Chapter 4

L^p Spaces

Let (X, μ) be a measure space. As in Chapter 3, we say a measurable function f belongs to $\mathcal{L}^1(X, \mu)$ provided

$$\|f\|_{L^1} = \int_X |f(x)|\, d\mu(x) < \infty. \tag{4.1}$$

Elements of $L^1(X, \mu)$ consist of equivalence classes of elements of $\mathcal{L}^1(X, \mu)$, where we say

$$f \sim \tilde{f} \Leftrightarrow f(x) = \tilde{f}(x) \text{ for } \mu\text{-almost every } x. \tag{4.2}$$

With a slight abuse of notation, we denote by f both a measurable function in $\mathcal{L}^1(X, \mu)$ and its equivalence class in $L^1(X, \mu)$. Also we say f, defined only almost everywhere on X, belongs to $L^1(X, \mu)$, if there exists $\tilde{f} \in \mathcal{L}^1(X, \mu)$, equal μ-almost everywhere to f. The quantity $\|f\|_{L^1}$ defined by (4.1) is called the L^1 *norm* of f.

In general, a normed linear space is a vector space equipped with a positive function $\|v\|$ having the properties

$$\begin{aligned} \|av\| &= |a| \cdot \|v\|, \text{ for } v \in V, \ a \in \mathbb{C} \text{ (or } \mathbb{R}), \\ \|v + w\| &\le \|v\| + \|w\|, \\ \|v\| &> 0, \text{ unless } v = 0. \end{aligned} \tag{4.3}$$

The second of these conditions is called the *triangle inequality.* Given a norm on V, setting $d(u, v) = \|u - v\|$ defines a distance function on V, making it a metric space.

It is easy to see that $\mathcal{L}^1(X, \mu)$ is a vector space and that $\|f\|_{L^1}$ satisfies the first two conditions in (4.3). However, $\|f\|_{L^1} = 0$ if and only if $f = 0$

almost everywhere. (Recall Exercise 4 of Chapter 3.) That is the reason we define $L^1(X,\mu)$ to consist of equivalence classes defined by (4.2), so $L^1(X,\mu)$ becomes a normed linear space.

Generally speaking, a sequence (v_j) in a normed linear space is said to be a Cauchy sequence if $\|v_j - v_k\| \to 0$ as $j,k \to \infty$. If every Cauchy sequence has a limit in V, then V is said to be complete; a complete normed linear space is called a Banach space.

Theorem 4.1. *$L^1(X,\mu)$ is a Banach space.*

The proof of completeness of $L^1(X,\mu)$ makes use of the following two lemmas, which are essentially restatements of the Monotone Convergence Theorem and the Dominated Convergence Theorem, respectively.

Lemma 4.2. *If $f_j \in \mathcal{L}^1(X,\mu)$, $0 \le f_1(x) \le f_2(x) \le \cdots$, and $\|f_j\|_{L^1} \le C < \infty$, then $\lim_{j\to\infty} f_j(x) = f(x)$, with $f \in L^1(X,\mu)$ and $\|f_j - f\|_{L^1} \to 0$ as $j \to \infty$.*

Proof. We know that $f \in \mathcal{M}^+(X)$. The Monotone Convergence Theorem implies $\int f_j\, d\mu \nearrow \int f\, d\mu$. Thus $\int f\, d\mu \le C$. Since $\|f_j - f\|_{L^1} = \int f\, d\mu - \int f_j\, d\mu$ in this case, the lemma follows.

Lemma 4.3. *If $f_j \in \mathcal{L}^1(X,\mu), \lim f_j(x) = f(x)$ μ-a.e., and if there is an $F \in \mathcal{L}^1(X,\mu)$ such that $|f_j(x)| \le F(x)$ μ-a.e., for all j, then $f \in L^1(X,\mu)$, and $\|f_j - f\|_{L^1} \to 0$.*

Proof. Apply the Dominated Convergence Theorem to $g_j = |f_j - f| \to 0$ a.e. Note that $|g_j| \le 2F$.

To show $L^1(X,\mu)$ is complete, suppose (f_n) is Cauchy in L^1. Passing to a subsequence, we can assume $\|f_{n+1} - f_n\|_{L^1} \le 2^{-n}$. Consider the infinite series

$$f_1(x) + \sum_{n=1}^{\infty} \big[f_{n+1}(x) - f_n(x)\big]. \tag{4.4}$$

Now the partial sums are dominated by

$$|f_1(x)| + \sum_{n=1}^{m} |f_{n+1}(x) - f_n(x)| = |f_1(x)| + G_m(x), \tag{4.5}$$

and since $0 \le G_1 \le G_2 \le \cdots$ and $\|G_m\|_{L^1} \le \sum 2^{-n} \le 1$, we deduce from Lemma 4.2 that $G_m \nearrow G$ μ-a.e. and in L^1-norm. Hence the infinite series

(4.4) is convergent a.e., to a limit $f(x)$, and via Lemma 4.3 we deduce that $f_n \to f$ in L^1-norm. This proves completeness.

Continuing with a description of L^p spaces, we define $\mathcal{L}^\infty(X,\mu)$ to consist of bounded measurable functions, $L^\infty(X,\mu)$ to consist of equivalence classes of such functions, via (4.2), and we define $\|f\|_{L^\infty}$ to be the smallest sup of $\tilde{f} \sim f$. It is easy to show that $L^\infty(X,\mu)$ is a Banach space.

For $p \in (1,\infty)$, we define $\mathcal{L}^p(X,\mu)$ to consist of measurable functions f such that

$$\Big[\int_X |f(x)|^p \, d\mu(x)\Big]^{1/p} \tag{4.6}$$

is finite. $L^p(X,\mu)$ consists of equivalence classes, via (4.2), and the L^p-norm $\|f\|_{L^p}$ is given by (4.6). This time it takes a little work to verify the triangle inequality. That this holds is the content of *Minkowski's inequality*:

$$\|f+g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p}. \tag{4.7}$$

One neat way to establish this is by the following characterization of the L^p-norm. Suppose p and q are related by

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{4.8}$$

We claim that, if $f \in L^p(X,\mu)$,

$$\|f\|_{L^p} = \sup\{\|fh\|_{L^1} : h \in L^q(X,\mu),\ \|h\|_{L^q} = 1\}. \tag{4.9}$$

We can apply (4.9) to $f+g$, which belongs to $L^p(X,\mu)$ if f and g do, since $|f+g|^p \le 2^p\big(|f|^p + |g|^p\big)$. Given this, (4.7) follows easily from the inequality $\|(f+g)h\|_{L^1} \le \|fh\|_{L^1} + \|gh\|_{L^1}$.

The identity (4.9) can be regarded as two inequalities. The "$\le$" part can be proved by choosing $h(x)$ to be an appropriate multiple $C|f(x)|^{p-1}$. We leave this as an exercise. The converse inequality, "$\ge$," is a consequence of *Hölder's inequality*:

$$\int |f(x)g(x)| \, d\mu(x) \le \|f\|_{L^p}\|g\|_{L^q}, \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{4.10}$$

Hölder's inequality can be proved via the following inequality for positive numbers:

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \quad \text{for } a,b > 0, \tag{4.11}$$

assuming that $p \in (1,\infty)$ and (4.8) holds. In fact, we claim that, given $a, b > 0,\ 1/p + 1/q = 1$,

$$\varphi(t) = \frac{a^p}{p}t^p + \frac{b^q}{q}t^{-q} \Longrightarrow \inf_{t>0} \varphi(t) = ab, \tag{4.12}$$

which implies (4.11) since the right side of (4.11) is $\varphi(1)$. As for (4.12), note that $\varphi(t) \to \infty$ as $t \searrow 0$ and as $t \nearrow \infty$, and the unique critical point occurs for $a^p t^p = b^q t^{-q}$, i.e., for $t = b^{1/p}/a^{1/q}$, giving the desired conclusion.

Applying (4.11) to the integrand in (4.10) gives

$$\int |f(x)g(x)|\, d\mu(x) \le \frac{1}{p}\|f\|_{L^p}^p + \frac{1}{q}\|g\|_{L^q}^q. \tag{4.13}$$

This looks weaker than (4.10), but now replace f by tf and g by $t^{-1}g$, so that the left side of (4.13) is dominated by

$$\frac{t^p}{p}\|f\|_{L^p}^p + \frac{1}{qt^q}\|g\|_{L^q}^q, \tag{4.14}$$

for all $t > 0$. Another application of (4.12) then gives Hölder's inequality. Consequently (4.6) defines a norm on $L^p(X,\mu)$. Completeness follows as in the $p = 1$ case discussed above.

In detail, given (f_n) Cauchy in $L^p(X,\mu)$, we can pass to the case $\|f_{n+1} - f_n\|_{L^p} \le 2^{-n}$ and define G_m as in (4.5). We have $\|G_m\|_{L^p} \le 1$, and hence (via the Monotone Convergence Theorem) deduce that $G_m \nearrow G$, μ-a.e., and in L^p-norm. Hence the series (4.4) converges, μ-a.e., to a limit $f(x)$. Since $|f - f_{m+1}| \le G - G_m$, we have by the Dominated Convergence Theorem that

$$\int_X |f - f_{m+1}|^p \, d\mu \le \int_X (G - G_m)^p \, d\mu \to 0,$$

as $m \to \infty$. Hence $L^p(X,\mu)$ is complete. To summarize, we have

Theorem 4.4. *For $p \in [1,\infty)$, $L^p(X,\mu)$, with norm given by* (4.6)*, is a Banach space.*

It is frequently useful to show that a certain linear subspace L of a Banach space V is *dense*. We give an important case of this here; $C(X)$ denotes the space of continuous functions on X.

Proposition 4.5. *If μ is a finite Borel measure on a compact metric space X, then $C(X)$ is dense in $L^p(X,\mu)$ for each $p \in [1,\infty)$.*

Proof. First, let K be any compact subset of X. The functions

$$f_{K,n}(x) = \big[1 + n\ \mathrm{dist}(x, K)\big]^{-1} \in C(X) \tag{4.15}$$

are all ≤ 1 and decrease monotonically to the characteristic function χ_K equal to 1 on K, 0 on $X \setminus K$. The Monotone Convergence Theorem gives $f_{K,n} \to \chi_K$ in $L^p(X,\mu)$ for $1 \le p < \infty$. Now let $A \subset X$ be any measurable set. Any Borel measure on a compact metric space is *regular*, i.e.,

$$\mu(A) = \sup\{\mu(K) : K \subset A,\ K \text{ compact}\}. \tag{4.16}$$

In case $X = I = [a,b]$ and $\mu = m$ is Lebesgue measure, this follows from (2.20) together with the consequence of Theorem 2.11, that all Borel sets in I are Lebesgue measurable. The general case follows from results that will be established in the next chapter; see (5.60).

Thus there exists an increasing sequence K_j of compact subsets of A such that $\mu(A \setminus \bigcup_j K_j) = 0$. Again, the Monotone Convergence Theorem implies $\chi_{K_j} \to \chi_A$ in $L^p(X,\mu)$ for $1 \le p < \infty$. Thus all *simple* functions on X are in the closure of $C(X)$ in $L^p(X,\mu)$ for $p \in [1,\infty)$. Construction of $L^p(X,\mu)$ directly shows that each $f \in L^p(X,\mu)$ is a norm limit of simple functions, so the result is proved.

Using a cut-off, we can easily deduce the following. Let $C_{00}(\mathbb{R})$ denote the space of continuous functions on $\mathbb{R}$ with compact support.

Corollary 4.6. *For $1 \le p < \infty$, the space $C_{00}(\mathbb{R})$ is dense in $L^p(\mathbb{R})$.*

The case $L^2(X,\mu)$ is special. In addition to the L^2-norm, there is an *inner product*, defined by

$$(f,g)_{L^2} = \int_X f(x)\overline{g(x)}\, d\mu(x). \tag{4.17}$$

This makes $L^2(X,\mu)$ into a *Hilbert space*. It is worthwhile to consider the general notion of Hilbert space in some detail. We devote the next few pages to this and then return to the specific consideration of $L^2(X,\mu)$.

Generally, a Hilbert space H is a complete inner product space. That is to say, first the space H is a linear space provided with an inner product, denoted (u,v), for u and v in H, satisfying the following defining conditions:

$$\begin{aligned} (au_1 + u_2, v) &= a(u_1,v) + (u_2,v), \\ (u,v) &= \overline{(v,u)}, \\ (u,u) &> 0 \text{ unless } u = 0. \end{aligned} \tag{4.18}$$

To such an inner product there is assigned a norm, denoted by

$$\|u\| = \sqrt{(u,u)}. \tag{4.19}$$

To establish that the triangle inequality holds for $\|u+v\|$, we can expand $\|u+v\|^2 = (u+v, u+v)$ and deduce that this is $\le \big[\|u\| + \|v\|\big]^2$, as a consequence of *Cauchy's inequality*:

$$|(u,v)| \le \|u\| \cdot \|v\|, \tag{4.20}$$

a result that can be proved as follows. The fact that $(u-v, u-v) \ge 0$ implies $2\ \mathrm{Re}\ (u,v) \le \|u\|^2 + \|v\|^2$; replacing u by $e^{i\theta}u$ with $e^{i\theta}$ chosen so that $e^{i\theta}(u,v)$ is real and positive, we get

$$|(u,v)| \le \frac{1}{2}\|u\|^2 + \frac{1}{2}\|v\|^2. \tag{4.21}$$

Now in (4.21) we can replace u by tu and v by $t^{-1}v$ to get

$$|(u,v)| \le \frac{t}{2}\|u\|^2 + \frac{1}{2t}\|v\|^2. \tag{4.22}$$

Minimizing over t gives (4.20). This establishes Cauchy's inequality, so we can deduce the triangle inequality. Thus (4.19) defines a norm on H. Note the parallel between this argument and the proof of (4.7), via (4.10). The completeness hypothesis on H is that, with this norm, H is a Banach space.

The nice properties of Hilbert spaces arise from their similarity with familiar Euclidean space, so a great deal of geometrical intuition is available. For example, we say u and v are orthogonal and write $u \perp v$, provided $(u,v) = 0$. Note that the Pythagorean Theorem holds on a general Hilbert space:

$$u \perp v \Longrightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2. \tag{4.23}$$

This follows directly from expanding $(u+v, u+v)$.

Another useful identity is the following, called the "parallelogram law," valid for all $u, v \in H$:

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{4.24}$$

This also follows directly by expanding $(u+v, u+v) + (u-v, u-v)$, observing some cancellations. One important application of this simple identity is to the following existence result.

Let K be any closed, convex subset of H. Convexity implies $x, y \in K \Rightarrow (x+y)/2 \in K$. Given $x \in H$, we define the distance from x to K to be

$$d(x,K) = \inf\ \{\|x-y\| : y \in K\}. \tag{4.25}$$

Proposition 4.7. *If $K \subset H$ is a nonempty, closed, convex set in a Hilbert space H and if $x \in H$, then there is a unique $z \in K$ such that $d(x, K) = \|x - z\|$.*

Proof. We can pick $y_n \in K$ such that $\|x - y_n\| \to d = d(x, K)$. It will suffice to show that (y_n) must be a Cauchy sequence. Use (4.24) with $u = y_m - x,\ v = x - y_n$, to get

$$\|y_m - y_n\|^2 = 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\Big\|x - \frac{1}{2}(y_n + y_m)\Big\|^2.$$

Since K is convex, $(y_n + y_m)/2 \in K$, so $\|x - (y_n + y_m)/2\| \geq d$. Therefore

$$\limsup_{m,n\to\infty} \|y_n - y_m\|^2 \leq 2d^2 + 2d^2 - 4d^2 \leq 0,$$

which implies convergence.

In particular, this result applies when K is a closed linear subspace of H. In this case, for $x \in H$, denote by $P_K x$ the point in K closest to x. We have

$$x = P_K x + (x - P_K x). \tag{4.26}$$

We claim that $x - P_K x$ belongs to the closed linear space $K^\perp$, called the orthogonal complement of K, defined by

$$K^\perp = \{u \in H : (u, v) = 0 \text{ for all } v \in K\}. \tag{4.27}$$

Indeed, take any $v \in K$. Then

$$\begin{aligned}\Delta(t) &= \|x - P_K x + tv\|^2 \\ &= \|x - P_K x\|^2 + 2t \operatorname{Re}\,(x - P_K x, v) + t^2\|v\|^2\end{aligned}$$

is minimal at $t = 0$, so $\Delta'(0) = 0$, i.e., $\operatorname{Re}\,(x - P_K x, v) = 0$, for all $v \in K$. Replacing v by iv shows that $(x - P_K x, v)$ also has vanishing imaginary part for any $v \in K$, so our claim is established. The decomposition (4.26) gives

$$x = x_1 + x_2, \quad x_1 \in K,\ x_2 \in K^\perp, \tag{4.28}$$

with $x_1 = P_K x,\ x_2 = x - P_K x$. Clearly such a decomposition is unique. This implies that H is an orthogonal direct sum of K and $K^\perp$; we write

$$H = K \oplus K^\perp. \tag{4.29}$$

From this it is clear that

$$(K^\perp)^\perp = K, \tag{4.30}$$

that

$$x - P_K x = P_{K^\perp} x, \tag{4.31}$$

and that P_K and $P_{K^\perp}$ are *linear* maps on H. We call P_K the *orthogonal projection* of H on K. Note that $P_K x$ is uniquely characterized by the condition

$$P_K x \in K, \ (P_K x, v) = (x, v) \text{ for all } v \in K. \tag{4.32}$$

We remark that if K is a linear subspace of H that is not closed, then $K^\perp$ coincides with $\overline{K}^\perp$, and (4.30) becomes $(K^\perp)^\perp = \overline{K}$.

Using the orthogonal projection discussed above, we can establish the following result.

Proposition 4.8. *If H is a Hilbert space and $\varphi : H \to \mathbb{C}$ is a continuous linear map, there exists a unique $f \in H$ such that*

$$\varphi(u) = (u, f) \textit{ for all } u \in H. \tag{4.33}$$

Proof. Consider $K = \text{Ker}\,\varphi = \{u \in H : \varphi(u) = 0\}$, a closed linear subspace of H. If $K = H$, $\varphi = 0$ and we can take $f = 0$. Otherwise, $K^\perp \neq 0$; select a nonzero $x_0 \in K^\perp$, such that $\varphi(x_0) = 1$. We claim $K^\perp$ is one dimensional in this case. Indeed, given any $y \in K^\perp$, $y - \varphi(y)x_0$ is annihilated by φ, so it belongs to K as well as to $K^\perp$, so it is zero. The result is now easily proved by setting $f = ax_0$ with $a \in \mathbb{C}$ chosen so that (4.33) works for $u = x_0$, namely $\overline{a}(x_0, x_0) = 1$, i.e., $a = \|x_0\|^{-2}$.

We note that the correspondence $\varphi \mapsto f$ gives a *conjugate linear* isomorphism

$$H' \to H, \tag{4.34}$$

where H' denotes the space of all continuous linear maps $\varphi : H \to \mathbb{C}$.

Recall that our interest in Hilbert spaces arises from our interest in $L^2(X, \mu)$. Let us record the content of Proposition 4.8 in that case.

Corollary 4.9. *If* $\varphi : L^2(X,\mu) \to \mathbb{C}$ *is a continuous linear map, there exists a unique* $f \in L^2(X,\mu)$ *such that*

$$\varphi(u) = \int u(x) f(x)\, d\mu(x), \quad \forall\, u \in L^2(X,\mu). \tag{4.35}$$

We can use orthogonal projection operators to construct an *orthonormal basis* of a Hilbert space H. Let us assume H is *separable*, i.e., H has a countable dense subset $S = \{v_j : j \geq 1\}$. See the exercises for results on separability. In such a case, pick a countable subset $\{w_j : j \geq 1\}$ of S such that each w_k is linearly independent of $\{w_j : j < k\}$ and such that the linear span of $\{w_j\}$ is dense in H. Let $L_k = \operatorname{span}\{w_j : 1 \leq j \leq k\}$. We define an orthonormal set $\{u_j : j \geq 1\}$ inductively, as follows. Set $u_1 = w_1/\|w_1\|$. Suppose you have $\{u_j : 1 \leq j \leq k\}$, an orthonormal basis of L_k. Then set

$$u_{k+1} = \frac{w_{k+1} - P_k w_{k+1}}{\|w_{k+1} - P_k w_{k+1}\|}, \tag{4.36}$$

where P_k is the orthogonal projection onto L_k. One does not need the construction involving Proposition 4.7 to get P_k here. We can simply set

$$P_k f = \sum_{j=1}^{k} (f, u_j) u_j. \tag{4.37}$$

The set $\{u_j : j \geq 1\}$ so constructed is orthonormal, i.e., $(u_k, u_\ell) = \delta_{k\ell}$. Also its linear span is dense in H, since the linear span coincides with that of $\{w_j : j \geq 1\}$.

We claim that, for each $f \in H$, $P_k f \to f$ as $k \to \infty$. To see this, note that

$$\|f\|^2 = \|P_k f\|^2 + \|f - P_k f\|^2 \geq \|P_k f\|^2 = \sum_{j=1}^{k} |(f, u_j)|^2. \tag{4.38}$$

We also see that, if $n \geq k$, then $\|P_n f - P_k f\|^2 = \sum_{k<j\leq n} |(f,u_j)|^2$, and hence that $(P_k f)$ is a Cauchy sequence in H, for each f. We claim that the limit is equal to f, hence that

$$f = \sum_j (f, u_j) u_j. \tag{4.39}$$

Indeed, we now know that the right side of (4.39) defines an element of H; call it g. Then, $f - g$ has inner product 0 with each u_j, hence with all elements of the linear span of $\{u_j\}$, hence with all elements of the closure, i.e., with all elements of H, so $f - g = 0$.

In the case $H = L^2(I, m)$, with $I = [-\pi, \pi]$ and $m = dx/2\pi$, an orthonormal basis is given by

$$e_k(x) = e^{ikx}, \quad k \in \mathbb{Z}. \tag{4.40}$$

See the exercises for a proof of this. In such a case, (4.39) is the expansion of a function in a *Fourier series.*

We use Corollary 4.9 to prove an important result known as the *Radon-Nikodym Theorem.* Let μ and ν be two finite measures on $(X, \mathfrak{F})$. Let

$$\alpha = \mu + 2\nu, \quad \omega = 2\mu + \nu. \tag{4.41}$$

On the Hilbert space $H = L^2(X, \alpha)$, consider the linear functional $\varphi : H \to \mathbb{C}$ given by

$$\varphi(f) = \int_X f(x)\, d\omega(x). \tag{4.42}$$

Note that $|\varphi(f)| \le 2 \int |f|\, d\alpha \le 2\sqrt{\alpha(X)} \|f\|_{L^2(X,\alpha)}$. By Corollary 4.9, there exists $g \in \mathcal{L}^2(X, \alpha)$ such that, for any $f \in \mathcal{L}^2(X, \alpha) = \mathcal{L}^2(X, \mu) \cap \mathcal{L}^2(X, \nu)$,

$$\int_X f(x)\, d\omega(x) = \int_X f(x) g(x)\, d\alpha(x). \tag{4.43}$$

In particular, this holds for any bounded measurable f. Note that this identity is equivalent to

$$\int f(2g - 1)\, d\nu = \int f(2 - g)\, d\mu, \tag{4.44}$$

for all $f \in \mathcal{L}^2(X, \alpha)$. If we let f be the characteristic function of $S_{1\ell} = \{x \in X : g(x) < 1/2 - 1/\ell\}$ or of $S_{2\ell} = \{x \in X : g(x) > 2 + 1/\ell\}$, we see that $\mu(S_{j\ell}) = \nu(S_{j\ell}) = 0$. As a consequence, we can arrange that $1/2 \le g(x) \le 2$, for all $x \in X$. We also see that $Z = \{x \in X : g(x) = 1/2\}$ must have μ-measure zero. (Similarly, $\{x : g(x) = 2\}$ has ν-measure zero.) Also, (4.44) holds for all $f \in \mathcal{M}^+(X)$, by the Monotone Convergence Theorem.

We say that ν is *absolutely continuous* with respect to μ and write $\nu << \mu$, provided

$$\mu(S) = 0 \Longrightarrow \nu(S) = 0. \tag{4.45}$$

In such a case, we see that $Z = \{x \in X : g(x) = 1/2\}$ has ν-measure zero. Given $F \in \mathcal{M}^+(X)$, we can set

$$f(x) = \frac{F(x)}{2g(x) - 1}, \quad h(x) = \frac{2 - g(x)}{2g(x) - 1} \tag{4.46}$$

and apply (4.44) to get

$$\int_X F(x)\, d\nu(x) = \int_X F(x)h(x)\, d\mu(x) \tag{4.47}$$

for all positive measurable F. Note that taking $F = 1$ gives $h \in L^1(X, \mu)$. The result we have just obtained is known as the Radon-Nikodym Theorem. We record a formal statement.

Theorem 4.10. *Let μ and ν be two finite measures on $(X, \mathfrak{F})$. If ν is absolutely continuous with respect to μ, then (4.47) holds for some nonnegative $h \in L^1(X, \mu)$ and every positive measurable F.*

We mention that (4.47) also holds for every bounded measurable F.

If we do not assume that $\nu << \mu$, we can still consider

$$h(x) = \begin{cases} \dfrac{2 - g(x)}{2g(x) - 1} & \text{if } g(x) \neq \dfrac{1}{2}, \\ 0 & \text{if } g(x) = \dfrac{1}{2}, \end{cases} \tag{4.48}$$

and we have

$$\int_Y F\, d\nu = \int_Y Fh\, d\mu \tag{4.49}$$

for any positive measurable F, where

$$Y = X \setminus Z = \Big\{x \in X : g(x) \neq \frac{1}{2}\Big\}. \tag{4.50}$$

Recall that $\mu(X \setminus Y) = 0$. We can define the measure λ on $(X, \mathfrak{F})$ by

$$\lambda(E) = \nu(Y \cap E). \tag{4.51}$$

Then we have

$$\int_X F\, d\lambda = \int_X Fh\, d\mu \tag{4.52}$$

for all positive measurable F. Write

$$\rho(E) = \nu(E \setminus Y) = \nu(E \cap Z), \tag{4.53}$$

so

$$\nu = \lambda + \rho. \tag{4.54}$$

Now the measure λ is *supported* on Y, i.e., $\lambda(X \setminus Y) = 0$. Similarly, ρ is supported on Z. Thus λ and ρ have disjoint supports. Generally, two measures with disjoint supports are said to be *mutually singular*. When two measures λ and ρ are mutually singular, we write $\lambda \perp \rho$.

We have the following result, known as the Lebesgue decomposition of ν with respect to μ.

Theorem 4.11. *If μ and ν are finite measures on $(X, \mathfrak{F})$, then we can write*

$$\nu = \lambda + \rho, \quad \lambda << \mu, \quad \rho \perp \mu. \tag{4.55}$$

This decomposition is unique.

Proof. The measures λ and ρ are given by (4.51) and (4.52). The fact that $\lambda << \mu$ is contained in (4.52). As we have noted, $\mu(X \setminus Y) = 0$, so μ is supported on Y, which is disjoint from Z, on which ρ is supported; hence $\rho \perp \mu$.

If also $\tilde{\lambda}$ and $\tilde{\rho}$ are measures such that $\nu = \tilde{\lambda} + \tilde{\rho}$, $\tilde{\lambda} << \mu$, $\tilde{\rho} \perp \mu$, we have $\widetilde{Z} \in \mathfrak{F}$ such that $\tilde{\rho}$ is supported on $\widetilde{Z}$ and $\mu(\widetilde{Z}) = 0$. Now $\mu(Z \cup \widetilde{Z}) = 0$, so $\lambda(Z \cup \widetilde{Z}) = 0$ and $\tilde{\lambda}(Z \cup \widetilde{Z}) = 0$, and, for $E \in \mathcal{M}$,

$$\begin{aligned} \lambda(E) &= \lambda(E \setminus \widetilde{Z}) = \nu\big(E \setminus (Z \cup \widetilde{Z})\big), \\ \tilde{\lambda}(E) &= \tilde{\lambda}(E \setminus Z) = \nu\big(E \setminus (Z \cup \widetilde{Z})\big). \end{aligned}$$

This gives uniqueness.

We say a measure μ on $(X, \mathfrak{F})$ is σ-finite if we can write X as a countable union $\bigcup_{j \ge 1} X_j$ where $X_j \in \mathfrak{F}$ and $\mu(X_j) < \infty$. A paradigm case is Lebesgue measure on $X = \mathbb{R}$. There are routine extensions of Theorems 4.10–4.11 to the case where μ and ν are σ-finite measures, which we leave to the reader.

Exercises

1. Let V and W be normed linear spaces. Suppose we have linear transformations

$$T_j : V \longrightarrow W, \quad \|T_j v\|_W \le C \|v\|_V, \tag{4.56}$$

with C independent of j. (We say $\{T_j\}$ is uniformly bounded.) Suppose also $T : V \to W$ satisfies this bound. Let L be a dense subspace of V. Then show that

$$T_j v \to Tv,\ \forall\, v \in L \Longrightarrow T_j v \to Tv,\ \forall\, v \in V. \tag{4.57}$$

2. Define

$$\tau_s : L^p(\mathbb{R}) \longrightarrow L^p(\mathbb{R}), \quad \tau_s f(x) = f(x-s). \tag{4.58}$$

Show that, for $p \in [1,\infty)$,

$$f \in L^p(\mathbb{R}) \Longrightarrow \tau_s f \to f \text{ in } L^p\text{-norm, as } s \to 0. \tag{4.59}$$

Hint. Apply Exercise 1, with $V = W = L^p(\mathbb{R})$, $L = C_{00}(\mathbb{R})$, as in Corollary 4.6. Note that $\|\tau_s f\|_{L^p} = \|f\|_{L^p}$.

One says a metric space is *separable* if it has a countable dense subset.

3. If $I = [a,b] \subset \mathbb{R}$ and $a \le \alpha < \beta \le b$, define

$$\varphi_{\alpha\beta}(x) = \operatorname{dist}\big(x, I \setminus [\alpha,\beta]\big).$$

Show that the linear span over $\mathbb{Q}$ of $\{\varphi_{\alpha\beta} : \alpha,\beta \in \mathbb{Q}\cap I\}$ is dense in $C(I)$, and deduce that $C(I)$ is separable. From the denseness and continuity of the inclusion

$$\iota : C(I) \longrightarrow L^p(I),$$

prove that $L^p(I)$ is separable, for $1 \le p < \infty$. Then prove that $L^p(\mathbb{R})$ is separable, for $1 \le p < \infty$.

4. Let X be a compact metric space; X has a countable dense subset $\{z_j : j \ge 1\}$. Given $0 < \rho < (\operatorname{diam} X)/2$, set

$$\psi_{j\rho}(x) = \operatorname{dist}\big(x, X \setminus B_\rho(z_j)\big).$$

Show that the algebra generated by $\{\psi_{j,\rho} : j \in \mathbb{Z}^+,\ \rho \in \mathbb{Q}^+\}$ and 1 is dense in $C(X)$ and deduce that $C(X)$ is separable. Conclude, from Proposition 4.5, that $L^p(X,\mu)$ is separable, for $p \in [1,\infty)$, if μ is a finite measure on the σ-algebra of Borel sets in X.

5. Let $\{u_j : j \ge 1\}$ be a countable orthonormal set in a Hilbert space H. Show that

$$\sum_{j=1}^{\infty} (f,u_j)u_j = P_V f, \tag{4.60}$$

where V is the closure of the linear span of $\{u_j\}$.

The Stone-Weierstrass Theorem states that, if X is a compact Hausdorff space and $\mathcal{A}$ an algebra of functions in $C_{\mathbb{R}}(X)$ (the space of real-valued continuous functions on X), such that $1 \in \mathcal{A}$, and if $\mathcal{A}$ has the property of separating points, i.e., for any two distinct $p, q \in X$, there exists $f \in \mathcal{A}$ such that $f(p) \neq f(q)$, then $\mathcal{A}$ is dense in $C_{\mathbb{R}}(X)$. If $\mathcal{A}$ is an algebra in $C_{\mathbb{C}}(X)$ with these properties, plus the property that $f \in \mathcal{A} \Rightarrow \overline{f} \in \mathcal{A}$, then $\mathcal{A}$ is dense in $C_{\mathbb{C}}(X)$. A proof is given in Appendix A.

6. Use the Stone-Weierstrass Theorem to show that, if $e_k(x) = e^{ikx}$, as in (4.40), then the linear span $\mathcal{E}$ of $\{e_k : k \in \mathbb{Z}\}$ is dense in $C(S^1)$, where $S^1 = \mathbb{R}/2\pi\mathbb{Z}$; hence $\mathcal{E}$ is dense in $L^p(S^1, dx/2\pi)$, for $p \in [1, \infty)$. Hence $\{e_k : k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(S^1, dx/2\pi)$.

7. For $f, u \in C_0^\infty(\mathbb{R})$, the space of smooth functions with compact support in $\mathbb{R}$, set

$$K_f u(x) = f * u(x) = \int f(y) u(x - y)\, dy. \tag{4.61}$$

Show that, for $1 \le p < \infty$, K_f has a unique bounded extension:

$$K_f : L^p(\mathbb{R}) \longrightarrow L^p(\mathbb{R}), \quad \|K_f u\|_{L^p} \le \|f\|_{L^1} \|u\|_{L^p}. \tag{4.62}$$

The operation in (4.61) is called *convolution*.
Hint. For $f, u \in C_0^\infty(\mathbb{R})$, if f is supported in $[a, b]$, show that

$$f * u(x) = \lim_{n \to \infty} \frac{b - a}{n} \sum_{j=0}^{n} f\Bigl(\frac{j}{n} b + \Bigl(1 - \frac{j}{n}\Bigr) a\Bigr)\, \tau_j u(x),$$

where $\tau_j u(x) = u(x - jb/n - (1 - j/n)a)$. Use the triangle inequality (4.7) to estimate norms.

8. Show that there is a unique extension from $f \in C_0^\infty(\mathbb{R})$ to $f \in L^1(\mathbb{R})$ of $K_f u$, with (4.62) continuing to hold, giving a continuous linear map $K : L^1(\mathbb{R}) \to \mathcal{L}(L^p(\mathbb{R}))$.

9. Let f_j be a sequence of nonnegative functions in $L^1(\mathbb{R})$ such that

$$\int f_j\, dx = 1, \quad \text{supp } f_j \subset \{x \in \mathbb{R} : |x| < 1/j\}. \tag{4.63}$$

Show that, for $1 \le p < \infty$,

$$u \in L^p(\mathbb{R}) \Longrightarrow f_j * u \to u \text{ in } L^p \text{ norm, as } j \to \infty. \tag{4.64}$$

Derive the same conclusion, upon weakening the second hypothesis in (4.63) to

$$\int\limits_{|x|<\varepsilon} f_j\,dx = \beta_j(\varepsilon) \to 1 \text{ as } j \to \infty, \quad \forall\, \varepsilon > 0.$$

Hint. First verify the conclusion for $u \in C_0^\infty(\mathbb{R})$. Then use Exercise 1.

10. Given $f \in L^1(S^1),\ 0 < r < 1$, define

$$P_r f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}(n) r^{|n|} e^{in\theta}, \quad \hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta) e^{-in\theta}\,d\theta. \tag{4.65}$$

Show that

$$P_r f(\theta) = p_r * f(\theta) = \frac{1}{2\pi}\int_0^{2\pi} p_r(\theta - \varphi) f(\varphi)\,d\varphi, \tag{4.66}$$

where

$$p_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}. \tag{4.67}$$

Show that

$$\frac{1}{2\pi}\int_0^{2\pi} p_r(\theta)\,d\theta = 1. \tag{4.68}$$

11. If $f \in L^p(S^1),\ 1 \le p < \infty$, show that

$$P_r f \to f, \quad \text{as} \quad r \nearrow 1, \tag{4.69}$$

in L^p-norm. If $f \in C(S^1)$, show that you have uniform convergence in (4.69). This is known as Abel convergence of Fourier series.
Hint. Use a variant of the analysis needed for Exercise 9.

12. Show that Exercises 10–11 provide an alternative proof of the conclusion of Exercise 6, that the linear span of $e_k(\theta) = e^{ik\theta},\ k \in \mathbb{Z}$, is dense in $C(S^1)$ and in $L^p(S^1)$, for $1 \le p < \infty$, and hence that $\{e_k : k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(S^1, d\theta/2\pi)$.

13. Suppose $(X, \overline{\mathfrak{F}}, \overline{\mu})$ is the completion of $(X, \mathfrak{F}, \mu)$. Show that $L^p(X, \mu)$ and $L^p(X, \overline{\mu})$ are identical.

Hint. Consult Exercise 6 of Chapter 3.

14. Show that if $f \in L^p(X, \mu)$ and $p \in [1, \infty)$, then, for $\lambda \in (0, \infty)$,

$$\mu\big(\{x \in X : |f(x)| > \lambda\}\big) \le \lambda^{-p} \, \|f\|_{L^p}^p. \tag{4.70}$$

Deduce that if $f_k \to 0$ in L^p-norm, for some $p \in [1, \infty)$, then $f_k \to 0$ in measure, as defined in Exercise 10 of Chapter 3.
Hint. Denote the set being measured in (4.70) by E_λ and note that $\int_{E_\lambda} |f|^p \, d\mu \ge \lambda^p \mu(E_\lambda)$.
The inequality (4.70) is called Tchebychev's inequality.