holds for each positive ε. As we noted in Sect. 2.2, the sequence $\{f_n\}$ converges to f *almost everywhere* if $f(x) = \lim_n f_n(x)$ holds at μ-almost every point x in X.

Examples 3.1.1. We should note that in general convergence in measure neither implies nor is implied by convergence almost everywhere.

(a) To see that convergence almost everywhere does not imply convergence in measure, consider the space $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ and the sequence whose nth term is the characteristic function of the interval $[n, +\infty)$. This sequence clearly converges to the zero function almost everywhere (in fact, everywhere) but not in measure.
(b) Next consider the interval $[0,1)$, together with the σ-algebra of Borel subsets of $[0,1)$ and Lebesgue measure. Let $\{f_n\}$ be the sequence whose first term is the characteristic function of $[0,1)$, whose next two terms are the characteristic functions of $[0,1/2)$ and $[1/2,1)$, whose next four terms are the characteristic functions of $[0,1/4)$, $[1/4,1/2)$, $[1/2,3/4)$, and $[3/4,1)$, and so on. Then $\{f_n\}$ converges to the zero function in measure, but for each x in $[0,1)$ the sequence $\{f_n(x)\}$ contains infinitely many ones and infinitely many zeros and so is not convergent. □

Nevertheless there are some useful relations, given by the following two propositions, between convergence in measure and convergence almost everywhere (see also Exercise 6).

Proposition 3.1.2. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and f_1, f_2, ... be real-valued $\mathscr{A}$-measurable functions on X. If μ is finite and if $\{f_n\}$ converges to f almost everywhere, then $\{f_n\}$ converges to f in measure.*

Proof. We must show that

$$\lim_n \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0$$

holds for each positive ε. So let ε be a positive number, and define sets A_1, A_2, ... and B_1, B_2, ... by

$$A_n = \{x \in X : |f_n(x) - f(x)| > \varepsilon\}$$

and $B_n = \cup_{k=n}^{\infty} A_k$. The sequence $\{B_n\}$ is decreasing, and its intersection is included in

$$\{x \in X : \{f_n(x)\} \text{ does not converge to } f(x)\}.$$

Thus $\mu(\cap_n B_n) = 0$, and so (Proposition 1.2.5) $\lim_n \mu(B_n) = 0$. Since $A_n \subseteq B_n$, it follows that

$$\lim_n \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = \lim_n \mu(A_n) = 0.$$

Thus $\{f_n\}$ converges to f in measure. □

Proposition 3.1.3. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and $f_1, f_2, \ldots$ be real-valued $\mathscr{A}$-measurable functions on X. If $\{f_n\}$ converges to f in measure, then there is a subsequence of $\{f_n\}$ that converges to f almost everywhere.*

Proof. The hypothesis that $\{f_n\}$ converges to f in measure means that

$$\lim_n \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0$$

holds for each positive ε. We use this relation to construct a sequence $\{n_k\}$ of positive integers, choosing n_1 so that

$$\mu(\{x \in X : |f_{n_1}(x) - f(x)| > 1\}) \leq \frac{1}{2},$$

and then choosing the remaining terms of $\{n_k\}$ inductively so that the relations $n_k > n_{k-1}$ and

$$\mu\left(\left\{x \in X : |f_{n_k}(x) - f(x)| > \frac{1}{k}\right\}\right) \leq \frac{1}{2^k}$$

hold for $k = 2, 3, \ldots$. Define sets A_k, $k = 1, 2, \ldots$, by

$$A_k = \left\{x \in X : |f_{n_k}(x) - f(x)| > \frac{1}{k}\right\}.$$

If $x \notin \cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k$, then there is a positive integer j such that $x \notin \cup_{k=j}^{\infty} A_k$ and hence such that $|f_{n_k}(x) - f(x)| \leq 1/k$ holds for $k = j, j+1, \ldots$. Thus $\{f_{n_k}\}$ converges to f at each x outside $\cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k$. Since

$$\mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) \leq \sum_{k=j}^{\infty} \frac{1}{2^k} = \frac{1}{2^{j-1}}$$

holds for each j, it follows that $\mu(\cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k) = 0$, and the proof is complete. □

Proposition 3.1.4 (Egoroff's Theorem). *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and $f_1, f_2, \ldots$ be real-valued $\mathscr{A}$-measurable functions on X. If μ is finite and if $\{f_n\}$ converges to f almost everywhere, then for each positive number ε there is a subset B of X that belongs to $\mathscr{A}$, satisfies $\mu(B^c) < \varepsilon$, and is such that $\{f_n\}$ converges to f uniformly on B.*

Proof. Let ε be a positive number, and for each n let $g_n = \sup_{j \geq n} |f_j - f|$. It is easy to check that each g_n is finite almost everywhere. The sequence $\{g_n\}$ converges to 0 almost everywhere, and so in measure (see Proposition 3.1.2 and the footnote at the beginning of this section). Hence for each positive integer k we can choose a positive integer n_k such that

$$\mu\left(\left\{x \in X : g_{n_k}(x) > \frac{1}{k}\right\}\right) < \frac{\varepsilon}{2^k}.$$

Define sets $B_1, B_2, \ldots$ by $B_k = \{x \in X : g_{n_k}(x) \le 1/k\}$, and let $B = \cap_k B_k$. The set B satisfies

$$\mu(B^c) = \mu\left(\bigcup_k B_k^c\right) \le \sum_k \mu(B_k^c) < \sum_k \frac{\varepsilon}{2^k} = \varepsilon.$$

If δ is a positive number and if k is a positive integer such that $1/k < \delta$, then, since $B \subseteq B_k$,

$$|f_n(x) - f(x)| \le g_{n_k}(x) \le \frac{1}{k} < \delta$$

holds for all x in B and all positive integers n such that $n \ge n_k$; thus $\{f_n\}$ converges to f uniformly on B. □

Egoroff's theorem provides motivation for the following definition. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and $f_1, f_2, \ldots$ be real-valued $\mathscr{A}$-measurable functions on X. Then $\{f_n\}$ converges to f *almost uniformly* if for each positive number ε there is a subset B of X that belongs to $\mathscr{A}$, satisfies $\mu(B^c) < \varepsilon$, and is such that $\{f_n\}$ converges to f uniformly on B. It is clear that if $\{f_n\}$ converges to f almost uniformly, then $\{f_n\}$ converges to f almost everywhere. It follows from this remark and Egoroff's theorem that on a finite measure space almost everywhere convergence is equivalent to almost uniform convergence.

Suppose that $(X, \mathscr{A}, \mu)$ is a measure space and that f and f_1, f_2, $\ldots$ belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$. Then $\{f_n\}$ converges to f *in mean* if

$$\lim_n \int |f_n - f| \, d\mu = 0.$$

Proposition 3.1.5. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and f_1, f_2, $\ldots$ belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$. If $\{f_n\}$ converges to f in mean, then $\{f_n\}$ converges to f in measure.*

Proof. This is an immediate consequence of the inequality

$$\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \le \frac{1}{\varepsilon} \int |f_n - f| \, d\mu$$

(see Proposition 2.3.10). □

Convergence in mean does not, however, imply convergence almost everywhere (see the example given above of a sequence that converges in measure but not almost everywhere). On the other hand, if $\{f_n\}$ converges to f in mean, then $\{f_n\}$ does have a subsequence that converges to f almost everywhere; this follows from Propositions 3.1.3 and 3.1.5 (or, alternatively, from Exercise 4).

Neither convergence almost everywhere nor convergence in measure implies convergence in mean. To see this, consider the space $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$, and define a sequence $\{f_n\}$ by letting f_n have value n on the interval $[0, 1/n]$ and value 0 elsewhere. Then $\{f_n\}$ converges to 0 almost everywhere and in measure, but not in mean (note that $\int |f_n - 0| \, d\lambda = 1$). There are, however, supplementary hypotheses under which convergence almost everywhere or in measure does imply convergence in mean; such hypotheses are given in the following proposition and in Exercise 4.2.16.

Proposition 3.1.6. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and f_1, f_2, ... belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$. If $\{f_n\}$ converges to f almost everywhere or in measure, and if there is a nonnegative extended real-valued integrable function g such that*

$$|f_n| \leq g \ \text{(for } n = 1, 2, \dots\text{) and } |f| \leq g \tag{1}$$

hold almost everywhere, then $\{f_n\}$ converges to f in mean.

Proof. First suppose that $\{f_n\}$ converges to f almost everywhere and hence that $\{|f_n - f|\}$ converges to 0 almost everywhere. Relation (1) implies that

$$|f_n - f| \leq |f_n| + |f| \leq 2g$$

holds almost everywhere. Thus we can use the dominated convergence theorem (Theorem 2.4.5) to conclude that $\lim_n \int |f_n - f| \, d\mu = 0$.

Now suppose that $\{f_n\}$ converges to f in measure and satisfies condition (1). Then every subsequence of $\{f_n\}$ has a subsequence that converges to f almost everywhere (Proposition 3.1.3), and so by what we have just proved, in mean. If the original sequence $\{f_n\}$ did not converge to f in mean, then there would be a positive number ε and a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $\int |f_{n_k} - f| \, d\mu \geq \varepsilon$ holds for each k. Since this subsequence could have no subsequence converging to f in mean, we have a contradiction. Thus $\{f_n\}$ must converge to f in mean. □

Exercises

1. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let A and $A_1, A_2, \dots$ belong to $\mathscr{A}$. Show that
   (a) $\{\chi_{A_n}\}$ converges to 0 in measure if and only if $\lim_n \mu(A_n) = 0$,
   (b) $\{\chi_{A_n}\}$ converges to 0 almost everywhere if and only if $\mu(\cap_{n=1}^\infty \cup_{k=n}^\infty A_k) = 0$, and
   (c) $\{\chi_{A_n}\}$ converges to χ_A almost everywhere if and only if the three sets A, $\cap_{n=1}^\infty \cup_{k=n}^\infty A_k$, and $\cup_{n=1}^\infty \cap_{k=n}^\infty A_k$ differ only by μ-null sets. (Hint: See Exercise 2.1.1.)
2. Let μ be counting measure on the σ-algebra of all subsets of $\mathbb{N}$, and let f and f_1, $f_2, \dots$ be real-valued functions on $\mathbb{N}$. Show that $\{f_n\}$ converges to f in measure if and only if it converges uniformly to f.
3. Let $(X, \mathscr{A}, \mu)$ be a measure space, let f and f_1, f_2, ... be real-valued $\mathscr{A}$-measurable functions on X, and let $g \colon \mathbb{R} \to \mathbb{R}$ be Borel measurable. Show that if $\{f_n\}$ converges to f almost everywhere and if g is continuous at $f(x)$ for almost every x, then $\{g \circ f_n\}$ converges to $g \circ f$ almost everywhere.

4. Suppose that $(X, \mathscr{A}, \mu)$ is a measure space and that f and f_1, f_2, ... belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$. Show that if $\{f_n\}$ converges to f in mean so fast that
$$\sum_n \int |f_n - f| \, d\mu < +\infty,$$
then $\{f_n\}$ converges to f almost everywhere.
5. Let μ be a measure on $(X, \mathscr{A})$, and let f, f_1, f_2, ... and g, g_1, g_2, ... be real-valued $\mathscr{A}$-measurable functions on X.
   (a) Show that if μ is finite, if $\{f_n\}$ converges to f in measure, and if $\{g_n\}$ converges to g in measure, then $\{f_n g_n\}$ converges to fg in measure.
   (b) Can the assumption that μ is finite be omitted in part (a)?
6. Let μ be a finite measure on $(X, \mathscr{A})$ and f and f_1, f_2, ... be real-valued $\mathscr{A}$-measurable functions on X. Show that $\{f_n\}$ converges to f in measure if and only if each subsequence of $\{f_n\}$ has a subsequence that converges to f almost everywhere.
7. Egoroff's theorem applies to *sequences* of measurable functions on a finite measure space. One can ask about the situation where one has a family $\{f_t\}_{t \in T}$ on a finite measure space $(X, \mathscr{A}, \mu)$, where T is a subinterval of $\mathbb{R}$ of the form $[t_0, +\infty)$. (The following results are due to Walter [125].)
   (a) For each n in $\mathbb{N}$ define g_n by $g_n(x) = \sup\{|f_t(x) - f(x)| : t \in [n, +\infty)\}$. Show that if each g_n is measurable, then the conclusion of Egoroff's theorem holds for the family $\{f_t\}_{t \in T}$.
   (b) Let $\{A_n\}$ be a sequence of disjoint subsets of $[0,1]$ that are not Lebesgue measurable and are such that all the A_n's have the same (strictly positive) Lebesgue outer measure. (See the discussion of nonmeasurable sets in Sect. 1.4.) Define a subset B of $[0,1] \times [1, +\infty)$ by
   $$B = \{(x,t) : x \in A_n \text{ and } t = x + n \text{ for some } n\},$$
   and for each t let f_t be the characteristic function of the set $\{x \in [0,1] : (x,t) \in B\}$. Show that each f_t is Borel measurable but that the conclusion of Egoroff's theorem fails for the family $\{f_t\}_{t \in [1, +\infty)}$

3.2 Normed Spaces

Let V be a vector space over $\mathbb{R}$ (or over $\mathbb{C}$). A *norm* on V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies

(a) $\|v\| \ge 0$,
(b) $\|v\| = 0$ if and only if $v = 0$,
(c) $\|\alpha v\| = |\alpha| \|v\|$, and
(d) $\|u + v\| \le \|u\| + \|v\|$

for each u and v in V and each α in $\mathbb{R}$ (or in $\mathbb{C}$). Condition (c) says that $\|\cdot\|$ is *homogeneous*, and condition (d) says that it satisfies the *triangle inequality*. If in condition (b) the words "if and only if" are replaced with the word "if," but conditions (a), (c), and (d) remain unchanged, then $\|\cdot\|$ is called a *seminorm*. Thus a norm is a seminorm for which 0 is the only vector that satisfies $\|v\| = 0$. A *normed vector space* (or a *normed linear space*) is a vector space together with a norm.

Examples 3.2.1. Let us consider a few examples.

(a) The function that assigns to each number its absolute value is a norm on $\mathbb{R}$ (or on $\mathbb{C}$). This is the norm that will be assumed whenever we deal with $\mathbb{R}$ or $\mathbb{C}$ as a normed space.

(b) The formula $\|(x_1,\dots,x_d)\|_2 = (\sum_{i=1}^d |x_i|^2)^{1/2}$ defines a norm on $\mathbb{R}^d$ and on $\mathbb{C}^d$ (the triangle inequality follows from Exercise 9 or from Minkowski's inequality (Proposition 3.3.3)).

(c) Let $(X,\mathscr{A},\mu)$ be a measure space and let $\mathscr{L}^1(X,\mathscr{A},\mu,\mathbb{R})$ be the set of all real-valued integrable functions on X. Then $\mathscr{L}^1(X,\mathscr{A},\mu,\mathbb{R})$ is a vector space over $\mathbb{R}$, and the formula

$$\|f\|_1 = \int |f|\,d\mu$$

defines a seminorm on $\mathscr{L}^1(X,\mathscr{A},\mu,\mathbb{R})$. If f is an $\mathscr{A}$-measurable function on X such that $f = 0$ holds almost everywhere but not everywhere, then f satisfies $\|f\|_1 = 0$ but not $f = 0$. Thus for many choices of $(X,\mathscr{A},\mu)$ the seminorm $\|\cdot\|_1$ is not a norm.

(d) Let $[a,b]$ be a closed bounded interval, and let $C[a,b]$ be the vector space of all continuous real-valued functions on $[a,b]$. The function $\|\cdot\|_1\colon C[a,b] \to \mathbb{R}$ defined by

$$\|f\|_1 = \int_a^b |f|\,d\lambda$$

is a norm (note that a continuous function on $[a,b]$ that vanishes almost everywhere must vanish everywhere).

(e) The function $\|\cdot\|_\infty\colon C[a,b] \to \mathbb{R}$ defined by the formula

$$\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$$

is a norm (the continuity of f and the compactness of $[a,b]$ imply that $\|f\|_\infty$ is finite; see Theorem C.12). It is called the *uniform norm* or the *sup* (for supremum) *norm* on $C[a,b]$.

(f) More generally, let X be an arbitrary nonempty set, and let V be a vector space of bounded real-valued (or complex-valued) functions on X. Then the formula

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}$$

defines a norm on V. □

Recall that a *metric* on a set S is a function $d\colon S \times S \to \mathbb{R}$ that satisfies

(a) $d(s,t) \geq 0$,
(b) $d(s,t) = 0$ if and only if $s = t$,
(c) $d(s,t) = d(t,s)$, and
(d) $d(r,t) \leq d(r,s) + d(s,t)$

for all r, s, and t in S. Condition (d) says that d satisfies the *triangle inequality*. If in condition (b) the words "if and only if" are replaced with the word "if," but conditions (a), (c), and (d) remain unchanged, then d is called a *semimetric*. A *metric space* is a set S together with a metric on S.

It is easy to check that if V is a vector space and if $\|\cdot\|$ is a norm (or a seminorm) on V, then the formula

$$d(u,v) = \|u - v\|$$

defines a metric (or a semimetric) on V.

Recall that if S is a metric space and if s and $s_1, s_2, \ldots$ are elements of S, then the sequence $\{s_n\}$ *converges* to s if $\lim_n d(s_n, s) = 0$; the point s is then called the *limit* of $\{s_n\}$ and is denoted by $\lim_n s_n$ (see Exercise 1). In particular, if V is a normed linear space and if v and v_1, v_2, ... are elements of V, then the sequence $\{v_n\}$ converges to v (with respect to the metric induced by the norm on V) if and only if $\lim_n \|v_n - v\| = 0$.

Examples 3.2.2. Let us return to some of the examples above. The metric induced on $\mathbb{R}^d$ by the norm defined in Example 3.2.1(b) is the usual one, stemming from the Pythagorean theorem. If $(X, \mathscr{A}, \mu)$ is a measure space and if f and f_1, f_2, ... belong to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$, then $\{f_n\}$ converges to f with respect to the seminorm[2] defined in Example 3.2.1(c) if and only if it converges to f in mean (see Sect. 3.1). Finally, if f and f_1, f_2, ... are continuous functions on $[a,b]$, then $\{f_n\}$ converges to f with respect to the norm defined in Example 3.2.1(e) if and only if it converges uniformly to f. □

Let d be a metric (or a semimetric) on a set S. Then a subset A of S is *dense* in S if for each s in S and each positive ε there is an element a of A that satisfies $d(s,a) < \varepsilon$. It is clear that A is dense in S if and only if for each s in S there is a sequence $\{a_n\}$ of elements of A such that $\lim_n d(a_n, s) = 0$. A metric (or semimetric) space is *separable* if it has a countable dense subset. For example, the rational numbers form a countable dense subset of $\mathbb{R}$, and so $\mathbb{R}$ is separable.

Now let S be an arbitrary metric space. A sequence $\{s_n\}$ of elements of S is a *Cauchy sequence* if for each positive number ε there is a positive integer N such that $d(s_m, s_n) < \varepsilon$ holds whenever $m \geq N$ and $n \geq N$. Of course, every convergent sequence is a Cauchy sequence (let s be the limit of $\{s_n\}$, and note that $d(s_m, s_n) \leq$

[2]Convergence with respect to a semimetric or a seminorm is defined in the same way as convergence with respect to a metric or a norm. Note, however, that a sequence that is convergent with respect to a semimetric or a seminorm might have several limits.

$d(s_m, s) + d(s, s_n)$). On the other hand, if every Cauchy sequence in S converges to a point in S, then S is called *complete*. A normed linear space that is complete (with respect to the metric induced by its norm) is called a *Banach space*.

It is a basic consequence of the axioms for the real number system that $\mathbb{R}$ is complete under the metric defined by $(x, y) \mapsto |x - y|$.[3] The proofs of completeness that we give for other spaces will depend ultimately on this fact.

Example 3.2.3. Let us show that $C[a, b]$ is complete under the uniform norm. Let $\{f_n\}$ be a Cauchy sequence in $C[a, b]$. For each x in $[a, b]$ the sequence $\{f_n(x)\}$ satisfies $|f_m(x) - f_n(x)| \le \|f_m - f_n\|_\infty$ and so is a Cauchy sequence of real numbers; thus it is convergent. Define a function $f\colon [a, b] \to \mathbb{R}$ by letting $f(x) = \lim_n f_n(x)$ hold at each x in $[a, b]$. We need to show that $\{f_n\}$ converges uniformly to f and that f is continuous. Let us begin by showing that the convergence of $\{f_n\}$ to f is uniform. Let ε be a positive number, and use the fact that $\{f_n\}$ is a Cauchy sequence to choose a positive integer N such that $\|f_m - f_n\|_\infty < \varepsilon$ holds whenever m and n satisfy $m \ge N$ and $n \ge N$. Then

$$|f_m(x) - f_n(x)| < \varepsilon$$

holds for all x in $[a, b]$ and all m and n satisfying $m \ge N$ and $n \ge N$, and so (take limits as m approaches infinity)

$$|f(x) - f_n(x)| \le \varepsilon$$

holds for all x in $[a, b]$ and all n satisfying $n \ge N$. Thus $\|f_n - f\|_\infty \le \varepsilon$ holds[4] when $n \ge N$. Since ε was arbitrary, we have shown that $\{f_n\}$ converges uniformly to f.

We turn to the continuity of f. Let x_0 belong to $[a, b]$, and let ε be an arbitrary positive number. Choose a positive integer N such that $\|f_n - f\|_\infty < \varepsilon/3$ holds whenever n satisfies $n \ge N$, and then use the continuity of f_N to choose a positive number δ such that $|f_N(x) - f_N(x_0)| < \varepsilon/3$ holds if x belongs to $[a, b]$ and satisfies $|x - x_0| < \delta$. It follows that if $x \in [a, b]$ and $|x - x_0| < \delta$, then

$$\begin{aligned} |f(x) - f(x_0)| &\le |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since ε and x_0 were arbitrary, the continuity of f follows. This finishes our proof of the completeness of $C[a, b]$ under $\|\cdot\|_\infty$. □

---

[3] See, for instance, Gleason [49], Hoffman [60], Rudin [104], or Thomson, Bruckner, and Bruckner [117].

[4] Actually, the norm here and in the following paragraph is the norm from Example 3.2.1(f). We can't say that it is the norm from $C[a, b]$ until we show that f is continuous.

Example 3.2.4. Let us also note an example of a normed linear space that is not complete. Consider the space $C[-1,1]$, together with the norm defined by $\|f\|_1 = \int_{-1}^{1} |f|\, d\lambda$. For each n define a function $f_n\colon [-1,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 0 & \text{if } -1 \le x \le 0, \\ nx & \text{if } 0 < x \le \frac{1}{n}, \\ 1 & \text{if } \frac{1}{n} < x \le 1. \end{cases}$$

It is easy to check that $\{f_n\}$ is a Cauchy sequence in $C[-1,1]$, but that there is no continuous function f such that $\lim_n \|f_n - f\|_1 = 0$. Hence $C[a,b]$ is not complete under $\|\cdot\|_1$. □

We close this section with a sometimes useful criterion for the completeness of a normed linear space. Let V be a normed linear space, and let $\sum_{k=1}^{\infty} v_k$ be an infinite series with terms in V. The series $\sum_{k=1}^{\infty} v_k$ is *convergent* if $\lim_n \sum_{k=1}^{n} v_k$ exists, and is *absolutely convergent* if the series $\sum_{k=1}^{\infty} \|v_k\|$ of real numbers is convergent. Recall that every absolutely convergent series of real numbers is convergent; for more general normed linear spaces we have the following result.

Proposition 3.2.5. *Let V be a normed linear space. Then V is complete if and only if every absolutely convergent series with terms in V is convergent.*

Proof. First suppose that V is complete, and let $\sum_{k=1}^{\infty} v_k$ be an absolutely convergent series in V. Let $\{s_n\}$ be the sequence of partial sums of the series $\sum_{k=1}^{\infty} v_k$, and let $\{t_n\}$ be the sequence of partial sums of the series $\sum_{k=1}^{\infty} \|v_k\|$; thus $s_n = \sum_{k=1}^{n} v_k$ and $t_n = \sum_{k=1}^{n} \|v_k\|$. Note that if $m < n$, then

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} v_k \right\| \le \sum_{k=m+1}^{n} \|v_k\| = t_n - t_m. \tag{1}$$

The convergence of $\sum_{k=1}^{\infty} \|v_k\|$ implies that $\{t_n\}$ is a Cauchy sequence and, in view of (1), that $\{s_n\}$ is a Cauchy sequence. Since V is complete, the sequence $\{s_n\}$, and hence the series $\sum_{k=1}^{\infty} v_k$, must converge.

Next suppose that every absolutely convergent series in V is convergent, and let $\{u_n\}$ be a Cauchy sequence in V. Since $\{u_n\}$ is a Cauchy sequence, we can choose (how?) a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that $\|u_{n_{k+1}} - u_{n_k}\| \le 1/2^{k+1}$ holds for each k. Define a series $\sum_{k=1}^{\infty} v_k$ by letting $v_1 = u_{n_1}$ and letting $v_k = u_{n_k} - u_{n_{k-1}}$ if $k > 1$; thus $\{u_{n_k}\}$ is the sequence of partial sums of the series $\sum_{k=1}^{\infty} v_k$. Since $\|v_k\| \le 1/2^k$ holds if $k > 1$, the series $\sum_{k=1}^{\infty} v_k$ is absolutely convergent and hence convergent. Thus the sequence $\{u_{n_k}\}$ converges, say to u. The inequality

$$\|u - u_n\| \le \|u - u_{n_k}\| + \|u_{n_k} - u_n\|$$

implies that $\|u - u_n\|$ can be made small by making n (and k) large, and so the original sequence $\{u_n\}$ also converges to u. The completeness of V follows. □

Exercises

1. Let S be a metric space, and let $\{s_n\}$ be a sequence of elements of S. Show that $\{s_n\}$ converges to at most one point in S. (Thus the expression "$\lim_n s_n$" makes sense.)
2. Let $C^1[0,1]$ consist of those functions $f\colon [0,1] \to \mathbb{R}$ such that f' is defined and continuous at each point in $[0,1]$ (of course $f'(0)$ and $f'(1)$ are to be interpreted as one-sided derivatives). Show that
   (a) the formula $\|f\| = \int_0^1 |f'(x)|\,dx$ defines a seminorm, but not a norm, on $C^1[0,1]$, and
   (b) the formula $\|f\| = |f(0)| + \int_0^1 |f'(x)|\,dx$ defines a norm on $C^1[0,1]$.
3. Let ℓ^∞ be the set of all bounded sequences of real numbers (of course ℓ^∞ is a vector space over $\mathbb{R}$.) Show that ℓ^∞ is complete under the norm defined in Example 3.2.1(f).
4. Let c_0 be the set of all sequences $\{x_n\}$ of real numbers for which $\lim_n x_n = 0$. Show that c_0 is a closed linear subspace of ℓ^∞ (see Exercise 3) and hence that c_0 is complete under the norm $\|\cdot\|_\infty$ defined by $\|\{x_n\}\|_\infty = \sup_n |x_n|$.
5. Let μ be a finite measure on $(X, \mathscr{A})$. Show that
   (a) the formula
   $$d(f,g) = \int \frac{|f-g|}{1+|f-g|}\,d\mu$$
   defines a semimetric on the collection of all real-valued $\mathscr{A}$-measurable functions on X, and
   (b) $\lim_n d(f_n, f) = 0$ holds if and only if $\{f_n\}$ converges to f in measure.
6. Now let us consider an analogous result for the space $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$. Suppose that $h\colon \mathbb{R} \to \mathbb{R}$ is defined by $h(t) = 1/(1+t^2)$. Show that
   (a) the formula
   $$d(f,g) = \int \frac{|f-g|}{1+|f-g|} h\,d\lambda$$
   defines a semimetric on the collection of all real-valued Borel measurable functions on $\mathbb{R}$, and
   (b) $\lim_n d(f_n, f) = 0$ holds if and only if $\{f_n\}$ converges to f in measure on each bounded subinterval of $\mathbb{R}$.
7. Let V be a vector space over $\mathbb{R}$. A function $(\cdot,\cdot)\colon V \times V \to \mathbb{R}$ is an *inner product* on V if

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

hold for all x, y, z in V and all α, β in $\mathbb{R}$.[5] An *inner product space* is a vector space, together with an inner product on it. The *norm* $\|\cdot\|$ associated to the inner product $(\cdot,\cdot)$ is defined by $\|x\| = \sqrt{(x,x)}$.

(a) Prove that an inner product satisfies the *Cauchy–Schwarz inequality*: if $x, y \in V$, then $|(x,y)| \leq \|x\|\|y\|$. (Hint: Define a function $p\colon \mathbb{R} \to \mathbb{R}$ by $p(t) = \|x\|^2 + 2t(x,y) + t^2\|y\|^2$, and note that $p(t) = \|x+ty\|^2 \geq 0$ holds for each real t; then recall that a quadratic polynomial $at^2 + bt + c$ is nonnegative for each t only if $b^2 - 4ac \leq 0$.)

(b) Verify that the norm associated to $(\cdot,\cdot)$ is indeed a norm. (Hint: Use the Cauchy–Schwarz inequality when checking the triangle inequality.)

8. Let $(\cdot,\cdot)$ be an inner product on the real vector space V, and let $\|\cdot\|$ be the associated norm. Show that

(a) $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ and

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

hold for all x, y in V. (The identity in part (a) is called the *parallelogram law*.)

9. (a) Check that the formula $(x,y) = \sum_{i=1}^{d} x_i y_i$ defines an inner product on $\mathbb{R}^d$ (here, x and y are the vectors $(x_1,\ldots,x_d)$ and $(y_1,\ldots,y_d)$).

(b) Conclude that the function $\|\cdot\|_2\colon \mathbb{R}^d \to \mathbb{R}$ defined by $\|x\|_2 = (\sum_{i=1}^{d} x_i^2)^{1/2}$ is indeed a norm. (See part (b) of Exercise 7.)

10. Let ℓ^2 be the set of all infinite sequences $\{x_n\}$ of real numbers for which $\sum_n x_n^2 < +\infty$.

(a) Show that ℓ^2 is a vector space over $\mathbb{R}$. (Hint: Note that $(x+y)^2 \leq 2x^2 + 2y^2$ holds for all real x and y.)

(b) Show that the formula $(\{x_n\},\{y_n\}) = \sum_n x_n y_n$ defines an inner product on ℓ^2 and hence (see part (b) of Exercise 7) that the formula $\|\{x_n\}\| = (\sum_n x_n^2)^{1/2}$ defines a norm on ℓ^2. (The issue is the convergence of $\sum_n x_n y_n$.)

(c) Show that ℓ^2 is complete under the norm defined in part (b) of this exercise.

11. A *Hilbert space* is an inner product space that is complete under the norm defined by $\|x\| = \sqrt{(x,x)}$. Show that if H is a Hilbert space and if C is a nonempty closed convex subset of H, then there is a unique point y in C that satisfies

$$\|y\| = \inf\{\|z\| : z \in C\}.$$

(Hint: Let $d = \inf\{\|z\| : z \in C\}$, and choose a sequence $\{z_n\}$ in C such that $\lim_n \|z_n\| = d$. Note that the convexity of C implies that $\frac{1}{2}(z_m + z_n) \in C$ and hence that $\|\frac{1}{2}(z_m + z_n)\| \geq d$. Use this inequality, together with part (a) of Exercise 8, to show that $\{z_n\}$ is a Cauchy sequence. Check that $\lim_n z_n$ is the required point y. To check the uniqueness of y, suppose that y' and y'' belong to C and satisfy $\|y'\| = \|y''\| = d$, and apply the preceding argument to the sequence $y', y'', y', y'', \ldots$.)

[5] An inner product on a complex vector space V is a complex-valued function $(\cdot,\cdot)$ on $V \times V$ that satisfies (i), (ii), (iv), and $(x,y) = \overline{(y,x)}$ for all x, y, z in V and all α, β in $\mathbb{C}$. In this book we will deal with inner products only on real vector spaces.

12. Let H be a Hilbert space, and let H_0 be a closed linear subspace of H.
    (a) Show that if $x \in H$, then there is a unique point y in H_0 such that
    $$\|x-y\| = \inf\{\|x-z\| : z \in H_0\}.$$
    (Hint: Apply Exercise 11 to the set $\{x-z : z \in H_0\}$.)
    (b) Show that if x and y are related as in part (a), then $x-y$ is *orthogonal* to H_0, in the sense that $(x-y,z) = 0$ holds for each z in H_0. (Hint: Let $f(t) = \|x-y-tz\|^2 = \|x-y\|^2 - 2t(x-y,z) + t^2\|z\|^2$. Then $f(t)$ is a quadratic polynomial in t, which, by our choice of y, is minimized when $t = 0$. Conclude that $(x-y,z) = 0$.)
13. Let V be a Banach space, and let v and v_1, v_2, ... belong to V. The series $\sum_{k=1}^{\infty} v_k$ is said to *converge unconditionally* to v if for each positive ε there is a finite subset F_ε of $\mathbb{N}$ such that $\|\sum_{k\in F} v_k - v\| < \varepsilon$ holds whenever F is a finite subset of $\mathbb{N}$ that includes F_ε.
    (a) Show that if $\sum_{k=1}^{\infty} v_k$ converges absolutely, then it converges unconditionally to some point in V.
    (b) Show that the converse of part (a) holds if $V = \mathbb{R}$.
    (c) Show that the converse of part (a) is not true in general. (Hint: Let V be c_0, ℓ^2, or ℓ^∞.)
14. Let V be the vector space of all real-valued Borel measurable functions on $[0,1]$. Show that convergence in measure (with respect to Lebesgue measure) is not given by a seminorm on V. That is, show that there is no seminorm $\|\cdot\|$ on V such that elements $f, f_1, f_2, \ldots$ of V satisfy $\lim_n \|f_n - f\| = 0$ if and only if $\{f_n\}$ converges to f in measure. (Hint: Show that if such a seminorm exists, then for each positive ε there are functions $g_1, \ldots, g_n$ in V such that $\|g_i\| \le \varepsilon$ holds for each i but $\frac{1}{n}\sum_{i=1}^{n} g_i$ is equal to the constant function 1. Derive a contradiction.)
15. Again, let V be the vector space of all real-valued Borel measurable functions on $[0,1]$. Show that convergence almost everywhere (with respect to Lebesgue measure) is not given by a semimetric on V. (Hint: Use Proposition 3.1.3 to show that if such a semimetric existed, then convergence in measure would imply convergence almost everywhere.)

3.3 Definition of $\mathscr{L}^p$ and L^p

Let $(X, \mathscr{A}, \mu)$ be a measure space, and let p satisfy $1 \le p < +\infty$ (p need not be an integer). Then $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ is the set of all $\mathscr{A}$-measurable functions $f : X \to \mathbb{R}$ such that $|f|^p$ is integrable, and $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$ is the set of all $\mathscr{A}$-measurable functions $f : X \to \mathbb{C}$ such that $|f|^p$ is integrable (see Exercise 2.6.2).

Note that if $a \in \mathbb{R}$ and $f \in \mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$, then $\alpha f \in \mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$, and that if $a \in \mathbb{C}$ and $f \in \mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$, then $\alpha f \in \mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$. Furthermore, if f and g belong to $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ or to $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$, then

$$
\begin{aligned}
|f(x)+g(x)|^p &\le (|f(x)|+|g(x)|)^p \le (2\max\{|f(x)|,|g(x)|\})^p \\
&\le 2^p|f(x)|^p + 2^p|g(x)|^p
\end{aligned}
$$

holds for each x in X, and so $f+g$ belongs to $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ or to $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$. Thus $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ is a vector space over $\mathbb{R}$, and $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$ is a vector space over $\mathbb{C}$.

We turn to the definition of $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ and $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$ in the case where $p = +\infty$. Let $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{R})$ be the set of all bounded real-valued $\mathscr{A}$-measurable functions on X, and let $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{C})$ be the set of all bounded[6] complex-valued $\mathscr{A}$-measurable functions on X. It is easy to see that $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{R})$ and $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{C})$ are vector spaces.

In discussions that are valid for both real- and complex-valued functions we will often use $\mathscr{L}^p(X, \mathscr{A}, \mu)$ to represent either $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{R})$ or $\mathscr{L}^p(X, \mathscr{A}, \mu, \mathbb{C})$.

Let us define, for each p, a function (in fact, a seminorm) $\|\cdot\|_p$ on $\mathscr{L}^p(X, \mathscr{A}, \mu)$. If $1 \le p < +\infty$, we define $\|\cdot\|_p$ by

$$
\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p}.
$$

For the case where $p = +\infty$ we need a few preliminaries. A subset N of X is *locally μ-null* (or simply *locally null*) if for each set A that belongs to $\mathscr{A}$ and satisfies $\mu(A) < +\infty$ the set $A \cap N$ is μ-null. A property of points of X is said to hold *locally almost everywhere* if the set of points at which it fails to hold is locally null. It is easy to check that

(a) every μ-null subset of X is locally μ-null,
(b) if $(X, \mathscr{A}, \mu)$ is σ-finite, then every locally μ-null subset of X is μ-null, and
(c) the union of a sequence of locally μ-null sets is locally μ-null.

See Exercises 5 and 6 for some examples of locally μ-null sets that are not μ-null.

We can define $\|\cdot\|_p$ in the case where $p = +\infty$ by letting $\|f\|_\infty$ be the infimum of those nonnegative numbers M such that $\{x \in X : |f(x)| > M\}$ is locally μ-null.[7] Note that if $f \in \mathscr{L}^\infty(X, \mathscr{A}, \mu)$, then $\{x \in X : |f(x)| > \|f\|_\infty\}$ is locally μ-null, for if $\{M_n\}$ is a nonincreasing sequence of real numbers such that $\|f\|_\infty = \lim_n M_n$ and

[6] Some authors define $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{R})$ and $\mathscr{L}^\infty(X, \mathscr{A}, \mu, \mathbb{C})$ to consist of functions f that are *essentially bounded*, which means that there is a nonnegative number M such that $\{x \in X : |f(x)| > M\}$ is locally μ-null (locally null sets are defined a bit later in this section). For most purposes, it does not matter which definition of $\mathscr{L}^\infty$ one uses. However, for the study of liftings (see Appendix F), the definition given here is the more convenient one.

[7] We use locally null sets, rather than null sets, here and in the construction of the L^∞ spaces given below in order to make Proposition 3.5.5, Theorem 7.5.4, and Theorem 9.4.8 true.

such that for each n the set $\{x \in X : |f(x)| > M_n\}$ is locally μ-null, then the set $\{x \in X : |f(x)| > \|f\|_\infty\}$ is the union of the sets $\{x \in X : |f(x)| > M_n\}$ and so is locally μ-null. Thus $\|f\|_\infty$ is not only the infimum of the set of numbers M such that $\{x \in X : |f(x)| > M\}$ is locally μ-null but is itself one of those numbers.

We need to derive some standard and important inequalities in order to prove that the functions $\|\cdot\|_p$ are seminorms. Let us begin by introducing some notation. Suppose that p satisfies $1 < p < +\infty$. Then $0 < 1/p < 1$, and so there is a real number q that satisfies $1/p + 1/q = 1$; q satisfies $1 < 1/q < +\infty$. The numbers p and q are sometimes called *conjugate exponents* (see the remarks following the proof of Proposition 3.5.5). The relation $1/p + 1/q = 1$ still holds if when $p = 1$ we let $q = +\infty$, and if when $p = +\infty$ we let $q = 1$; thus we can deal with all p that satisfy $1 \leq p \leq +\infty$. Note that the relation $1/p + 1/q = 1$ implies that $p + q = pq$, and for finite p and q implies that $p = q(p-1)$ and $q = p(q-1)$.

We turn to the necessary inequalities.

Lemma 3.3.1. *Let p satisfy $1 < p < +\infty$, let q be defined by $1/p + 1/q = 1$, and let x and y be nonnegative real numbers. Then*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Proof. Since it is clear that the required inequality holds if either $x = 0$ or $y = 0$, we can assume that both x and y are positive. It is enough to prove that

$$u^{1/p} v^{1/q} \leq \frac{u}{p} + \frac{v}{q}$$

holds for all positive u and v (let $u = x^p$ and $v = y^q$), and for this it is enough to prove that

$$t^{1/p} \leq \frac{t}{p} + \frac{1}{q}$$

holds for all positive t (let $t = u/v$, and then multiply by v). However, this last inequality is easy, since according to elementary calculus the function defined for positive t by

$$t \mapsto \frac{t}{p} + \frac{1}{q} - t^{1/p}$$

has a minimum value of 0. □

Proposition 3.3.2 (Hölder's Inequality). *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let p and q satisfy $1 \leq p \leq +\infty$, $1 \leq q \leq +\infty$, and $1/p + 1/q = 1$. If $f \in \mathscr{L}^p(X, \mathscr{A}, \mu)$ and $g \in \mathscr{L}^q(X, \mathscr{A}, \mu)$, then fg belongs to $\mathscr{L}^1(X, \mathscr{A}, \mu)$ and satisfies $\int |fg|\, d\mu \leq \|f\|_p \|g\|_q$.*

Proof. First suppose that $p = 1$ and $q = +\infty$. If f and g belong to $\mathscr{L}^1(X, \mathscr{A}, \mu)$ and $\mathscr{L}^\infty(X, \mathscr{A}, \mu)$, respectively, then the set $\{x \in X : f(x) \neq 0\}$ is σ-finite under

μ (see Corollary 2.3.11) and the set $\{x \in X : |g(x)| > \|g\|_\infty\}$ is locally μ-null. The intersection of these two sets in thus a μ-null set, and so the inequality

$$|f(x)g(x)| \le \|g\|_\infty |f(x)|$$

holds at almost every x in X. It follows that $fg \in \mathscr{L}^1(X,\mathscr{A},\mu)$ and that

$$\int |fg|\,d\mu \le \int \|g\|_\infty |f|\,d\mu = \|g\|_\infty \|f\|_1.$$

A similar argument applies in case $p = +\infty$ and $q = 1$.

Now suppose that $1 < p < +\infty$ and hence that $1 < q < +\infty$. Let f belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and g belong to $\mathscr{L}^q(X,\mathscr{A},\mu)$. Lemma 3.3.1 implies that

$$|f(x)g(x)| \le \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

holds for each x in X; hence if $\|f\|_p = 1$ and $\|g\|_q = 1$, then fg belongs to $\mathscr{L}^1(X,\mathscr{A},\mu)$ and satisfies

$$\int |fg|\,d\mu \le \frac{1}{p}\int |f|^p\,d\mu + \frac{1}{q}\int |g|^q\,d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

In case neither $\|f\|_p$ nor $\|g\|_q$ is 0, we can use this inequality, with f and g replaced by $f/\|f\|_p$ and $g/\|g\|_q$, to conclude that

$$\frac{1}{\|f\|_p\|g\|_q}\int |fg|\,d\mu \le 1$$

and hence that

$$\int |fg|\,d\mu = \|f\|_p\|g\|_q. \tag{1}$$

Since inequality (1) is clear if $\|f\|_p = 0$ or $\|g\|_q = 0$ (for then fg vanishes almost everywhere), the proof is complete. □

Proposition 3.3.3 (Minkowski's Inequality). *Let $(X,\mathscr{A},\mu)$ be a measure space, and let p satisfy $1 \le p \le +\infty$. If f and g belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$, then $f+g$ belongs to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and $\|f+g\|_p \le \|f\|_p + \|g\|_p$.*

Proof. First suppose that $p = +\infty$. Define N_1 and N_2 by $N_1 = \{x \in X : |f(x)| > \|f\|_\infty\}$ and $N_2 = \{x \in X : |g(x)| > \|g\|_\infty\}$. Then N_1 and N_2 are locally μ-null, and the inequality

$$|f(x)+g(x)| \le |f(x)| + |g(x)| \le \|f\|_\infty + \|g\|_\infty$$

holds at each x outside the locally μ-null set $N_1 \cup N_2$. Thus $f+g \in \mathscr{L}^\infty(X,\mathscr{A},\mu)$ and $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

Next suppose that $p = 1$. Then the inequality $|f(x)+g(x)| \le |f(x)|+|g(x)|$ holds at each x in X, and so $f+g \in \mathscr{L}^1(X,\mathscr{A},\mu)$ and

$$\|f+g\|_1 = \int |f+g|\,d\mu \le \int |f|\,d\mu + \int |g|\,d\mu = \|f\|_1 + \|g\|_1.$$

Now consider the case where $1 < p < +\infty$. We checked that $f+g \in \mathscr{L}^p(X,\mathscr{A},\mu)$ earlier in this section. Define q by $1/p + 1/q = 1$. Since $p+q = pq$, it follows that $(|f+g|^{p-1})^q = |f+g|^p$ and hence that $|f+g|^{p-1} \in \mathscr{L}^q(X,\mathscr{A},\mu)$. Thus if we use the fact that $|f+g|^p \le (|f|+|g|)|f+g|^{p-1}$ and then apply Hölder's inequality (Proposition 3.3.2) to the functions f and $|f+g|^{p-1}$ and to the functions g and $|f+g|^{p-1}$, we can conclude that

$$\begin{aligned}\int |f+g|^p\,d\mu &\le \int (|f|+|g|)|f+g|^{p-1}\,d\mu \\ &= \int |f|\,|f+g|^{p-1}\,d\mu + \int |g|\,|f+g|^{p-1}\,d\mu \\ &\le \|f\|_p \||f+g|^{p-1}\|_q + \|g\|_p \||f+g|^{p-1}\|_q \\ &= (\|f\|_p + \|g\|_p)\left(\int |f+g|^p\,d\mu\right)^{1/q}.\end{aligned}$$

If $\int |f+g|^p\,d\mu \ne 0$, we can divide the terms of this inequality by $(\int |f+g|^p\,d\mu)^{1/q}$, obtaining

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{2}$$

Since inequality (2) is clear if $\int |f+g|^p\,d\mu = 0$, the proof is complete. □

Corollary 3.3.4. *Let $(X,\mathscr{A},\mu)$ be a measure space, and let p satisfy $1 \le p \le +\infty$. Then $\mathscr{L}^p(X,\mathscr{A},\mu)$ is a vector space, and $\|\cdot\|_p$ is a seminorm on $\mathscr{L}^p(X,\mathscr{A},\mu)$.*

Proof. We have already verified that $\mathscr{L}^p(X,\mathscr{A},\mu)$ is a vector space. The triangle inequality for $\|\cdot\|_p$ is the only other nontrivial thing to check, and it is given by Proposition 3.3.3. □

Example 3.3.5. Suppose that $1 \le p_1 < p_2 < +\infty$. Then each sequence $\{a_n\}$ that satisfies $\sum |a_n|^{p_1} < +\infty$ also satisfies $\sum |a_n|^{p_2} < +\infty$. Thus if μ is counting measure on the σ-algebra $\mathscr{A}$ of all subsets of $\mathbb{N}$, then $\mathscr{L}^{p_1}(\mathbb{N},\mathscr{A},\mu) \subseteq \mathscr{L}^{p_2}(\mathbb{N},\mathscr{A},\mu)$. The inclusion is reversed for finite measures: if μ is a finite measure on a measurable space $(X,\mathscr{A})$, then $\mathscr{L}^{p_2}(X,\mathscr{A},\mu) \subseteq \mathscr{L}^{p_1}(X,\mathscr{A},\mu)$. See Exercise 9. □

Note that if there are nonempty subsets A of X that belong to $\mathscr{A}$ and satisfy $\mu(A) = 0$, then there are nonzero functions f that belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and satisfy $\|f\|_p = 0$. Thus for many common measure spaces, the seminorms $\|\cdot\|_p$ are not norms. Since norms and metrics are often easier to deal with than are seminorms and semimetrics, the following construction of normed spaces $L^p(X,\mathscr{A},\mu)$ from the spaces $\mathscr{L}^p(X,\mathscr{A},\mu)$ proves useful.

Let $(X,\mathscr{A},\mu)$ be a measure space, and let $\mathscr{N}^p(X,\mathscr{A},\mu)$ be the subset of $\mathscr{L}^p(X,\mathscr{A},\mu)$ that consists of those functions f that belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and satisfy $\|f\|_p = 0$. Thus if $1 \le p < +\infty$, then $\mathscr{N}^p(X,\mathscr{A},\mu)$ consists of the $\mathscr{A}$-measurable functions on X that satisfy $\int |f|^p \, d\mu = 0$ (that is, that vanish almost everywhere), and if $p = +\infty$, then $\mathscr{N}^p(X,\mathscr{A},\mu)$ consists of the bounded $\mathscr{A}$-measurable functions on X that vanish locally almost everywhere. It is clear that $\mathscr{N}^p(X,\mathscr{A},\mu)$ is a linear subspace of the vector space $\mathscr{L}^p(X,\mathscr{A},\mu)$. The space $L^p(X,\mathscr{A},\mu)$ is defined to be the quotient $\mathscr{L}^p(X,\mathscr{A},\mu)/\mathscr{N}^p(X,\mathscr{A},\mu)$. Recall that this means that $L^p(X,\mathscr{A},\mu)$ is the collection of cosets of $\mathscr{N}^p(X,\mathscr{A},\mu)$ in $\mathscr{L}^p(X,\mathscr{A},\mu)$; these cosets[8] are by definition the equivalence classes induced by the equivalence relation $\sim$, where $f \sim g$ holds if and only if $f - g$ belongs to $\mathscr{N}^p(X,\mathscr{A},\mu)$. Note that if $1 \le p < +\infty$, then $f \sim g$ holds if and only if f and g are equal almost everywhere, and so $L^p(X,\mathscr{A},\mu)$ is formed by identifying functions in $\mathscr{L}^p(X,\mathscr{A},\mu)$ that agree almost everywhere. Likewise, $L^\infty(X,\mathscr{A},\mu)$ is formed by identifying functions in $\mathscr{L}^\infty(X,\mathscr{A},\mu)$ that agree locally almost everywhere.

For f in $\mathscr{L}^p(X,\mathscr{A},\mu)$ let $\langle f \rangle$ be the coset of $\mathscr{N}^p(X,\mathscr{A},\mu)$ to which f belongs. It is easy to check that the formulas $\langle f \rangle + \langle g \rangle = \langle f + g \rangle$ and $\alpha\langle f \rangle = \langle \alpha f \rangle$ define operations that make $L^p(X,\mathscr{A},\mu)$ into a vector space. Furthermore, if f and g are functions that belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and satisfy $f \sim g$, then $\|f\|_p = \|g\|_p$ (check this). Thus we can define a function, again called $\|\cdot\|_p$, on $L^p(X,\mathscr{A},\mu)$ by means of the formula $\|\langle f \rangle\|_p = \|f\|_p$. It is easy to check that $\|\cdot\|_p$ is a norm on $L^p(X,\mathscr{A},\mu)$ (see Corollary 3.3.4).

We will, of course, write $L^p(X,\mathscr{A},\mu,\mathbb{R})$ or $L^p(X,\mathscr{A},\mu,\mathbb{C})$ when the real and complex cases must be distinguished from one another.

It is often convenient to act as though the elements of $L^p(X,\mathscr{A},\mu)$ were functions, rather than equivalence classes of functions. In fact, some authors use the same symbol for $\mathscr{L}^p(X,\mathscr{A},\mu)$ and $L^p(X,\mathscr{A},\mu)$. We will try to avoid this identification of functions and classes of functions, since it can lead to confusion (especially in the study of stochastic processes). However to simplify notation we will often deal with $\mathscr{L}^p(X,\mathscr{A},\mu)$ when proving theorems about $L^p(X,\mathscr{A},\mu)$. For example, in the next section we will prove that $L^p(X,\mathscr{A},\mu)$ is complete by showing that if $\sum_k f_k$ is a series in $\mathscr{L}^p(X,\mathscr{A},\mu)$ such that $\sum_k \|f_k\|_p < +\infty$, then there is a function f in $\mathscr{L}^p(X,\mathscr{A},\mu)$ such that $\lim_n \|f - \sum_{k=1}^n f_k\|_p = 0$ (recall Proposition 3.2.5). This will imply the completeness of $L^p(X,\mathscr{A},\mu)$ and yet avoid the cumbersome notation associated with equivalence classes.

We close this section with a definition. Let $(X,\mathscr{A},\mu)$ be a measure space, let p satisfy $1 \le p < +\infty$, and let f and $f_1, f_2, \ldots$ belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$. Then $\{f_n\}$ converges to f *in pth mean* (or *in L^p-norm*) if $\lim_n \int |f_n - f|^p \, d\mu = 0$, or, equivalently, if $\lim_n \|f_n - f\|_p = 0$. There are a number of results relating

[8]Equivalently, for each f in $\mathscr{L}^p(X,\mathscr{A},\mu)$ the coset to which f belongs is the set $f + \mathscr{N}^p(X,\mathscr{A},\mu)$ and hence the set $\{f + g : g \in \mathscr{N}^p(X,\mathscr{A},\mu)\}$.

convergence in pth mean to convergence in measure and convergence almost everywhere; the reader would do well to formulate and prove some of them, using the corresponding results in Sect. 3.1 as models (see also Exercise 9).

Exercises

1. Use the inequality $(x-y)^2 \geq 0$ to give an alternate proof of Lemma 3.3.1 in the case where $p = q = 2$.
2. Give an alternate proof of Lemma 3.3.1 by noting that x^p/p and y^q/q are the areas of the shaded regions in Fig. 3.1. (The curve in Fig. 3.1 represents the graph of the equation $t = s^{p-1}$, or, equivalently, of the equation $s = t^{q-1}$.)
3. Let $(X, \mathscr{A}, \mu)$ be a measure space. Check that the formula

$$(\langle f \rangle, \langle g \rangle) = \int fg \, d\mu$$

defines an inner product on $L^2(X, \mathscr{A}, \mu, \mathbb{R})$ and that the norm associated with this inner product is the usual norm on $L^2(X, \mathscr{A}, \mu, \mathbb{R})$.
4. Let $\mathscr{B}$ be the σ-algebra of Borel subsets of $[0,1]$ and let λ be the restriction of Lebesgue measure to $\mathscr{B}$. Show that if $1 \leq p < 2$ or $2 < p \leq +\infty$ then there is no inner product on $L^p([0,1], \mathscr{B}, \lambda, \mathbb{R})$ that induces the usual norm on $L^p([0,1], \mathscr{B}, \lambda, \mathbb{R})$. (Hint: A norm that comes from an inner product must satisfy the identity in part (a) of Exercise 3.2.8.)

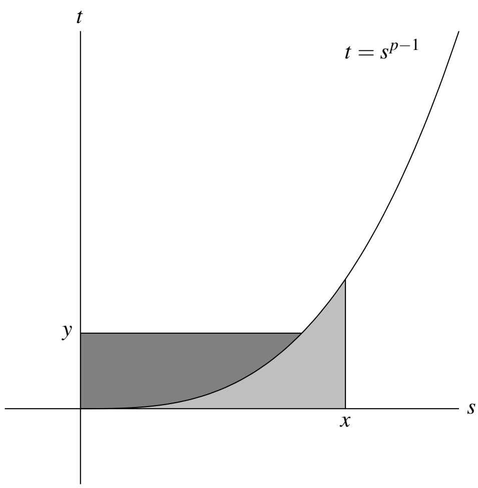

Fig. 3.1 Region used in Exercise 2 for proof of Lemma 3.3.1

5. Let X be a nonempty set, let $\mathscr{A} = \{\varnothing, X\}$, and let $\mu\colon \mathscr{A} \to [0, +\infty]$ be defined by

$$\mu(A) = \begin{cases} 0 & \text{if } A = \varnothing, \\ +\infty & \text{if } A = X. \end{cases}$$

Show that X is locally μ-null but not μ-null.

6. Suppose that for each subset A of $\mathbb{R}^2$ and each real number x we denote the set $\{y \in \mathbb{R} : (x,y) \in A\}$ by A_x. Let $\mathscr{A}$ consist of those subsets A of $\mathbb{R}^2$ that satisfy $A_x \in \mathscr{B}(\mathbb{R})$ for each x in $\mathbb{R}$, and define $\mu\colon \mathscr{A} \to [0, +\infty]$ by

$$\mu(A) = \begin{cases} \sum_x \lambda(A_x) & \text{if } A_x \neq \varnothing \text{ for only countably many } x, \\ +\infty & \text{otherwise.} \end{cases}$$

   (a) Show that $\mathscr{A}$ is a σ-algebra on $\mathbb{R}^2$ and that μ is a measure on $(\mathbb{R}^2, \mathscr{A})$.
   (b) Show that $\{(x,y) \in \mathbb{R}^2 : y = 0\}$ is locally μ-null but not μ-null.

7. Let $(X, \mathscr{A}, \mu)$ be a finite measure space, and let f be an $\mathscr{A}$-measurable real- or complex-valued function on X.

   (a) Show that f belongs to $\mathscr{L}^\infty(X, \mathscr{A}, \mu)$ if and only if

      (i) f belongs to $\mathscr{L}^p(X, \mathscr{A}, \mu)$ for each p in $[1, +\infty)$, and
      (ii) $\sup\{\|f\|_p : 1 \le p < +\infty\}$ is finite.

   (b) Show that if these conditions hold, then $\|f\|_\infty = \lim_{p \to +\infty} \|f\|_p$.

8. (Jensen's inequality.) Let $(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $(X, \mathscr{A})$ such that $\mu(X) = 1$. Suppose that $\varphi\colon \mathbb{R} \to \mathbb{R}$ is convex, in the sense that $\varphi(tx + (1-t)y) \le t\varphi(x) + (1-t)\varphi(y)$ holds for all x, y in $\mathbb{R}$ and all t in $[0,1]$.
   (a) Show that φ is continuous, and hence Borel measurable.
   (b) Show that if f belongs to $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$, then

$$\varphi\Big(\int f\,d\mu\Big) \le \int \varphi \circ f\,d\mu.$$

   In particular, the integral of $\varphi \circ f$ exists, either as a real number or as $+\infty$. (Hint: Show that for each x_0 in $\mathbb{R}$ there is a straight line (say given by the equation $y = ax + b$) that passes through the point $(x_0, \varphi(x_0))$ and never goes above the graph of $y = \varphi(x)$. Then note that for a suitable such line we have $\varphi(\int f\,d\mu) = \int (af + b)\,d\mu \le \int \varphi \circ f\,d\mu$.)

9. Let $(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $(X, \mathscr{A})$ such that $\mu(X) = 1$. Suppose that $1 \le p_1 < p_2 < +\infty$.

   (a) Show that if f belongs to $\mathscr{L}^{p_2}(X, \mathscr{A}, \mu)$, then f belongs to $\mathscr{L}^{p_1}(X, \mathscr{A}, \mu)$ and satisfies $\|f\|_{p_1} \le \|f\|_{p_2}$. (Hint: Use Hölder's inequality or Jensen's inequality.)
   (b) Show that if f and $f_1, f_2, \ldots$ belong to $\mathscr{L}^{p_2}(X, \mathscr{A}, \mu)$ and if $\{f_n\}$ converges to f in p_2th mean, then $\{f_n\}$ converges to f in p_1th mean.

3.4 Properties of $\mathscr{L}^p$ and L^p

This section is devoted to some basic properties of the L^p spaces.

Theorem 3.4.1. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let p satisfy $1 \leq p \leq +\infty$. Then $L^p(X, \mathscr{A}, \mu)$ is complete under the norm $\|\cdot\|_p$.*

Proof. According to Proposition 3.2.5, we need only show that each absolutely convergent series in $L^p(X, \mathscr{A}, \mu)$ is convergent. We do this by considering functions (not equivalence classes) in $\mathscr{L}^p(X, \mathscr{A}, \mu)$, as outlined near the end of Sect. 3.3.

First suppose that $p = +\infty$ and that $\{f_k\}$ is a sequence of functions that belong to $\mathscr{L}^\infty(X, \mathscr{A}, \mu)$ and satisfy $\sum_k \|f_k\|_\infty < +\infty$. For each positive integer k let $N_k = \{x \in X : |f_k(x)| > \|f_k\|_\infty\}$. Then the series $\sum_k f_k(x)$ converges at each x outside $\cup_k N_k$, and the function f defined by

$$f(x) = \begin{cases} \sum_k f_k(x) & \text{if } x \notin \bigcup_k N_k, \\ 0 & \text{if } x \in \bigcup_k N_k \end{cases}$$

is bounded and $\mathscr{A}$-measurable. Since $\cup_k N_k$ is locally null, the inequality

$$\left\| f - \sum_{k=1}^{n} f_k \right\|_\infty \leq \sum_{k=n+1}^{\infty} \|f_k\|_\infty$$

holds for each n, and so

$$\lim_n \left\| f - \sum_{k=1}^{n} f_k \right\|_\infty \leq \lim_n \sum_{k=n+1}^{\infty} \|f_k\|_\infty = 0.$$

Thus $L^\infty(X, \mathscr{A}, \mu)$ is complete.

Now suppose that $1 \leq p < +\infty$ and that $\{f_k\}$ is a sequence of functions that belong to $\mathscr{L}^p(X, \mathscr{A}, \mu)$ and satisfy $\sum_k \|f_k\|_p < +\infty$. Define $g \colon X \to [0, +\infty]$ by

$$g(x) = \left(\sum_{k=1}^{\infty} |f_k(x)| \right)^p$$

(of course $(+\infty)^p = +\infty$). Minkowski's inequality (Proposition 3.3.3), applied to the functions $|f_k|$, implies that

$$\left(\int \left(\sum_{k=1}^{n} |f_k| \right)^p d\mu \right)^{1/p} = \left\| \sum_{k=1}^{n} |f_k| \right\|_p \leq \sum_{k=1}^{n} \|f_k\|_p$$

holds for each n, and so it follows from the monotone convergence theorem that

$$\int g \, d\mu = \lim_n \int \left(\sum_{k=1}^{n} |f_k| \right)^p d\mu \leq \left(\sum_{k=1}^{\infty} \|f_k\|_p \right)^p;$$

thus g is integrable. Consequently $g(x)$ is finite for almost every x (Corollary 2.3.14), and the series $\sum_k f_k(x)$ is absolutely convergent, and hence convergent, for almost every x. Define a function f on X by

$$f(x) = \begin{cases} \sum_{k=1}^{\infty} f_k(x) & \text{if } g(x) < +\infty, \\ 0 & \text{otherwise.} \end{cases}$$

Then f is measurable and satisfies $|f|^p \le g$, and so it belongs to $\mathscr{L}^p(X,\mathscr{A},\mu)$. Since $\lim_n |\sum_{k=1}^n f_k(x) - f(x)| = 0$ and $|\sum_{k=1}^n f_k(x) - f(x)|^p \le g(x)$ hold for almost every x, the dominated convergence theorem implies that $\lim_n \|\sum_{k=1}^n f_k - f\|_p = 0$. The completeness of $L^p(X,\mathscr{A},\mu)$ follows. □

Let $(X,\mathscr{A},\mu)$ be a measure space. We will say that a function f in $\mathscr{L}^p(X,\mathscr{A},\mu)$ *determines* the equivalence class $\langle f\rangle$ in $L^p(X,\mathscr{A},\mu)$ to which it belongs. Likewise, if S is a subset of $\mathscr{L}^p(X,\mathscr{A},\mu)$ and if T is a subset of $L^p(X,\mathscr{A},\mu)$, then we will say that S *determines* T if T consists of the equivalence classes in $L^p(X,\mathscr{A},\mu)$ determined by the elements of S. This terminology will allow us to avoid a fair amount of pedantic notation. (See also the next-to-the-last paragraph in Sect. 3.3.)

Proposition 3.4.2. *Let $(X,\mathscr{A},\mu)$ be a measure space, and let p satisfy $1 \le p \le +\infty$. Then the simple functions in $\mathscr{L}^p(X,\mathscr{A},\mu)$ form a dense subspace of $\mathscr{L}^p(X,\mathscr{A},\mu)$ and so determine a dense subspace of $L^p(X,\mathscr{A},\mu)$.*

Proof. We will consider only real-valued functions. The corresponding results for $L^p(X,\mathscr{A},\mu,\mathbb{C})$ can be proved by separating a complex-valued function into its real and imaginary parts.

Let us first consider the case where $1 \le p < +\infty$. Let f belong to $\mathscr{L}^p(X,\mathscr{A},\mu,\mathbb{R})$. Choose nondecreasing sequences $\{g_k\}$ and $\{h_k\}$ of nonnegative simple $\mathscr{A}$-measurable functions such that $f^+ = \lim_k g_k$ and $f^- = \lim_k h_k$ (Proposition 2.1.8), and define a sequence $\{f_k\}$ by $f_k = g_k - h_k$. Then each f_k is a simple $\mathscr{A}$-measurable function that satisfies $|f_k| \le |f|$ and so belongs to $\mathscr{L}^p(X,\mathscr{A},\mu,\mathbb{R})$. Since these functions satisfy $|f_k(x) - f(x)| \le |f(x)|$ and $\lim_k |f_k(x) - f(x)| = 0$ at each x in X, the dominated convergence theorem, applied to the pth powers of the functions $|f_k - f|$, implies that $\lim_k \|f_k - f\|_p = 0$. With this the proof is complete in the case where $1 \le p < +\infty$.

Now suppose that $p = +\infty$. Let f belong to $\mathscr{L}^\infty(X,\mathscr{A},\mu,\mathbb{R})$, and let ε be a positive number. Choose real numbers $a_0, a_1, \ldots, a_n$ such that

$$a_0 < a_1 < \ldots < a_n$$

and such that the intervals $(a_{i-1},a_i]$ cover the interval $[-\|f\|_\infty, \|f\|_\infty]$ and have lengths at most ε. Let $A_i = f^{-1}((a_{i-1},a_i])$ for $i = 1, \ldots, n$, and let $f_\varepsilon = \sum_{i=1}^n a_i \chi_{A_i}$. Then f_ε is a simple $\mathscr{A}$-measurable function that satisfies $\|f - f_\varepsilon\|_\infty \le \varepsilon$. Since f and ε are arbitrary, the proof is complete. □

We now turn to some results concerning Lebesgue measure on $\mathbb{R}$. Let $[a,b]$ be a closed bounded interval. A real- or complex-valued function f on $[a,b]$ is a *step function* if there are real numbers $a_0, \ldots, a_n$ such that

(a) $a = a_0 < a_1 < \ldots < a_n = b$, and
(b) f is constant on each interval (a_{i-1}, a_i).

We will use $\mathscr{L}^p([a,b])$ and $L^p([a,b])$ as abbreviations for $\mathscr{L}^p([a,b], \mathscr{B}([a,b]), \lambda)$ and $L^p([a,b], \mathscr{B}([a,b]), \lambda)$, where $\mathscr{B}([a,b])$ is the σ-algebra of Borel subsets of $[a,b]$ and λ is the restriction of Lebesgue measure to $\mathscr{A}$.

The following two propositions are often useful, since step functions and continuous functions are usually easier to deal with than are more general functions.

Proposition 3.4.3. *Suppose that $[a,b]$ is a closed bounded interval and that p satisfies $1 \le p < +\infty$. Then the subspace of $L^p([a,b])$ determined by the step functions on $[a,b]$ is dense in $L^p([a,b])$.*

Proof. Of course, each step function on $[a,b]$ belongs to $\mathscr{L}^p([a,b])$. Since the Borel measurable simple functions on $[a,b]$ determine a dense subspace of $L^p([a,b])$ (Proposition 3.4.2), it is enough to show that if f is a Borel measurable simple function and if ε is a positive number, then there is a step function g such that $\|f - g\|_p < \varepsilon$, and for this it is enough to check that if χ_A is the characteristic function of a Borel subset A of $[a,b]$, then there are step functions g that make $\|\chi_A - g\|_p$ arbitrarily small. So let A be a Borel subset of $[a,b]$ and let δ be a positive number. Use the construction of Lebesgue outer measure (or Proposition C.4 and the regularity of Lebesgue measure) to choose a sequence $\{(a_n, b_n)\}$ of open intervals such that $A \subseteq \cup_{n=1}^{\infty}(a_n, b_n)$ and $\sum_{n=1}^{\infty}(b_n - a_n) < \lambda(A) + \delta$, and then choose a positive integer N such that $\sum_{n=N+1}^{\infty}(b_n - a_n) < \delta$. Let g be the characteristic function of $[a,b] \cap (\cup_1^N (a_n, b_n))$ and let h be the characteristic function of $[a,b] \cap (\cup_1^{\infty}(a_n, b_n))$. Then g is a step function, and

$$\begin{aligned}
\|\chi_A - g\|_p &\le \|\chi_A - h\|_p + \|h - g\|_p \\
&\le \left(\lambda(\bigcup_1^{\infty}(a_n, b_n) - A)\right)^{1/p} + \left(\lambda\left(\bigcup_{N+1}^{\infty}(a_n, b_n)\right)\right)^{1/p} \\
&< \delta^{1/p} + \delta^{1/p} = 2\delta^{1/p}.
\end{aligned}$$

Since δ is arbitrary, the proof is complete. □

Proposition 3.4.4. *Suppose that $[a,b]$ is a closed bounded interval and that p satisfies $1 \le p < +\infty$. Then the subspace of $L^p([a,b])$ determined by the continuous functions on $[a,b]$ is dense in $L^p([a,b])$.*

Proof. Of course, each continuous function on $[a,b]$ belongs to $\mathscr{L}^p([a,b])$. Since the step functions on $[a,b]$ determine a dense subspace of $L^p([a,b])$ (Proposition 3.4.3), it is enough to prove that for each step function f and each positive number ε there is a continuous function g that satisfies $\|f - g\|_p < \varepsilon$. So let f be a step function on $[a,b]$, let $M = \sup\{|f(x)| : x \in [a,b]\}$, and let δ be a positive number. It is easy to construct a continuous function g on $[a,b]$ such that $\sup\{|g(x)| : x \in [a,b]\} \le M$ and $\lambda(\{x \in [a,b] : f(x) \ne g(x)\}) < \delta$ (take a suitable piecewise linear function). Then

$$\int_a^b |f-g|^p \, d\lambda \le (2M)^p \lambda(\{x : f(x) \ne g(x)\}) \le (2M)^p \delta,$$

and so $\|f-g\|_p \le 2M\delta^{1/p}$. Since δ is arbitrary and M depends only on f, the proof is complete. □

The reader should note that Propositions 3.4.3 and 3.4.4 would fail if p were allowed to be infinite (see Exercises 3 and 4).

Let us call a function on $\mathbb{R}$ a *step function* if for each interval $[a,b]$ its restriction to $[a,b]$ is a step function. Analogues of Propositions 3.4.3 and 3.4.4 hold for $L^p(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ if we replace the set of step functions on $[a,b]$ with the set of step functions on $\mathbb{R}$ that vanish outside some bounded interval and if we replace the set of continuous functions on $[a,b]$ with the set of continuous functions on $\mathbb{R}$ that vanish outside some bounded interval. The details are left to the reader. (See also Proposition 7.4.3.)

Let $\mathscr{A}$ be a σ-algebra on the set X. Then $\mathscr{A}$ is *countably generated* if there is a countable subfamily $\mathscr{C}$ of $\mathscr{A}$ such that $\mathscr{A} = \sigma(\mathscr{C})$. For example, the σ-algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is countably generated (see Exercise 1.1.2).

Proposition 3.4.5. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let p satisfy $1 \le p < +\infty$. If μ is σ-finite and $\mathscr{A}$ is countably generated, then $L^p(X, \mathscr{A}, \mu)$ is separable.*

The proof will depend on the following two lemmas.

Lemma 3.4.6. *Let $(X, \mathscr{A}, \mu)$ be a finite measure space, and let $\mathscr{A}_0$ be an algebra of subsets of X such that $\mathscr{A} = \sigma(\mathscr{A}_0)$. Then $\mathscr{A}_0$ is dense in $\mathscr{A}$, in the sense that for each A in $\mathscr{A}$ and each positive number ε there is a set A_0 that belongs to $\mathscr{A}_0$ and satisfies $\mu(A \triangle A_0) < \varepsilon$.*

Proof. Let $\mathscr{F}$ be the collection consisting of those sets A in $\mathscr{A}$ such that for each positive ε there is a set A_0 that belongs to $\mathscr{A}_0$ and satisfies $\mu(A \triangle A_0) < \varepsilon$. Of course $\mathscr{A}_0 \subseteq \mathscr{F}$, and so $X \in \mathscr{F}$. The identity $A^c \triangle A_0^c = A \triangle A_0$ implies that if $A \in \mathscr{F}$, then $A^c \in \mathscr{F}$; hence $\mathscr{F}$ is closed under complementation. Now let $\{A_n\}$ be an infinite sequence of sets in $\mathscr{F}$, let $A = \cup_n A_n$, and let ε be a positive number. Choose a positive integer N such that $\mu(A - \cup_1^N A_n) < \varepsilon/2$ (see Proposition 1.2.5), and for $n = 1, 2, \ldots, N$ choose a set B_n that belongs to $\mathscr{A}_0$ and satisfies $\mu(A_n \triangle B_n) < \varepsilon/2N$. The set B defined by $B = \cup_1^N B_n$ then belongs to $\mathscr{A}_0$ and satisfies

$$\begin{aligned}\mu(A \triangle B) &\le \mu\left(A \triangle \left(\bigcup_1^N A_n\right)\right) + \mu\left(\left(\bigcup_1^N A_n\right) \triangle B\right) \\ &\le \mu\left(A \triangle \left(\bigcup_1^N A_n\right)\right) + \sum_1^N \mu(A_n \triangle B_n) \\ &< \frac{\varepsilon}{2} + \sum_1^N \frac{\varepsilon}{2N} = \varepsilon.\end{aligned}$$

Since we can produce such a set B for each positive ε, it follows that $A \in \mathscr{F}$. Consequently $\mathscr{F}$ is closed under the formation of countable unions and so is a σ-algebra. Since in addition $\mathscr{A}_0 \subseteq \mathscr{F} \subseteq \mathscr{A} = \sigma(\mathscr{A}_0)$, $\mathscr{F}$ must be equal to $\mathscr{A}$. Thus every set in $\mathscr{A}$ can be approximated with sets in $\mathscr{A}_0$. □

Lemma 3.4.7. *Let $(X, \mathscr{A}, \mu)$ be a measure space. Suppose that $\mathscr{A}_0$ is an algebra of subsets of X such that*

(a) *$\sigma(\mathscr{A}_0) = \mathscr{A}$, and*
(b) *X is the union of a sequence of sets that belong to $\mathscr{A}_0$ and have finite measure under μ.*

Then for each positive ε and each set A that belongs to $\mathscr{A}$ and satisfies $\mu(A) < +\infty$ there is a set A_0 that belongs to $\mathscr{A}_0$ and satisfies $\mu(A \triangle A_0) < \varepsilon$.

Proof. Choose a sequence $\{B_n\}$ of sets that belong to $\mathscr{A}_0$, have finite measure under μ, and satisfy $X = \cup_n B_n$. By replacing B_n with $\cup_{k=1}^n B_k$, we can assume that the sequence $\{B_n\}$ is increasing.

Let A belong to $\mathscr{A}$ and satisfy $\mu(A) < +\infty$. Proposition 1.2.5, applied to the sequence $\{A \cap B_n\}$, implies that there is a positive integer N such that $\mu(A \cap B_N) > \mu(A) - \varepsilon/2$. Since the measure $C \mapsto \mu(C \cap B_N)$ is finite, we can use Lemma 3.4.6 to obtain a set A_0 that belongs to $\mathscr{A}_0$ and satisfies $\mu((A \triangle A_0) \cap B_N) < \varepsilon/2$. Then $A_0 \cap B_N$ belongs to $\mathscr{A}_0$ and satisfies

$$\begin{aligned}\mu(A \triangle (A_0 \cap B_N)) &\le \mu(A \triangle (A \cap B_N)) + \mu((A \cap B_N) \triangle (A_0 \cap B_N)) \\ &= \mu(A - (A \cap B_N)) + \mu((A \triangle A_0) \cap B_N) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\end{aligned}$$

and the proof of the lemma is complete. □

Proof of Proposition 3.4.5. We can choose a countable subfamily $\mathscr{C}$ of $\mathscr{A}$ that generates $\mathscr{A}$ and contains sets B_n, $n = 1, 2, \ldots$, that have finite measure under μ and satisfy $X = \cup_n B_n$. Let $\mathscr{C}^+$ consist of the sets in $\mathscr{C}$, together with their complements, and let $\mathscr{A}_0$ be the algebra (not the σ-algebra) generated by $\mathscr{C}$. Then $\mathscr{A}_0$ is the set of finite unions of sets that have the form

$$C_1 \cap C_2 \cap \cdots \cap C_N$$

for some choice of N and some choice of sets $C_1, \ldots, C_N$ in $\mathscr{C}^+$. Clearly $\mathscr{A}_0$ is countable and satisfies the hypotheses of Lemma 3.4.7.

Let $\mathscr{S}$ be the collection of all finite sums

$$\sum_j d_j \chi_{D_j},$$

where each d_j is a rational number[9] and each D_j belongs to $\mathscr{A}_0$ and satisfies $\mu(D_j) < +\infty$. The set $\mathscr{S}$ is countable and is included in $\mathscr{L}^p(X,\mathscr{A},\mu)$; we will show that it determines a dense subset of $L^p(X,\mathscr{A},\mu)$.

Let f belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$, and let ε be a positive number. Then there is a simple function g that belongs to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and satisfies $\|f-g\|_p < \varepsilon$ (Proposition 3.4.2). Suppose that the simple function g has the form $\sum_j a_j \chi_{A_j}$, where each A_j belongs to $\mathscr{A}$ and satisfies $\mu(A_j) < +\infty$. We can choose rational numbers d_j such that

$$\left\| \sum_j a_j \chi_{A_j} - \sum_j d_j \chi_{A_j} \right\|_p \le \sum_j |a_j - d_j| \|\chi_{A_j}\|_p < \varepsilon,$$

and then we can produce sets D_j in $\mathscr{A}_0$ such that $\|\sum_j d_j \chi_{A_j} - \sum_j d_j \chi_{D_j}\|_p < \varepsilon$ (use Lemma 3.4.7). Since f and ε are arbitrary, and since $\sum_j d_j \chi_{D_j}$ belongs to $\mathscr{S}$ and satisfies

$$\begin{aligned} \left\| f - \sum_j d_j \chi_{D_j} \right\|_p \le \|f-g\|_p + \left\| g - \sum_j d_j \chi_{A_j} \right\|_p \\ + \left\| \sum_j d_j \chi_{A_j} - \sum_j d_j \chi_{D_j} \right\|_p < 3\varepsilon, \end{aligned}$$

the proof is complete. □

Exercises

1. Use Proposition 3.4.3 to show that if $1 \le p < +\infty$, then $L^p([a,b])$ is separable.
2. Show that $L^\infty([a,b])$ is not separable. (Hint: Consider the elements of $L^\infty([a,b])$ determined by the characteristic functions of the sets $[a,c]$, where $a < c < b$.)

[9] When dealing with the complex L^p spaces, let each d_j be a complex number whose real and imaginary parts are rational.

3. Show that Proposition 3.4.3 would be false if p were allowed to be infinite. (Hint: Construct a Borel subset A of $[a,b]$ such that $\|\chi_A - f\|_\infty \geq 1/2$ holds whenever f is a step function.)
4. Show that Proposition 3.4.4 would be false if p were allowed to be infinite. (Hint: Let $A = [a,c]$, where $a < c < b$. How small can $\|\chi_A - f\|_\infty$ be if f is continuous?)
5. Suppose that for each function $f\colon \mathbb{R} \to \mathbb{R}$ and each x in $\mathbb{R}$ we define a function $f_x\colon \mathbb{R} \to \mathbb{R}$ by $f_x(t) = f(t-x)$. (A similar definition applies to complex-valued functions on $\mathbb{R}$.) Show that if $1 \leq p < +\infty$ and if f belongs to $\mathscr{L}^p(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$, then

$$\lim_{x \to x_0} \|f_x - f_{x_0}\|_p = 0$$

holds for each x_0 in $\mathbb{R}$. (Hint: First, consider the case where f is a step function that vanishes outside some bounded interval. Then use Proposition 3.4.3 (see the remarks following the proof of Proposition 3.4.4).)
6. Show that the hypothesis of σ-finiteness cannot be omitted in Proposition 3.4.5. (Hint: Consider counting measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.)
7. Show that in Lemma 3.4.7 condition (b) cannot be replaced with the assumption that μ is σ-finite. (Hint: Let $\mathscr{A}_0$ be the algebra on $\mathbb{R}$ defined in Example 1.1.1(g), let $\{r_n\}$ be an enumeration of $\mathbb{Q}$, and let μ be the Borel measure on $\mathbb{R}$ defined by $\mu = \sum_n \delta_{r_n}$.)

3.5 Dual Spaces

Recall that if V_1 and V_2 are vector spaces over $\mathbb{R}$ (or over $\mathbb{C}$), then a function $T\colon V_1 \to V_2$ is a *linear operator* (or *linear transformation*) if for each v and w in V_1 and each α in $\mathbb{R}$ (or in $\mathbb{C}$) it satisfies $T(v+w) = T(v) + T(w)$ and $T(\alpha v) = \alpha T(v)$. Recall also that if S_1 and S_2 are metric spaces, say with metrics d_1 and d_2, then a function $f\colon S_1 \to S_2$ is *continuous* if for each point a in S_1 and each positive number ε there is a positive number δ such that $d_2(f(s), f(a)) < \varepsilon$ holds whenever s belongs to S_1 and satisfies $d_1(s,a) < \delta$. Thus if V_1 and V_2 are normed linear spaces, say with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, then a function $f\colon V_1 \to V_2$ is continuous if and only if for each a in V_1 and each positive number ε there is a positive number δ such that $\|f(v) - f(a)\|_2 < \varepsilon$ holds whenever v belongs to V_1 and satisfies $\|v - a\|_1 < \delta$.

When dealing with several normed spaces, we will often use the symbol $\|\cdot\|$ to denote each of the norms involved. This will of course be done only when there seems to be little chance of confusion.

Proposition 3.5.1. *Let V_1 and V_2 be normed linear spaces, and let $T\colon V_1 \to V_2$ be a linear operator. Then T is continuous if and only if there is a nonnegative number M such that*

$$\|T(v)\| \leq M\|v\| \tag{1}$$

holds for each v in V_1.

Proof. First suppose that there is a nonnegative number M such that inequality (1) holds for each v in V_1. Then for each v and a in V_1 we have

$$\|T(v) - T(a)\| = \|T(v - a)\| \leq M\|v - a\|;$$

hence if ε is a positive number and if we define δ by $\delta = \varepsilon/M$ (let δ be an arbitrary positive number if $M = 0$), then $\|T(v) - T(a)\| < \varepsilon$ holds whenever $\|v - a\| < \delta$. Thus T is continuous.

Now suppose that T is continuous, and choose a positive number δ such that $\|T(v)\| = \|T(v) - T(0)\| < 1$ if $\|v\| = \|v - 0\| < \delta$. Note that (1) holds if $v = 0$, whatever value we use for M. Now suppose that $v \neq 0$ and let $w = v/\|v\|$. It follows that if $0 < t < \delta$, then we have $\|tw\| < \delta$ and $t\|T(w)\| < 1$, from which we get

$$\|T(v)\| < \frac{1}{t}\|v\|.$$

Since t can be chosen arbitrarily close to $1/\delta$, it follows that $\|T(v)\| \leq \frac{1}{\delta}\|v\|$. Thus inequality (1) holds, with M equal to $1/\delta$. □

Let V_1 and V_2 be normed linear spaces, and let $T\colon V_1 \to V_2$ be linear. A nonnegative number A such that $\|T(v)\| \leq A\|v\|$ holds for each v in V_1 is called a *bound* for T, and the operator T is called *bounded* if there is a bound for it (see also Exercises 3 and 4). Thus Proposition 3.5.1 says that a linear operator is continuous if and only if it is bounded. It is easy to check that if the operator T is bounded, then the infimum of the set of bounds for T is a bound for T. This smallest bound for T is called the *norm* of T and is written $\|T\|$. It is not hard to check that $\|\cdot\|$ is a norm on the vector space of all bounded linear operators from V_1 to V_2.

We turn to a few special cases. Suppose that V_1 and V_2 are normed linear spaces and that $T\colon V_1 \to V_2$ is linear. Then T is an *isometry* if $\|T(v)\| = \|v\|$ holds for each v in V_1. Note that if T is an isometry and if v and w belong to V_1, then

$$\|T(v) - T(w)\| = \|T(v - w)\| = \|v - w\|,$$

and so T preserves distances. The linear operator T is an *isometric isomorphism* if it is an isometry that is surjective (note that an isometry is necessarily injective and so is bijective if and only if it is surjective). Thus an isometric isomorphism is a bijection that preserves both linear and metric structure.

Let V be a normed linear space. Recall that a *linear functional* on V is a linear operator on V whose values lie in $\mathbb{R}$ (if V is a vector space over $\mathbb{R}$) or in $\mathbb{C}$ (if V is a vector space over $\mathbb{C}$). We will be particularly concerned with the bounded, that is, continuous, linear functionals on V. It is easy to check that the set of all continuous linear functionals on V is a subspace of the vector space of all linear functionals on V; this subspace is called the *dual space* (or *conjugate space*) of V and is denoted by V^*. The space V^* is sometimes called the *topological dual space* of V in order to distinguish it from the space of *all* linear functionals on V (which is then called the *algebraic dual space* of V).

Note that the function $\|\cdot\|: V^* \to \mathbb{R}$ that assigns to each functional in V^* its norm (as defined above) is in fact a norm on the vector space V^*; for instance, the calculation

$$|(F+G)(v)| \le |F(v)| + |G(v)| \le \|F\|\|v\| + \|G\|\|v\| = (\|F\| + \|G\|)\|v\|$$

shows that $\|F\| + \|G\|$ is a bound for $F+G$ and so implies that $\|F+G\| \le \|F\| + \|G\|$.

Example 3.5.2. Let $[a,b]$ be a closed bounded subinterval of $\mathbb{R}$, and let μ be a finite Borel measure on $[a,b]$. Define $F: C[a,b] \to \mathbb{R}$ by letting

$$F(f) = \int f\,d\mu \qquad (2)$$

hold for each f in $C[a,b]$. It is clear that F is a linear functional and that F is *positive*, in the sense that each nonnegative[10] f in $C[a,b]$ satisfies $F(f) \ge 0$. We will see that every positive linear functional on $C[a,b]$ arises in this way (Theorem 7.2.8). □

Example 3.5.3. Now suppose that $C[a,b]$ is given the norm $\|\cdot\|_\infty$ defined by

$$\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$$

(see Example 3.2.1(e) above). Then the functional F defined by (2) satisfies

$$|F(f)| = \left|\int f\,d\mu\right| \le \int |f|\,d\mu \le \|f\|_\infty \mu([a,b]),$$

and so is bounded and hence continuous. Likewise, if μ_1 and μ_2 are finite Borel measures on $[a,b]$, then the linear functional G defined by

$$G(f) = \int f\,d\mu_1 - \int f\,d\mu_2$$

is continuous. We will see that every continuous linear functional on $C[a,b]$ arises in this way (Theorem 7.3.6). These facts and their generalizations form the basis for many of the applications of measure theory.[11] □

Example 3.5.4. Suppose that $(X, \mathscr{A}, \mu)$ is an arbitrary measure space, that p satisfies $1 \le p < +\infty$, and that q is defined by $1/p + 1/q = 1$. Let g belong

[10] The function f is called nonnegative if $f(x) \ge 0$ holds at each x in $[a,b]$.

[11] The usefulness of these results seems to be attributable to two facts:

(a) If a linear functional on a space of functions can be represented as an integral, then the limit theorems of Sect. 2.4 are applicable.
(b) The methods available for decomposing and analyzing measures are often easier to visualize than those that apply directly to linear functionals.

to $\mathscr{L}^q(X,\mathscr{A},\mu)$. Then fg is integrable whenever f belongs to $\mathscr{L}^p(X,\mathscr{A},\mu)$ (Proposition 3.3.2), and so the formula

$$T_g(f) = \int fg\,d\mu$$

defines a linear functional T_g on $\mathscr{L}^p(X,\mathscr{A},\mu)$. It is clear that if f_1 and f_2 belong to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and agree almost everywhere, then $T_g(f_1) = T_g(f_2)$; thus we can use the formula $T_g(\langle f\rangle) = T_g(f)$ to define a functional, also called T_g, on $L^p(X,\mathscr{A},\mu)$. Hölder's inequality (Proposition 3.3.2) implies that $|T_g(f)| \le \|g\|_q\|f\|_p$ holds for each f in $\mathscr{L}^p(X,\mathscr{A},\mu)$. Thus T_g is continuous on $L^p(X,\mathscr{A},\mu)$, and $\|T_g\| \le \|g\|_q$. We'll see in the following proposition that $\|T_g\| = \|g\|_q$. □

We will denote by T the map from $\mathscr{L}^q(X,\mathscr{A},\mu)$ to $(L^p(X,\mathscr{A},\mu))^*$ that takes the function g to the functional T_g defined above.

Proposition 3.5.5. *Let $(X,\mathscr{A},\mu)$ be a measure space, let p satisfy $1 \le p < +\infty$, and let q be defined by $1/p + 1/q = 1$. Then the map $T\colon \mathscr{L}^q(X,\mathscr{A},\mu) \to (L^p(X,\mathscr{A},\mu))^*$ defined above induces an isometry of $L^q(X,\mathscr{A},\mu)$ into $(L^p(X,\mathscr{A},\mu))^*$.*

Note that Proposition 3.5.5 says that T is an isometry *into* $(L^p(X,\mathscr{A},\mu))^*$; it does not say that T is surjective. Example 4.5.2 in the next chapter gives a case in which T is not surjective. Later we will see that the map T is a surjection, and hence an isometric isomorphism, if

(a) $1 < p < +\infty$ and $(X,\mathscr{A},\mu)$ is arbitrary,
(b) $p = 1$ and μ is σ-finite, or
(c) $p = 1$ and $(X,\mathscr{A},\mu)$ arises through certain topological constructions

(see Theorems 4.5.1, 7.5.4, and 9.4.8). It is because of this relationship between $L^q(X,\mathscr{A},\mu)$ and $(L^p(X,\mathscr{A},\mu))^*$ that numbers p and q satisfying $1/p + 1/q = 1$ are called conjugate exponents.

We need a bit of notation for the proof of Proposition 3.5.5. Recall that if z is a complex number, say $z = x + iy$, then $\bar{z}$ (the *complex conjugate* of z) and $\operatorname{sgn}(z)$ are defined by $\bar{z} = x - iy$ and

$$\operatorname{sgn}(z) = \begin{cases} \frac{z}{|z|} & \text{if } z \ne 0 \\ 0 & \text{if } z = 0. \end{cases}$$

It is easy to check that $z\bar{z} = |z|^2$ and $z\,\overline{\operatorname{sgn}(z)} = |z|$ hold for each z and that $|\operatorname{sgn}(z)| = 1$ holds for each nonzero z. If f is a complex-valued function on a set S, then $\bar{f}$ and $\operatorname{sgn}(f)$ are the functions whose values at the point s are $\overline{f(s)}$ and $\operatorname{sgn}(f(s))$.

In the following proof we will assume that the functions involved are complex-valued. The details are essentially the same for real-valued functions (then $\bar{z} = z$ and $\operatorname{sgn}(z)$ is 1, 0, or -1).

Proof of Proposition 3.5.5. It is clear that if g_1 and g_2 are equal almost everywhere (or, in case $q = +\infty$, locally almost everywhere), then $T_{g_1} = T_{g_2}$. Thus T_g depends on g only through the equivalence class $\langle g \rangle$ to which g belongs, and we can define a map, again called T, from $L^q(X,\mathscr{A},\mu)$ to $(L^p(X,\mathscr{A},\mu))^*$ by means of the formula $T_{\langle g \rangle} = T_g$. It is clear that T is linear. Since we have already seen that $\|T_g\| \le \|g\|_q$ holds for each g in $\mathscr{L}^q(X,\mathscr{A},\mu)$, we need only verify the reverse inequality. Let us consider two cases.

First suppose that $p = 1$ and hence that $q = +\infty$. Let g be an element of $\mathscr{L}^\infty(X,\mathscr{A},\mu)$ such that $\|g\|_\infty \neq 0$, and let ε be a positive number. Since $\{x \in X: |g(x)| > \|g\|_\infty - \varepsilon\}$ is not locally μ-null,[12] there is a set A that belongs to $\mathscr{A}$, has finite measure under μ, and is such that the set B defined by

$$B = A \cap \{x \in X : |g(x)| > \|g\|_\infty - \varepsilon\}$$

has nonzero measure. Let $f = \overline{\operatorname{sgn}(g)}\chi_B$. Then $f \in \mathscr{L}^1(X,\mathscr{A},\mu)$,

$$\|f\|_1 = \int |\overline{\operatorname{sgn}(g)}\chi_B|\, d\mu \le \int \chi_B \, d\mu = \mu(B),$$

and

$$T_g(f) = \int g\,\overline{\operatorname{sgn}(g)}\chi_B \, d\mu = \int |g|\chi_B \, d\mu \ge (\|g\|_\infty - \varepsilon)\mu(B).$$

It is clear that $|T_g(f)| = T_g(f)$, and so the preceding calculations, together with the inequality $|T_g(f)| \le \|T_g\|\|f\|_1$, imply that $\|g\|_\infty - \varepsilon \le \|T_g\|$. Since ε can be made arbitrarily close to 0, it follows that $\|g\|_\infty \le \|T_g\|$. Thus $\|T_g\| = \|g\|_\infty$.

Now suppose that $1 < p < +\infty$ and hence that $1 < q < +\infty$. Let g belong to $\mathscr{L}^q(X,\mathscr{A},\mu)$, and define a function f by $f = \overline{\operatorname{sgn}(g)}|g|^{q-1}$. The relation $q = p(q-1)$ implies that $|f|^p = |g|^q$; thus f belongs to $\mathscr{L}^p(X,\mathscr{A},\mu)$ and satisfies $\|f\|_p = (\int |g|^q \, d\mu)^{1/p}$. Furthermore

$$T_g(f) = \int \overline{\operatorname{sgn}(g)}|g|^{q-1} g \, d\mu = \int |g|^q \, d\mu.$$

Consequently it follows from the relation $|T_g(f)| \le \|T_g\|\|f\|_p$ that

$$\int |g|^q \, d\mu \le \|T_g\| \Big(\int |g|^q \, d\mu\Big)^{1/p} \tag{3}$$

and hence that $\|g\|_q \le \|T_g\|$ (this is clear if $\|g\|_q = 0$; otherwise divide both sides of (3) by $(\int |g|^q \, d\mu)^{1/p}$ and recall that $1 - 1/p = 1/q$.) Thus $\|T_g\| = \|g\|_q$, and the proof is complete. □

---

[12] We are here assuming that the space X is not locally null. If X is locally null, then $L^1(X,\mathscr{A},\mu)$ and $L^\infty(X,\mathscr{A},\mu)$ contain only 0, and the proposition is true (but uninteresting).

Exercises

1. Let V_1, V_2, and V_3 be normed linear spaces, and let $S\colon V_1 \to V_2$ and $T\colon V_2 \to V_3$ be bounded linear operators. Show that $T \circ S\colon V_1 \to V_3$ is bounded and that $\|T \circ S\| \le \|T\|\|S\|$.
2. Suppose that V_1 and V_2 are normed linear spaces and that $T\colon V_1 \to V_2$ is an invertible linear operator such that T and T^{-1} are both bounded.
   (a) Show that $1 \le \|T\|\|T^{-1}\|$. (Hint: See Exercise 1.)
   (b) Show by example that equality need not hold in part (a).
3. Let V_1 and V_2 be normed linear spaces, and let $T\colon V_1 \to V_2$ be a linear operator. Show that the subset $T(V_1)$ of V_2 is bounded if and only if T is the zero operator. Thus to say that a linear operator is bounded is *not* to say that its values form a bounded set.
4. Let V_1 and V_2 be normed linear spaces, and let $T\colon V_1 \to V_2$ be a linear operator.
   (a) Show that T is bounded if and only if the set

   $$\{\|T(v)\| : v \in V_1 \text{ and } \|v\| \le 1\}$$

   is bounded above.
   (b) Show that if T is bounded, then

   $$\|T\| = \sup\{\|T(v)\| : v \in V_1 \text{ and } \|v\| \le 1\}.$$

5. Suppose that V_1 and V_2 are normed linear spaces and that $T\colon V_1 \to V_2$ is a linear operator. Show that if T is bounded, then T is uniformly continuous.
6. Let V be a normed linear space. Show that the dual V^* of V is complete under the norm $\|\cdot\|$ defined above. (Hint: Let $\{F_n\}$ be a Cauchy sequence in V^*. Show that for each v in V the sequence $\{F_n(v)\}$ is a Cauchy sequence in $\mathbb{R}$ (or in $\mathbb{C}$) and so is convergent. Then show that the formula $F(v) = \lim_n F_n(v)$ defines a bounded linear functional on V and that $\lim_n \|F_n - F\| = 0$.)
7. Let V be an inner product space, and for each y in V define $F_y\colon V \to \mathbb{R}$ by $F_y(x) = (x, y)$.
   (a) Show that F_y belongs to V^* and satisfies $\|F_y\| = \|y\|$. (Hint: Use the Cauchy–Schwarz inequality; see Exercise 3.2.7. To check that $\|F_y\|$ is equal to (rather than less than) $\|y\|$, consider $F_y(y)$.)
   (b) Show that if $y \ne y'$, then $F_y \ne F_{y'}$.
   (c) Show that if the inner product space V is a Hilbert space and if F belongs to V^*, then there is an element y of V such that $F = F_y$. (Hint: Let $y = 0$ if $F = 0$. Otherwise choose a nonzero element v of V such that $(u, v) = 0$ holds whenever $F(u) = 0$ (see Exercise 3.2.12), and check that a suitable multiple of v works.)
8. (This exercise depends on the Hahn–Banach theorem, which is stated without proof in Appendix E.) Let V be the subspace of ℓ^∞ consisting of those sequences $\{x_n\}$ for which $\lim_n x_n$ exists, and let $F_0\colon V \to \mathbb{R}$ be defined by $F_0(\{x_n\}) = \lim_n x_n$.

(a) Show that F_0 is a bounded linear functional on V and that $\|F_0\| = 1$.
(b) Let F be a bounded linear functional on ℓ^∞ that satisfies $\|F\| = 1$ and agrees with F_0 on V (see Theorem E.7). Show that if $\{x_n\}$ is a nonnegative element of ℓ^∞ (that is, if $\{x_n\}$ belongs to ℓ^∞ and satisfies $x_n \geq 0$ for each n), then $F(\{x_n\}) \geq 0$. (Hint: Consider the sequence $\{x'_n\}$ defined by $x'_n = x_n - c$, where c is a suitably chosen constant.)
(c) For each subset A of $\mathbb{N}$ let $\{\chi_{A,n}\}_{n=1}^\infty$ be the sequence defined by

$$\chi_{A,n} = \begin{cases} 1 & \text{if } n \in A, \\ 0 & \text{if } n \notin A. \end{cases}$$

Show that the function $\mu \colon \mathscr{P}(\mathbb{N}) \to \mathbb{R}$ defined by $\mu(A) = F(\{\chi_{A,n}\})$ is a finitely additive measure, but is not countably additive.

Notes

Kolmogorov and Fomin [73] and Simmons [109] are useful elementary sources of information on metric spaces and normed linear spaces. The basic properties of the L^p-spaces can be found in virtually every book on integration theory.

Chapter 4
Signed and Complex Measures

In this chapter we study signed and complex measures, which are defined to be the countably additive functions from a σ-algebra to $[-\infty, +\infty]$ or to $\mathbb{C}$ that have value 0 on the empty set. We begin in Sect. 4.1 with some basic definitions and facts. Section 4.2 is devoted to the main result of this chapter, the Radon–Nikodym theorem. Let μ be a σ-finite positive measure on a measurable space $(X, \mathscr{A})$. The Radon–Nikodym theorem characterizes those finite positive, signed, or complex measures ν whose values can be computed by integrating some μ-integrable function—in other words, it characterizes those ν for which there is a μ-integrable f such that $\nu(A) = \int_A f \, d\mu$ holds for all A in $\mathscr{A}$. The last part of the chapter is devoted to the relation of the material in the early parts of the chapter to the classical concepts of bounded variation and absolute continuity (Sect. 4.4) and to the use of the Radon–Nikodym theorem to compute the dual spaces of a number of the L^p spaces (Sect. 4.5).

4.1 Signed and Complex Measures

Let $(X, \mathscr{A})$ be a measurable space, and let μ be a function on $\mathscr{A}$ with values in $[-\infty, +\infty]$. The function μ is *finitely additive* if the identity

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

holds for each finite sequence $\{A_i\}_{i=1}^{n}$ of disjoint sets in $\mathscr{A}$ and is *countably additive* if the identity

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

D.L. Cohn, *Measure Theory: Second Edition*, Birkhäuser Advanced Texts Basler Lehrbücher, DOI 10.1007/978-1-4614-6956-8_4,