H *onto* $\ell^2(A)$ is known as the *Riesz-Fischer theorem.*

4.18 Theorem *Let* $\{u_\alpha: \alpha \in A\}$ *be an orthonormal set in* H. *Each of the following four conditions on* $\{u_\alpha\}$ *implies the other three*:

(i) $\{u_\alpha\}$ *is a maximal orthonormal set in* H.
(ii) *The set* P *of all finite linear combinations of members of* $\{u_\alpha\}$ *is dense in* H.
(iii) *The equality*

$$\sum_{\alpha \in A} |\hat{x}(\alpha)|^2 = \|x\|^2$$

holds for every $x \in H$.

(iv) *The equality*

$$\sum_{\alpha \in A} \hat{x}(\alpha)\overline{\hat{y}(\alpha)} = (x, y)$$

holds for all $x \in H$ *and* $y \in H$.

The last formula is known as *Parseval's identity*. Observe that $\hat{x}$ and $\hat{y}$ are in $\ell^2(A)$, hence $\hat{x}\bar{\hat{y}}$ is in $\ell^1(A)$, so that the sum in (iv) is well defined. Of course, (iii) is the special case $x = y$ of (iv).

Maximal orthonormal sets are often called *complete orthonormal sets* or *orthonormal bases*.

PROOF To say that $\{u_\alpha\}$ is maximal means simply that no vector of H can be adjoined to $\{u_\alpha\}$ in such a way that the resulting set is still orthonormal. This happens precisely when there is no $x \neq 0$ in H that is orthogonal to every u_α.

We shall prove that (i) $\to$ (ii) $\to$ (iii) $\to$ (iv) $\to$ (i).

If P is not dense in H, then its closure $\bar{P}$ is not all of H, and the corollary to Theorem 4.11 implies that $P^\perp$ contains a nonzero vector. Thus $\{u_\alpha\}$ is not maximal when P is not dense, and (i) implies (ii).

If (ii) holds, so does (iii), by Theorem 4.17.

The implication (iii) $\to$ (iv) follows from the easily proved Hilbert space identity (sometimes called the "polarization identity")

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2$$

which expresses the inner product (x, y) in terms of norms and which is equally valid with $\hat{x}$, $\hat{y}$ in place of x, y, simply because $\ell^2(A)$ is also a Hilbert space. (See Exercise 19 for other identities of this type.) Note that the sums in (iii) and (iv) are $\|\hat{x}\|_2^2$ and $(\hat{x}, \hat{y})$, respectively.

Finally, if (i) is false, there exists $u \neq 0$ in H so that $(u, u_\alpha) = 0$ for all $\alpha \in A$. If $x = y = u$, then $(x, y) = \|u\|^2 > 0$ but $\hat{x}(\alpha) = 0$ for all $\alpha \in A$, hence (iv) fails. Thus (iv) implies (i), and the proof is complete. ////

4.19 Isomorphisms Speaking informally, two algebraic systems of the same nature are said to be *isomorphic* if there is a one-to-one mapping of one onto the other which preserves all relevant properties. For instance, we may ask whether two groups are isomorphic or whether two fields are isomorphic. Two vector spaces are isomorphic if there is a one-to-one *linear* mapping of one onto the other. The linear mappings are the ones which preserve the relevant concepts in a vector space, namely, addition and scalar multiplication.

In the same way, two Hilbert spaces H_1 and H_2 are isomorphic if there is a one-to-one linear mapping Λ of H_1 onto H_2 which also preserves inner products: $(\Lambda x, \Lambda y) = (x, y)$ for all x and $y \in H_1$. Such a Λ is an isomorphism (or, more specifically, a *Hilbert space isomorphism*) of H_1 onto H_2. Using this terminology, Theorems 4.17 and 4.18 yield the following statement:

If $\{u_\alpha: \alpha \in A\}$ is a maximal orthonormal set in a Hilbert space H, and if $\hat{x}(\alpha) = (x, u_\alpha)$, then the mapping $x \to \hat{x}$ is a Hilbert space isomorphism of H onto $\ell^2(A)$.

One can prove (we shall omit this) that $\ell^2(A)$ and $\ell^2(B)$ are isomorphic if and only if the sets A and B have the same cardinal number. But we shall prove that every nontrivial Hilbert space (this means that the space does not consist of 0 alone) is isomorphic to some $\ell^2(A)$, by proving that every such space contains a maximal orthonormal set (Theorem 4.22). The proof will depend on a property of partially ordered sets which is equivalent to the axiom of choice.

4.20 Partially Ordered Sets A set $\mathscr{P}$ is said to be *partially ordered* by a binary relation $\leq$ if

(*a*) $a \leq b$ and $b \leq c$ implies $a \leq c$.
(*b*) $a \leq a$ for every $\alpha \in \mathscr{P}$.
(c) $a \leq b$ and $b \leq a$ implies $a = b$.

A subset $\mathscr{Q}$ of a partially ordered set $\mathscr{P}$ is said to be *totally ordered* (or *linearly ordered*) if every pair $a, b \in \mathscr{Q}$ satisfies either $\alpha \leq b$ or $b \leq a$.

For example, every collection of subsets of a given set is partially ordered by the inclusion relation $\subset$.

To give a more specific example, let $\mathscr{P}$ be the collection of all open subsets of the plane, partially ordered by set inclusion, and let $\mathscr{Q}$ be the collection of all open circular discs with center at the origin. Then $\mathscr{Q} \subset \mathscr{P}$, $\mathscr{Q}$ is totally ordered by $\subset$, and $\mathscr{Q}$ is a *maximal* totally ordered subset of $\mathscr{P}$. This means that if any member of $\mathscr{P}$ not in $\mathscr{Q}$ is adjoined to $\mathscr{Q}$, the resulting collection of sets is no longer totally ordered by $\subset$.

4.21 The Hausdorff Maximality Theorem *Every nonempty partially ordered set contains a maximal totally ordered subset.*

This is a consequence of the axiom of choice and is, in fact, equivalent to it; another (very similar) form of it is known as Zorn's lemma. We give the proof in the Appendix.

If now H is a nontrivial Hilbert space, then there exists a $u \in H$ with $\|u\| = 1$, so that there is a nonempty orthonormal set in H. The existence of a maximal orthonormal set is therefore a consequence of the following theorem:

4.22 Theorem *Every orthonormal set B in a Hilbert space H is contained in a maximal orthonormal set in H.*

PROOF Let $\mathscr{P}$ be the class of all orthonormal sets in H which contain the given set B. Partially order $\mathscr{P}$ by set inclusion. Since $B \in \mathscr{P}$, $\mathscr{P} \neq \varnothing$. Hence $\mathscr{P}$ contains a maximal totally ordered class Ω. Let S be the union of all members of Ω. It is clear that $B \subset S$. We claim that S is a maximal orthonormal set:

If u_1 and $u_2 \in S$, then $u_1 \in A_1$ and $u_2 \in A_2$ for some A_1 and $A_2 \in \Omega$. Since Ω is total ordered, $A_1 \subset A_2$ (or $A_2 \subset A_1$), so that $u_1 \in A_2$ and $u_2 \in A_2$. Since A_2 is orthonormal, $(u_1, u_2) = 0$ if $u_1 \neq u_2$, $(u_1, u_2) = 1$ if $u_1 = u_2$. Thus S is an orthonormal set.

Suppose S is not maximal. Then S is a proper subset of an orthonormal set S^*. Clearly, $S^* \notin \Omega$, and S^* contains every member of Ω. Hence we may adjoin S^* to Ω and still have a total order. This contradicts the maximality of Ω. ////

Trigonometric Series

4.23 Definitions Let T be the unit circle in the complex plane, i.e., the set of all complex numbers of absolute value 1. If F is any function on T and if f is defined on R^1 by

$$f(t) = F(e^{it}), \tag{1}$$

then f is a periodic function of period 2π. This means that $f(t + 2\pi) = f(t)$ for all real t. Conversely, if f is a function on R^1, with period 2π, then there is a function F on T such that (1) holds. Thus we may identify functions on T with 2π-periodic functions on R^1; and, for simplicity of notation, we shall sometimes write $f(t)$ rather than $f(e^{it})$, even if we think of f as being defined on T.

With these conventions in mind, we define $L^p(T)$, for $1 \le p < \infty$, to be the class of all complex, Lebesgue measurable, 2π-periodic functions on R^1 for which the norm

$$\|f\|_p = \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^p\, dt\right\}^{1/p} \tag{2}$$

is finite.

In other words, we are looking at $L^p(\mu)$, where μ is Lebesgue measure on $[0, 2\pi]$ (or on T), divided by 2π. $L^\infty(T)$ will be the class of all 2π-periodic members of $L^\infty(R^1)$, with the essential supremum norm, and $C(T)$ consists of all continuous complex functions on T (or, equivalently, of all continuous, complex, 2π-periodic functions on R^1), with norm

$$\|f\|_\infty = \sup_t |f(t)|, \tag{3}$$

The factor $1/(2\pi)$ in (2) simplifies the formalism we are about to develop. For instance, the L^p-norm of the constant function 1 is 1.

A *trigonometric polynomial* is a finite sum of the form

$$f(t) = a_0 + \sum_{n=1}^{N} (a_n \cos nt + b_n \sin nt) \qquad (t \in R^1) \tag{4}$$

where $a_0, a_1, \ldots, a_N$ and $b_1, \ldots, b_N$ are complex numbers. On account of the Euler identities, (4) can also be written in the form

$$f(t) = \sum_{n=-N}^{N} c_n e^{int} \tag{5}$$

which is more convenient for most purposes. It is clear that every trigonometric polynomial has period 2π.

We shall denote the set of all integers (positive, zero, and negative) by Z, and put

$$u_n(t) = e^{int} \qquad (n \in Z). \tag{6}$$

If we define the inner product in $L^2(T)$ by

$$(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt \tag{7}$$

[note that this is in agreement with (2)], an easy computation shows that

$$(u_n, u_m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)t}\, dt = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{if } n \ne m. \end{cases} \tag{8}$$

Thus $\{u_n\colon n \in Z\}$ is an orthonormal set in $L^2(T)$, usually called the *trigonometric system*. We shall now prove that this system is maximal, and shall then derive concrete versions of the abstract theorems previously obtained in the Hilbert space context.

4.24 The Completeness of the Trigonometric System Theorem 4.18 shows that the maximality (or completeness) of the trigonometric system will be proved as soon as we can show that the set of all trigonometric polynomials is dense in $L^2(T)$. Since $C(T)$ is dense in $L^2(T)$, by Theorem 3.14 (note that T is compact), it suffices to show that to every $f \in C(T)$ and to every $\epsilon > 0$ there is a trigonometric polynomial P such that $\|f - P\|_2 < \epsilon$. Since $\|g\|_2 \le \|g\|_\infty$ for every $g \in C(T)$, the estimate $\|f - P\|_2 < \epsilon$ will follow from $\|f - P\|_\infty < \epsilon$, and it is this estimate which we shall prove.

Suppose we had trigonometric polynomials $Q_1, Q_2, Q_3, \ldots$, with the following properties:

(*a*) $\quad Q_k(t) \ge 0$ for $t \in R^1$.

(*b*) $$\frac{1}{2\pi}\int_{-\pi}^{\pi} Q_k(t)\, dt = 1.$$

(*c*) *If* $\eta_k(\delta) = \sup\{Q_k(t)\colon \delta \le |t| \le \pi\}$, *then*

$$\lim_{k\to\infty} \eta_k(\delta) = 0$$

for every $\delta > 0$.

Another way of stating (*c*) is to say: for every $\delta > 0$, $Q_k(t) \to 0$ uniformly on $[-\pi, -\delta] \cup [\delta, \pi]$.

To each $f \in C(T)$ we associate the functions P_k defined by

$$P_k(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t - s)Q_k(s)\, ds \qquad (k = 1, 2, 3, \ldots). \tag{1}$$

If we replace s by $-s$ (using Theorem 2.20(*e*)) and then by $s - t$, the periodicity of f and Q_k shows that the value of the integral is not affected. Hence

$$P_k(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)Q_k(t-s)\,ds \qquad (k = 1, 2, 3, \ldots). \tag{2}$$

Since each Q_k is a trigonometric polynomial, Q_k is of the form

$$Q_k(t) = \sum_{n=-N_k}^{N_k} a_{n,k}e^{int}, \tag{3}$$

and if we replace t by $t - s$ in (3) and substitute the result in (2), we see that each P_k is a trigonometric polynomial.

Let $\epsilon > 0$ be given. Since f is uniformly continuous on T, there exists a $\delta > 0$ such that $|f(t) - f(s)| < \epsilon$ whenever $|t - s| < \delta$. By (*b*), we have

$$P_k(t) - f(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\{f(t-s) - f(t)\}Q_k(s)\,ds,$$

and (*a*) implies, for all t, that

$$|P_k(t) - f(t)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t-s) - f(t)|\,Q_k(s)\,ds = A_1 + A_2,$$

where A_1 is the integral over $[-\delta, \delta]$ and A_2 is the integral over $[-\pi, -\delta] \cup [\delta, \pi]$. In A_1, the integrand is less than $\epsilon Q_k(s)$, so $A_1 < \epsilon$, by (*b*). In A_2, we have $Q_k(s) \le \eta_k(\delta)$, hence

$$A_2 \le 2\|f\|_\infty \cdot \eta_k(\delta) < \epsilon \tag{4}$$

for sufficiently large k, by (*c*). Since these estimates are independent of t, we have proved that

$$\lim_{k\to\infty}\|f - P_k\|_\infty = 0. \tag{5}$$

It remains to construct the Q_k. This can be done in many ways. Here is a simple one. Put

$$Q_k(t) = c_k\left\{\frac{1+\cos t}{2}\right\}^k, \tag{6}$$

where c_k is chosen so that (*b*) holds. Since (*a*) is clear, we only need to show (*c*). Since Q_k is even, (*b*) shows that

$$1 = \frac{c_k}{\pi}\int_0^{\pi}\left\{\frac{1+\cos t}{2}\right\}^k dt > \frac{c_k}{\pi}\int_0^{\pi}\left\{\frac{1+\cos t}{2}\right\}^k \sin t\,dt = \frac{2c_k}{\pi(k+1)}.$$

Since Q_k is decreasing on $[0, \pi]$, it follows that

$$Q_k(t) \le Q_k(\delta) \le \frac{\pi(k+1)}{2}\left(\frac{1+\cos\delta}{2}\right)^k \qquad (0 < \delta \le |t| \le \pi). \tag{7}$$

This implies (*c*), since $1 + \cos\delta < 2$ if $0 < \delta \le \pi$.

We have proved the following important result:

4.25 Theorem *If $f \in C(T)$ and $\epsilon > 0$, there is a trigonometric polynomial P such that*

$$|f(t) - P(t)| < \epsilon$$

for every real t.

A more precise result was proved by Fejér (1904): *The arithmetic means of the partial sums of the Fourier series of any $f \in C(T)$ converge uniformly to f.* For a proof (quite similar to the above) see Theorem 3.1 of [45], or p. 89 of [36], vol. I.

4.26 Fourier Series For any $f \in L^1(T)$, we define the *Fourier coefficients* of f by the formula

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt \qquad (n \in Z), \tag{1}$$

where, we recall, Z is the set of all integers. We thus associate with each $f \in L^1(T)$ a function $\hat{f}$ on Z. The *Fourier series* of f is

$$\sum_{-\infty}^{\infty}\hat{f}(n)e^{int}, \tag{2}$$

and its *partial sums* are

$$s_N(t) = \sum_{-N}^{N}\hat{f}(n)e^{int} \qquad (N = 0, 1, 2, \ldots). \tag{3}$$

Since $L^2(T) \subset L^1(T)$, (1) can be applied to every $f \in L^2(T)$. Comparing the definitions made in Secs. 4.23 and 4.13, we can now restate Theorems 4.17 and 4.18 in concrete terms:

The *Riesz-Fischer theorem* asserts that if $\{c_n\}$ is a sequence of complex numbers such that

$$\sum_{n=-\infty}^{\infty}|c_n|^2 < \infty, \tag{4}$$

then there exists an $f \in L^2(T)$ such that

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt \qquad (n \in Z). \tag{5}$$

The *Parseval theorem* asserts that

$$\sum_{n=-\infty}^{\infty}\hat{f}(n)\overline{\hat{g}(n)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt \tag{6}$$

whenever $f \in L^2(T)$ and $g \in L^2(T)$; the series on the left of (6) converges absolutely; and if s_N is as in (3), then

$$\lim_{N\to\infty} \|f - s_N\|_2 = 0, \tag{7}$$

since a special case of (6) yields

$$\|f - s_N\|_2^2 = \sum_{|n|>N} |\hat{f}(n)|^2. \tag{8}$$

Note that (7) says that every $f \in L^2(T)$ is the L^2-limit of the partial sums of its Fourier series; i.e., the Fourier series of f converges to f, in the L^2-sense. Pointwise convergence presents a more delicate problem, as we shall see in Chap. 5.

The Riesz-Fischer theorem and the Parseval theorem may be summarized by saying that the mapping $f \leftrightarrow \hat{f}$ is a Hilbert space isomorphism of $L^2(T)$ onto $\ell^2(Z)$.

The theory of Fourier series in other function spaces, for instance in $L^1(T)$, is much more difficult than in $L^2(T)$, and we shall touch only a few aspects of it.

Observe that the crucial ingredient in the proof of the Riesz-Fischer theorem is the fact that L^2 is complete. This is so well recognized that the name "Riesz-Fischer theorem" is sometimes given to the theorem which asserts the completeness of L^2, or even of any L^p.

Exercises

In this set of exercises, H always denotes a Hilbert space.

1 If M is a closed subspace of H, prove that $M = (M^\perp)^\perp$. Is there a similar true statement for subspaces M which are not necessarily closed?

2 Let $\{x_n: n = 1, 2, 3, \ldots\}$ be a linearly independent set of vectors in H. Show that the following construction yields an orthonormal set $\{u_n\}$ such that $\{x_1, \ldots, x_N\}$ and $\{u_1, \ldots, u_N\}$ have the same span for all N.

Put $u_1 = x_1/\|x_1\|$. Having $u_1, \ldots, u_{n-1}$ define

$$v_n = x_n - \sum_{i=1}^{n-1} (x_n, u_i)u_i, \qquad u_n = v_n/\|v_n\|.$$

Note that this leads to a proof of the existence of a maximal orthonormal set in separable Hilbert spaces which makes no appeal to the Hausdorff maximality principle. (A space is *separable* if it contains a countable dense subset.)

3 Show that $L^p(T)$ is separable if $1 \le p < \infty$, but that $L^\infty(T)$ is not separable.

4 Show that H is separable if and only if H contains a maximal orthonormal system which is at most countable.

5 If $M = \{x: Lx = 0\}$, where L is a continuous linear functional on H, prove that $M^\perp$ is a vector space of dimension 1 (unless $M = H$).

6 Let $\{u_n\}$ $(n = 1, 2, 3, \ldots)$ be an orthonormal set in H. Show that this gives an example of a closed and bounded set which is not compact. Let Q be the set of all $x \in H$ of the form

$$x = \sum_1^\infty c_n u_n \qquad \left(\text{where } |c_n| \le \frac{1}{n}\right).$$

Prove that Q is compact. (Q is called the Hilbert cube.)

More generally, let $\{\delta_n\}$ be a sequence of positive numbers, and let S be the set of all $x \in H$ of the form

$$x = \sum_1^\infty c_n u_n \qquad (\text{where } |c_n| \le \delta_n).$$

Prove that S is compact if and only if $\sum_1^\infty \delta_n^2 < \infty$.

Prove that H is not locally compact.

7 Suppose $\{a_n\}$ is a sequence of positive numbers such that $\sum a_n b_n < \infty$ whenever $b_n \ge 0$ and $\sum b_n^2 < \infty$. Prove that $\sum a_n^2 < \infty$.

Suggestion: If $\sum a_n^2 = \infty$ then there are disjoint sets E_k $(k = 1, 2, 3, \ldots)$ so that

$$\sum_{n \in E_k} a_n^2 > 1.$$

Define b_n so that $b_n = c_k a_n$ for $n \in E_k$. For suitably chosen c_k, $\sum a_n b_n = \infty$ although $\sum b_n^2 < \infty$.

8 If H_1 and H_2 are two Hilbert spaces, prove that one of them is isomorphic to a subspace of the other. (Note that every closed subspace of a Hilbert space is a Hilbert space.)

9 If $A \subset [0, 2\pi]$ and A is measurable, prove that

$$\lim_{n\to\infty} \int_A \cos nx\, dx = \lim_{n\to\infty} \int_A \sin nx\, dx = 0.$$

10 Let $n_1 < n_2 < n_3 < \cdots$ be positive integers, and let E be the set of all $x \in [0, 2\pi]$ at which $\{\sin n_k x\}$ converges. Prove that $m(E) = 0$. *Hint*: $2 \sin^2 \alpha = 1 - \cos 2\alpha$, so $\sin n_k x \to \pm 1/\sqrt{2}$ a.e. on E, by Exercise 9.

11 Find a nonempty closed set E in $L^2(T)$ that contains no element of smallest norm.

12 The constants c_k in Sec. 4.24 were shown to be such that $k^{-1}c_k$ is bounded. Estimate the relevant integral more precisely and show that

$$0 < \lim_{k\to\infty} k^{-1/2} c_k < \infty.$$

13 Suppose f is a continuous function on R^1, with period 1. Prove that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(n\alpha) = \int_0^1 f(t)\, dt$$

for every irrational real number α. *Hint*: Do it first for

$$f(t) = \exp(2\pi i k t), \qquad k = 0, \pm 1, \pm 2, \ldots.$$

14 Compute

$$\min_{a,b,c} \int_{-1}^{1} |x^3 - a - bx - cx^2|^2\, dx$$

and find

$$\max \int_{-1}^{1} x^3 g(x)\, dx,$$

where g is subject to the restrictions

$$\int_{-1}^{1} g(x)\, dx = \int_{-1}^{1} xg(x)\, dx = \int_{-1}^{1} x^2 g(x)\, dx = 0; \qquad \int_{-1}^{1} |g(x)|^2\, dx = 1.$$

15 Compute

$$\min_{a,b,c} \int_0^\infty |x^3 - a - bx - cx^2|^2 e^{-x}\, dx.$$

State and solve the corresponding maximum problem, as in Exercise 14.

16 If $x_0 \in H$ and M is a closed linear subspace of H, prove that

$$\min \{\|x - x_0\| : x \in M\} = \max \{|(x_0, y)| : y \in M^\perp, \|y\| = 1\}.$$

17 Show that there is a continuous one-to-one mapping γ of $[0, 1]$ into H such that $\gamma(b) - \gamma(a)$ is orthogonal to $\gamma(d) - \gamma(c)$ whenever $0 \le a \le b \le c \le d \le 1$. ($\gamma$ may be called a "curve with orthogonal increments.") *Hint*: Take $H = L^2$, and consider characteristic functions of certain subsets of $[0, 1]$.

18 Define $u_s(t) = e^{ist}$ for all $s \in R^1$, $t \in R^1$. Let X be the complex vector space consisting of all finite linear combinations of these functions u_s. If $f \in X$ and $g \in X$, show that

$$(f, g) = \lim_{A\to\infty} \frac{1}{2A} \int_{-A}^{A} f(t)\overline{g(t)}\, dt$$

exists. Show that this inner product makes X into a unitary space whose completion is a nonseparable Hilbert space H. Show also that $\{u_s : s \in R^1\}$ is a maximal orthonormal set in H.

19 Fix a positive integer N, put $\omega = e^{2\pi i/N}$, prove the orthogonality relations

$$\frac{1}{N}\sum_{n=1}^{N} \omega^{nk} = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } 1 \le k \le N-1 \end{cases}$$

and use them to derive the identities

$$(x, y) = \frac{1}{N}\sum_{n=1}^{N} \|x + \omega^n y\|^2 \omega^n$$

that hold in every inner product space if $N \ge 3$. Show also that

$$(x, y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \|x + e^{i\theta} y\|^2 e^{i\theta}\, d\theta.$$

CHAPTER

FIVE

EXAMPLES OF BANACH SPACE TECHNIQUES

Banach Spaces

5.1 In the preceding chapter we saw how certain analytic facts about trigonometric series can be made to emerge from essentially goemetric considerations about general Hilbert spaces, involving the notions of convexity, subspaces, orthogonality, and completeness. There are many problems in analysis that can be attacked with greater ease when they are placed within a suitably chosen abstract framework. The theory of Hilbert spaces is not always suitable since orthogonality is something rather special. The class of all Banach spaces affords greater variety. In this chapter we shall develop some of the basic properties of Banach spaces and illustrate them by applications to concrete problems.

5.2 Definition A complex vector space X is said to be a *normed linear space* if to each $x \in X$ there is associated a nonnegative real number $\|x\|$, called the *norm* of x, such that

(*a*) $\|x + y\| \le \|x\| + \|y\|$ for all x and $y \in X$,
(*b*) $\|\alpha x\| = |\alpha|\, \|x\|$ if $x \in X$ and α is a scalar,
(*c*) $\|x\| = 0$ implies $x = 0$.

By (*a*), the triangle inequality

$$\|x - y\| \le \|x - z\| + \|z - y\| \qquad (x, y, z \in X)$$

holds. Combined with (*b*) (take $\alpha = 0$, $\alpha = -1$) and (*c*) this shows that every normed linear space may be regarded as a metric space, the distance between x and y being $\|x - y\|$.

A *Banach space* is a normed linear space which is *complete* in the metric defined by its norm.

For instance, every Hilbert space is a Banach space, so is every $L^p(\mu)$ normed by $\|f\|_p$ (provided we identify functions which are equal a.e.) if $1 \le p \le \infty$, and so is $C_0(X)$ with the supremum norm. The simplest Banach space is of course the complex field itself, normed by $\|x\| = |x|$.

One can equally well discuss *real* Banach spaces; the definition is exactly the same, except that all scalars are assumed to be real.

5.3 Definition Consider a linear transformation Λ from a normed linear space X into a normed linear space Y, and define its *norm* by

$$\|\Lambda\| = \sup\{\|\Lambda x\| : x \in X, \|x\| \le 1\}. \tag{1}$$

If $\|\Lambda\| < \infty$, then Λ is called a *bounded linear transformation.*

In (1), $\|x\|$ is the norm of x in X, $\|\Lambda x\|$ is the norm of Λx in Y; it will frequently happen that several norms occur together, and the context will make it clear which is which.

Observe that we could restrict ourselves to *unit vectors* x in (1), i.e., to x with $\|x\| = 1$, without changing the supremum, since

$$\|\Lambda(\alpha x)\| = \|\alpha \Lambda x\| = |\alpha| \, \|\Lambda x\|. \tag{2}$$

Observe also that $\|\Lambda\|$ is the smallest number such that the inequality

$$\|\Lambda x\| \le \|\Lambda\| \, \|x\| \tag{3}$$

holds for *every* $x \in X$.

The following geometric picture is helpful: Λ maps the *closed unit ball* in X, i.e., the set

$$\{x \in X : \|x\| \le 1\}, \tag{4}$$

into the closed ball in Y with center at 0 and radius $\|\Lambda\|$.

An important special case is obtained by taking the complex field for Y; in that case we talk about *bounded linear functionals.*

5.4 Theorem *For a linear transformation Λ of a normed linear space X into a normed linear space Y, each of the following three conditions implies the other two:*

(*a*) *Λ is bounded.*
(*b*) *Λ is continuous.*
(*c*) *Λ is continuous at one point of X.*

PROOF Since $\|\Lambda(x_1 - x_2)\| \le \|\Lambda\| \, \|x_1 - x_2\|$, it is clear that (*a*) implies (*b*), and (*b*) implies (*c*) trivially. Suppose Λ is continuous at x_0. To each $\epsilon > 0$ one can then find a $\delta > 0$ so that $\|x - x_0\| < \delta$ implies $\|\Lambda x - \Lambda x_0\| < \epsilon$. In other words, $\|x\| < \delta$ implies

$$\|\Lambda(x_0 + x) - \Lambda x_0\| < \epsilon.$$

But then the linearity of Λ shows that $\|\Lambda x\| < \epsilon$. Hence $\|\Lambda\| \le \epsilon/\delta$, and (*c*) implies (*a*). ////

Consequences of Baire's Theorem

5.5 The manner in which the completeness of Banach spaces is frequently exploited depends on the following theorem about complete metric spaces, which also has many applications in other parts of mathematics. It implies two of the three most important theorems which make Banach spaces useful tools in analysis, the *Banach-Steinhaus theorem* and the *open mapping theorem.* The third is the *Hahn-Banach extension theorem,* in which completeness plays no role.

5.6 Baire's Theorem *If X is a complete metric space, the intersection of every countable collection of dense open subsets of X is dense in X.*

In particular (except in the trivial case $X = \varnothing$), the intersection is not empty. This is often the principal significance of the theorem.

PROOF Suppose $V_1, V_2, V_3, \dots$ are dense and open in X. Let W be any open set in X. We have to show that $\bigcap V_n$ has a point in W if $W \ne \varnothing$.

Let ρ be the metric of X; let us write

$$S(x, r) = \{y \in X : \rho(x, y) < r\} \tag{1}$$

and let $\bar{S}(x, r)$ be the closure of $S(x, r)$. [*Note*: There exist situations in which $\bar{S}(x, r)$ does *not* contain all y with $\rho(x, y) \le r$!]

Since V_1 is dense, $W \cap V_1$ is a nonempty open set, and we can therefore find x_1 and r_1 such that

$$\bar{S}(x_1, r_1) \subset W \cap V_1 \quad \text{and} \quad 0 < r_1 < 1. \tag{2}$$

If $n \ge 2$ and x_{n-1} and r_{n-1} are chosen, the denseness of V_n shows that $V_n \cap S(x_{n-1}, r_{n-1})$ is not empty, and we can therefore find x_n and r_n such that

$$\bar{S}(x_n, r_n) \subset V_n \cap S(x_{n-1}, r_{n-1}) \quad \text{and} \quad 0 < r_n < \frac{1}{n}. \tag{3}$$

By induction, this process produces a sequence $\{x_n\}$ in X. If $i > n$ and $j > n$, the construction shows that x_i and x_j both lie in $S(x_n, r_n)$, so that $\rho(x_i, x_j) < 2r_n < 2/n$, and hence $\{x_n\}$ is a Cauchy sequence. Since X is complete, there is a point $x \in X$ such that $x = \lim_{n\to\infty} x_n$.

Since x_i lies in the closed set $\bar{S}(x_n, r_n)$ if $i > n$, it follows that x lies in each $\bar{S}(x_n, r_n)$, and (3) shows that x lies in each V_n. By (2), $x \in W$. This completes the proof. ////

Corollary *In a complete metric space, the intersection of any countable collection of dense G_δ's is again a dense G_δ.*

This follows from the theorem, since every G_δ is the intersection of a countable collection of open sets, and since the union of countably many countable sets is countable.

5.7 Baire's theorem is sometimes called the *category theorem*, for the following reason.

Call a set $E \subset X$ *nowhere dense* if its closure $\bar{E}$ contains no nonempty open subset of X. Any countable union of nowhere dense sets is called a set of the *first category*; all other subsets of X are of the *second category* (Baire's terminology). Theorem 5.6 is equivalent to the statement that *no complete metric space is of the first category*. To see this, just take complements in the statement of Theorem 5.6.

5.8 The Banach-Steinhaus Theorem *Suppose X is a Banach space, Y is a normed linear space, and $\{\Lambda_\alpha\}$ is a collection of bounded linear transformations of X into Y, where α ranges over some index set A. Then either there exists an $M < \infty$ such that*

$$\|\Lambda_\alpha\| \le M \tag{1}$$

for every $\alpha \in A$, or

$$\sup_{\alpha \in A} \|\Lambda_\alpha x\| = \infty \tag{2}$$

for all x belonging to some dense G_δ in X.

In geometric terminology, the alternatives are as follows: *Either* there is a ball B in Y (with radius M and center at 0) such that every Λ_α maps the unit ball of X into B, *or* there exist $x \in X$ (in fact, a whole dense G_δ of them) such that no ball in Y contains $\Lambda_\alpha x$ for all α.

The theorem is sometimes referred to as the *uniform boundedness principle*.

PROOF Put

$$\varphi(x) = \sup_{\alpha \in A} \|\Lambda_\alpha x\| \qquad (x \in X) \tag{3}$$

and let

$$V_n = \{x : \varphi(x) > n\} \qquad (n = 1, 2, 3, \ldots). \tag{4}$$

Since each Λ_α is continuous and since the norm of Y is a continuous function on Y (an immediate consequence of the triangle inequality, as in the proof of Theorem 4.6), each function $x \to \|\Lambda_\alpha x\|$ is continuous on X. Hence φ is lower semicontinuous, and each V_n is open.

If one of these sets, say V_N, fails to be dense in X, then there exist an $x_0 \in X$ and an $r > 0$ such that $\|x\| \le r$ implies $x_0 + x \notin V_N$; this means that $\varphi(x_0 + x) \le N$, or

$$\|\Lambda_\alpha(x_0 + x)\| \le N \tag{5}$$

for all $\alpha \in A$ and all x with $\|x\| \le r$. Since $x = (x_0 + x) - x_0$, we then have

$$\|\Lambda_\alpha x\| \le \|\Lambda_\alpha(x_0 + x)\| + \|\Lambda_\alpha x_0\| \le 2N, \tag{6}$$

and it follows that (1) holds with $M = 2N/r$.

The other possibility is that every V_n is dense in X. In that case, $\bigcap V_n$ is a dense G_δ in X, by Baire's theorem; and since $\varphi(x) = \infty$ for every $x \in \bigcap V_n$, the proof is complete. ////

5.9 The Open Mapping Theorem *Let U and V be the open unit balls of the Banach spaces X and Y. To every bounded linear transformation Λ of X onto Y there corresponds a $\delta > 0$ so that*

$$\Lambda(U) \supset \delta V. \tag{1}$$

Note the word "onto" in the hypothesis. The symbol δV stands for the set $y: y \in V\}$, i.e., the set of all $y \in Y$ with $\|y\| < \delta$.

It follows from (1) and the linearity of Λ that the image of every open ball in , with center at x_0, say, contains an open ball in Y with center at Λx_0. Hence e image of every open set is open. This explains the name of the theorem.

Here is another way of stating (1): *To every y with $\|y\| < \delta$ there corresponds x with $\|x\| < 1$ so that $\Lambda x = y$.*

PROOF Given $y \in Y$, there exists an $x \in X$ such that $\Lambda x = y$; if $\|x\| < k$, it follows that $y \in \Lambda(kU)$. Hence Y is the union of the sets $\Lambda(kU)$, for $k = 1, 2, 3, \ldots$. Since Y is complete, the Baire theorem implies that there is a nonempty open set W in the closure of some $\Lambda(kU)$. This means that every point of W is the limit of a sequence $\{\Lambda x_i\}$, where $x_i \in kU$; from now on, k and W are fixed.

Choose $y_0 \in W$, and choose $\eta > 0$ so that $y_0 + y \in W$ if $\|y\| < \eta$. For any such y there are sequences $\{x_i'\}$, $\{x_i''\}$ in kU such that

$$\Lambda x_i' \to y_0, \qquad \Lambda x_i'' \to y_0 + y \qquad (i \to \infty). \tag{2}$$

Setting $x_i = x_i'' - x_i'$, we have $\|x_i\| < 2k$ and $\Lambda x_i \to y$. Since this holds for every y with $\|y\| < \eta$, the linearity of Λ shows that the following is true, if $\delta = \eta/2k$:

To each $y \in Y$ and to each $\epsilon > 0$ there corresponds an $x \in X$ such that

$$\|x\| \le \delta^{-1}\|y\| \quad \textit{and} \quad \|y - \Lambda x\| < \epsilon. \tag{3}$$

This is almost the desired conclusion, as stated just before the start of the proof, except that there we had $\epsilon = 0$.

Fix $y \in \delta V$, and fix $\epsilon > 0$. By (3) there exists an x_1 with $\|x_1\| < 1$ and

$$\|y - \Lambda x_1\| < \tfrac{1}{2}\delta\epsilon. \tag{4}$$

Suppose $x_1, \ldots, x_n$ are chosen so that

$$\|y - \Lambda x_1 - \cdots - \Lambda x_n\| < 2^{-n}\delta\epsilon. \tag{5}$$

Use (3), with y replaced by the vector on the left side of (5), to obtain an x_{n+1} so that (5) holds with $n + 1$ in place of n, and

$$\|x_{n+1}\| < 2^{-n}\epsilon \qquad (n = 1, 2, 3, \ldots). \tag{6}$$

If we set $s_n = x_1 + \cdots + x_n$, (6) shows that $\{s_n\}$ is a Cauchy sequence in X. Since X is complete, there exists an $x \in X$ so that $s_n \to x$. The inequality $\|x_1\| < 1$, together with (6), shows that $\|x\| < 1 + \epsilon$. Since Λ is continuous, $\Lambda s_n \to \Lambda x$. By (5), $\Lambda s_n \to y$. Hence $\Lambda x = y$.

We have now proved that

$$\Lambda((1 + \epsilon)U) \supset \delta V, \tag{7}$$

or

$$\Lambda(U) \supset (1 + \epsilon)^{-1}\delta V, \tag{8}$$

for every $\epsilon > 0$. The union of the sets on the right of (8), taken over all $\epsilon > 0$, is δV. This proves (1). ////

5.10 Theorem *If X and Y are Banach spaces and if Λ is a bounded linear transformation of X onto Y which is also one-to-one, then there is a $\delta > 0$ such that*

$$\|\Lambda x\| \geq \delta\|x\| \qquad (x \in X). \tag{1}$$

In other words, Λ^{-1} is a bounded linear transformation of Y onto X.

PROOF If δ is chosen as in the statement of Theorem 5.9, the conclusion of that theorem, combined with the fact that Λ is now one-to-one, shows that $\|\Lambda x\| < \delta$ implies $\|x\| < 1$. Hence $\|x\| \geq 1$ implies $\|\Lambda x\| \geq \delta$, and (1) is proved.

The transformation Λ^{-1} is defined on Y by the requirement that $\Lambda^{-1}y = x$ if $y = \Lambda x$. A trivial verification shows that Λ^{-1} is linear, and (1) implies that $\|\Lambda^{-1}\| \leq 1/\delta$. ////

Fourier Series of Continuous Functions

5.11 A Convergence Problem *Is it true for every $f \in C(T)$ that the Fourier series of f converges to $f(x)$ at every point x?*

Let us recall that the nth partial sum of the Fourier series of f at the point x is given by

$$s_n(f; x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(x - t)\, dt \qquad (n = 0, 1, 2, \ldots), \tag{1}$$

where

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt}. \tag{2}$$

This follows directly from formulas 4.26(1) and 4.26(3).

The problem is to determine whether

$$\lim_{n\to\infty} s_n(f; x) = f(x) \tag{3}$$

for every $f \in C(T)$ and for every real x. We observed in Sec. 4.26 that the partial sums do converge to f in the L^2-norm, and Theorem 3.12 implies therefore that each $f \in L^2(T)$ [hence also each $f \in C(T)$] is the pointwise limit a.e. of some subsequence of the full sequence of the partial sums. But this does not answer the present question.

We shall see that the Banach-Steinhaus theorem answers the question *negatively*. Put

$$s^*(f; x) = \sup_n |s_n(f; x)|. \tag{4}$$

To begin with, take $x = 0$, and define

$$\Lambda_n f = s_n(f; 0) \qquad (f \in C(T);\ n = 1, 2, 3, \ldots). \tag{5}$$

We know that $C(T)$ is a Banach space, relative to the supremum norm $\|f\|_\infty$. It follows from (1) that each Λ_n is a bounded linear functional on $C(T)$, of norm

$$\|\Lambda_n\| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt = \|D_n\|_1. \tag{6}$$

We claim that

$$\|\Lambda_n\| \to \infty \qquad \text{as } n \to \infty. \tag{7}$$

This will be proved by showing that equality holds in (6) and that

$$\|D_n\|_1 \to \infty \qquad \text{as } n \to \infty. \tag{8}$$

Multiply (2) by $e^{it/2}$ and by $e^{-it/2}$ and subtract one of the resulting two equations from the other, to obtain

$$D_n(t) = \frac{\sin\,(n + \frac{1}{2})t}{\sin\,(t/2)}. \tag{9}$$

Since $|\sin x| \le |x|$ for all real x, (9) shows that

$$\|D_n\|_1 > \frac{2}{\pi}\int_0^\pi \left|\sin\left(n+\frac{1}{2}\right)t\right| \frac{dt}{t} = \frac{2}{\pi}\int_0^{(n+1/2)\pi} |\sin t| \frac{dt}{t}$$

$$> \frac{2}{\pi}\sum_{k=1}^{n} \frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi} |\sin t|\, dt = \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \to \infty,$$

which proves (8).

Next, fix n, and put $g(t) = 1$ if $D_n(t) \ge 0$, $g(t) = -1$ if $D_n(t) < 0$. There exist $f_j \in C(T)$ such that $-1 \le f_j \le 1$ and $f_j(t) \to g(t)$ for every t, as $j \to \infty$. By the dominated convergence theorem,

$$\lim_{j\to\infty} \Lambda_n(f_j) = \lim_{j\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} f_j(-t)D_n(t)\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(-t)D_n(t)\, dt = \|D_n\|_1.$$

Thus equality holds in (6), and we have proved (7).

Since (7) holds, the Banach-Steinhaus theorem asserts now that $s^*(f; 0) = \infty$ for every f in some dense G_δ-set in $C(T)$.

We chose $x = 0$ just for convenience. It is clear that the same result holds for every other x:

To each real number x there corresponds a set $E_x \subset C(T)$ which is a dense G_δ in $C(T)$, such that $s^*(f; x) = \infty$ for every $f \in E_x$.

In particular, the Fourier series of each $f \in E_x$ diverges at x, and we have a negative answer to our question. (Exercise 22 shows the answer is positive if mere continuity is replaced by a somewhat stronger smoothness assumption.)

It is interesting to observe that the above result can be strengthened by another application of Baire's theorem. Let us take countably many points x_i, and let E be the intersection of the corresponding sets

$$E_{x_i} \subset C(T).$$

By Baire's theorem, E is a dense G_δ in $C(T)$. Every $f \in E$ has

$$s^*(f; x_i) = \infty$$

at every point x_i.

For each f, $s^*(f; x)$ is a lower semicontinuous function of x, since (4) exhibits it as the supremum of a collection of continuous functions. Hence $\{x: s^*(f; x) = \infty\}$ is a G_δ in R^1, for each f. If the above points x_i are taken so that their union is dense in $(-\pi, \pi)$, we obtain the following result:

5.12 Theorem *There is a set $E \subset C(T)$ which is a dense G_δ in $C(T)$ and which has the following property: For each $f \in E$, the set*

$$Q_f = \{x: s^*(f; x) = \infty\}$$

is a dense G_δ in R^1.

This gains in interest if we realize that E, as well as each Q_f, is an *uncountable* set:

5.13 Theorem *In a complete metric space X which has no isolated points, no countable dense set is a G_δ.*

PROOF Let x_k be the points of a countable dense set E in X. Assume that E is a G_δ. Then $E = \bigcap V_n$, where each V_n is dense and open. Let

$$W_n = V_n - \bigcup_{k=1}^{n} \{x_k\}.$$

Then each W_n is still a dense open set, but $\bigcap W_n = \varnothing$, in contradiction to Baire's theorem. ////

Note: A slight change in the proof of Baire's theorem shows actually that every dense G_δ contains a perfect set if X is as above.

Fourier Coefficients of L^1-functions

5.14 As in Sec. 4.26, we associate to every $f \in L^1(T)$ a function $\hat{f}$ on Z defined by

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\, dt \qquad (n \in Z). \tag{1}$$

It is easy to prove that $\hat{f}(n) \to 0$ as $|n| \to \infty$, for every $f \in L^1$. For we know that $C(T)$ is dense in $L^1(T)$ (Theorem 3.14) and that the trigonometric polynomials are dense in $C(T)$ (Theorem 4.25). If $\epsilon > 0$ and $f \in L^1(T)$, this says that there is a $g \in C(T)$ and a trigonometric polynomial P such that $\|f - g\|_1 < \epsilon$ and $\|g - P\|_\infty < \epsilon$. Since

$$\|g - P\|_1 \le \|g - P\|_\infty$$

if follows that $\|f - P\|_1 < 2\epsilon$; and if $|n|$ is large enough (depending on P), then

$$|\hat{f}(n)| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} \{f(t) - P(t)\}e^{-int}\, dt\right| \le \|f - P\|_1 < 2\epsilon. \tag{2}$$

Thus $\hat{f}(n) \to 0$ as $n \to \pm\infty$. This is known as the Riemann-Lebesgue lemma.

The question we wish to raise is whether the converse is true. That is to say, if $\{a_n\}$ is a sequence of complex numbers such that $a_n \to 0$ as $n \to \pm\infty$, does it follow that there is an $f \in L^1(T)$ such that $\hat{f}(n) = a_n$ for all $n \in Z$? In other words, is something like the Riesz-Fischer theorem true in this situation?

This can easily be answered (negatively) with the aid of the open mapping theorem.

Let c_0 be the space of all complex functions φ on Z such that $\varphi(n) \to 0$ as $n \to \pm\infty$, with the supremum norm

$$\|\varphi\|_\infty = \sup \{|\varphi(n)|: n \in Z\}. \tag{3}$$

Then c_0 is easily seen to be a Banach space. In fact, if we declare every subset of Z to be open, then Z is a locally compact Hausdorff space, and c_0 is nothing but $C_0(Z)$.

The following theorem contains the answer to our question:

5.15 Theorem *The mapping $f \to \hat{f}$ is a one-to-one bounded linear transformation of $L^1(T)$ into (but not onto) c_0.*

PROOF Define Λ by $\Lambda f = \hat{f}$. It is clear that Λ is linear. We have just proved that Λ maps $L^1(T)$ into c_0, and formula 5.14(1) shows that $|\hat{f}(n)| \le \|f\|_1$, so that $\|\Lambda\| \le 1$. (Actually, $\|\Lambda\| = 1$; to see this, take $f = 1$.) Let us now prove that Λ is one-to-one. Suppose $f \in L^1(T)$ and $\hat{f}(n) = 0$ for every $n \in Z$. Then

$$\int_{-\pi}^{\pi} f(t)g(t)\,dt = 0 \tag{1}$$

if g is any trigonometric polynomial. By Theorem 4.25 and the dominated convergence theorem, (1) holds for every $g \in C(T)$. Apply the dominated convergence theorem once more, in conjunction with the Corollary to Lusin's theorem, to conclude that (1) holds if g is the characteristic function of any measurable set in T. Now Theorem 1.39(b) shows that $f = 0$ a.e.

If the range of Λ were all of c_0, Theorem 5.10 would imply the existence of a $\delta > 0$ such that

$$\|\hat{f}\|_\infty \ge \delta \|f\|_1 \tag{2}$$

for every $f \in L^1(T)$. But if $D_n(t)$ is defined as in Sec. 5.11, then $D_n \in L^1(T)$, $\|\hat{D}_n\|_\infty = 1$ for $n = 1, 2, 3, \ldots$, and $\|D_n\|_1 \to \infty$ as $n \to \infty$. Hence there is no $\delta > 0$ such that the inequalities

$$\|\hat{D}_n\|_\infty \ge \delta \|D_n\|_1 \tag{3}$$

hold for every n.

This completes the proof. ////

The Hahn-Banach Theorem

5.16 Theorem *If M is a subspace of a normed linear space X and if f is a bounded linear functional on M, then f can be extended to a bounded linear functional F on X so that $\|F\| = \|f\|$.*

Note that M need not be closed.

Before we turn to the proof, some comments seem called for. First, to say (in the most general situation) that a function F is an *extension* of f means that the domain of F includes that of f and that $F(x) = f(x)$ for all x in the domain of f. Second, the norms $\|F\|$ and $\|f\|$ are computed relative to the domains of F and f; explicitly,

$$\|f\| = \sup \{|f(x)|: x \in M, \|x\| \le 1\}, \qquad \|F\| = \sup \{|F(x)|: x \in X, \|x\| \le 1\},$$

The third comment concerns the field of scalars. So far everything has been stated for complex scalars, but the complex field could have been replaced by the real field without any changes in statements or proofs. The Hahn-Banach theorem is also true in both cases; nevertheless, it appears to be essentially a "real" theorem. The fact that the complex case was not yet proved when Banach wrote his classical book "Opérations linéaires" may be the main reason that real scalars are the only ones considered in his work.

It will be helpful to introduce some temporary terminology. Recall that V is a complex (real) vector space if $x + y \in V$ for x and $y \in V$, and if $\alpha x \in V$ for all complex (real) numbers α. It follows trivially that *every complex vector space is also a real vector space.* A complex function φ on a complex vector space V is a *complex-linear functional* if

$$\varphi(x + y) = \varphi(x) + \varphi(y) \quad \text{and} \quad \varphi(\alpha x) = \alpha\varphi(x) \tag{1}$$

for all x and $y \in V$ and all *complex* α. A real-valued function φ on a complex (real) vector space V is a *real-linear functional* if (1) holds for all *real* α.

If u is the real part of a complex-linear functional f, i.e., if $u(x)$ is the real part of the complex number $f(x)$ for all $x \in V$, it is easily seen that u is a real-linear functional. The following relations hold between f and u:

5.17 Proposition *Let V be a complex vector space.*

(a) *If u is the real part of a complex-linear functional f on V, then*

$$f(x) = u(x) - iu(ix) \qquad (x \in V). \tag{1}$$

(b) *If u is a real-linear functional on V and if f is defined by (1), then f is a complex-linear functional on V.*

(c) *If V is a normed linear space and f and u are related as in (1), then $\|f\| = \|u\|$.*

PROOF If α and β are real numbers and $z = \alpha + i\beta$, the real part of iz is $-\beta$. This gives the identity

$$z = \text{Re } z - i \text{ Re } (iz) \tag{2}$$

for all complex numbers z. Since

$$\text{Re } (if(x)) = \text{Re } f(ix) = u(ix), \tag{3}$$

(1) follows from (2) with $z = f(x)$.

Under the hypotheses (*b*), it is clear that $f(x+y) = f(x) + f(y)$ and that $f(\alpha x) = \alpha f(x)$ for all real α. But we also have

$$f(ix) = u(ix) - iu(-x) = u(ix) + iu(x) = if(x), \tag{4}$$

which proves that f is complex-linear.

Since $|u(x)| \le |f(x)|$, we have $\|u\| \le \|f\|$. On the other hand, to every $x \in V$ there corresponds a complex number α, $|\alpha| = 1$, so that $\alpha f(x) = |f(x)|$. Then

$$|f(x)| = f(\alpha x) = u(\alpha x) \le \|u\| \cdot \|\alpha x\| = \|u\| \cdot \|x\|, \tag{5}$$

which proves that $\|f\| \le \|u\|$. ////

5.18 Proof of Theorem 5.16 We first assume that X is a real normed linear space and, consequently, that f is a real-linear bounded functional on M. If $\|f\| = 0$, the desired extension is $F = 0$. Omitting this case, there is no loss of generality in assuming that $\|f\| = 1$.

Choose $x_0 \in X$, $x_0 \notin M$, and let M_1 be the vector space spanned by M and x_0. Then M_1 consists of all vectors of the form $x + \lambda x_0$, where $x \in M$ and λ is a real scalar. If we define $f_1(x + \lambda x_0) = f(x) + \lambda\alpha$, where α is any fixed real number, it is trivial to verify that an extension of f to a linear functional on M_1 is obtained. The problem is to choose α so that the extended functional still has norm 1. This will be the case provided that

$$|f(x) + \lambda\alpha| \le \|x + \lambda x_0\| \qquad (x \in M,\ \lambda \text{ real}). \tag{1}$$

Replace x by $-\lambda x$ and divide both sides of (1) by $|\lambda|$. The requirement is then that

$$|f(x) - \alpha| \le \|x - x_0\| \qquad (x \in M), \tag{2}$$

i.e., that $A_x \le \alpha \le B_x$ for all $x \in M$, where

$$A_x = f(x) - \|x - x_0\| \quad \text{and} \quad B_x = f(x) + \|x - x_0\|. \tag{3}$$

There exists such an α if and only if all the intervals $[A_x, B_x]$ have a common point, i.e., if and only if

$$A_x \le B_y \tag{4}$$

for all x and $y \in M$. But

$$f(x) - f(y) = f(x - y) \le \|x - y\| \le \|x - x_0\| + \|y - x_0\|, \tag{5}$$

and so (4) follows from (3).

We have now proved that there exists a norm-preserving extension f_1 of f on M_1.

Let $\mathscr{P}$ be the collection of all ordered pairs (M', f'), where M' is a subspace of X which contains M and where f' is a real-linear extension of f to M', with $\|f'\| = 1$. Partially order $\mathscr{P}$ by declaring $(M', f') \le (M'', f'')$ to mean that $M' \subset M''$ and $f''(x) = f'(x)$ for all $x \in M'$. The axioms of a partial order are clearly satisfied, $\mathscr{P}$ is not empty since it contains (M, f), and so the Hausdorff maximality theorem asserts the existence of a maximal totally ordered subcollection Ω of $\mathscr{P}$.

Let Φ be the collection of all M' such that $(M', f') \in \Omega$. Then Φ is totally ordered, by set inclusion, and *therefore* the union $\tilde{M}$ of all members of Φ is a subspace of X. (Note that in general the union of two subspaces is not a subspace. An example is two planes through the origin in R^3.) If $x \in \tilde{M}$, then $x \in M'$ for some $M' \in \Phi$; define $F(x) = f'(x)$, where f' is the function which occurs in the pair $(M', f') \in \Omega$. Our definition of the partial order in Ω shows that it is immaterial which $M' \in \Phi$ we choose to define $F(x)$, as long as M' contains x.

It is now easy to check that F is a linear functional on $\tilde{M}$, with $\|F\| = 1$. If $\tilde{M}$ were a proper subspace X, the first part of the proof would give us a further extension of F, and this would contradict the maximality of Ω. Thus $\tilde{M} = X$, and the proof is complete for the case of real scalars.

If now f is a complex-linear functional on the subspace M of the complex normed linear space X, let u be the real part of f, use the real Hahn-Banach theorem to extend u to a real-linear functional U on X, with $\|U\| = \|u\|$, and define

$$F(x) = U(x) - iU(ix) \qquad (x \in X). \tag{6}$$

By Proposition 5.17, F is a complex-linear extension of f, and

$$\|F\| = \|U\| = \|u\| = \|f\|.$$

This completes the proof. ////

Let us mention two important consequences of the Hahn-Banach theorem:

5.19 Theorem *Let M be a linear subspace of a normed linear space X, and let $x_0 \in X$. Then x_0 is in the closure $\bar{M}$ of M if and only if there is no bounded linear functional f on X such that $f(x) = 0$ for all $x \in M$ but $f(x_0) \ne 0$.*

Proof If $x_0 \in \bar{M}$, f is a bounded linear functional on X, and $f(x) = 0$ for all $x \in M$, the continuity of f shows that we also have $f(x_0) = 0$.

Conversely, suppose $x_0 \notin \bar{M}$. Then there exists a $\delta > 0$ such that $\|x - x_0\| > \delta$ for all $x \in M$. Let M' be the subspace generated by M and x_0, and define $f(x + \lambda x_0) = \lambda$ if $x \in M$ and λ is a scalar. Since

$$\delta|\lambda| \le |\lambda|\,\|x_0 + \lambda^{-1}x\| = \|\lambda x_0 + x\|,$$

we see that f is a linear functional on M' whose norm is at most δ^{-1}. Also $f(x) = 0$ on M, $f(x_0) = 1$. The Hahn-Banach theorem allows us to extend this f from M' to X. ////

5.20 Theorem *If X is a normed linear space and if $x_0 \in X$, $x_0 \ne 0$, there is a bounded linear functional f on X, of norm 1, so that $f(x_0) = \|x_0\|$.*

PROOF Let $M = \{\lambda x_0\}$, and define $f(\lambda x_0) = \lambda \|x_0\|$. Then f is a linear functional of norm 1 on M, and the Hahn-Banach theorem can again be applied. ////

5.21 Remarks If X is a normed linear space, let X^* be the collection of all bounded linear functionals on X. If addition and scalar multiplication of linear functionals are defined in the obvious manner, it is easy to see that X^* is again a normed linear space. In fact, X^* is a Banach space; this follows from the fact that the field of scalars is a complete metric space. We leave the verification of these properties of X^* as an exercise.

One of the consequences of Theorem 5.20 is that X^* is not the trivial vector space (i.e., X^* consists of more than 0) if X is not trivial. In fact, X^* *separates points on* X. This means that if $x_1 \neq x_2$ in X there exists an $f \in X^*$ such that $f(x_1) \neq f(x_2)$. To prove this, merely take $x_0 = x_2 - x_1$ in Theorem 5.20.

Another consequence is that, for $x \in X$,

$$\|x\| = \sup \{|f(x)| : f \in X^*, \|f\| = 1\}.$$

Hence, for fixed $x \in X$, the mapping $f \to f(x)$ is a bounded linear functional on X^*, of norm $\|x\|$.

This interplay between X and X^* (the so-called "dual space" of X) forms the basis of a large portion of that part of mathematics which is known as *functional analysis*.

An Abstract Approach to the Poisson Integral

5.22 Successful applications of the Hahn-Banach theorem to concrete problems depend of course on a knowledge of the bounded linear functionals on the normed linear space under consideration. So far we have only determined the bounded linear functionals on a Hilbert space (where a much simpler proof of the Hahn-Banach theorem exists; see Exercise 6), and we know the positive linear functionals on $C_c(X)$.

We shall now describe a general situation in which the last-mentioned functionals occur naturally.

Let K be a compact Hausdorff space, let H be a compact subset of K, and let A be a subspace of $C(K)$ such that $1 \in A$ (1 denotes the function which assigns the number 1 to each $x \in K$) and such that

$$\|f\|_K = \|f\|_H \qquad (f \in A). \tag{1}$$

Here we used the notation

$$\|f\|_E = \sup \{|f(x)| : x \in E\}. \tag{2}$$

Because of the example discussed in Sec. 5.23, H is sometimes called a *boundary* of K, corresponding to the space A.

If $f \in A$ and $x \in K$, (1) says that

$$|f(x)| \leq \|f\|_H. \tag{3}$$

In particular, if $f(y) = 0$ for every $y \in H$, then $f(x) = 0$ for all $x \in K$. Hence if f_1 and $f_2 \in A$ and $f_1(y) = f_2(y)$ for every $y \in H$, then $f_1 = f_2$; to see this, put $f = f_1 - f_2$.

Let M be the set of all functions on H that are restrictions to H of members of A. It is clear that M is a subspace of $C(H)$. The preceding remark shows that each member of M has a unique extension to a member of A. Thus we have a natural one-to-one correspondence between M and A, which is also norm-preserving, by (1). Hence it will cause no confusion if we use the same letter to designate a member of A and its restriction to H.

Fix a point $x \in K$. The inequality (3) shows that the mapping $f \to f(x)$ is a bounded linear functional on M, of norm 1 [since equality holds in (3) if $f = 1$]. By the Hahn-Banach theorem there is a linear functional Λ on $C(H)$, of norm 1, such that

$$\Lambda f = f(x) \qquad (f \in M). \tag{4}$$

We claim that the properties

$$\Lambda 1 = 1, \qquad \|\Lambda\| = 1 \tag{5}$$

imply that Λ is a *positive* linear functional on $C(H)$.

To prove this, suppose $f \in C(H)$, $0 \leq f \leq 1$, put $g = 2f - 1$, and put $\Lambda g = \alpha + i\beta$, where α and β are real. Note that $-1 \leq g \leq 1$, so that $|g + ir|^2 \leq 1 + r^2$ for every real constant r. Hence (5) implies that

$$(\beta + r)^2 \leq |\alpha + i(\beta + r)|^2 = |\Lambda(g + ir)|^2 \leq 1 + r^2. \tag{6}$$

Thus $\beta^2 + 2r\beta \leq 1$ for every real r, which forces $\beta = 0$. Since $\|g\|_H \leq 1$, we have $|\alpha| \leq 1$; hence

$$\Lambda f = \tfrac{1}{2}\Lambda(1 + g) = \tfrac{1}{2}(1 + \alpha) \geq 0. \tag{7}$$

Now Theorem 2.14 can be applied. It shows that there is a regular positive Borel measure μ_x on H such that

$$\Lambda f = \int_H f \, d\mu_x \qquad (f \in C(H)). \tag{8}$$

In particular, we get the representation formula

$$f(x) = \int_H f \, d\mu_x \qquad (f \in A). \tag{9}$$

What we have proved is that *to each* $x \in K$ *there corresponds a positive measure* μ_x *on the "boundary"* H *which "represents"* x *in the sense that* (9) *holds for every* $f \in A$.

Note that Λ determines μ_x uniquely; but there is no reason to expect the Hahn-Banach extension to be unique. Hence, in general, we cannot say much about the uniqueness of the representing measures. Under special circumstances we do get uniqueness, as we shall see presently.

5.23 To see an example of the preceding situation, let $U = \{z: |z| < 1\}$ be the open unit disc in the complex plane, put $K = \bar{U}$ (the closed unit disc), and take for H the boundary T of U. We claim that every polynomial f, i.e., every function of the form

$$f(z) = \sum_{n=0}^{N} a_n z^n, \tag{1}$$

where $a_0, \ldots, a_N$ are complex numbers, satisfies the relation

$$\|f\|_U = \|f\|_T. \tag{2}$$

(Note that the continuity of f shows that the supremum of $|f|$ over U is the same as that over $\bar{U}$.)

Since $\bar{U}$ is compact, there exists a $z_0 \in \bar{U}$ such that $|f(z_0)| \geq |f(z)|$ for all $z \in \bar{U}$. Assume $z_0 \in U$. Then

$$f(z) = \sum_{n=0}^{N} b_n(z - z_0)^n, \tag{3}$$

and if $0 < r < 1 - |z_0|$, we obtain

$$\sum_{n=0}^{N} |b_n|^2 r^{2n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(z_0 + re^{i\theta})|^2\, d\theta \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(z_0)|^2\, d\theta = |b_0|^2,$$

so that $b_1 = b_2 = \cdots = b_N = 0$; i.e., f is constant. Thus $z_0 \in T$ for every nonconstant polynomial f, and this proves (2).

(We have just proved a special case of the *maximum modulus theorem*; we shall see later that this is an important property of all holomorphic functions.)

5.24 The Poisson Integral Let A be any subspace of $C(\bar{U})$ (where $\bar{U}$ is the closed unit disc, as above) such that A contains all polynomials and such that

$$\|f\|_U = \|f\|_T \tag{1}$$

holds for every $f \in A$. We do not exclude the possibility that A consists of precisely the polynomials, but A might be larger.

The general result obtained in Sec. 5.22 applies to A and shows that to each $z \in U$ there corresponds a positive Borel measure μ_z on T such that

$$f(z) = \int_T f\, d\mu_z \qquad (f \in A). \tag{2}$$

(This also holds for $z \in T$, but is then trivial: μ_z is simply the unit mass concentrated at the point z.)

We now fix $z \in U$ and write $z = re^{i\theta}$, $0 \leq r < 1$, θ real.

If $u_n(w) = w^n$, then $u_n \in A$ for $n = 0, 1, 2, \ldots$; hence (2) shows that

$$r^n e^{in\theta} = \int_T u_n\, d\mu_z \qquad (n = 0, 1, 2, \ldots). \tag{3}$$

Since $u_{-n} = \bar{u}_n$ on T, (3) leads to

$$\int_T u_n\, d\mu_z = r^{|n|} e^{in\theta} \qquad (n = 0, \pm 1, \pm 2, \ldots). \tag{4}$$

This suggests that we look at the real function

$$P_r(\theta - t) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta - t)} \qquad (t \text{ real}), \tag{5}$$

since

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta - t) e^{int}\, dt = r^{|n|} e^{in\theta} \qquad (n = 0, \pm 1, \pm 2, \ldots). \tag{6}$$

Note that the series (5) is dominated by the convergent geometric series $\sum r^{|n|}$, so that it is legitimate to insert the series into the integral (6) and to integrate term by term, which gives (6). Comparison of (4) and (6) gives

$$\int_T f\, d\mu_z = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it}) P_r(\theta - t)\, dt \tag{7}$$

for $f = u_n$, hence for every trigonometric polynomial f, and Theorem 4.25 now implies that (7) holds for every $f \in C(T)$. [This shows that μ_z was uniquely determined by (2). Why?]

In particular, (7) holds if $f \in A$, and then (2) gives the representation

$$f(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it}) P_r(\theta - t)\, dt \qquad (f \in A). \tag{8}$$

The series (5) can be summed explicitly, since it is the real part of

$$1 + 2\sum_{1}^{\infty} (ze^{-it})^n = \frac{e^{it} + z}{e^{it} - z} = \frac{1 - r^2 + 2ir \sin(\theta - t)}{|1 - ze^{-it}|^2}.$$

Thus

$$P_r(\theta - t) = \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2}. \tag{9}$$

This is the so-called "Poisson kernel." Note that $P_r(\theta - t) \geq 0$ if $0 \leq r < 1$.

We now summarize what we have proved:

5.25 Theorem *Suppose A is a vector space of continuous complex functions on the closed unit disc $\bar{U}$. If A contains all polynomials, and if*

$$\sup_{z \in U} |f(z)| = \sup_{z \in T} |f(z)| \tag{1}$$

for every $f \in A$ (where T is the unit circle, the boundary of U), then the Poisson integral representation

$$f(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2} f(e^{it})\, dt \qquad (z = re^{i\theta}) \tag{2}$$

is valid for every $f \in A$ and every $z \in U$.

Exercises

1 Let X consist of two points a and b, put $\mu(\{a\}) = \mu(\{b\}) = \frac{1}{2}$, and let $L^p(\mu)$ be the resulting *real* L^p-space. Identify each real function f on X with the point $(f(a), f(b))$ in the plane, and sketch the unit balls of $L^p(\mu)$, for $0 < p \le \infty$. Note that they are convex if and only if $1 \le p \le \infty$. For which p is this unit ball a square? A circle? If $\mu(\{a\}) \ne \mu(b)$, how does the situation differ from the preceding one?

2 Prove that the unit ball (open or closed) is convex in every normed linear space.

3 If $1 < p < \infty$, prove that the unit ball of $L^p(\mu)$ is *strictly convex*; this means that if

$$\|f\|_p = \|g\|_p = 1, \qquad f \ne g, \qquad h = \tfrac{1}{2}(f + g),$$

then $\|h\|_p < 1$. (Geometrically, the surface of the ball contains no straight lines.) Show that this fails in every $L^1(\mu)$, in every $L^\infty(\mu)$, and in every $C(X)$. (Ignore trivialities, such as spaces consisting of only one point.)

4 Let C be the space of all continuous functions on $[0, 1]$, with the supremum norm. Let M consist of all $f \in C$ for which

$$\int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt = 1.$$

Prove that M is a closed convex subset of C which contains no element of minimal norm.

5 Let M be the set of all $f \in L^1([0, 1])$, relative to Lebesgue measure, such that

$$\int_0^1 f(t)\, dt = 1.$$

Show that M is a closed convex subset of $L^1([0, 1])$ which contains infinitely many elements of minimal norm. (Compare this and Exercise 4 with Theorem 4.10.)

6 Let f be a bounded linear functional on a subspace M of a Hilbert space H. Prove that f has a *unique* norm-preserving extension to a bounded linear functional on H, and that this extension vanishes on $M^\perp$.

7 Construct a bounded linear functional on some subspace of some $L^1(\mu)$ which has two (hence infinitely many) distinct norm-preserving linear extensions to $L^1(\mu)$.

8 Let X be a normed linear space, and let X^* be its dual space, as defined in Sec. 5.21, with the norm

$$\|f\| = \sup\{|f(x)| : \|x\| \le 1\}.$$

(*a*) Prove that X^* is a Banach space.

(*b*) Prove that the mapping $f \to f(x)$ is, for each $x \in X$, a bounded linear functional on X^*, of norm $\|x\|$. (This gives a natural imbedding of X in its "second dual" X^{**}, the dual space of X^*.)

(*c*) Prove that $\{\|x_n\|\}$ is bounded if $\{x_n\}$ is a sequence in X such that $\{f(x_n)\}$ is bounded for every $f \in X^*$.

9 Let c_0, ℓ^1, and ℓ^∞ be the Banach spaces consisting of all complex sequences $x = \{\xi_i\}$, $i = 1, 2, 3, \ldots$, defined as follows:

$$x \in \ell^1 \text{ if and only if } \|x\|_1 = \sum |\xi_i| < \infty.$$

$$x \in \ell^\infty \text{ if and only if } \|x\|_\infty = \sup |\xi_i| < \infty.$$

c_0 is the subspace of ℓ^∞ consisting of all $x \in \ell^\infty$ for which $\xi_i \to 0$ as $i \to \infty$.

Prove the following four statements.

(*a*) If $y = \{\eta_i\} \in \ell^1$ and $\Lambda x = \sum \xi_i \eta_i$ for every $x \in c_0$, then Λ is a bounded linear functional on c_0, and $\|\Lambda\| = \|y\|_1$. Moreover, every $\Lambda \in (c_0)^*$ is obtained in this way. In brief, $(c_0)^* = \ell^1$.

(More precisely, these two spaces are not equal; the preceding statement exhibits an isometric vector space isomorphism between them.)

(*b*) In the same sense, $(\ell^1)^* = \ell^\infty$.

(*c*) Every $y \in \ell^1$ induces a bounded linear functional on ℓ^∞, as in (*a*). However, this does *not* give all of $(\ell^\infty)^*$, since $(\ell^\infty)^*$ contains nontrivial functionals that vanish on all of c_0.

(*d*) c_0 and ℓ^1 are separable but ℓ^∞ is not.

10 If $\sum \alpha_i \xi_i$ converges for every sequence (ξ_i) such that $\xi_i \to 0$ as $i \to \infty$, prove that $\sum |\alpha_i| < \infty$.

11 For $0 < \alpha \le 1$, let Lip α denote the space of all complex functions f on $[a, b]$ for which

$$M_f = \sup_{s \ne t} \frac{|f(s) - f(t)|}{|s - t|^\alpha} < \infty.$$

Prove that Lip α is a Banach space, if $\|f\| = |f(a)| + M_f$; also, if

$$\|f\| = M_f + \sup_x |f(x)|.$$

(The members of Lip α are said to satisfy a *Lipschitz condition* of order α.)

12 Let K be a triangle (two-dimensional figure) in the plane, let H be the set consisting of the vertices of K, and let A be the set of all real functions f on K, of the form

$$f(x, y) = \alpha x + \beta y + \gamma \qquad (\alpha, \beta, \text{ and } \gamma \text{ real}).$$

Show that to each $(x_0, y_0) \in K$ there corresponds a unique measure μ on H such that

$$f(x_0, y_0) = \int_H f\, d\mu.$$

(Compare Sec. 5.22.)

Replace K by a square, let H again be the set of its vertices, and let A be as above. Show that to each point of K there still corresponds a measure on H, with the above property, but that uniqueness is now lost.

Can you extrapolate to a more general theorem? (Think of other figures, higher dimensional spaces.)

13 Let $\{f_n\}$ be a sequence of continuous complex functions on a (nonempty) complete metric space X, such that $f(x) = \lim f_n(x)$ exists (as a complex number) for every $x \in X$.

(*a*) Prove that there is an open set $V \ne \varnothing$ and a number $M < \infty$ such that $|f_n(x)| < M$ for all $x \in V$ and for $n = 1, 2, 3, \ldots$.

(*b*) If $\epsilon > 0$, prove that there is an open set $V \ne \varnothing$ and an integer N such that $|f(x) - f_n(x)| \le \epsilon$ if $x \in V$ and $n \ge N$.

Hint for (*b*): For $N = 1, 2, 3, \ldots$, put

$$A_N = \{x : |f_m(x) - f_n(x)| \le \epsilon \text{ if } m \ge N \text{ and } n \ge N\}.$$

Since $X = \bigcup A_N$, some A_N has a nonempty interior.

14 Let C be the space of all real continuous functions on $I = [0, 1]$ with the supremum norm. Let X_n be the subset of C consisting of those f for which there exists a $t \in I$ such that $|f(s) - f(t)| \le n|s - t|$ for all $s \in I$. Fix n and prove that each open set in C contains an open set which does not intersect X_n. (Each $f \in C$ can be uniformly approximated by a zigzag function g with very large slopes, and if $\|g - h\|$ is small, $h \notin X_n$.) Show that this implies the existence of a dense G_δ in C which consists entirely of nowhere differentiable functions.

15 Let $A = (a_{ij})$ be an infinite matrix with complex entries, where $i, j = 0, 1, 2, \ldots$. A associates with each sequence $\{s_j\}$ a sequence $\{\sigma_i\}$, defined by

$$\sigma_i = \sum_{j=0}^{\infty} a_{ij} s_j \qquad (i = 1, 2, 3, \ldots),$$

provided that these series converge.

Prove that A transforms every convergent sequence $\{s_j\}$ to a sequence $\{\sigma_i\}$ which converges to the same limit if and only if the following conditions are satisfied:

(a) $$\lim_{i \to \infty} a_{ij} = 0 \qquad \text{for each } j.$$

(b) $$\sup_i \sum_{j=0}^{\infty} |a_{ij}| < \infty.$$

(c) $$\lim_{i \to \infty} \sum_{j=0}^{\infty} a_{ij} = 1.$$

The process of passing from $\{s_j\}$ to $\{\sigma_i\}$ is called a *summability method*. Two examples are:

$$a_{ij} = \begin{cases} \dfrac{1}{i+1} & \text{if } 0 \le j \le i, \\ 0 & \text{if } i < j, \end{cases}$$

and
$$a_{ij} = (1 - r_i)r_i^j, \qquad 0 < r_i < 1, \qquad r_i \to 1.$$

Prove that each of these also transforms some divergent sequences $\{s_j\}$ (even some unbounded ones) to convergent sequences $\{\sigma_i\}$.

16 Suppose X and Y are Banach spaces, and suppose Λ is a linear mapping of X into Y, with the following property: For every sequence $\{x_n\}$ in X for which $x = \lim x_n$ and $y = \lim \Lambda x_n$ exist, it is true that $y = \Lambda x$. Prove that Λ is continuous.

This is the so-called "closed graph theorem." *Hint*: Let $X \oplus Y$ be the set of all ordered pairs (x, y), $x \in X$ and $y \in Y$, with addition and scalar multiplication defined componentwise. Prove that $X \oplus Y$ is a Banach space, if $\|(x, y)\| = \|x\| + \|y\|$. The graph G of Λ is the subset of $X \oplus Y$ formed by the pairs $(x, \Lambda x)$, $x \in X$. Note that our hypothesis says that G is closed; hence G is a Banach space. Note that $(x, \Lambda x) \to x$ is continuous, one-to-one, and linear and maps G onto X.

Observe that there exist *nonlinear* mappings (of R^1 onto R^1, for instance) whose graph is closed although they are not continuous: $f(x) = 1/x$ if $x \ne 0$, $f(0) = 0$.

17 If μ is a positive measure, each $f \in L^\infty(\mu)$ defines a multiplication operator M_f on $L^2(\mu)$ into $L^2(\mu)$, such that $M_f(g) = fg$. Prove that $\|M_f\| \le \|f\|_\infty$. For which measures μ is it true that $\|M_f\| = \|f\|_\infty$ for all $f \in L^\infty(\mu)$? For which $f \in L^\infty(\mu)$ does M_f map $L^2(\mu)$ onto $L^2(\mu)$?

18 Suppose $\{\Lambda_n\}$ is a sequence of bounded linear transformations from a normed linear space X to a Banach space Y, suppose $\|\Lambda_n\| \le M < \infty$ for all n, and suppose there is a dense set $E \subset X$ such that $\{\Lambda_n x\}$ converges for each $x \in E$. Prove that $\{\Lambda_n x\}$ converges for each $x \in X$.

19 If s_n is the nth partial sum of the Fourier series of a function $f \in C(T)$, prove that $s_n/\log n \to 0$ uniformly, as $n \to \infty$, for each $f \in C(T)$. That is, prove that

$$\lim_{n \to \infty} \frac{\|s_n\|_\infty}{\log n} = 0.$$

On the other hand, if $\lambda_n/\log n \to 0$, prove that there exists an $f \in C(T)$ such that the sequence $\{s_n(f; 0)/\lambda_n\}$ is unbounded. *Hint*: Apply the reasoning of Exercise 18 and that of Sec. 5.11, with a better estimate of $\|D_n\|_1$ than was used there.

20 (a) Does there exist a sequence of continuous positive functions f_n on R^1 such that $\{f_n(x)\}$ is unbounded if and only if x is rational?

(b) Replace "rational" by "irrational" in (a) and answer the resulting question.

(c) Replace "$\{f_n(x)\}$ is unbounded" by "$f_n(x) \to \infty$ as $n \to \infty$" and answer the resulting analogues of (a) and (b).

21 Suppose $E \subset R^1$ is measurable, and $m(E) = 0$. Must there be a translate $E + x$ of E that does not intersect E? Must there be a homeomorphism h of R^1 onto R^1 so that $h(E)$ does not intersect E?

22 Suppose $f \in C(T)$ and $f \in \text{Lip}\,\alpha$ for some $\alpha > 0$. (See Exercise 11.) Prove that the Fourier series of f converges to $f(x)$, by completing the following outline: It is enough to consider the case $x = 0$, $f(0) = 0$. The difference between the partial sums $s_n(f; 0)$ and the integrals

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin nt}{t}\, dt$$

tends to 0 as $n \to \infty$. The function $f(t)/t$ is in $L^1(T)$. Apply the Riemann-Lebesgue lemma. More careful reasoning shows that the convergence is actually uniform on T.