Chapter 8

L^p Classes

1. Definition of L^p

If E is a measurable subset of $\mathbf{R}^n$ and p satisfies $0 < p < \infty$, then $L^p(E)$ denotes the collection of measurable f for which $\int_E |f|^p$ is finite, that is,

$$L^p(E) = \left\{f\colon \int_E |f|^p < +\infty\right\},\ 0 < p < \infty.$$

Here, f may be complex-valued. (See Exercise 3 of Chapter 4 for the definition of measurability of vector-valued functions.) In this case, if $f = f_1 + if_2$ for measurable real-valued f_1 and f_2, we have $|f|^2 = f_1^2 + f_2^2$, so that

$$|f_1|,|f_2| \le |f| \le |f_1| + |f_2|.$$

It follows that $f \in L^p(E)$ if and only if both $f_1, f_2 \in L^p(E)$. (See Exercise 1.)

We shall write

$$\|f\|_{p,E} = \left(\int_E |f|^p\right)^{1/p} \qquad (0 < p < \infty);$$

thus, $L^p(E)$ is the class of measurable f for which $\|f\|_{p,E}$ is finite. Whenever it is clear from context what E is, we will write L^p for $L^p(E)$ and $\|f\|_p$ for $\|f\|_{p,E}$. Note that $L^1 = L$.

In order to define $L^\infty(E)$, let f be real-valued and measurable on a set E of positive measure. Define the *essential supremum* of f on E as follows: If $|\{\mathbf{x} \in E : f(\mathbf{x}) > \alpha\}| > 0$ for all real α, let $\operatorname{ess}_E \sup f = +\infty$; otherwise, let

$$\operatorname*{ess\,sup}_E f = \inf\{\alpha : |\{\mathbf{x} \in E : f(\mathbf{x}) > \alpha\}| = 0\}.$$

Since the distribution function $\omega(\alpha) = |\{\mathbf{x} \in E : f(\mathbf{x}) > \alpha\}|$ is continuous from the right [see (5.39)], it follows that $\omega(\operatorname{ess}_E \sup f) = 0$ if $\operatorname{ess}_E \sup f$ is finite. Therefore, $\operatorname{ess}_E \sup f$ is the smallest number M, $-\infty \le M \le +\infty$, such that $f(x) \le M$ except for a subset of E of measure zero.

A real- or complex-valued measurable f is said to be *essentially bounded*, or simply *bounded*, on E if $\text{ess}_E\sup |f|$ is finite. The class of all functions that are essentially bounded on E is denoted by $L^\infty(E)$. Clearly, f belongs to $L^\infty(E)$ if and only if its real and imaginary parts do. We shall write

$$\|f\|_\infty = \|f\|_{\infty,E} = \operatorname*{ess\,sup}_E |f|.$$

Thus, $\|f\|_\infty$ is the smallest M such that $|f| \le M$ a.e. in E, and

$$L^\infty = L^\infty(E) = \{f : \|f\|_\infty < +\infty\}.$$

The following theorem gives some motivation for this notation.

(8.1) Theorem *If* $|E| < +\infty$, *then* $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.

Proof. Let $M = \|f\|_\infty$. If $M' < M$, then the set $A = \{\mathbf{x} \in E: |f(\mathbf{x})| > M'\}$ has positive measure. Moreover, $\|f\|_p \ge (\int_A |f|^p)^{1/p} \ge M'|A|^{1/p}$. Since $|A|^{1/p} \to 1$ as $p \to \infty$, it follows that $\liminf_{p\to\infty} \|f\|_p \ge M'$, so that $\liminf_{p\to\infty} \|f\|_p \ge M$. However, we also have $\|f\|_p \le (\int_E M^p)^{1/p} = M|E|^{1/p}$. Therefore, $\limsup_{p\to\infty} \|f\|_p \le M$, which completes the proof.

Remark: This result may fail if $|E| = +\infty$. Consider, for example, the constant function $f(x) = c$, $c \ne 0$, in $(0,\infty)$. Clearly, $f \in L^\infty$ but $f \notin L^p$ for $0 < p < \infty$.

We will now study some basic properties of the L^p classes.

(8.2) Theorem *If* $0 < p_1 < p_2 \le \infty$ *and* $|E| < +\infty$, *then* $L^{p_2} \subset L^{p_1}$.

Proof. Write $E = E_1 \cup E_2$, E_1 being the set where $|f| \le 1$, and E_2 the set where $|f| > 1$. Then

$$\int_E |f|^p = \int_{E_1} |f|^p + \int_{E_2} |f|^p, \qquad 0 < p < \infty.$$

The first term on the right is majorized by $|E_1|$; the second increases with p since its integrand exceeds 1. It follows that if $f \in L^{p_2}$, $p_2 < \infty$, then $f \in L^{p_1}$, $p_1 < p_2$. If $p_2 = \infty$, then f is a bounded function on a set of finite measure, and so belongs to L^{p_1}.

Remarks:

(i) The hypothesis above that E have finite measure cannot be omitted: for example, x^{-1/p_1} belongs to $L^{p_2}(1,\infty)$ if $p_2 > p_1$, but does not belong to $L^{p_1}(1,\infty)$. Again, any nonzero constant is in L^∞, but is not in $L^{p_1}(E)$ if $|E| = +\infty$ and $p_1 < \infty$.

(ii) A function may belong to all L^{p_1} with $p_1 < p_2$ and yet not belong to L^{p_2}. In fact, if $p_2 < \infty$, x^{-1/p_2} belongs to $L^{p_1}(0,1)$, $p_1 < p_2$, but does not belong to $L^{p_2}(0,1)$; $\log 1/x$ is in $L^{p_1}(0,1)$ for $p_1 < \infty$, but is not in $L^\infty(0,1)$.

(iii) We leave it to the reader to show that any function which is bounded on

E ($|E| < +\infty$ or not) and which belongs to L^{p_1} also belongs to L^{p_2}, $p_2 > p_1$.

The next theorem states that the L^p classes are vector (i.e., linear) spaces. Its proof is left as an exercise.

(8.3) Theorem *If $f,g \in L^p(E)$, $p > 0$, then $f + g \in L^p(E)$ and $cf \in L^p(E)$ for any constant c.*

2. Hölder's Inequality; Minkowski's Inequality

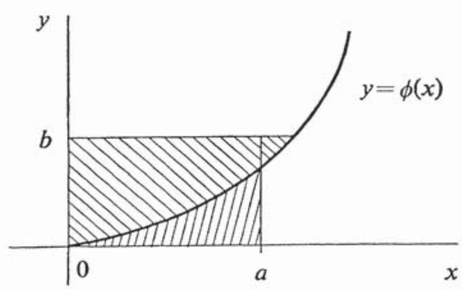
In order to discuss the integrability of the product of two functions, we will use the following basic result.

(8.4) Theorem (*Young's Inequality*) *Let $y = \phi(x)$ be continuous, real-valued, and strictly increasing for $x \geq 0$, and let $\phi(0) = 0$. If $x = \psi(y)$ is the inverse of ϕ, then for $a,b > 0$,*

$$ab \leq \int_0^a \phi(x)\,dx + \int_0^b \psi(y)\,dy.$$

Equality holds if and only if $b = \phi(a)$.

Proof. A geometric proof is immediate if we interpret each term as an area and remember that the graph of ϕ also serves as that of ψ if we interchange the x and y axes. Equality holds if and only if the point (a,b) lies on the graph of ϕ.



If $\phi(x) = x^\alpha$, $\alpha > 0$, then $\psi(y) = y^{1/\alpha}$, and Young's inequality becomes $ab \leq a^{1+\alpha}/(1 + \alpha) + b^{1+1/\alpha}/(1 + 1/\alpha)$. Setting $p = \alpha + 1$ and $p' = 1 + 1/\alpha$, we obtain

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'} \text{ if } a,b \geq 0,\ 1 < p < \infty, \text{ and } \frac{1}{p} + \frac{1}{p'} = 1. \tag{8.5}$$

Two numbers p and p' which satisfy $1/p + 1/p' = 1$, $p,p' > 1$, are called *conjugate exponents*. Note that $p' = p/(p - 1)$, and that 2 is self-conjugate. We will adopt the conventions that $p' = \infty$ if $p = 1$, and $p' = 1$ if $p = \infty$.

(8.6) Theorem (*Hölder's Inequality*) *If* $1 \le p \le \infty$ *and* $1/p + 1/p' = 1$, *then* $\|fg\|_1 \le \|f\|_p \|g\|_{p'}$; *that is,*

$$\int_E |fg| \le \left(\int_E |f|^p\right)^{1/p} \left(\int_E |g|^{p'}\right)^{1/p'} \qquad (1 < p < \infty);$$

$$\int_E |fg| \le (\operatorname*{ess\,sup}_E |f|) \int_E |g|.$$

Proof. The last inequality, which corresponds to the case $p = \infty$, is obvious. Let us suppose then that $1 < p < \infty$. In case $\|f\|_p = \|g\|_{p'} = 1$, (8.5) implies that

$$\int_E |fg| \le \int_E \left(\frac{|f|^p}{p} + \frac{|g|^{p'}}{p'}\right) = \frac{\|f\|_p^p}{p} + \frac{\|g\|_{p'}^{p'}}{p'}$$
$$= \frac{1}{p} + \frac{1}{p'} = 1 = \|f\|_p \|g\|_{p'}.$$

For the general case, we may assume that neither $\|f\|_p$ nor $\|g\|_{p'}$ is zero; otherwise fg is zero a.e. in E, and the result is immediate. We may also assume that neither $\|f\|_p$ nor $\|g\|_{p'}$ is infinite. If we set $f_1 = f/\|f\|_p$ and $g_1 = g/\|g\|_{p'}$, then $\|f_1\|_p = \|g_1\|_{p'} = 1$. Therefore, by the case above, we have $\int_E |f_1 g_1| \le 1$; i.e., $\int_E |fg| \le \|f\|_p \|g\|_{p'}$, as desired.

The case $p = p' = 2$ is a classical inequality.

(8.7) Corollary (*Schwarz's Inequality*)

$$\int_E |fg| \le \left(\int_E |f|^2\right)^{1/2} \left(\int_E |g|^2\right)^{1/2}.$$

The theorem which follows is usually referred to as the "converse of Hölder's inequality." (See also Exercise 15 in Chapter 10.)

(8.8) Theorem *Let f be real-valued and measurable on E, let $1 \le p \le \infty$ and $1/p + 1/p' = 1$. Then*

(8.9) $$\|f\|_p = \sup \int_E fg,$$

where the supremum is taken over all real-valued g such that $\|g\|_{p'} \le 1$ and $\int_E fg$ exists.

Proof. That the left-hand side of (8.9) majorizes the right-hand side follows from Hölder's inequality. To show the opposite inequality, let us consider first the case of $f \ge 0$, $1 < p < \infty$.

If $\|f\|_p = 0$, then $f = 0$ a.e. in E, and the result is obvious. If $0 < \|f\|_p <$

$+\infty$, we may further assume that $\|f\|_p = 1$ by dividing both sides of (8.9) by $\|f\|_p$. Now let $g = f^{p/p'}$. It is easy to verify that $\|g\|_{p'} = 1$ and $\int_E fg = 1$, which completes the proof in this case.

If $\|f\|_p = +\infty$, define functions f_k on E by setting

$$f_k(\mathbf{x}) = 0 \text{ if } |\mathbf{x}| > k, \qquad f_k(\mathbf{x}) = \min\,[f(\mathbf{x}),k] \text{ if } |\mathbf{x}| \le k.$$

Then each f_k belongs to L^p and $\|f_k\|_p \to \|f\|_p = +\infty$. By the case already considered, we have $\|f_k\|_p = \int_E f_k g_k$ for some $g_k \ge 0$ with $\|g_k\|_{p'} = 1$. Since $f \ge f_k$, it follows that

$$\int_E f g_k \ge \int_E f_k g_k \to +\infty.$$

This shows that

$$\sup_{\|g\|_{p'}=1} \int_E fg = +\infty = \|f\|_p.$$

To dispose of the restriction $f \ge 0$, apply the result above to $|f|$. Thus, there exist g_k with $\|g_k\|_{p'} = 1$ such that

$$\|f\|_p = \lim \int_E |f| g_k = \lim \int_E f \tilde{g}_k,$$

where $\tilde{g}_k = g_k\,(\operatorname{sign} f)$. (By $\operatorname{sign} x$, we mean the function equal to $+1$ for $x > 0$ and to -1 for $x < 0$.) Since $\|\tilde{g}_k\|_{p'} = 1$, the result follows.

The cases $p = 1$ and ∞ are left as exercises.

We have already observed that the sum of two L^p functions is again in L^p. The next theorem gives a more specific result when $1 \le p \le \infty$.

(8.10) Theorem (*Minkowski's Inequality*) *If* $1 \le p \le \infty$, *then* $\|f + g\|_p \le \|f\|_p + \|g\|_p$; *that is,*

$$\left(\int_E |f + g|^p\right)^{1/p} \le \left(\int_E |f|^p\right)^{1/p} + \left(\int_E |g|^p\right)^{1/p} \qquad (1 \le p < \infty),$$

$$\operatorname*{ess\,sup}_E |f + g| \le \operatorname*{ess\,sup}_E |f| + \operatorname*{ess\,sup}_E |g|.$$

Proof. If $p = 1$, the result is obvious. If $p = \infty$, we have $|f| \le \|f\|_\infty$ a.e. in E and $|g| \le \|g\|_\infty$ a.e. in E. Therefore, $|f + g| \le \|f\|_\infty + \|g\|_\infty$ a.e. in E, so that $\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

For $1 < p < \infty$,

$$\begin{aligned}\|f + g\|_p^p = \int_E |f + g|^p &= \int_E |f + g|^{p-1}|f + g| \\ &\le \int_E |f + g|^{p-1}|f| + \int_E |f + g|^{p-1}|g|.\end{aligned}$$

In the last integral, apply Hölder's inequality to $|f + g|^{p-1}$ and $|g|$ with exponents $p' = p/(p-1)$ and p, respectively. This gives

$$\int_E |f + g|^{p-1}|g| \leq \|f + g\|_p^{p-1}\|g\|_p.$$

Since a similar result holds for $\int_E |f + g|^{p-1}|f|$, we obtain $\|f + g\|_p^p \leq \|f + g\|_p^{p-1}(\|f\|_p + \|g\|_p)$, and the theorem follows by dividing both sides by $\|f + g\|_p^{p-1}$. [Note that if $\|f + g\|_p = 0$, there is nothing to prove; if $\|f + g\|_p = +\infty$, then either $\|f\|_p = +\infty$ or $\|g\|_p = +\infty$ by (8.3), and the result is obvious again.]

Remark: Minkowski's inequality fails for $0 < p < 1$. To see this, take $E = (0,1)$, $f = \chi_{(0,\frac{1}{2})}$ and $g = \chi_{(\frac{1}{2},1)}$. Then $\|f + g\|_p = 1$, while $\|f\|_p + \|g\|_p = 2^{-1/p} + 2^{-1/p} = 2^{1-1/p} < 1$. [See also (8.17).]

3. Classes l^p

Let $a = \{a_k\}$ be a sequence of real or complex numbers, and let

$$\|a\|_p = (\sum_k |a_k|^p)^{1/p}, \ 0 < p < \infty; \qquad \|a\|_\infty = \sup_k |a_k|.$$

Then a is said to belong to l^p, $0 < p < \infty$, if $\|a\|_p < +\infty$, and to belong to l^∞ if $\|a\|_\infty < +\infty$.

Let us show that if $0 < p_1 < p_2 \leq \infty$, then $l^{p_1} \subset l^{p_2}$. [The opposite inclusion holds for $L^p(E)$, $|E| < +\infty$, by (8.2).] For $p_2 = \infty$, this is clear, and for $p_2 < \infty$, it follows from the fact that if $|a_k| \leq 1$, then $|a_k|^{p_2} \leq |a_k|^{p_1}$. An example of a sequence which is in l^{p_2} for a given $p_2 < \infty$ but which is not in l^{p_1} for $p_1 < p_2$ is $\{(1/k \log^2 k)^{1/p_2} : k \geq 2\}$. Any constant sequence $\{a_k\}$, $a_k = c \neq 0$, belongs to l^∞ but not to l^p for $p < \infty$. The same is true for $\{1/\log k : k \geq 2\}$, whose terms even tend to zero.

(8.11) Theorem *If $a = \{a_k\}$ belongs to l^p for some $p < \infty$, then $\lim_{p\to\infty} \|a\|_p = \|a\|_\infty$.*

Proof. If $a \in l^{p_0}$, then $a \in l^p$ for $p_0 \leq p \leq \infty$. Since $|a_k| \to 0$, there is a largest $|a_k|$, say $|a_{k_0}|$. Thus, $\|a\|_\infty = |a_{k_0}|$. Write $\sum |a_k|^p = |a_{k_0}|^p \sum |a_k/a_{k_0}|^p$. Since $|a_k/a_{k_0}| \leq 1$, we see that $\sum |a_k/a_{k_0}|^p$ decreases (and so is bounded) as $p \to \infty$. Hence, there is a constant c such that $|a_{k_0}|^p \leq \|a\|_p^p \leq c|a_{k_0}|^p$. Since $c^{1/p} \to 1$ as $p \to \infty$, the theorem follows.

The next two results are analogues for series of Hölder's and Minkowski's inequalities. Their proofs are left as exercises. If $a = \{a_k\}$ and $b = \{b_k\}$, we use the notation

$$ab = \{a_k b_k\}, \qquad a + b = \{a_k + b_k\}, \text{ etc.}$$

(8.12) Theorem (*Hölder's Inequality*) *Suppose that $1 \leq p \leq \infty$, $1/p + 1/p'$*

$= 1$, $a = \{a_k\}$, $b = \{b_k\}$, *and* $ab = \{a_k b_k\}$. *Then* $\|ab\|_1 \le \|a\|_p \|b\|_{p'}$; *that is,*

$$\sum |a_k b_k| \le \left(\sum |a_k|^p\right)^{1/p} \left(\sum |b_k|^{p'}\right)^{1/p'} \qquad (1 < p < \infty);$$
$$\sum |a_k b_k| \le (\sup |a_k|) \sum |b_k|.$$

(8.13) Theorem (*Minkowski's Inequality*) *Suppose that* $1 \le p \le \infty$, $a = \{a_k\}, b = \{b_k\}$, *and* $a + b = \{a_k + b_k\}$. *Then* $\|a + b\|_p \le \|a\|_p + \|b\|_p$; *that is,*

$$\left(\sum |a_k + b_k|^p\right)^{1/p} \le \left(\sum |a_k|^p\right)^{1/p} + \left(\sum |b_k|^p\right)^{1/p} \qquad (1 \le p < \infty);$$
$$\sup |a_k + b_k| \le \sup |a_k| + \sup |b_k|.$$

Even though Minkowski's inequality fails when $p < 1$ (see Exercise 3), l^p is still a vector space for $0 < p < 1$; that is, $a + b \in l^p$ and $\alpha a = \{\alpha a_k\} \in l^p$ if $a,b \in l^p$ and α is any constant.

4. Banach and Metric Space Properties

We now define a notion which incorporates the main properties of L^p and l^p when $p \ge 1$. A set X is called a *Banach space over the complex numbers* if it satisfies the following conditions:

(B_1) X is a *linear space* over the complex numbers $\mathbf{C}$; that is, if $x,y \in X$ and $\alpha \in \mathbf{C}$, then $x + y \in X$ and $\alpha x \in X$.

(B_2) X is a *normed space*; that is, for every $x \in X$ there is a non-negative number $\|x\|$ such that

(a) $\|x\| = 0$ if and only if x is the zero element of X.
(b) $\|\alpha x\| = |\alpha| \|x\|$ for $\alpha \in \mathbf{C}$; $x \in X$.
(c) $\|x + y\| \le \|x\| + \|y\|$.

If these conditions are fulfilled, $\|x\|$ is called the *norm* of x.

(B_3) X is *complete* with respect to its norm; that is, every Cauchy sequence in X converges in X, or if $\|x_k - x_m\| \to 0$ as $k,m \to \infty$, then there is an $x \in X$ such that $\|x_k - x\| \to 0$.

A set X which satisfies (B_1) and (B_2), but not necessarily (B_3), is called a *normed linear space* over the complex numbers. A sequence $\{x_k\}$ such that $\|x_k - x\| \to 0$ as $k \to \infty$ is said to *converge in norm* to x.

Restricting the scalars α in (B_1) and (B_2) to be real numbers, we obtain definitions for a Banach space over the real numbers and for a normed linear space over the real numbers. Unless specifically stated to the contrary, we will take the scalar field to be the complex numbers.

If X is a Banach space, define $d(x,y) = \|x - y\|$ to be the *distance between x and y*. Then,

(M_1) $d(x,y) \geq 0$; $d(x,y) = 0$ if and only if $x = y$.
(M_2) $d(x,y) = d(y,x)$.
(M_3) $d(x,y) \leq d(x,z) + d(z,y)$ (triangle inequality).

Any set which has a distance function $d(x,y)$ satisfying (M_1), (M_2), and (M_3) is called a *metric space* with *metric d*. Therefore, a Banach space is a metric space whose metric is the norm. Moreover, by (B_3), a Banach space X is a complete metric space; that is, if $d(x_k,x_m) \to 0$ as $k,m \to \infty$, then there is an $x \in X$ such that $d(x_k,x) \to 0$.

(8.14) Theorem *For* $1 \leq p \leq \infty$, $L^p(E)$ *is a Banach space with norm* $\|f\| = \|f\|_{p,E}$.

Proof. Parts (B_1) and (B_2) in the definition of a Banach space are clearly fulfilled by $L^p(E)$, parts (a) and (c) of (B_2) being (5.11) and Minkowski's inequality, respectively. [Regarding part (a), we do not distinguish between two L^p functions which are equal a.e.; thus, the zero element of $L^p(E)$ means any function equal to zero a.e. in E.]

To verify (B_3), suppose that $\{f_k\}$ is a Cauchy sequence in $L^p(E)$. If $p = \infty$, then $|f_k - f_m| \leq \|f_k - f_m\|_\infty$ except for a set $Z_{k,m}$ of measure zero. If $Z = \bigcup_{k,m} Z_{k,m}$, then Z has measure zero, and $|f_k - f_m| \leq \|f_k - f_m\|_\infty$ outside Z for all k and m. Hence, $\{f_k\}$ converges uniformly outside Z to a bounded limit f, and it follows that $\|f_k - f\|_\infty \to 0$. (Note that convergence in L^∞ is equivalent to uniform convergence outside a set of measure zero.)

In case $1 \leq p < \infty$, Tschebyshev's inequality (5.49) implies that

$$|\{\mathbf{x} \in E : |f_k(\mathbf{x}) - f_m(\mathbf{x})| > \varepsilon\}| \leq \varepsilon^{-p} \int_E |f_k - f_m|^p.$$

Hence, $\{f_k\}$ is a Cauchy sequence in measure. By (4.22) and (4.23), there is a subsequence $\{f_{k_j}\}$ and a function f such that $f_{k_j} \to f$ a.e. in E. Given $\varepsilon > 0$, there is a K such that

$$\left(\int_E |f_{k_j} - f_k|^p \right)^{1/p} = \|f_{k_j} - f_k\|_p < \varepsilon \text{ if } k_j,k > K.$$

Letting $k_j \to \infty$, we obtain by Fatou's lemma that $\|f - f_k\|_p \leq \varepsilon$ if $k > K$. Hence, $\|f - f_k\|_p \to 0$ as $k \to \infty$. Finally, since $\|f\|_p \leq \|f - f_k\|_p + \|f_k\|_p < +\infty$, it follows that $f \in L^p(E)$, which completes the proof.

A metric space X is said to be *separable* if it has a countable dense subset; that is, X is separable if there exists a countable set $\{x_k\}$ in X with the property that for every $x \in X$ and every $\varepsilon > 0$, there is an x_k with $d(x,x_k) < \varepsilon$. In the next theorem, we will show that L^p is separable if $1 \leq p < \infty$. Note that L^∞ is not separable: take $L^\infty(0,1)$, for example, and consider the functions $f_t(x) = \chi_{(0,t)}(x)$, $0 < t < 1$. There are an uncountable number of these, and $\|f_t - f_{t'}\|_\infty = 1$ if $t \neq t'$. See also Exercise 10.

(8.15) Theorem *If $1 \le p < \infty$, $L^p(E)$ is separable.*

Proof. Suppose first that $E = \mathbf{R}^n$, and consider a class of dyadic cubes in $\mathbf{R}^n$. Let D be the set of all (finite) linear combinations of characteristic functions of these cubes, the coefficients being complex numbers with rational real and imaginary parts. Then D is a countable subset of $L^p(\mathbf{R}^n)$. To see that D is dense, use the method of successively approximating more and more general functions: First, consider characteristic functions of open sets [every open set is the countable union of nonoverlapping dyadic cubes by (1.11)], of G_δ sets, and of measurable sets with finite measure; then consider simple functions whose supports have finite measure, nonnegative functions in $L^p(\mathbf{R}^n)$, and, finally, arbitrary functions in $L^p(\mathbf{R}^n)$. The details are left to the reader [cf. lemma (7.3)]. This proves the case $E = \mathbf{R}^n$.

For an arbitrary measurable E, let D' denote the restrictions to E of the functions in D. Then D' is dense in $L^p(E)$, $1 \le p < \infty$. In fact, given p and $f \in L^p(E)$, let $f_1 = f$ on E and $f_1 = 0$ off E. Then $f_1 \in L^p(\mathbf{R}^n)$, so that given $\varepsilon > 0$, there exists $g \in D$ with $(\int_{\mathbf{R}^n} |f_1 - g|^p \, d\mathbf{x})^{1/p} < \varepsilon$. Therefore, $(\int_E |f - g|^p d\mathbf{x})^{1/p} < \varepsilon$. This shows that D' is dense in $L^p(E)$ and completes the proof.

As we have already noted, Minkowski's inequality fails when $0 < p < 1$. Therefore, $\|\cdot\|_{p,E}$ is not a norm for such p. However, we still have the following facts.

(8.16) Theorem *If $0 < p < 1$, $L^p(E)$ is a complete, separable metric space, with distance defined by*

$$d(f,g) = \|f - g\|_{p,E}^p.$$

Proof. With $d(f,g)$ so defined, properties (M_1) and (M_2) of a metric space are clear. To verify (M_3), which is the triangle inequality, we first claim that

$$(a + b)^p \le a^p + b^p \text{ if } a,b \ge 0,\ 0 < p \le 1.$$

If both a and b are zero, this is obvious. If, say, $a \neq 0$, then dividing by a^p, we reduce the inequality to $(1 + t)^p \le 1 + t^p$, $t > 0$ $(t = b/a)$. This is clear since both sides are equal when $t = 0$ and the derivative of the right side majorizes that of the left for $t > 0$.

It follows that $|f(\mathbf{x}) - g(\mathbf{x})|^p \le |f(\mathbf{x}) - h(\mathbf{x})|^p + |h(\mathbf{x}) - g(\mathbf{x})|^p$ if $0 < p \le 1$. Integrating, we obtain $\|f - g\|_p^p \le \|f - h\|_p^p + \|h - g\|_p^p$, which is just the triangle inequality. The proofs that L^p is complete and separable with respect to $\|\cdot\|_p$ are the same as in (8.14) and (8.15).

It is worth noting that the triangle inequality is equivalent to the basic estimate

$$\|f + g\|_p^p \le \|f\|_p^p + \|g\|_p^p \qquad (0 < p \le 1). \tag{8.17}$$

The analogous results for series are listed in the following theorem.

(8.18) Theorem

(i) *If* $1 \le p \le \infty$, l^p *is a Banach space with* $\|a\| = \|a\|_p$. *For* $1 \le p < \infty$, l^p *is separable*; l^∞ *is not separable.*

(ii) *If* $0 < p < 1$, l^p *is a complete, separable metric space, with distance* $d(a,b) = \|a - b\|_p^p$.

Proof. We will show that l^p is complete and separable when $1 \le p < \infty$ and that l^∞ is not separable. The rest of the proof of (i) and the proof of (ii) are left to the reader.

Suppose that $1 \le p < \infty$, $a^{(i)} = \{a_k^{(i)}\} \in l^p$ for $i = 1, 2, \ldots,$ and $\|a^{(i)} - a^{(j)}\|_p \to 0$ as $i,j \to \infty$. Since $\|a^{(i)} - a^{(j)}\|_p \ge |a_k^{(i)} - a_k^{(j)}|$ for every k, it follows that $|a_k^{(i)} - a_k^{(j)}| \to 0$ for every k as $i,j \to \infty$. Let $a_k = \lim_{i\to\infty} a_k^{(i)}$ and $a = \{a_k\}$. We will show that $a \in l^p$ and $\|a^{(i)} - a\|_p \to 0$. Given $\varepsilon > 0$, there exists N such that

$$(\sum_k |a_k^{(i)} - a_k^{(j)}|^p)^{1/p} = \|a^{(i)} - a^{(j)}\|_p < \varepsilon \quad \text{if} \quad i,j > N.$$

Restricting the summation to $k \le M$ and letting $j \to \infty$, we obtain

$$\left(\sum_{k=1}^{M} |a_k^{(i)} - a_k|^p\right)^{1/p} \le \varepsilon \text{ for any } M, \text{ if } i > N.$$

Letting $M \to \infty$, we get $\|a^{(i)} - a\|_p \le \varepsilon$ if $i > N$; that is, $\|a^{(i)} - a\|_p \to 0$. The fact that $\|a\|_p \le \|a - a^{(i)}\|_p + \|a^{(i)}\|_p$ shows that $a \in l^p$.

To prove that l^p is separable when $p < \infty$, let D be the set of all sequences $\{d_k\}$ such that (a) the real and imaginary parts of d_k are rational, and (b) $d_k = 0$ for $k \ge N$ (N may vary from sequence to sequence). Then D is a countable subset of l^p. If $a = \{a_k\} \in l^p$ and $\varepsilon > 0$, choose N so that $\sum_{k=N+1}^{\infty} |a_k|^p < \varepsilon/2$. Choose $d_1, \ldots, d_N$ with rational real and imaginary parts such that $\sum_{k=1}^{N} |a_k - d_k|^p < \varepsilon/2$. Then $d = \{d_1, \ldots, d_N, 0, \ldots\}$ belongs to D and $\|a - d\|_p^p < \varepsilon$. It follows that D is dense in l^p, and therefore that l^p is separable.

To see that l^∞ is not separable, consider the subclass of sequences $a = \{a_k\}$ for which each a_k is 0 or 1. The number of such sequences is uncountable, and $\|a - a'\|_\infty = 1$ for any two different such sequences. Hence, l^∞ cannot be separable.

We know from Lusin's theorem that measurable functions have continuity properties. The next theorem gives a useful continuity property of functions in L^p.

(8.19) Theorem (*Continuity in L^p*) *If* $f \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$, *then*

$$\lim_{|\mathbf{h}|\to 0} \|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})\|_p = 0.$$

Proof. Let C_p denote the class of $f \in L^p$ such that $\|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})\|_p \to 0$ as $|\mathbf{h}| \to 0$. We claim that (a) a finite linear combination of functions in C_p is in C_p, and (b) if $f_k \in C_p$ and $\|f_k - f\|_p \to 0$, then $f \in C_p$. Both of these facts follow easily from Minkowski's inequality; for (b), for example, note that

$$\begin{aligned}
&\|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})\|_p \\
&\leq \|f(\mathbf{x} + \mathbf{h}) - f_k(\mathbf{x} + \mathbf{h})\|_p + \|f_k(\mathbf{x} + \mathbf{h}) - f_k(\mathbf{x})\|_p + \|f_k(\mathbf{x}) - f(\mathbf{x})\|_p \\
&= \|f_k(\mathbf{x} + \mathbf{h}) - f_k(\mathbf{x})\|_p + 2\|f_k(\mathbf{x}) - f(\mathbf{x})\|_p.
\end{aligned}$$

Since $f_k \in C_p$, we have $\limsup_{|\mathbf{h}| \to 0} \|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})\|_p \leq 2\|f_k(\mathbf{x}) - f(\mathbf{x})\|_p$, and (b) follows by letting $k \to \infty$.

Clearly, the characteristic function of a cube belongs to C_p. Hence, in view of the fact that linear combinations of characteristic functions of cubes are dense in $L^p(\mathbf{R}^n)$ [see (8.15)], it follows from (a) and (b) that L^p is contained in C_p.

We remark without proof that this result is also true for $0 < p < 1$. (Use the same ideas for $\|\cdot\|_p^p$.) It fails, however, for $p = \infty$, as shown by the function $\chi = \chi_{(0,\infty)}(x)$ on $(-\infty, +\infty)$. In fact, $\chi \in L^\infty(-\infty, +\infty)$ but $\|\chi(x + h) - \chi(x)\|_\infty = 1$ for all $h \neq 0$.

5. The Space L^2; Orthogonality

For complex-valued measurable f, $f = f_1 + if_2$ with f_1 and f_2 real-valued and measurable, we have $\int_E f = \int_E f_1 + i\int_E f_2$ (see p. 76). We will use the fact that $|\int_E f| \leq \int_E |f|$. (See Exercise 1.)

Among the L^p spaces, L^2 has the special property that the product of any two of its elements is integrable (Schwarz's inequality). This simple fact leads to some important extra structure in L^2 which we will now discuss.

Consider $L^2 = L^2(E)$, where E is a fixed subset of $\mathbf{R}^n$ of positive measure, and write $\|f\| = \|f\|_{2,E}$, $\int_E f = \int f$, etc. For $f, g \in L^2$, define the *inner product* of f and g by

(8.20) $$\langle f,g \rangle = \int f\bar{g},$$

where $\bar{g}$ denotes the complex conjugate of g. Note that by Schwarz's inequality,

$$|\langle f,g \rangle| \leq \|f\|\|g\|.$$

Moreover, the inner product has the following properties:

(a) $\langle g,f \rangle = \overline{\langle f,g \rangle}$.
(b) $\langle f_1 + f_2, g \rangle = \langle f_1,g \rangle + \langle f_2,g \rangle$, $\langle f, g_1 + g_2 \rangle = \langle f,g_1 \rangle + \langle f,g_2 \rangle$.
(c) $\langle \alpha f,g \rangle = \alpha\langle f,g \rangle$, $\langle f,\alpha g \rangle = \bar{\alpha}\langle f,g \rangle$, $\alpha \in \mathbf{C}$.
(d) $\langle f,f \rangle^{1/2} = \|f\|$.

If $\langle f,g \rangle = 0$, then f and g are said to be *orthogonal*. A set $\{\phi_\alpha\}_{\alpha \in A}$ is *orthogonal* if any two of its elements are orthogonal; $\{\phi_\alpha\}$ is *orthonormal* if it is orthogonal and $\|\phi_\alpha\| = 1$ for all α. Note that if $\{\phi_\alpha\}$ is orthogonal and

$\|\phi_\alpha\| \neq 0$ for every α, then $\{\phi_\alpha/\|\phi_\alpha\|\}$ is orthonormal. Henceforth, we will assume that $\|\phi_\alpha\| \neq 0$ for all α for an orthogonal system $\{\phi_\alpha\}$. This implies that no element is zero and that no two elements are equal.

(8.21) Theorem *Any orthogonal system $\{\phi_\alpha\}$ in L^2 is countable.*

Proof. We may assume that $\{\phi_\alpha\}$ is orthonormal. Then for $\alpha \neq \beta$, we have

$$\|\phi_\alpha - \phi_\beta\|^2 = \int (\phi_\alpha - \phi_\beta)(\bar{\phi}_\alpha - \bar{\phi}_\beta) = \|\phi_\alpha\|^2 + \|\phi_\beta\|^2 = 2,$$

so that $\|\phi_\alpha - \phi_\beta\| = \sqrt{2}$. Since L^2 is separable, it follows that $\{\phi_\alpha\}$ must be countable.

A collection $\psi_1, \ldots, \psi_N$ is said to be *linearly independent* if $\sum_{k=1}^N a_k\psi_k(\mathbf{x}) = 0$ (a.e.) implies that every a_k is zero. An infinite collection of functions is called *linearly independent* if each finite subcollection is. No function in a linearly independent set can be zero a.e.

(8.22) Theorem *If $\{\psi_k\}$ is orthogonal, it is linearly independent.*

Proof. Suppose that $a_1\psi_{k_1} + \cdots + a_N\psi_{k_N} = 0$. Multiplying both sides by $\bar{\psi}_{k_1}$ and integrating, we obtain by orthogonality that $a_1 = 0$. Similarly, $a_2 = \cdots = a_N = 0$.

The converse of (8.22) is not true. However, the next result shows that if $\{\psi_k\}$ is linearly independent, then the system formed from suitable linear combinations of its elements is orthogonal.

(8.23) Theorem (*Gram-Schmidt Process*) *If $\{\psi_k\}$ is linearly independent, then the system $\{\phi_k\}$ defined by*

$$\begin{aligned}
\phi_1 &= \psi_1 \\
\phi_2 &= a_{21}\psi_1 + \psi_2 \\
&\vdots \\
\phi_k &= a_{k1}\psi_1 + \cdots + a_{k,k-1}\psi_{k-1} + \psi_k \\
&\vdots
\end{aligned}$$

is orthogonal for proper selection of the a_{ij}.

Proof. Having $\phi_1 = \psi_1$, we proceed by induction, assuming that $\phi_1, \ldots, \phi_{k-1}$ have been chosen as required. We will determine constants $b_{k1}, \ldots, b_{k,k-1}$ so that the function ϕ_k defined by

$$\phi_k = b_{k1}\phi_1 + \cdots + b_{k,k-1}\phi_{k-1} + \psi_k$$

is orthogonal to $\phi_1, \ldots, \phi_{k-1}$. If $j < k$,

$$\langle \phi_k, \phi_j \rangle = b_{kj}\langle \phi_j, \phi_j \rangle + \langle \psi_k, \phi_j \rangle$$

by orthogonality. Since $\langle \phi_j, \phi_j \rangle \neq 0$, b_{kj} can be chosen so that $\langle \phi_k, \phi_j \rangle = 0$,

$j < k$. Since each ϕ_j with $j < k$ is a linear combination of $\psi_1, \ldots, \psi_j$, the theorem follows.

When the ϕ_k are selected by the Gram-Schmidt process, we shall say that they are *generated* from the ψ_k. Note that the triangular character of the matrix in (8.23) means that each ψ_k can also be written as a linear combination of the ϕ_j, $j \le k$.

We call an orthogonal system $\{\phi_k\}$ *complete* if the only function which is orthogonal to every ϕ_k is zero; that is, $\{\phi_k\}$ is complete if $\langle f,\phi_k\rangle = 0$ for all k implies that $f = 0$ a.e. Thus, a complete orthogonal system is one which is maximal in the sense that it is not properly contained in any larger orthogonal system.

The *span* of a set of functions $\{\psi_k\}$ is the collection of all finite linear combinations of the ψ_k. In speaking of the span of $\{\psi_k\}$, we may always assume that $\{\psi_k\}$ is orthogonal by discarding any dependent functions and applying the Gram-Schmidt process to the resulting linearly independent set.

A set $\{\psi_k\}$ is called a *basis* for L^2 if its span is dense in L^2; that is, $\{\psi_k\}$ is a basis if given $f \in L^2$ and $\varepsilon > 0$, there exist N and $\{a_k\}$ such that $\|f - \sum_{k=1}^N a_k\psi_k\| < \varepsilon$. The a_k can always be chosen with rational real and imaginary parts. Any countable dense set in L^2 is of course a basis. It follows that L^2 has an orthogonal basis.

(8.24) Theorem *Any orthogonal basis in L^2 is complete. In particular, there exists a complete orthonormal basis for L^2.*

Proof. Let $\{\psi_k\}$ be an orthogonal basis for L^2. We may assume that $\{\psi_k\}$ is orthonormal. To show that it is complete, let $\langle f,\psi_k\rangle = 0$ for all k. Then $\langle f,f\rangle = \langle f, f - \sum_{k=1}^N a_k\psi_k\rangle$ for any N and a_k. By Schwarz's inequality, $|\langle f,f\rangle| \le \|f\|\|f - \sum_{k=1}^N a_k\psi_k\|$, and so, since the term on the right can be chosen arbitrarily small, $\langle f,f\rangle = 0$. Therefore, $f = 0$ a.e., which completes the proof.

6. Fourier Series; Parseval's Formula

Let $\{\phi_k\}$ be any orthonormal system for L^2. If $f \in L^2$, the numbers defined by

$$c_k = c_k(f) = \langle f,\phi_k\rangle = \int_E f\bar{\phi}_k$$

are called the *Fourier coefficients* of f with respect to $\{\phi_k\}$. The series $\sum_k c_k\phi_k$ is called the *Fourier series* of f with respect to $\{\phi_k\}$, and denoted $S[f] = \sum_k c_k\phi_k$. We also write

$$f \sim \sum_k c_k\phi_k.$$

The first question we ask is how well $S[f]$, or more precisely, the sequence of its partial sums, approximates f. Fix N and let $L = \sum_{k=1}^{N} \gamma_k \phi_k$ be a linear combination of $\phi_1, \ldots, \phi_N$. We wish to know what choice of $\gamma_1, \ldots, \gamma_N$ makes $\|f - L\|$ a minimum. Note that since $\{\phi_k\}$ is orthonormal, $\|L\|^2 = \langle L,L\rangle = \sum_{k=1}^{N} |\gamma_k|^2$. Hence,

$$\begin{aligned}\|f - L\|^2 &= \int\left(f - \sum_{k=1}^{N} \gamma_k\phi_k\right)\left(\bar{f} - \sum_{k=1}^{N} \bar{\gamma}_k\bar{\phi}_k\right)\\ &= \|f\|^2 - \sum_{k=1}^{N} (\bar{\gamma}_k c_k + \gamma_k \bar{c}_k) + \sum_{k=1}^{N} |\gamma_k|^2,\end{aligned}$$

where the c_k are the Fourier coefficients of f. Since

$$|c_k - \gamma_k|^2 = (c_k - \gamma_k)(\bar{c}_k - \bar{\gamma}_k) = |c_k|^2 - (\bar{\gamma}_k c_k + \gamma_k \bar{c}_k) + |\gamma_k|^2,$$

we obtain

$$\|f - L\|^2 = \|f\|^2 + \sum_{k=1}^{N} |c_k - \gamma_k|^2 - \sum_{k=1}^{N} |c_k|^2.$$

Therefore,

$$\textbf{(8.25)} \qquad \min_{\gamma_1,\cdots,\gamma_N} \|f - L\|^2 = \|f\|^2 - \sum_{k=1}^{N} |c_k|^2;$$

that is, the minimum is achieved when $\gamma_k = c_k$ for $k = 1, \ldots, N$, or equivalently, when L is the Nth partial sum of $S[f]$. Writing $s_N = s_N(f) = \sum_{k=1}^{N} c_k\phi_k$, we have from (8.25) that

$$\textbf{(8.26)} \qquad \|f - s_N\|^2 = \|f\|^2 - \sum_{k=1}^{N} |c_k|^2.$$

(8.27) Theorem *Let $\{\phi_k\}$ be an orthonormal system in L^2, and let $f \in L^2$.*

(i) *Of all linear combinations $\sum_1^N \gamma_k\phi_k$ with N fixed, the best L^2-approximation to f is given by the partial sum $s_N = \sum_1^N c_k\phi_k$ of the Fourier series of f.*

(ii) (*Bessel's inequality*) *The sequence $\{c_k\}$ of Fourier coefficients of f belongs to l^2 and*

$$\left(\sum_{k=1}^{\infty} |c_k|^2\right)^{1/2} \le \|f\|.$$

Proof. Part (i) has been proved. Note that since $\|f - s_N\|^2 \ge 0$, Bessel's inequality follows from (8.26) by letting $N \to \infty$.

If f is a function for which equality holds in Bessel's inequality, that is, if

$$\textbf{(8.28)} \qquad \left(\sum_{k=1}^{\infty} |c_k|^2\right)^{1/2} = \|f\|,$$

then f is said to satisfy *Parseval's formula*. From (8.26), we immediately obtain the next result.

(8.29) Theorem *Parseval's formula holds for f if and only if $\|s_N - f\| \to 0$, i.e., if and only if $S[f]$ converges to f in L^2 norm.*

The following theorem is of great importance.

(8.30) Theorem (*Riesz-Fischer Theorem*) *Let $\{\phi_k\}$ be any orthonormal system, and let $\{c_k\}$ be any sequence in l^2. Then there is an $f \in L^2$ such that $S[f] = \sum c_k\phi_k$, i.e., such that $\{c_k\}$ is the sequence of Fourier coefficients of f with respect to $\{\phi_k\}$. Moreover, f can be chosen to satisfy Parseval's formula.*

Proof. Let $t_N = \sum_{k=1}^N c_k\phi_k$. Then if $M < N$,

$$\|t_N - t_M\|^2 = \left\| \sum_{M+1}^N c_k\phi_k \right\|^2 = \sum_{M+1}^N |c_k|^2.$$

The fact that $\{c_k\} \in l^2$ implies that $\{t_N\}$ is a Cauchy sequence in L^2. Since L^2 is complete, there is an $f \in L^2$ such that $\|f - t_N\| \to 0$. If $N \geq k$,

$$\int f\overline{\phi}_k = \int (f - t_N)\overline{\phi}_k + \int t_N\overline{\phi}_k = \int (f - t_N)\overline{\phi}_k + c_k.$$

Since the integral on the right is bounded in absolute value by $\|f - t_N\|\|\phi_k\| = \|f - t_N\|$, we obtain by letting $N \to \infty$ that $\int f\overline{\phi}_k = c_k$. Thus, $S[f] = \sum_k c_k\phi_k$, so that $t_N = s_N(f)$, and it follows from (8.29) that Parseval's formula holds for f. This completes the proof.

There is no guarantee that the Fourier coefficients of a function uniquely determine the function. However, if $\{\phi_k\}$ is complete, we can show that the correspondence between a function and its Fourier coefficients is unique; that is, if f and g have the same Fourier coefficients with respect to a complete system, then $f = g$ a.e. This is simple, since the vanishing of all the Fourier coefficients of $f - g$ implies that $f - g = 0$ a.e. An important related fact is the following.

(8.31) Theorem *Let $\{\phi_k\}$ be an orthonormal system. Then $\{\phi_k\}$ is complete if and only if Parseval's formula holds for every $f \in L^2$.*

Proof. Suppose that $\{\phi_k\}$ is complete. If $f \in L^2$, Bessel's inequality implies that its Fourier coefficients $\{c_k\}$ belong to l^2. Hence, by the Riesz-Fischer theorem, there exists a g in L^2 with $S[g] = \sum c_k\phi_k$ and $\|g\| = (\sum |c_k|^2)^{1/2}$. Since f and g have the same Fourier coefficients and $\{\phi_k\}$ is complete, we see that $f = g$ a.e. Hence, $\|f\| = \|g\| = (\sum |c_k|^2)^{1/2}$, which is Parseval's formula.

Conversely, suppose that Parseval's formula holds with respect to $\{\phi_k\}$ for every $f \in L^2$. If $\langle f,\phi_k \rangle = 0$ for all k, then $\|f\| = (\sum |\langle f,\phi_k \rangle|^2)^{1/2} = 0$. Therefore, $f = 0$ a.e., which proves that $\{\phi_k\}$ is complete.

Suppose that $\{\phi_k\}$ is orthonormal and complete and that $f,g \in L^2$. Let

$c_k = \langle f,\phi_k \rangle$, $d_k = \langle g,\phi_k \rangle$, $c = \{c_k\}$, $d = \{d_k\}$, $\|c\| = (\sum |c_k|^2)^{1/2}$ and $(c,d) = \sum c_k \bar{d}_k$. We claim that

$$\langle f,g \rangle = (c,d). \tag{8.32}$$

To prove this, observe that by Parseval's formula, $\langle f + g, f + g \rangle = (c + d, c + d)$, or

$$\langle f,f \rangle + \langle g,g \rangle + 2 \operatorname{Re} \langle f,g \rangle = (c,c) + (d,d) + 2 \operatorname{Re} (c,d),$$

where $\operatorname{Re} z$ denotes the real part of z. Cancelling equal terms, we see that $\operatorname{Re} \langle f,g \rangle = \operatorname{Re} (c,d)$. Applying this to the function if gives $\operatorname{Re} \langle if,g \rangle = \operatorname{Re} (ic,d)$. But $\operatorname{Re} \langle if,g \rangle = \operatorname{Re} i\langle f,g \rangle = -\operatorname{Im} \langle f,g \rangle$. Similarly, $\operatorname{Re} (ic,d) = -\operatorname{Im} (c,d)$. Therefore, $\operatorname{Im} \langle f,g \rangle = \operatorname{Im} (c,d)$, and (8.32) is proved.

Another corollary of (8.31) is given in the next result. First, we make several definitions. Let X_1 and X_2 be metric spaces with metrics d_1 and d_2, respectively. Then X_1 and X_2 are said to be *isometric* if there is a mapping T of X_1 onto X_2 such that

$$d_1(f,g) = d_2(Tf,Tg)$$

for all $f,g \in X_1$. Such a T is called an *isometry*. Thus, an isometry is a mapping which preserves distances. An isometry is automatically one-to-one, and two isometric metric spaces may be regarded as the same space with a relabeling of the points. For example, two L^2 spaces, $L^2(E)$ and $L^2(E')$, are isometric if there is a mapping T of $L^2(E)$ onto $L^2(E')$ such that $\|f - g\|_{2,E} = \|Tf - Tg\|_{2,E'}$ for all $f,g \in L^2(E)$. The isometries we shall encounter will be *linear*, i.e., will satisfy

$$T(\alpha f + \beta g) = \alpha Tf + \beta Tg \text{ for all scalars } \alpha, \beta.$$

If T is a linear map of $L^2(E)$ onto $L^2(E')$, then since $Tf - Tg = T(f - g)$, it follows that T is an isometry if and only if

$$\|f\|_{2,E} = \|Tf\|_{2,E'}$$

for all $f \in L^2(E)$. Similarly, a linear map of $L^2(E)$ onto l^2 is an isometry if and only if $\|f\|_{2,E} = \|Tf\|_{l^2}$ for all $f \in L^2(E)$.

(8.33) Theorem *All spaces $L^2(E)$ are linearly isometric with l^2, and so with one another.*

Proof. For a given E, define a linear correspondence between $L^2(E)$ and l^2 by choosing a complete orthonormal system $\{\phi_k\}$ in $L^2(E)$ and mapping an $f \in L^2(E)$ onto the sequence $\{\langle f,\phi_k \rangle\}$ of its Fourier coefficients. This mapping is onto all of l^2 by the Riesz-Fischer theorem, and is an isometry by (8.31).

7. Hilbert Spaces

A set H is called a *Hilbert space over the complex numbers* **C** if it satisfies the following three conditions:

(H_1) H is a vector space over **C**; that is, if $f,g \in H$ and $\alpha \in \mathbf{C}$, then $f + g \in H$ and $\alpha f \in H$.

(H_2) For every pair $f,g \in H$, there is a complex number (f,g), called the *inner product* of f and g, which satisfies

(a) $(g,f) = \overline{(f,g)}$.
(b) $(f_1 + f_2, g) = (f_1,g) + (f_2,g)$.
(c) $(\alpha f,g) = \alpha(f,g)$ for $\alpha \in \mathbf{C}$.
(d) $(f,f) \geq 0$, and $(f,f) = 0$ if and only if $f = 0$.

Notice that (a), (b) and (c) imply that $(f, g_1 + g_2) = (f,g_1) + (f,g_2)$ and $(f,\alpha g) = \bar{\alpha}(f,g)$. Define

$$\|f\| = (f,f)^{1/2}.$$

Before stating the third condition, we claim that

(8.34) $$|(f,g)| \leq \|f\|\|g\| \qquad (\textit{Schwarz's inequality}).$$

If $\|g\| = 0$, this is obvious. Otherwise, letting $\lambda = -(f,g)/\|g\|^2$, we obtain

$$0 \leq (f + \lambda g, f + \lambda g) = \|f\|^2 - 2\frac{|(f,g)|^2}{\|g\|^2} + \frac{|(f,g)|^2}{\|g\|^2} = \|f\|^2 - \frac{|(f,g)|^2}{\|g\|^2},$$

and Schwarz's inequality follows at once.

We will show that $\|\cdot\|$ is a norm on H by proving the triangle inequality. In fact,

$$\|f + g\|^2 = (f + g, f + g) = \|f\|^2 + 2\,\mathrm{Re}\,(f,g) + \|g\|^2.$$

Since $|\mathrm{Re}\,(f,g)| \leq |(f,g)| \leq \|f\|\|g\|$, it follows that the right side is at most $(\|f\| + \|g\|)^2$. Taking square roots, we obtain $\|f + g\| \leq \|f\| + \|g\|$, as desired. Hence, H is a normed linear space.

We also require

(H_3) H is complete with respect to $\|\cdot\|$.

In particular, a Hilbert space is a Banach space.

As for L^2 spaces, a linear map T of a Hilbert space H onto a Hilbert space H' is an isometry if and only if $\|f\|_H = \|Tf\|_{H'}$ for all $f \in H$.

A Hilbert space is called *infinite dimensional* if it cannot be spanned by a finite number of elements; hence, an infinite dimensional Hilbert space has an infinite linearly independent subset. The space L^2 with inner product

$(f,g) = \int f\bar{g}$ and the space l^2 with $(c,d) = \sum_k c_k\bar{d}_k$ are examples of separable infinite dimensional Hilbert spaces. In fact, there are essentially no other examples, as the following theorem shows.

(8.35) Theorem *All separable infinite dimensional Hilbert spaces are linearly isometric with l^2, and so with one another.*

Proof. The proof is a repetition of the ideas leading to (8.33), so we shall be brief. Let H be a separable infinite dimensional Hilbert space, and let $\{e_k'\}$ be a countable dense subset. Discarding those e_k' which are spanned by other e_i', we obtain a linearly independent set $\{e_k\}$ with the same dense span as $\{e_k'\}$. Since H is infinite dimensional, $\{e_k\}$ is infinite. Using the Gram-Schmidt process, we may assume that $\{e_k\}$ is orthonormal: $(e_i,e_k) = 0$ for $i \neq k$ and $\|e_k\| = 1$ for all k. It follows that $\{e_k\}$ is complete; in fact, if $(f,e_k) = 0$ for all k, then

$$\left\| f - \sum_{k=1}^{N} a_k e_k \right\|^2 = \|f\|^2 + \sum_{k=1}^{N} |a_k|^2 \geq \|f\|^2.$$

If f were not zero, the span of the e_k could not be dense. Hence, $f = 0$, which shows that H has a complete orthonormal system $\{e_k\}$.

Next, we will show that Bessel's inequality and the Riesz-Fischer theorem hold for $\{e_k\}$. If $f \in H$, let $c_k = (f,e_k)$. Then

$$0 \leq \left\| f - \sum_{k=1}^{N} c_k e_k \right\|^2 = \|f\|^2 - \sum_{k=1}^{N} |c_k|^2.$$

Letting $N \to \infty$, we obtain Bessel's inequality $(\sum |c_k|^2)^{1/2} \leq \|f\|$. In particular, $\{c_k\}$ belongs to l^2.

To derive the Riesz-Fischer theorem, let $\{\gamma_k\}$ be a sequence in l^2, and set $t_N = \sum_{k=1}^{N} \gamma_k e_k$. Then

$$\|t_M - t_N\|^2 = \sum_{k=M+1}^{N} |\gamma_k|^2 \to 0 \text{ as } M,N \to \infty,\ M < N.$$

Since H is complete, there is a $g \in H$ such that $\|g - t_N\| \to 0$. We have

$$(g,e_k) = (g - t_N, e_k) + (t_N,e_k) = (g - t_N, e_k) + \gamma_k \qquad (k \leq N).$$

Letting $N \to \infty$, it follows from Schwarz's inequality that $(g,e_k) = \gamma_k$. Hence, $t_N = s_N(g)$ and $\|g - t_N\|^2 = \|g\|^2 - \sum_{k=1}^{N} |\gamma_k|^2$. Letting $N \to \infty$ in the last equation, we see that g satisfies Parseval's formula $\|g\| = (\sum |\gamma_k|^2)^{1/2}$. This gives the analogue of the Riesz-Fischer theorem.

Now let $f \in H$, and let $c_k = (f,e_k)$. Taking $\gamma_k = c_k$ above, we see by the completeness of $\{e_k\}$ that $g = f$, so that Parseval's formula holds: $\|f\| = (\sum |c_k|^2)^{1/2}$. The fact that H is linearly isometric with l^2 now follows as in the proof of (8.33).

Exercises

1. For complex-valued, measurable f, $f = f_1 + if_2$ with f_1 and f_2 real-valued and measurable, we have $\int_E f = \int_E f_1 + i\int_E f_2$. Prove that $\int_E f$ is finite if and only if $\int_E |f|$ is finite, and $|\int_E f| \le \int_E |f|$. (Note that $|\int_E f| = [(\int_E f_1)^2 + (\int_E f_2)^2]^{1/2}$, and use the fact that $(a^2+b^2)^{1/2} = a\cos\alpha + b\sin\alpha$ for an appropriate α, while $(a^2+b^2)^{1/2} \ge |a\cos\alpha + b\sin\alpha|$ for all α.)
2. Prove the converse of Hölder's inequality for $p = 1$ and ∞. Show also that for real-valued $f \notin L^p(E)$, there exists a function $g \in L^{p'}(E)$, $1/p + 1/p' = 1$, such that $fg \notin L^1(E)$. (Construct g of the form $\sum a_k g_k$ for appropriate a_k and g_k, with g_k satisfying $\int_E fg_k \to +\infty$.)
3. Prove Theorems (8.12) and (8.13). Show that Minkowski's inequality for series fails when $p < 1$.
4. Let f and g be real-valued, and let $1 < p < \infty$. Prove that equality holds in the inequality $|\int fg| \le \|f\|_p \|g\|_{p'}$ if and only if fg has constant sign a.e. and $|f|^p$ is a multiple of $|g|^{p'}$ a.e.

   If $\|f+g\|_p = \|f\|_p + \|g\|_p$ and $g \ne 0$ in Minkowski's inequality, show that f is a multiple of g.
5. For $1 \le p < \infty$ and $0 < |E| < +\infty$, define
$$N_p[f] = \left(\frac{1}{|E|}\int_E |f|^p\right)^{1/p}.$$
   Prove that if $p_1 < p_2$, then $N_{p_1}[f] \le N_{p_2}[f]$.

   Prove also that $N_p[f+g] \le N_p[f] + N_p[g]$, $(1/|E|)\int_E |fg| \le N_p[f]N_{p'}[g]$, $1/p + 1/p' = 1$, and that $\lim_{p\to\infty} N_p[f] = \|f\|_\infty$. Thus, N_p behaves like $\|\cdot\|_p$, but has the advantage of being monotone in p.
6. Prove the following generalization of Hölder's inequality. If $\sum_{i=1}^k 1/p_i = 1/r$, $p_i, r \ge 1$, then
$$\|f_1 \cdots f_k\|_r \le \|f_1\|_{p_1} \cdots \|f_k\|_{p_k}.$$
   (See also Exercise 12 of Chapter 7.)
7. Show that when $0 < p < 1$, the neighborhoods $\{f : \|f\|_p < \varepsilon\}$ of zero in $L^p(0,1)$ are not convex. (Let $f = \chi_{(0,\varepsilon^p)}$ and $g = \chi_{(\varepsilon^p, 2\varepsilon^p)}$. Show that $\|f\|_p = \|g\|_p = \varepsilon$, but that $\|\frac{1}{2}f + \frac{1}{2}g\|_p > \varepsilon$.)
8. Prove the following *integral version of Minkowski's inequality* for $1 \le p < \infty$:
$$\left[\int\left|\int f(x,y)\,dx\right|^p dy\right]^{1/p} \le \int\left[\int |f(x,y)|^p\,dy\right]^{1/p} dx.$$
   [For $1 < p < \infty$, note that the pth power of the left-hand side is majorized by $\iint [\int |f(z,y)|\,dz]^{p-1} |f(x,y)|\,dx\,dy$. Integrate first with respect to y and apply Hölder's inequality.]
9. If f is real-valued and measurable on E, define its *essential infimum* on E by
$$\operatorname*{ess\,inf}_E f = \sup\{\alpha : |\{x \in E : f(x) < \alpha\}| = 0\}.$$
   If $f \ge 0$, show that $\mathrm{ess}_E \inf f = (\mathrm{ess}_E \sup 1/f)^{-1}$.

10. Prove that $L^\infty(E)$ is not separable for any E with $|E| > 0$. (Construct a sequence of decreasing subsets of E whose measures strictly decrease. Consider the characteristic functions of the class of sets obtained by taking all possible unions of the differences of these subsets.)
11. If $f_k \to f$ in L^p, $1 \le p < \infty$, $g_k \to g$ pointwise, and $\|g_k\|_\infty \le \mathrm{M}$ for all k, prove that $f_k g_k \to fg$ in L^p.
12. Let $f,\{f_k\} \in L^p$. Show that if $\|f - f_k\|_p \to 0$, then $\|f_k\|_p \to \|f\|_p$. Conversely, if $f_k \to f$ a.e. and $\|f_k\|_p \to \|f\|_p$, $1 \le p < \infty$, show that $\|f - f_k\|_p \to 0$.
13. Suppose that $f_k \to f$ a.e. and that $f_k, f \in L^p$, $1 < p < \infty$. If $\|f_k\|_p \le M < +\infty$, show that $\int f_k g \to \int fg$ for all $g \in L^{p'}$, $1/p + 1/p' = 1$.
14. Verify that the following systems are orthogonal:
    (a) $\{\frac{1}{2}, \cos x, \sin x, \ldots, \cos kx, \sin kx, \ldots\}$ on any interval of length 2π;
    (b) $\{e^{2\pi ikx/(b-a)}; k = 0, \pm 1, \pm 2, \ldots\}$ on (a,b).
15. If $f \in L^2(0,2\pi)$, show that

$$\lim_{k\to\infty} \int_0^{2\pi} f(x) \cos kx\, dx = \lim_{k\to\infty} \int_0^{2\pi} f(x) \sin kx\, dx = 0.$$

Prove that the same is true if $f \in L^1(0,2\pi)$ (This last statement is the *Riemann-Lebesgue lemma*. To prove it, approximate f in L^1 norm by L^2 functions. See (12.21).)
16. A sequence $\{f_k\}$ in L^p is said to *converge weakly* to a function f in L^p if $\int f_k g \to \int fg$ for all $g \in L^{p'}$. Prove that if $f_k \to f$ in L^p norm, $1 \le p \le \infty$, then $\{f_k\}$ converges weakly to f in L^p. Note by Exercise 15 that the converse is not true.
17. Suppose that $f_k, f \in L^2$ and that $\int f_k g \to \int fg$ for all $g \in L^2$ (that is, $\{f_k\}$ converges weakly to f in L^2). If $\|f_k\|_2 \to \|f\|_2$, show that $f_k \to f$ in L^2 norm.
18. Prove the *parallelogram law* for L^2:

$$||f + g||^2 + ||f - g||^2 = 2||f||^2 + 2||g||^2.$$

Is this true for L^p when $p \ne 2$? The geometric interpretation is that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of the sides.
19. Prove that a finite dimensional Hilbert space is isometric with $\mathbf{R}^n$ for some n.
20. Construct a function in $L^1(-\infty,+\infty)$ which is not in $L^2(a,b)$ for any $a < b$. (Let $g(x) = x^{-1/2}$ on $(0,1)$ and $g(x) = 0$ elsewhere, so that $\int_{-\infty}^{+\infty} g = 2$. Consider the function $f(x) = \sum a_k g(x - r_k)$, where $\{r_k\}$ is the rational numbers and $\{a_k\}$ satisfies $a_k > 0$, $\sum a_k < +\infty$.)
21. If $f \in L^p(\mathbf{R}^n)$, $0 < p < \infty$, show that

$$\lim_{Q \searrow x} \frac{1}{|Q|} \int_Q |f(\mathbf{y}) - f(\mathbf{x})|^p\, d\mathbf{y} = 0 \text{ a.e.}$$

Note by Exercise 5 that this condition for a given p implies it for all smaller p.
22. Let $\{\phi_k\}$ be a complete orthonormal system in L^2, and let $m = \{m_k\}$ be a given sequence of numbers. If $f \in L^2$, $f \sim \sum c_k\phi_k$, define Tf by $Tf \sim \sum m_k c_k \phi_k$. Show that T is bounded on L^2, i.e., that there is a constant c independent of f such that $||Tf||_2 \le c||f||_2$ for all $f \in L^2$, if and only if $m \in l^\infty$. Show that the smallest possible choice for c is $\|m\|_{l^\infty}$.