CHAPTER 8

Introduction to Banach Spaces and L^p Space

1. Uniform and Absolute Convergence

As a preparation we begin by reviewing some familiar properties of Cauchy sequences and uniform limits in the setting of metric spaces.

DEFINITION 1.1. A *metric space* is a pair (X, ρ), where X is a set and ρ is a real-valued function on $X \times X$ which satisfies that, for any x, y, $z \in X$,

(a) $\rho(x, y) \geq 0$ and $\rho(x, y) = 0$ if and only if $x = y$,
(b) $\rho(x, y) = \rho(y, x)$,
(c) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$. (Triangle inequality)

The function ρ is called the *metric* on X.

Any metric space has a natural topology induced from its metric. A subset U of X is said to be *open* if for any $x \in U$ there exists some $r > 0$ such that $B_r(x) \subset U$. Here $B_r(x) = \{y \in X : \rho(x, y) < r\}$ is the open ball of radius r centered at x. It is an easy exercise to show that open balls are indeed open and the collection of open sets is indeed a topology, called the *metric topology*.

DEFINITION 1.2. A sequence $\{x_n\}$ in a metric space (X, ρ) is said to be a *Cauchy Sequence* if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \ \text{ such that } \rho(x_n, x_m) < \varepsilon \text{ whenever } n, m \geq N.$$

The metric space (X, ρ) is said to be *complete* if every Cauchy sequence is convergent.

DEFINITION 1.3. Let (X, ρ) be a metric space. For any nonempty set $A \subset X$, the *diameter* of the set A is defined by

$$\operatorname{diam}(A) \ = \ \sup\{\rho(x, y) : \ x, y \in A\}.$$

The set A is said to be *bounded* if its diameter is finite. Otherwise, we say it is *unbounded*.

Let S be a nonempty set. We say a function $f : S \to X$ is *bounded* if its image $f(S)$ is a bounded set. Equivalently, it is bounded if for any $x \in X$, there exists $M > 0$ such that $\rho(f(s), x) \leq M$ for any $s \in S$. We say f is *unbounded* if it is not bounded.

DEFINITION 1.4. Given a sequence $\{f_n\}$ of functions from nonempty set S to metric space (X, ρ). We say $\{f_n\}$ *converges pointwise* to a function $f : S \to X$ if $\lim_{n\to\infty} f_n(s) = f(s)$ for any $s \in S$; that is,

$$\forall \, s \in S, \forall \, \varepsilon > 0, \exists \, N_s \in \mathbb{N} \quad \text{such that} \quad \rho(f_n(s), f(s)) < \varepsilon, \ \forall \, n \geq N_s.$$

In this case, the function f is called the *pointwise limit*.

We say $\{f_n\}$ *converges uniformly* to a function $f : S \to X$ if the above N_s is independent of s; that is,

$$\forall \, \varepsilon > 0, \exists \, N \in \mathbb{N} \quad \text{such that} \quad \rho(f_n(s), f(s)) < \varepsilon, \ \forall \, n \geq N, \ \forall \, s \in S.$$

The function f is called the *uniform limit* of $\{f_n\}$.

The next two theorems highlight some important features of Cauchy sequences and uniform convergence.

THEOREM 1.1. (Cauchy Sequences) *Consider sequences in a metric space* (X, ρ).

(a) *Any convergent sequence is a Cauchy sequence.*
(b) *Any Cauchy sequence is bounded.*
(c) *If a subsequence of a Cauchy sequence converges, then the Cauchy sequence converges to the same limit.*

PROOF. (a) Suppose $\{x_n\}$ converges to x. Given $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $\rho(x_n, x) < \varepsilon/2$ for any $n \geq N$. The sequence $\{x_n\}$ is Cauchy because

$$\rho(x_n, x_m) < \rho(x_n, x) + \rho(x, x_m) < \varepsilon \quad \text{for any } n, m \geq N.$$

(b) Let $\{x_n\}$ be a Cauchy sequence. Choose $N \in \mathbb{N}$ such that $\rho(x_n, x_m) < 1$ for all $n, m \geq N$. Then for any $x \in X$,

$$\begin{aligned} \rho(x_n, x) &\leq \rho(x_n, x_N) + \rho(x_N, x) \\ &< \max\{\rho(x_1, x_N), \rho(x_2, x_N), \cdots, \rho(x_{N-1}, x_N), 1\} + \rho(x_N, x), \end{aligned}$$

where the last equation is a finite bound independent of n.

(c) Let $\{x_n\}$ be a Cauchy sequence with a subsequence $\{x_{n_k}\}$ converging to x. Given $\varepsilon > 0$, choose K, $N \in \mathbb{N}$ such that

$$\begin{aligned} \rho(x_{n_k}, x) &< \frac{\varepsilon}{2} \quad \text{for any } k \geq K, \\ \rho(x_n, x_m) &< \frac{\varepsilon}{2} \quad \text{for any } n, m \geq N. \end{aligned}$$

Taking n_k such that $k \geq K$ and $n_k \geq N$, then

$$\rho(x_n, x) \leq \rho(x_n, x_{n_k}) + \rho(x_{n_k}, x) < \varepsilon \quad \text{for any } n \geq N.$$

This shows that $\{x_n\}$ converges to x. □

THEOREM 1.2. (Uniform Convergence) *Given a sequence of functions* $\{f_n\}$ *from a nonempty set* S *to a metric space* (X, ρ). *Suppose* $\{f_n\}$ *converges uniformly to a function* $f : S \to X$.

(a) *If each* f_n *is bounded, then so is* f.
(b) *Assume* S *is a topological space,* $E \subset S$. *If each* f_n *is continuous on* E, *then so is* f.

PROOF. (a) Choose $N \in \mathbb{N}$ such that $\rho(f(s), f_n(s)) < 1$ for any $n \geq N$ and $s \in S$. Given $x \in X$, choose $M > 0$ such that $\rho(f_N(s), x) < M$ for any $s \in S$. Then f is bounded since

$$\rho(f(s), x) \leq \rho(f(s), f_N(s)) + \rho(f_N(s), x) \leq 1 + M \quad \text{for any } s \in S.$$

(b) Given $\varepsilon > 0$, $e \in E$. Choose $N \in \mathbb{N}$ such that $\rho(f(s), f_n(s)) < \varepsilon/3$ for any $n \geq N$ and $s \in S$. For this particular N, f_N is continuous at e, and so there is a neighborhood U of e such that $\rho(f_N(e), f_N(u)) < \varepsilon/3$ whenever $u \in U$. Then

$$\rho(f(e), f(u)) \;\leq\; \rho(f(e), f_N(e)) + \rho(f_N(e), f_N(u)) + \rho(f_N(u), f(u)) \;<\; \varepsilon \quad \forall u \in U.$$

Therefore f is continuous at e, and is continuous on E since $e \in E$ is arbitrary. □

DEFINITION 1.5. A vector space V over field $\mathbb{F}$ is called a *normed vector space* (or *normed space*) if there is a real-valued function $\|\cdot\|$ on V, called the *norm*, such that for any $x, y \in V$ and any $\alpha \in \mathbb{F}$,

(a) $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$.
(b) $\|\alpha x\| = |\alpha|\|x\|$.
(c) $\|x + y\| \leq \|x\| + \|y\|$. (Triangle inequality)

A norm $\|\cdot\|$ of V defines a metric ρ on V via $\rho(x, y) = \|x - y\|$. All concepts from metric and topological spaces are applicable to normed spaces.

There are multiple ways of choosing norms once a norm is selected. A trivial one is to multiply the original norm by a positive constant. Concepts like neighborhood, convergence, boundedness, and completeness are independent of the choice of these two norms, and so we shall consider them equivalent norms. A more precise characterization of equivalent norms is as follows.

Definition 1.6. Let V be a vector space with two norms $\|\cdot\|$, $\|\cdot\|'$. We say these two norms are *equivalent* if there exists some constant $c > 0$ such that

$$\frac{1}{c}\|x\|' \;\leq\; \|x\| \;\leq\; c\|x\|' \quad \text{for any } x \in V.$$

Example 1.1. The Euclidean space $\mathbb{F}^d$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, with the standard norm $\|\cdot\|$ defined by

$$\|x\| \;=\; (|x_1|^2 + \cdots + |x_d|^2)^{\frac{1}{2}}$$

is a normed space. Now consider two other norms of $\mathbb{F}^d$ defined by

$$\begin{aligned} \|x\|_\infty &= \max\{|x_1|, \cdots, |x_d|\}, \quad \text{called the } \textit{sup norm}; \\ \|x\|_1 &= |x_1| + \cdots + |x_d|, \quad \text{called the } \textit{1-norm}. \end{aligned}$$

Verifications for axioms of norms are easy exercises.

In the case of sup norm, "balls" in $\mathbb{R}^d$ are actually cubes in $\mathbb{R}^d$ with faces parallel to coordinate axes. In the case of 1-norm, "balls" in $\mathbb{R}^d$ are cubes in $\mathbb{R}^d$ with vertices on coordinate axes. These norms are equivalent since

$$\|x\|_\infty \leq \|x\| \leq \|x\|_1 \leq d\|x\|_\infty.$$

Definition 1.7. A complete normed vector space is called a *Banach space*.

Example 1.2. Consider the Euclidean space $\mathbb{F}^d$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, with the standard norm $\|\cdot\|$. The normed space $(\mathbb{R}^d, \|\cdot\|)$ is complete since every Cauchy sequence is bounded and every bounded sequence has a convergent subsequence with limit in $\mathbb{R}^d$ (the Bolzano-Weierstrass theorem). The spaces $(\mathbb{R}^d, \|\cdot\|_1)$ and $(\mathbb{R}^d, \|\cdot\|_\infty)$ are also Banach spaces since these norms are equivalent.

Example 1.3. Given a nonempty set X and a normed space $(Y, \|\cdot\|)$ over field $\mathbb{F}$. The space of functions from X to Y form a vector space over $\mathbb{F}$, where addition and scalar multiplication are defined in a trivial manner: Given two functions f, g, and two scalars α, $\beta \in \mathbb{F}$, define $\alpha f + \beta g$ by

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x), \quad x \in X.$$

Let $b(X, Y)$ be the subspace consisting of bounded functions from X to Y. Define a real-valued function $\|\cdot\|_\infty$ on $b(X, Y)$ by

$$\|f\|_\infty = \sup_{x \in X} \|f(x)\|.$$

It is clearly a norm on $b(X, Y)$, also called the *sup norm*. Convergence with respect to the sup norm means uniform convergence.

If $(Y, \|\cdot\|)$ is a Banach space, then any Cauchy sequence $\{f_n\}$ in $b(X, Y)$ converges pointwise to some function $f : X \to Y$, since $\{f_n(x)\}$ is a Cauchy sequence in Y for any fixed $x \in X$. In fact, the convergence $f_n \to f$ is uniform. To see this, let $\varepsilon > 0$ be arbitrary. Choose $N \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty < \varepsilon/2$ whenever $n, m \geq N$. For any $x \in X$, there exists some $m_x \geq N$ such that $\|f_{m_x}(x) - f(x)\| < \varepsilon/2$. Then for any $n \geq N$,

$$\|f_n(x) - f(x)\| \leq \|f_n(x) - f_{m_x}(x)\| + \|f_{m_x}(x) - f(x)\| < \varepsilon.$$

This proves that the convergence $f_n \to f$ is uniform. By Theorem 1.2(a), $f \in b(X, Y)$, and so the space $b(X, Y)$ with the sup norm is a Banach space.

Example 1.4. Let $(X, \mathcal{T})$ be a topological space and let $(Y, \|\cdot\|)$ be a Banach space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Denote by $C(X, Y)$ the space of continuous functions from X to Y. Let $C_b(X, Y) = C(X, Y) \cap b(X, Y)$, the space of bounded continuous functions from X to Y. Given any sequence $\{f_n\}$ in $C_b(X, Y)$ which converges uniformly to $f \in b(X, Y)$. By Theorem 1.2(b), $f \in C_b(X, Y)$. This shows that $C_b(X, Y)$ is a closed subspace of $b(X, Y)$. Note that a subspace of a Banach space is complete if and only if it is closed (see Exercise 1.1), so $C_b(X, Y)$ is Banach space.

Definition 1.8. A series $\sum_{k=1}^\infty a_k$ in a normed space X is said to be *convergent* (or *summable*) if its partial sum $\sum_{k=1}^n a_k$ converges to some $s \in X$ as $n \to \infty$. We say $\sum_{k=1}^\infty a_k$ is *absolutely convergent* (or *absolutely summable*) if $\sum_{k=1}^\infty \|a_k\| < \infty$.

In the following we prove some useful criteria for completeness and uniform convergence of series.

Theorem 1.3. *A normed space X is complete if and only if every absolutely convergent series is convergent.*

Proof. Suppose X is complete, $\sum_{k=1}^\infty a_k$ is absolutely convergent. We need to show the convergence of $s_n = \sum_{k=1}^n a_k$. Given $\varepsilon > 0$, choose $N \in \mathbb{N}$ such that $\sum_{k=N}^\infty \|a_k\| < \varepsilon$, then $\|s_n - s_m\| < \varepsilon$ whenever $n, m \geq N$. Thus $\{s_n\}_{n=1}^\infty$ is a Cauchy sequence, and so it converges.

Conversely, suppose every absolutely convergent series in X converges. Let $\{s_n\}_{n=1}^\infty$ be a Cauchy sequence in X. By Theorem 1.1 it suffices to show that $\{s_n\}_{n=1}^\infty$ has a convergent subsequence $\{s_{n_k}\}_{k=1}^\infty$. Now choose n_k such that

$$n_k < n_{k+1}, \quad \left\|s_{n_k} - s_{n_{k+1}}\right\| < \frac{1}{2^k} \quad \text{for any } k \in \mathbb{N}.$$

Then the series $s_{n_1} + \sum_{k=1}^\infty \left(s_{n_{k+1}} - s_{n_k}\right)$ converges absolutely, so that it converges to some $s \in X$. This implies that $s_{n_k} = s_{n_1} + \sum_{j=1}^{k-1} \left(s_{n_{j+1}} - s_{n_j}\right)$ converges to s as $k \to \infty$, completing the proof. $\square$

Corollary 1.4. (Weierstrass M-test)
Let $b(X, Y)$ be the space bounded functions from a nonempty set X to a Banach space $(Y, \|\cdot\|)$. Given a sequence of functions $\{f_n\}_{n=1}^\infty$ in $b(X, Y)$. If $\|f_n\|_\infty \leq M_n$ for any $n \in \mathbb{N}$ and $\sum_{n=1}^\infty M_n$ converges, then $\sum_{n=1}^\infty f_n$ converges uniformly.

Proof. The assumption says that $\sum_{n=1}^\infty f_n$ is absolutely convergent. By Theorem 1.3 (and Example 1.3), the series converges in $b(X, Y)$, implying that the convergence is uniform. $\square$

Exercises.

1.1. Show that a subset of a complete metric space is complete if and only if it is closed.

1.2. Let $\Sigma_2 = \{0, 1\}^{\mathbb{N}}$, the space of infinite sequences of $\{0, 1\}$; that is,

$$\Sigma_2 = \{(a_1, a_2, ...) : \ a_k = 0 \text{ or } 1 \text{ for each } k\}.$$

Given $\lambda > 1$, $a, b \in \Sigma_2$, let

$$\rho_\lambda(a,b) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{\lambda^k}.$$

Show that (Σ_2, ρ_λ) is a complete metric space.

1.3. Consider the space of real sequences s. Let

$$\rho(a,b) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{2^k(1+|a_k - b_k|)}.$$

Show that (s, ρ) is a complete metric space.

1.4. Show that all norms in $\mathbb{F}^d$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, are equivalent. Consequently, $\mathbb{F}^d$ with any norm is complete.

1.5. Consider the space $C^0[a,b] \cap C^k(a,b)$, $k \in \mathbb{N}$, and let

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty + \cdots + \|f^{(k)}\|_\infty.$$

Show that $(C^0[a,b] \cap C^k(a,b), \|\cdot\|_{C^k})$ is a Banach space.

1.6. Consider the space $BV[a,b]$ of functions on $[a,b]$ with bounded variations. Let

$$\|f\|_{BV} = |f(a)| + V_a^b(f).$$

Show that $(BV[a,b], \|\cdot\|_{BV})$ is a Banach space.

2. The ℓ^p Space

Definition 2.1. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Given $0 < p < \infty$, define

$$\ell^p = \{a = (a_1, a_2, \cdots) : a_k \in \mathbb{F} \text{ for any } k, \sum_k |a_k|^p < \infty\}, \qquad \|a\|_p = \left(\sum_k |a_k|^p\right)^{\frac{1}{p}},$$
$$\ell^\infty = \{a = (a_1, a_2, \cdots) : a_k \in \mathbb{F} \text{ for any } k, \sup_k |a_k| < \infty\}, \qquad \|a\|_\infty = \sup_k |a_k|.$$

The space ℓ^∞ consists of bounded sequences in $\mathbb{F}$. Addition and multiplication of sequences are defined componentwise:

$$\begin{aligned}(a_1, a_2, \cdots) + (b_1, b_2, \cdots) &= (a_1 + b_1, a_2 + b_2, \cdots) \\ (a_1, a_2, \cdots) \cdot (b_1, b_2, \cdots) &= (a_1 b_1, a_2 b_2, \cdots).\end{aligned}$$

Clearly ℓ^p with any $0 < p \le \infty$ is a vector space, since $\|\alpha a\|_p = |\alpha| \|a\|_p$ for any $\alpha \in \mathbb{F}$ and

$$\begin{aligned} a, b \in \ell^p \quad &\Rightarrow \quad \sum_k |a_k + b_k|^p \le \sum_k (2\max\{|a_k|, |b_k|\})^p \le 2^p \sum_k (|a_k|^p + |b_k|^p) < \infty, \\ a, b \in \ell^\infty \quad &\Rightarrow \quad \sup_k |a_k + b_k| \le \sup_k |a_k| + \sup_k |b_k| < \infty. \end{aligned}$$

.

Theorem 2.1. (Young's inequality) *Given $u, v \ge 0$, $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$uv \le \frac{u^p}{p} + \frac{v^q}{q}.$$

PROOF. It is convenient to write $q = \frac{p}{p-1}$ ($q = \infty$ if $p = 1$, $q = 1$ if $p = \infty$). The curve $y = x^{p-1}$ can be alternatively written $x = y^{q-1}$. The desired inequality follows easily by observing

$$U = \int_0^u x^{p-1}dx = \frac{u^p}{p}, \quad V = \int_0^v y^{q-1}dy = \frac{v^q}{q}, \quad U + V \geq uv.$$

□

THEOREM 2.2. *Given $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let a, b be sequences of complex numbers.*

(a) (HÖLDER'S INEQUALITY FOR ℓ^p) *If $a \in \ell^p$, $b \in \ell^q$, then $ab \in \ell^1$ and*

$$\|ab\|_1 \leq \|a\|_p \|b\|_q.$$

(b) (MINKOWSKI'S INEQUALITY FOR ℓ^p)

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p.$$

PROOF. In (a), the cases $p = 1$, $q = \infty$ and $p = \infty$, $q = 1$ are obvious. Consider $1 < p, q < \infty$. The cases $\|a\|_p = 0$ or $\|b\|_q = 0$ are also obvious, so we assume $0 < \|a\|_p$, $\|b\|_q < \infty$.

Let $A = a/\|a\|_p$, $B = b/\|b\|_q$. By Young's inequality,

$$\begin{aligned}\sum_k |A_k B_k| &\leq \sum_k \left(\frac{|A_k|^p}{p} + \frac{|B_k|^q}{q}\right) = \frac{\|A\|_p^p}{p} + \frac{\|B\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1, \\ \|ab\|_1 &= \sum_k |a_k b_k| = \|a\|_p \|b\|_q \sum_k |A_k B_k| \leq \|a\|_p \|b\|_q.\end{aligned}$$

The cases $p = 1$ and $p = \infty$ for (b) are obvious. For $1 < p < \infty$, (c) follows easily from (b):

$$\begin{aligned}\|a + b\|_p^p &= \sum_k |a_k + b_k|^p \\ &\leq \sum_k |a_k + b_k|^{p-1}|a_k| + \sum_k |a_k + b_k|^{p-1}|b_k| \\ &\leq \left(\sum_k |a_k + b_k|^p\right)^{\frac{p-1}{p}} \left(\sum_k |a_k|^p\right)^{\frac{1}{p}} + \left(\sum_k |a_k + b_k|^p\right)^{\frac{p-1}{p}} \left(\sum_k |b_k|^p\right)^{\frac{1}{p}} \\ &= \|a + b\|_p^{p-1}(\|a\|_p + \|b\|_p).\end{aligned}$$

Thus $\|a + b\|_p \leq \|a\|_p + \|b\|_p$. □

When $1 \leq p, q \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, we say p, q are *conjugate exponents*. It follows from Hölder's inequality that $(\ell^p, \|\cdot\|_p)$ is a normed space when $1 \leq p \leq \infty$.

THEOREM 2.3. *For any $1 \leq p \leq \infty$, $(\ell^p, \|\cdot\|_p)$ is a Banach space.*

PROOF. Completeness of ℓ^∞ is a special case of Example 1.3. Consider $1 \leq p < \infty$. Let $\{a^{(n)}\}_{i=1}^\infty$ be a Cauchy sequence in ℓ^p. For each k, $\{a_k^{(n)}\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{F}$ since

$$\left|a_k^{(n)} - a_k^{(m)}\right| \leq \left(\sum_{j=1}^\infty |a_j^{(n)} - a_j^{(m)}|^p\right)^{\frac{1}{p}} = \left\|a^{(n)} - a^{(m)}\right\|_p.$$

Then there is a sequence $a = (a_1, a_2, \cdots)$ such that, for each k,

$$a_k^{(n)} \to a_k \in \mathbb{R} \quad \text{as } n \to \infty.$$

Fix arbitrary $M \in \mathbb{N}$. Given $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that

$$\left(\sum_{k=1}^{M} \left| a_k^{(n)} - a_k^{(m)} \right|^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{\infty} \left| a_k^{(n)} - a_k^{(m)} \right|^p \right)^{\frac{1}{p}} < \varepsilon \quad \forall\, n > m \geq N.$$

Let $m \to \infty$, then let $M \to \infty$, we find

$$\|a^{(n)} - a\|_p = \left(\sum_{k=1}^{\infty} \left| a_k^{(n)} - a_k \right|^p \right)^{\frac{1}{p}} \leq \varepsilon \quad \text{for any } n \geq N.$$

Thus $\|a^{(n)} - a\|_p \to 0$ as $n \to \infty$, and so $\|a\|_p \leq \|a - a^{(n)}\|_p + \|a^{(n)}\|_p < \infty$, $a \in \ell^p$. This verifies completeness of ℓ^p. □

THEOREM 2.4. *The space ℓ^p is separable if $1 \leq p < \infty$, and the space ℓ^∞ is not separable.*

PROOF. The space ℓ^∞ is not separable because it has an uncountable subset

$$s = \{a = (a_1, a_2, \ldots) \in \ell^\infty : a_n = 0 \text{ or } 1\ \forall\ n\}$$

with $\|a - b\|_\infty = 1$ for any $a \neq b \in s$.

Consider $1 \leq p < \infty$. Let $\mathcal{D}$ be the set of finite sequences with rational coordinates. Clearly $\mathcal{D}$ is countable. Given $a \in \ell^p$ and any $\varepsilon > 0$, we can choose $N \in \mathbb{N}$ such that $\sum_{k=N+1}^{\infty} |a_k|^p < \varepsilon^p/2$. Now choose $b_1, \cdots, b_N \in \mathbb{Q}$ such that $\sum_{k=1}^{N} |a_k - b_k|^p < \varepsilon^p/2$. Let $b = (b_1, \cdots, b_N, 0, 0, \cdots) \in \mathcal{D}$. Then

$$\|a - b\|_p^p \;=\; \sum_{k=1}^{\infty} |a_k - b_k|^p \;=\; \sum_{k=1}^{N} |a_k - b_k|^p + \sum_{k=N+1}^{\infty} |a_k|^p \;<\; \varepsilon^p.$$

Thus $\|a - b\|_p < \varepsilon$. This shows that $\mathcal{D}$ is dense since $\varepsilon > 0$ is arbitrary. □

Exercises.

2.1. Consider the ℓ^p space with $0 < p < 1$. Verify that $\rho_p(a, b) = \sum_{k=1}^{\infty} |a_k - b_k|^p$ is a metric on ℓ^p. Prove that (ℓ^p, ρ_p) is a complete separable metric space.

2.2. Show that

$$x^p \;\geq\; 1 + p(x - 1) \quad \text{for any } x \geq 0,\ p \geq 1.$$

Then use it to give an alternative proof for Young's inequality.

2.3. Prove the following generalization of Young's inequality: Given $1 < p_1, \cdots, p_n < \infty$ with $\sum_{k=1}^{n} \frac{1}{p_k} = 1$. If $u_1, \cdots, u_n \geq 0$, then

$$u_1 \cdots u_n \leq \frac{u_1^{p_1}}{p_1} + \cdots + \frac{u_n^{p_n}}{p_n}.$$

Use it to formulate a generalization of Hölder's inequality for ℓ^p.

2.4. Consider sequences of real numbers. Show that the space c_0 of sequences converging to zero with sup norm is a Banach space, and for any $1 \le p < q \le \infty$, $a \in \ell^p$,

$$\ell^p \subsetneq \ell^q \subsetneq c_0, \quad \|a\|_\infty \le \|a\|_q \le \|a\|_p.$$

Are these norms on ℓ^p equivalent?

2.5. Explain why the set s in the proof of Theorem 2.4 is uncountable.

3. The L^p Space

In this section we consider a space $L^p(E)$ which resembles ℓ^p on many aspects. After general concepts of measure and integral were introduced, we will see that these two spaces can be viewed as special cases of a more general L^p space.

Definition 3.1. Given a measurable set $E \subset \mathbb{R}^d$. For $0 < p < \infty$, define the space $L^p(E)$ and the real-valued function $\|\cdot\|_p$ on $L^p(E)$ by

$$L^p(E) = \{f : f \text{ is measurable on } E \text{ and } \int_E |f|^p < \infty\}, \qquad \|f\|_p = \left(\int_E |f|^p\right)^{\frac{1}{p}}.$$

The *essential supremum* of a measurable function f on E is defined by

$$\operatorname*{ess\,sup}_E f = \inf\{\alpha \in (-\infty, \infty] : m(\{f > \alpha\}) = 0\}.$$

The space $L^\infty(E)$ and the real-valued function $\|\cdot\|_\infty$ on $L^\infty(E)$ are given by

$$L^\infty(E) = \{f : f \text{ is measurable on } E \text{ and } \operatorname*{ess\,sup}_E |f| < \infty\}, \qquad \|f\|_\infty = \operatorname{ess\,sup} |f|.$$

Functions in $L^\infty(E)$ are said to be *essentially bounded.*

The measurable function f in the definition of $L^p(E)$ for $0 < p < \infty$ can be complex-valued, but functions in $L^\infty(E)$ are assumed to be real-valued. We leave it to the readers to check that $m(\{f > \operatorname{ess\,sup}_E f\}) = 0$ for any $f \in L^\infty(E)$ (Exercise 3.1). In other words, $f \le \operatorname{ess\,sup}_E f$ and $|f| \le \|f\|_\infty$ almost everywhere.

Being "essentially bounded" is different from being "bounded" in that the former is regardless of measure zero sets. For instance, the function

$$f(x) = \begin{cases} x^3 & x \in \mathbb{Q} \\ \sin x & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is unbounded but essential bounded with $\operatorname{ess\,sup} f = 1$.

For any $0 < p \le \infty$, two functions f_1, f_2 in $L^p(E)$ are considered equivalent if $f_1 = f_2$ almost everywhere on E. The space of equivalence classes, still denoted by $L^p(E)$, are called $L^p(E)$ *classes* or $L^p(E)$ *spaces.*

Similar to ℓ^p, the space $L^p(E)$ is a vector space for any $0 < p \le \infty$. Indeed, $\|\alpha f\|_p = |\alpha|\|f\|_p$ for any scalar α, $\|\alpha f\|_p = |\alpha|\|f\|_p$ and

$$\begin{aligned} f, g \in L^p(E) \quad &\Rightarrow \quad |f+g|^p \le (2\max\{|f|, |g|\})^p \le 2^p(|f|^p + |g|^p), \\ f, g \in L^\infty(E) \quad &\Rightarrow \quad \|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty. \end{aligned}$$

The second line follows by observing that

$$\left.\begin{array}{l} |f| \le \|f\|_\infty \text{ a.e.} \\ |g| \le \|g\|_\infty \text{ a.e.} \end{array}\right\} \Rightarrow |f+g| \le \|f\|_\infty + \|g\|_\infty \text{ a.e.}$$

When $1 \le p \le \infty$, the function $\|\cdot\|_p$ is a norm on $L^p(E)$. This follows from the theorem below, the proof for which is similar to that of ℓ^p.

THEOREM 3.1. *Given* $1 \le p, q \le \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$. *Let* f, g *be measurable functions on* $E \subset \mathbb{R}^d$.

(a) (HÖLDER'S INEQUALITY FOR L^p) *If* $f \in L^p(E)$, $g \in L^q(E)$, *then* $fg \in L^1(E)$ *and*

$$\|fg\|_1 \le \|f\|_p \|g\|_q.$$

(b) (MINKOWSKI'S INEQUALITY FOR L^p)

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

PROOF. (a) The cases $p = 1$, $q = \infty$ and $p = \infty$, $q = 1$ are obvious. Consider $1 < p, q < \infty$. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $fg = 0$ almost everywhere on E, and the asserted inequality is obvious. We may now assume $0 < \|f\|_p, \|g\|_q < \infty$.

Let $F = f/\|f\|_p$, $G = g/\|g\|_q$. By Young's inequality,

$$\begin{aligned} \int_E |FG| &\le \int_E \frac{|F|^p}{p} + \frac{|G|^q}{q} = \frac{\|F\|_p^p}{p} + \frac{\|G\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1, \\ \|fg\|_1 &= \int_E |fg| = \|f\|_p \|g\|_q \int_E |FG| \le \|f\|_p \|g\|_q. \end{aligned}$$

(b) The case $p = 1$ is obvious, and the case $p = \infty$ has been proved. Now we consider $1 < p < \infty$. Note that $q = \frac{p}{p-1}$. Minkowski's inequality follows easily from (a):

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

Thus $\|f+g\|_p \le \|f\|_p + \|g\|_p$. $\square$

The special case $p = q = 2$ of the Hölder inequality is also known as the *Cauchy-Schwarz inequality*. The assumption $1 \le p \le \infty$ is necessary. For example, let $E = [0,1]$, $f = \chi_{[0,\frac{1}{2}]}$, $g = \chi_{[\frac{1}{2},1]}$. Then for $0 < p < 1$,

$$\|f\|_p + \|g\|_p = \left(\frac{1}{2}\right)^{\frac{1}{p}} + \left(\frac{1}{2}\right)^{\frac{1}{p}} = 2^{1-\frac{1}{p}} < 1 = \|f+g\|_p.$$

COROLLARY 3.2. *Suppose* $0 < p < q < \infty$, $m(E) < \infty$. *Then*

$$\left(\frac{1}{m(E)} \int_E |f|^p\right)^{\frac{1}{p}} \le \left(\frac{1}{m(E)} \int_E |f|^q\right)^{\frac{1}{q}}.$$

In particular, $L^q(E) \subset L^p(E)$.

PROOF. Let $r = \frac{q}{q-p}$, then $\frac{1}{q/p} + \frac{1}{r} = 1$. Therefore,

$$\int_E |f|^p \le \left(\int_E (|f|^p)^{\frac{q}{p}}\right)^{\frac{p}{q}} \left(\int_E 1^r\right)^{\frac{1}{r}} = \left(\int_E |f|^q\right)^{\frac{p}{q}} m(E)^{\frac{q-p}{q}}.$$

Then the corollary follows from

$$\|f\|_p = \left(\int_E |f|^p\right)^{\frac{1}{p}} \le \|f\|_q \, m(E)^{\frac{q-p}{qp}} = \|f\|_q \, m(E)^{\frac{1}{p}-\frac{1}{q}}.$$

□

COROLLARY 3.3. *Suppose* $0 < p < r < q < \infty$. *Then* $L^p(E) \cap L^q(E) \subset L^r(E)$.

PROOF. Given $f \in L^p(E) \cap L^q(E)$. Let $r = \theta p + (1-\theta)q$, $\theta \in (0,1)$. Then

$$|f|^{\theta p} \in L^{1/\theta}(E), \quad |f|^{(1-\theta)q} \in L^{1/(1-\theta)}(E).$$

Apply Hölder's inequality with conjugate exponents $1/\theta$, $1/(1-\theta)$, we find

$$\begin{aligned} \int_E |f|^r &= \int_E |f|^{\theta p} \, |f|^{(1-\theta)q} \\ &\le \left(\int_E |f|^p\right)^{\theta} \left(\int_E |f|^q\right)^{1-\theta} \end{aligned}$$

Thus $f \in L^r(E)$. □

EXAMPLE 3.1. Consider $f(x) = x^r$, $r \ne 0$, defined on $[0, \infty)$.

$$\begin{aligned} r < 0 \quad &\Rightarrow \quad f \in L^p[1,\infty) \text{ if and only if } p > -1/r, \\ & \quad f \in L^p(0,1) \text{ if and only if } 0 < p < -1/r. \\ r > 0 \quad &\Rightarrow \quad f \notin L^p[1,\infty) \text{ for any } p > 0, \\ & \quad f \in L^p(0,1) \text{ for any } p > 0. \end{aligned}$$

This example shows that the assumption $m(E) < \infty$ is necessary in the above corollary, and $L^q(E) \subsetneq L^p(E)$ if $0 < p < q < \infty$ and $E = (0,1)$.

Obviously, $L^\infty(E) \subset L^p(E)$ when $m(E) < \infty$. The inclusion is generally proper. To see this, consider the function $g(x) = \log x$, it belongs to $L^p(0,1)$ for any $0 < p < \infty$ but it is not in $L^\infty(0,1)$.

THEOREM 3.4. (RIESZ-FISHER) *For any* $1 \le p \le \infty$, *the space* $(L^p(E), \|\cdot\|_p)$ *is a Banach space.*

PROOF. Consider $p = \infty$ first. Note that convergence in $L^\infty(E)$ means uniform convergence outside a set of measure zero.

Let $\{f_n\}$ be a Cauchy sequence in $L^\infty(E)$. For each $n, m \in \mathbb{N}$, $|f_n - f_m| \le \|f_n - f_m\|_\infty$ except on a set $Z_{n,m}$ of measure zero. Let $Z = \bigcup_{n,m\in\mathbb{N}} Z_{n,m}$, then Z has measure zero and

$$|f_n - f_m| \le \|f_n - f_m\|_\infty \quad \text{on } E \setminus Z$$

In particular, for any $x \in E \setminus Z$, $\{f_n(x)\}$ converges. Let $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in E \setminus Z$ and set $f(x) = 0$ on Z. Then (similar to arguments in Example 1.3)

$$f_n \to f \text{ uniformly on } E \setminus Z.$$

This implies that f_n converges to f in $L^\infty(E)$, and so $L^\infty(E)$ is complete.

Now we consider $1 \leq p < \infty$. By Theorem 1.3, we only have to show that every absolutely convergent series converges to some element in $L^p(E)$.

Let $\sum_{k=1}^{\infty} f_k$ be an absolutely convergent series. Then $\sum_{k=1}^{\infty} \|f_k\|_p = M$ is finite. Let

$$g_n = \sum_{k=1}^{n} |f_k|, \quad s_n = \sum_{k=1}^{n} f_k.$$

By Minkowski's inequality, $\|g_n\|_p \leq \sum\limits_{k=1}^{n} \|f_k\|_p \leq M$. Thus $\int_E g_n^p \leq M^p$ for any n. For any $x \in E$, the function $g_n(x)$ is increasing in n, and so g_n converges pointwise to some function $g : E \to [0, \infty]$. The function g is measurable and, by Fatou's lemma,

$$\int_E g^p \leq \liminf_{n\to\infty} \int_E g_n^p \leq M^p.$$

Therefore g is finite almost everywhere and $g \in L^p(E)$. When $g(x)$ is finite, $\sum_{k=1}^{\infty} f_k(x)$ is absolutely convergent. Let $s(x)$ be its value, and set $s(x) = 0$ elsewhere. Then the function s is defined everywhere, measurable on E, and

$$\sum_{k=1}^{n} f_k = s_n \to s \quad \text{almost everywhere on } E.$$

Since $|s_n(x)| \leq g(x)$ for all n, we have $|s(x)| \leq g(x)$, where hence $s \in L^p(E)$ and $|s_n(x) - s(x)| \leq 2g(x) \in L^p(E)$. By the Lebesgue dominated convergence theorem,

$$\int_E |s_n - s|^p \to 0 \text{ as } n \to \infty.$$

This proves that $\sum_{k=1}^{\infty} f_k$ converges to $s \in L^p(E)$, and thus proves completeness of $L^p(E)$. □

THEOREM 3.5. *If $1 \leq p < \infty$, then $L^p(E)$ is separable.*

PROOF. Consider $E = \mathbb{R}^d$. Consider the collection of dyadic cubes; namely, consider cubes of the form $[k_1, k_1 + 1] \times \ldots \times [k_d, k_d + 1]$, $k_1 \ldots, k_d \in \mathbb{Z}$, bisect each of these cubes into 2^d congruent subcubes, and repeat this process. Let $\mathcal{D}$ be the set of finite linear combinations of characteristic functions on these dyadic cubes with rational coefficients. Clearly $\mathcal{D}$ is countable. We need to show that its closure $\overline{\mathcal{D}}$ is exactly $L^p(\mathbb{R}^d)$.

It follows easily from Minkowski's inequality that $\overline{\mathcal{D}}$ is a linear subspace of $L^p(\mathbb{R}^d)$. Any characteristic function $\chi_{\mathcal{O}}$ on bounded open set $\mathcal{O}$ belongs to $\overline{\mathcal{D}}$ since it can be expressed $\sum_{k=1}^{\infty} \chi_{c_k}$ a.e. for some dyadic cubes c_k.

Given bounded measurable set E, choose bounded open sets $\mathcal{O}_k \supset E$ such that $m(\mathcal{O}_n \setminus E) < 1/n$ for each n, then by the Lebesgue dominated convergence theorem

$$\int_{\mathbb{R}^d} |\chi_{\mathcal{O}_k} - \chi_E|^p \to 0 \quad \text{as } k \to \infty.$$

Therefore $\chi_E \in \overline{\mathcal{D}}$, then so are simple functions $\sum_{k=1}^{N} a_k \chi_{E_k}$ with bounded measurable E_k's. This includes all nonnegative simple functions with compact supports.

For any nonnegative $f \in L^p(\mathbb{R}^d)$, choose an increasing sequence $\{f_n\}$ of nonnegative simple functions with compact supports such that $f_n \nearrow f$ as $n \to \infty$. By the monotone convergence theorem,

$$\int_{\mathbb{R}^d} |f_n - f|^p \to 0 \quad \text{as } n \to \infty.$$

Therefore $f \in \overline{\mathcal{D}}$. This implies $\overline{\mathcal{D}} = L^p(\mathbb{R}^d)$ since any function in $L^p(\mathbb{R}^d)$ is the difference of two nonnegative functions in $L^p(\mathbb{R}^d)$.

Now consider arbitrary measurable set $E \subset \mathbb{R}^d$. Let $\mathcal{D}' = \{g \cdot \chi_E : g \in \mathcal{D}\}$. Then $\mathcal{D}'$ is a countable set consisting of finite linear combinations of characteristic functions on dyadic cubes which intersect with E and with rational coefficients.

Given $f \in L^p(E)$. Let $\tilde{f} = f$ on E, $\tilde{f} = 0$ on $\mathbb{R}^d \backslash E$. Choose $\{f_k\} \subset \mathcal{D}$ such that $\int_{\mathbb{R}^d} |f_k - \tilde{f}|^p \to 0$ as $k \to \infty$. Then

$$\int_E |f_k \cdot \chi_E - f|^p \ \le \ \int_{\mathbb{R}^d} |f_k - \tilde{f}|^p \ \longrightarrow \ 0 \ \text{ as } k \to \infty.$$

This proves that $\mathcal{D}'$ is dense in $L^p(E)$. □

Given $h \in \mathbb{R}^d$. Let $\tau_h f(x) = f(x+h)$ be the translation operator. Similar to the case $L^1(\mathbb{R}^d)$, we have continuity of variable translations with respect to $\|\cdot\|_p$:

THEOREM 3.6. *If $1 \le p < \infty$ and $f \in L^p(\mathbb{R}^d)$, then*

$$\lim_{h\to\infty} \|\tau_h f - f\|_p = 0.$$

PROOF. Let $\mathcal{C}$ be the collection of $L^p(\mathbb{R}^d)$ functions satisfying this property. It follows easily from the Minkowski inequality that it is a linear subspace of $L^p(\mathbb{R}^d)$. Note that $\mathcal{C}$ is closed. To see this, take any sequence $\{f_n\}$ in $\mathcal{C}$ which converges to f in $L^p(\mathbb{R}^d)$, we have

$$\begin{aligned} \|\tau_h f - f\|_p &\le \|\tau_h f - \tau_h f_n\|_p + \|\tau_h f_n - f_n\|_p + \|f_n - f\|_p \\ &\le \|\tau_h f_n - f_n\|_p + 2\|f - f_n\|_p. \end{aligned}$$

Let $h \to 0$, then let $n \to \infty$, we find that $\limsup_{h\to 0} \|\tau_h f - f\|_p = 0$, so $f \in \mathcal{C}$.

We proceed as the proof of the previous theorem. Clearly characteristic functions on cubes are in $\mathcal{C}$, so characteristic functions on bounded measurable sets are also in $\mathcal{C}$, then it follows that simple functions with compact supports are in $\mathcal{C}$ as well.

Suppose $f \in L^p(\mathbb{R}^d)$ is nonnegative. Choose nonnegative simple functions f_n with compact supports such that $f_n \nearrow f$ as $n \to \infty$. Then $f_n \in L^p(\mathbb{R}^d)$ and, by the monotone convergence theorem, $\|f_n - f\|_p \to 0$ as $n \to \infty$. Therefore $f \in \mathcal{C}$. This implies $\mathcal{C} = L^p(\mathbb{R}^d)$ since any $f \in L^p(\mathbb{R}^d)$ is the difference of two nonnegative measurable functions in $L^p(\mathbb{R}^d)$. □

REMARK 3.1. Theorem 3.6 is false for $p = \infty$. This can be easily seen by considering, for example, the characteristic function on the unit ball.

Exercises.

3.1. Given any $f \in L^\infty(E)$. Show that $m(\{f > \operatorname{ess\,sup}_E f\}) = 0$.

3.2. Use the generalized Young's inequality in Exercise 2.3 to formulate a generalization of Hölder's inequality for $L^p(E)$.

3.3. Suppose $m(E) < \infty$. Show that $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$. How about if $m(E) = \infty$?

3.4. Let f be a real-valued measurable function on E, $m(E) > 0$. Define the *essential infimum* on E by

$$\operatorname*{ess\,inf}_E f = \sup\{\alpha \in (-\infty, \infty] : m(\{f < \alpha\}) = 0\}.$$

Show that, if $f \geq 0$, then $\operatorname{ess\,inf}_E f = 1/\operatorname{ess\,sup}_E(1/f)$.

3.5. Consider $L^p(E)$ with $0 < p < 1$. Verify that $\rho_p(f, g) = \int_E |f - g|^p$ is a metric on $L^p(E)$. Prove that $(L^p(E), \rho_p)$ is a complete separable metric space.

3.6. Given $1 \leq p < \infty$, $f_n, f \in L^p(E)$. Suppose f_n converges to f almost everywhere. Prove that each of the following conditions implies $\|f_n - f\|_p \to 0$ as $n \to \infty$.

(a) There exists some $g \in L^p(E)$ such that $|f_n| \leq g$ for any n.
(b) $\|f_n\|_p \to \|f\|_p$ as $n \to \infty$.

3.7. For what kind of $f \in L^p(E)$ and $g \in L^q(E)$, $\frac{1}{p} + \frac{1}{q} = 1$, do we have equality for the Hölder inequality? For what kind of $f, g \in L^p(E)$ do we have equality for the Minkowski inequality?

3.8. Given $0 < p < q < \infty$, $f \in L^p(E) \cap L^q(E)$. From Corollary 3.3 we know $f \in L^r(E)$ for any $r \in (p, q)$. Show that $\|f\|_r$ as function of r is absolutely continuous on any compact subinterval of (p, q). This implies, in particular, that this function is differentiable a.e. on (p, q).

3.9. Consider $1 < p < \infty$. Give a proof for the Minkowski inequality using convexity of x^p, and without using Hölder's inequality.

3.10. Show that $L^\infty(E)$ is not separable whenever $m(E) > 0$.

4. Dense Subspaces of L^p

In the proof of Theorem 3.5 we constructed a countable collection of step functions which is dense in $L^p(E)$. These step functions are linear combinations of characteristic functions on some dyadic cubes. This implies that the space of simple functions is also dense in $L^p(\mathbb{R}^d)$. In this section we prove that the space of smooth functions with compact supports, and the space of functions with rapidly decreasing derivatives are also dense in $L^p(\mathbb{R}^d)$.

Let us begin by recalling that $(L^1(\mathbb{R}^d), *)$ is a commutative algebra without identity, but there are "good kernels" $\{K_\delta\}$ which approximate the identity in the sense that, for any $f \in L^1(\mathbb{R}^d)$,

$$f * K_\delta \to f \quad \text{a.e. and} \quad \|f * K_\delta - f\|_1 \to 0 \quad \text{as } \delta \searrow 0.$$

For example, kernels satisfying the following conditions are approximations to the identity:

(1) $\int K_\delta = 1$ for any $\delta > 0$.
(2) There exist some $C > 0$ such that $|K_\delta| \leq \frac{C}{\delta^d}$ for any $\delta > 0$.
(3) There exist some $C' > 0$ such that $|K_\delta(x)| \leq \frac{C'\delta}{|x|^{d+1}}$ for any $\delta > 0$, $x \in \mathbb{R}^d \setminus \{0\}$.

Conditions (2) and (3) imply

(4) There exist some $C > 0$ such that $\int K_\delta = 1$ and $\int |K_\delta| < C$ for any $\delta > 0$.
(5) For any $\eta > 0$, $\int_{|x| \geq \eta} |K_\delta(x)| dx \to 0$ as $\delta \searrow 0$.

Recall also that, for any $f, g \in L^1(\mathbb{R}^d)$, $f * g \in L^1(\mathbb{R}^d)$ and

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

We may extend these results to $L^p(\mathbb{R}^d)$ for $1 \leq p < \infty$, based on two auxiliary inequalities:

LEMMA 4.1. *Given* $1 \le p \le \infty$.

(a) *If* $f \in L^p(\mathbb{R}^d)$, $g \in L^1(\mathbb{R}^d)$, *then* $f * g \in L^p(\mathbb{R}^d)$ *and*

$$\|f * g\|_p \le \|f\|_p \|g\|_1.$$

(b) (MINKOWSKI'S INTEGRAL INEQUALITY)
If $1 \le p < \infty$, $f \in L^p(\mathbb{R}^d \times \mathbb{R}^d)$, *then*

$$\left[\int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} f(x,y)dx\right|^p dy\right]^{\frac{1}{p}} \le \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} |f(x,y)|^p dy\right]^{\frac{1}{p}} dx.$$

Both inequalities are simple applications of Hölder's inequality. The first one is a special case of *Young's convolution inequality* (see exercise 4.1). We leave details and proofs as exercises.

LEMMA 4.2. *Given* $1 \le p \le \infty$. *If* $f \in L^p(\mathbb{R}^d)$, $g \in C_0^m(\mathbb{R}^d)$, $0 \le m \le \infty$, *then* $f * g \in C^m(\mathbb{R}^d)$ *and* $D^\alpha(f * g) = f * D^\alpha g$ *for any multi-index* α *with* $|\alpha| \le m$.

PROOF. Consider $1 < p \le \infty$. Let $q = \frac{p}{p-1}$ be the conjugate exponent of p, then $1 \le q < \infty$. Consider $m = 0$. Given $h \in \mathbb{R}^d$,

$$\begin{aligned} |(f * g)(x+h) - (f * g)(x)| &= \left|\int_{\mathbb{R}^d} f(x+h-y)g(y)dy - \int_{\mathbb{R}^d} f(x-y)g(y)dy\right| \\ &= \left|\int_{\mathbb{R}^d} f(x-y)g(y+h)dy - \int_{\mathbb{R}^d} f(x-y)g(y)dy\right| \\ &= \left|\int_{\mathbb{R}^d} f(x-y)\left(g(y+h) - g(y)\right)dy\right| \\ &\le \|f\|_p \|\tau_h g - g\|_q \quad \text{(by Hölder's inequality).} \end{aligned}$$

By Theorem 3.6, the last term converges to zero as $h \to 0$. Note that when $p = 1$, the term $\|\tau_h g - g\|_q$ converges to zero as $h \to 0$ since g is uniformly continuous. This proves $f * g \in C^0(\mathbb{R}^d)$.

Consider $m = 1$. Given $t > 0$ and $i \in \{1, \cdots, d\}$. By the mean-value theorem, there exist $s \in [0, t]$ such that

$$\begin{aligned} (f * g)(x + te_i) - (f * g)(x) &= \int_{\mathbb{R}^d} f(y)\left(g(x + te_i - y) - g(x-y)\right)dy \\ &= \int_{\mathbb{R}^d} f(y)\frac{\partial}{\partial x_i} g(x + se_i - y)t\,dy \\ &= f * \frac{\partial}{\partial x_i} g(x + se_i)t. \end{aligned}$$

Since $\frac{\partial}{\partial x_i} g \in C_0^0(\mathbb{R}^d)$, $f * \frac{\partial}{\partial x_i} g \in C^0(\mathbb{R}^d)$, we see that $\frac{\partial}{\partial x_i}(f * g)$ exists and equals $f * \frac{\partial}{\partial x_i} g$. This implies $f * g \in C^1(\mathbb{R}^d)$ since i is arbitrary. The proof for general m follows by induction. □

THEOREM 4.3. *Given* $1 \le p < \infty$. *Let* $\{K_\delta\}$ *be kernels satisfying (1),(2),(3). Then for any* $f \in L^p(\mathbb{R}^d)$,

$$\|f * K_\delta - f\|_p \to 0 \quad \text{as} \quad \delta \searrow 0.$$

PROOF. Observe that

$$|f * K_\delta(x) - f(x)| \le \int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dy.$$

By Minkowski's integral inequality,

$$\begin{aligned}
\|f * K_\delta - f\|_p &\leq \left[\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dy\right]^p dx\right]^{\frac{1}{p}} \\
&\leq \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} |f(x-y) - f(x)|^p |K_\delta(y)|^p dx\right]^{\frac{1}{p}} dy \\
&= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} |f(x-y) - f(x)|^p dx\right]^{\frac{1}{p}} |K_\delta(y)|dy \\
&= \int_{\mathbb{R}^d} \|\tau_{-y} f - f\|_p |K_\delta(y)|dy.
\end{aligned}$$

Given $\varepsilon > 0$, choose $\eta > 0$ such that $\|\tau_{-y} f - f\|_p < \varepsilon$ whenever $|y| < \eta$. Then

$$\|f * K_\delta - f\|_p \leq \varepsilon \int_{|y|<\eta} |K_\delta(y)|dy + \int_{|y|\geq\eta} 2\|f\|_p |K_\delta(y)|dy.$$

The theorem follows by letting $\delta \to 0$ and then $\varepsilon \to 0$. □

Recall that the support $\operatorname{supp}(f)$ of a measurable function f is the largest closed subset for which every open neighborhood of every point from it has positive measure. Alternatively, it can be expressed as the intersection of sets in

$$\{K \subset \mathbb{R}^d : K \text{ is closed and } f = 0 \text{ a.e. on } K^c\}.$$

Lemma 4.4. *Given measurable functions f and g. We have $f * g = 0$ on $(\operatorname{supp}(f) + \operatorname{supp}(g))^c$. In particular, $\operatorname{supp}(f * g) \subset \operatorname{supp}(f) + \operatorname{supp}(g)$.*

Proof. Observe that

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy = \int_{\operatorname{supp}(g)} f(x-y)g(y)dy = 0$$

if $x - y \notin \operatorname{supp}(f)$ for every $y \in \operatorname{supp}(g)$. That is, $f * g = 0$ on $(\operatorname{supp}(f) + \operatorname{supp}(g))^c$. □

Theorem 4.5. *$C_0^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$, $1 \leq p < \infty$.*

Proof. Given $f \in L^p(\mathbb{R}^d)$. For any $N \in \mathbb{N}$, let

$$f_N = \begin{cases} f(x) & \text{if } |f(x)|, \|x\| < N \\ 0 & \text{otherwise.} \end{cases}$$

Clearly f_N converges almost everywhere to f as $N \to \infty$, and $|f(x) - f_N(x)|^p \leq 2^p |f(x)|^p$. By the Lebesgue dominated convergence theorem, $\|f - f_N\|_p \to 0$ as $N \to \infty$. From this observation, it is sufficient to consider the case of bounded f with compact support.

Choose a nonnegative function $K \in C_0^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} K = 1$. Let

$$K_\delta(x) = \frac{1}{\delta^n} K\left(\frac{x_1}{\delta}, \cdots, \frac{x_d}{\delta}\right) \quad \text{for } \delta > 0.$$

Then the family $\{K_\delta\}$ satisfies conditions (1), (2), (3) stated at the beginning of this section (check it!). By Lemma 4.2 and Lemma 4.4, $f * K_\delta \in C_0^\infty(\mathbb{R}^d)$. By Theorem 4.3, $\|f * K_\delta - f\|_p \to 0$ as $\delta \searrow 0$. This completes the proof. □

DEFINITION 4.1. The *Schwartz class* $\mathcal{S}(\mathbb{R}^d)$ is defined by

$$\mathcal{S}(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) \mid \sup_{x\in\mathbb{R}^d} |x^\alpha D^\beta f(x)| < \infty \quad \text{for any multi-indices } \alpha,\, \beta\}.$$

Roughly speaking, the Schwartz class consists of smooth functions whose derivatives decrease to zero faster than the inverse of any polynomial. This function space is of special importance in Fourier analysis and distribution theory.

COROLLARY 4.6. *For any $1 \le p < \infty$, the Schwartz class $\mathcal{S}(\mathbb{R}^d)$ is a dense subspace of $L^p(\mathbb{R}^d)$.*

PROOF. It follows easily from Theorem 4.5, $C_0^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$, and the observation that $\mathcal{S}(\mathbb{R}^d)$ is a subspace of $L^p(\mathbb{R}^d)$. □

Exercises.

4.1. Prove the Lemma 4.1(a). More generally, given $1 \le p, q, r \le \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, and given $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$, prove the following *Young's convolution inequality*:

$$\|f * g\|_r \le \|f\|_p \|g\|_q.$$

Hint: Find suitable p', q' such that $\frac{1}{p'} + \frac{1}{q} + \frac{1}{r} = 1$, then apply generalized Hölder's inequality.

4.2. Use Hölder's inequality to prove Minkowski's integral inequality.

4.3. Verify that kernels $\{K_\delta\}$ in the proof of Theorem 4.5 satisfies (1), (2), (3) stated at the beginning of this section.

4.4. Suppose $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Show that if $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$, then $f * g \in C^0(\mathbb{R}^d)$.

4.5. Show that $f(x) = e^{-\frac{1}{x^2} - x^2}\chi_{(0,\infty)}$ belongs to $\mathcal{S}(\mathbb{R})$, and $g(x) = f(x-a)f(b-x) \in C_0^\infty(\mathbb{R})$, where $a < b$ are fixed.

4.6. Given bounded open sets $G_1 \subset G_2$ such that $\overline{G}_1 \subset G_2$. Construct a function $f \in C_0^\infty$ such that $f = 1$ on G_1 and $f = 0$ on G_2^c.

5. Linear Transformations

This section is brief introduction to the concepts of bounded linear operators and dual spaces. Those who unacquainted with undefined concepts here are referred to any standard textbook on linear algebra. Those who discontented with brevity of discussions herein are referred to any standard textbook on functional analysis. For concepts intimately related to our discussions for L^p spaces, we prove them here.

In this section we consider only normed spaces, even though many concepts can be extended to more general topological vector spaces. For convenience, we shall use the same notation $\|\cdot\|$ for norms of various spaces, and we shall simply write X for $(X, \|\cdot\|)$, for instance. It is often evident which norm we are referring to. When it is necessary to avoid confusion, we denote the norm of the space X by $\|\cdot\|_X$.

DEFINITION 5.1. Given two normed spaces X and Y, a linear transformation (operator) $T : X \to Y$ is said to be *bounded* if there exists some $M > 0$ such that

$$\|Tx\| \;\le\; M\|x\| \quad \text{for any } x \in X.$$

Denote the space of bounded linear operators from X to Y by $B(X, Y)$.

Theorem 5.1. *A linear operator $T : X \to Y$ is bounded if and only if it is continuous.*

Proof. Clearly any bounded linear operator is Lipschitz continuous. If T is continuous, then $T^{-1}(B_1(0)) \supset B_\delta(0)$ for some $\delta > 0$. Then, whenever $\|x\| \le 1$, we have

$$\|Tx\| = \frac{2}{\delta}\left\|T\left(\frac{\delta}{2}x\right)\right\| \le \frac{2}{\delta}.$$

For general x, we have

$$\|Tx\| = \left\|T\left(\frac{x}{\|x\|}\right)\right\| \|x\| \le \frac{2}{\delta}\|x\|.$$

Thus T is bounded. □

It is a simple exercise to show that

$$\sup_{\|x\|=1} \|Tx\| \ = \ \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \ = \ \sup_{\|x\|\le 1} \|Tx\| \ = \ \inf\{M : \|Tx\| \le M\|x\| \ \forall x \in X\}.$$

Their common value is denoted by $\|T\|$. This notation is justified in the following theorem.

Theorem 5.2. *$B(X, Y)$ with the function $T \mapsto \|T\|$ is a normed space. If Y is a Banach space, then so is $B(X, Y)$.*

Proof. We only verify the triangle inequality for $\|\cdot\|$, for the other two axioms of norm are obvious. Given S, $T \in B(X, Y)$, and $x \in X$.

$$\|(S + T)x\| = \|Sx + Tx\| \le \|Sx\| + \|Tx\| \le \|S\| \, \|x\| + \|T\| \, \|x\|.$$

This implies the triangle inequality $\|S + T\| \le \|S\| + \|T\|$.

Given a Cauchy sequence $\{T_n\}$ in $B(X, Y)$. For any $x \in X$, the sequence $\{T_n x\}$ is a Cauchy sequence in Y, and so it converges. Let $T : X \to Y$ be defined by $Tx = \lim_{n\to\infty} T_n x$. Clearly T is linear. Given $\varepsilon > 0$, choose N such that $\|T_n - T_m\| < \varepsilon$ for all $n, m \ge N$. Fix $n \ge N$ and $x \in X$,

$$\|T_n x - Tx\| = \lim_{m\to\infty} \|T_n x - T_m x\| \le \varepsilon\|x\|.$$

Therefore,

$$\|Tx\| \le \|Tx - T_N x\| + \|T_N x\| \le (\varepsilon + \|T_N\|)\|x\|.$$

This shows that T is bounded. Furthermore,

$$\|T_n - T\| = \sup_{\|x\|=1} \|T_n x - Tx\| \le \varepsilon \quad \text{for any } n \ge N.$$

This implies that $\|T_n - T\| \to 0$ as $n \to \infty$, since $\varepsilon > 0$ is arbitrary. □

Definition 5.2. A *linear functional* f on a normed space X over $\mathbb{F}$ ($= \mathbb{R}$ or $\mathbb{C}$) is a linear transformation from X to $\mathbb{F}$. The space $B(X, \mathbb{F})$ of bounded linear functionals on X is called the *dual space* of X. It is usually denoted by X^*.

Corollary 5.3. *The dual space X^* of any normed space X over $\mathbb{R}$ or $\mathbb{C}$ is a Banach space.*

Observe that any composition of bounded linear operators is a bounded linear operator. Indeed, if $S \in B(X, Y)$, $T \in B(Y, Z)$, then $TS \in B(X, Z)$ and $\|TS\| \le \|T\| \, \|S\|$ since

$$\|TSx\| \le \|T\| \, \|Sx\| \le \|T\| \, \|S\| \, \|x\| \quad \text{for any } x \in X.$$

In particular, the space $B(X, X)$ of bounded linear operators on X (often denoted by $B(X)$ for simplicity) with composition is an algebra with identity.

EXAMPLE 5.1. Given $1 \leq q \leq \infty$ and let p be its conjugate exponent. Given $g \in L^q(\mathbb{R}^d)$. By Hölder's inequality, the linear map $G : L^p(\mathbb{R}^d) \to \mathbb{R}$ defined by $G(f) = \int_{\mathbb{R}^d} fg$ is bounded and $\|G\| \leq \|g\|_q$.

EXAMPLE 5.2. Given $g \in L^1(\mathbb{R}^d)$, $1 \leq p \leq \infty$. Define the convolution operator $G : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ by $G(f) = f * g$. Lemma 4.1 tells us that G is a bounded linear operator and $\|G\| \leq \|g\|_1$. More generally, if $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, $g \in L^q(\mathbb{R}^d)$, then Young's convolution inequality (exercise 4.1) tells us that $G : L^p(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$ is a bounded linear operator and $\|G\| \leq \|g\|_q$.

DEFINITION 5.3. Given $T \in B(X, Y)$. The *adjoint* of T, denoted by T^*, is a linear operator from Y^* to X^* defined by

$$(T^*y^*)(x) = y^*(Tx),$$

where $y^* \in Y^*$, $x \in X$. In other words, $T^*y^* = y^* \circ T$.

It is straightforward to verify that T^* is linear. T^*y^* is indeed a bounded linear functional on X since it is simply composition of y^* and T. Moreover, $T^* \in B(Y^*, X^*)$ since

$$\|T^*y^*\| = \|y^* \circ T\| = \sup_{\|x\|=1} \|y^*(Tx)\| \leq \sup_{\|x\|=1} \|y^*\| \, \|Tx\| \leq \|T\| \|y^*\| \quad \text{for any } y^* \in Y^*.$$

A commonly used notation for $x^*(x)$ is $\langle x^*, x\rangle$, where $x^* \in X^*$ and $x \in X$.

REMARK 5.1. If $S \in B(X, Y)$, $T \in B(Y, Z)$, then $(TS)^* = S^*T^*$. This follows trivially from the definition of adjoint:

$$(TS)^*z^* = z^* \circ (TS) = (z^* \circ T) \circ S = S^*(T^*z^*) = (S^*T^*)z^* \quad \text{for any } z^* \in Z^*.$$

DEFINITION 5.4. Two normed spaces X, Y are said to be *isomorphic* if there exists a bijective $T \in B(X, Y)$ with inverse $T^{-1} \in B(Y, X)$. In this case we say T is *invertible*. Such an operator T is called an *isomorphism*. We say X, Y are *isometrically isomorphic* if there exists an isomorphism $T : X \to Y$ which is also an isometry.

Two normed spaces are considered equivalent if they are isometrically isomorphic. This clearly defines an equivalence relation on the class of Banach spaces.

REMARK 5.2. We remark here that T is invertible implies T^* is invertible, and $(T^*)^{-1} = (T^{-1})^*$. Indeed, given any $x^* \in X^*$,

$$T^*(T^{-1})^*(x^*) = T^*(x^* \circ T^{-1}) = x^*(T \circ T^{-1}) = x^*.$$

Thus $(T^{-1})^*$ is a right inverse of T^*. The proof for being a left inverse is similar.

DEFINITION 5.5. Given any normed space X, there exists a canonical isometric isomorphism τ from X to a subspace of X^{**}. It is defined by $\tau(x)(x^*) = x^*(x)$. The mapping τ is often called the *canonical embedding* from X to X^{**}. For convenience, the term $\tau(x)$ is often written x, thereby treating elements in X as elements in X^{**}.

We say a normed space X is *reflexive* if the canonical embedding $\tau : X \to X^{**}$ is an isomorphism.

Note that, by Corollary 5.3, any reflexive normed space is a Banach space. The following theorem can be served as incentives to consider isometric isomorphisms and separable Banach spaces.

THEOREM 5.4. (BANACH-MAZUR THEOREM)
Every separable Banach is isometrically isomorphic to a closed subspace of $C^0[0, 1]$.

We omit the proof of the theorem. Note that the space $\mathbb{Q}[t]$ of polynomials with rational coefficients is a countable dense subset of $C^0[0,1]$, according to either the Weierstrass approximation theorem or the Stone-Weierstrass theorem. The Banach-Mazur theorem tells us that $C^0[0,1]$ is the "largest" separable Banach space.

Exercises.

5.1. Verify identities above Theorem 5.2.

5.2. Determine the operator norm $\|G\|$ in Examples 5.1 (without using any theorem in the next section), and Example 5.2 for the case $q = 1$.[1]

5.3. Given $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Suppose $K \in L^q(\mathbb{R}^d \times \mathbb{R}^d)$. Consider the integral operator T defined by

$$T(f)(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\, dy, \quad f \in L^p(\mathbb{R}^d).$$

(a) Given $f \in L^p(\mathbb{R}^d)$. Show that for almost every x, the function $K(x,y)f(y)$ is integrable with respect to y.
(b) Show that T is a bounded operator from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$, and $\|T\| \le \|K\|_q$.

5.4. Assume the following axiom[2]:

For any linear subspace X_0 of X, $\ell \in X_0^*$, there exists $\tilde{\ell} \in X^*$ such that $\tilde{\ell}\,|_{X_0} = \ell$ and $\|\tilde{\ell}\| = \|\ell\|$.

Now prove the following statements under this assumption.

(a) For any $x \in X$, $\|x\| = \sup\{|x^*(x)| : x^* \in X^*,\ \|x^*\| \le 1\}$.
(*Hint*: Consider a function which sends αx to $\alpha\|x\|$, where α is a scalar.)
(b) For any $T \in B(X,Y)$, $\|T\| = \sup\{|y^*(Tx)| : y^* \in Y^*,\ \|y^*\| \le 1,\ x \in X,\ \|x\| \le 1\}$.
(c) $\|T\| = \|T^*\|$ for any $T \in B(X,Y)$.

6. Dual Space of L^p

In this section we characterize the dual space of $L^p(I)$, where I is an interval. A more general theorem will be proved after general measures and integrals were introduced.

The key ingredients of the proof are the following two lemmas. The first one says that, roughly speaking, any bounded linear functional on $L^p(I)$ induces a canonical indefinite integral. Generalization of this lemma to higher dimensional spaces requires further knowledge about indefinite integrals.

Lemma 6.1. *Let I be a bounded interval, $\bar{I} = [a,b]$, $1 \le p < \infty$. For any $G \in L^p(I)^*$, there exists $g \in L^1(I)$ such that $G(\chi_A) = \int_A g$ for any measurable set $A \subset I$.*

Proof. Without loss of generality, assume $I = [a,b]$. First note that $G(\chi_A) = G(\chi_B)$ if $A \Delta B$ has measure zero. Let $\phi(s) = G(\chi_{[a,s]})$, $s \in I$. Then for any subinterval $[s,t] \subset I$,

$$\phi(t) - \phi(s) = G(\chi_{[a,t]}) - G(\chi_{[a,s]}) = G(\chi_{[a,t]} - \chi_{[a,s]}) = G(\chi_{[s,t]}).$$

[1] Determination of the operator norm for Example 5.2 amounts to finding sharp form of Young's convolution inequality. The case $q = 1$ is elementary, but the case $q \ne 1$ is challenging and requires further tools. See: W. Beckner, Inequalities in Fourier Analysis, *Ann. Math.* (1975).

[2] It is a special and simple version of a more general proposition known as the *Hahn-Banach theorem*, which is a consequence of the axiom of choice.

Given almost disjoint subintervals $\{[a_i, b_i] : i = 1, \cdots, n\}$ of I with total length δ.

$$\begin{aligned}
\sum_{i=1}^{n} |\phi(b_i) - \phi(a_i)| &= \sum_{i=1}^{n} (\phi(b_i) - \phi(a_i)) \operatorname{sgn}(\phi(b_i) - \phi(a_i)) \\
&= \sum_{i=1}^{n} G(\chi_{[a_i,b_i]}) \operatorname{sgn}(\phi(b_i) - \phi(a_i)) \\
&= G\left(\sum_{i=1}^{n} \operatorname{sgn}(\phi(b_i) - \phi(a_i))\, \chi_{[a_i,b_i]}\right) \\
&\leq \|G\| \left\|\sum_{i=1}^{n} \operatorname{sgn}(\phi(b_i) - \phi(a_i))\, \chi_{[a_i,b_i]}\right\|_p \leq \|G\|\delta^{1/p}.
\end{aligned}$$

This implies that $\phi \in AC[a, b]$. By the fundamental theorem of calculus, $g = \phi' \in L^1(I)$ and

$$G(\chi_{[s,t]}) = \phi(t) - \phi(s) = \int_s^t g(u)\, du \quad \text{for any subinterval } [s, t] \subset I.$$

It follows that $G(\chi_A) = \int_A g$ for any measurable set A in I (exercise 6.1). $\square$

Lemma 6.2. *Let $1 \leq p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Let $E \subset \mathbb{R}$ be measurable. Suppose $g \in L^1(E)$ and for some $M > 0$,*

$$\left|\int_E \varphi g\right| \leq M\|\varphi\|_p$$

for any simple function φ. Then $g \in L^q(E)$ and $\|g\|_q \leq M$.

Proof. Assume $1 < p < \infty$. Let ψ_n be a sequence of nonnegative simple functions with compact support such that $\psi_n \nearrow |g|^q$. Let $\varphi_n = \psi_n^{\frac{1}{p}} \operatorname{sgn}(g)$. Then φ_n is a simple function and

$$\varphi_n g = \psi_n^{\frac{1}{p}} |g| \geq \psi_n^{\frac{1}{p}+\frac{1}{q}} = \psi_n.$$

Then,

$$\int_E \psi_n \leq \int_E \varphi_n g \leq M\|\varphi_n\|_p = M\left(\int_E \psi_n\right)^{\frac{1}{p}},$$

implying that $\int_E \psi_n \leq M^q$. By the monotone convergence theorem, $\int_E |g|^q \leq M^q$. The case $p = 1$ is left to the reader (exercise 6.2). $\square$

Theorem 6.3. (Riesz Representation Theorem for $L^p(I)$, $I \subset \mathbb{R}$)
Suppose $1 \leq p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $I \subset \mathbb{R}$ is an interval. For any $G \in L^p(I)^*$, there exists unique $g \in L^q(I)$ such that

$$G(f) = \int_E fg \quad \textit{for any } f \in L^p(I).$$

Moreover, $\|G\| = \|g\|_q$. The map $G \mapsto g$ is an isometric isomorphism from $L^p(I)^*$ to $L^q(I)$.

Proof. Denote the σ-algebra of measurable sets in I by $\mathcal{B}$. Consider the case $\bar{I} = [a, b]$. By Lemma 6.1, there exists $g \in L^1(I)$ such that $G(\chi_A) = \int_A g$ for any $A \in \mathcal{B}$. Given any simple function $\varphi = \sum_{i=1}^{n} a_i \chi_{E_i}$, where $\{E_i\} \subset \mathcal{B}$ are disjoint.

$$G(\varphi) = \sum_{i=1}^{n} a_i G(\chi_{E_i}) = \sum_{i=1}^{n} a_i \int_{E_i} g = \sum_{i=1}^{n} a_i \int_{E} \chi_{E_i} g = \int_I \varphi g.$$

Therefore,

$$\left| \int_I \varphi g \right| = |G(\varphi)| \leq \|G\| \, \|\varphi\|_p.$$

By Lemma 6.2, $g \in L^q(I)$ and $\|g\|_q \leq \|G\|$. Consider the linear functional $\tilde{G}$ on $L^p(I)$ defined by $\tilde{G}(f) = \int_I fg$. It is continuous, by Hölder's inequality. Since simple functions are dense in $L^p(I)$, we see that $G = \tilde{G}$. Hölder's inequality also tells us that $\|G\| \leq \|g\|_q$, and so $\|g\|_q = \|G\|$. The function $g \in L^q(I)$ is unique, for if $G(f) = \int_E f\tilde{g}$ for all $f \in L^p(I)$, we would have

$$\int_I f(g - \tilde{g}) = 0 \quad \text{for all } f \in L^p(I).$$

Then $g - \tilde{g}$ gives the zero functional on $L^p(I)$, so that $\|g - \tilde{g}\|_q = 0$. That is, $g = \tilde{g}$ in $L^q(I)$.

Now we consider unbounded interval I. Let $\{I_n\}$ be an increasing sequence of bounded intervals such that $\bigcup_{n=1}^{\infty} I_n = I$. Then for any n, there exists unique $g_n \in L^q(I)$ such that $g_n = 0$ on I_n^c and

$$G(f) = \int_I f g_n \quad \text{for any } f \in L^p(I) \text{ with } f = 0 \text{ on } I_n^c.$$

Moreover, $\|g_n\|_q \leq \|G\|$. By uniqueness, $g_{n+1} = g_n$ a.e. on I_n for each n. We may assume without loss of generality that $g_{n+1} = g_n$ on I_n. Then the function

$$g(x) := g_n(x), \quad x \in I_n$$

is well-defined on I. Furthermore, $|g_n(x)| \nearrow |g(x)|$ as $n \to \infty$. By the monotone convergence theorem,

$$\int_I |g|^q = \lim_{n\to\infty} \int_I |g_n|^q \leq \|G\|^q.$$

In particular, $g \in L^q(I)$, $\|g\|_q \leq \|G\|$. For any $f \in L^p(I)$, let $f_n = f\,\chi_{I_n}$. Then $|f_n| \leq |f|$, $f_n \to f$ pointwise on I as $n \to \infty$. It follows that $\|g_n - g\|_q \to 0$, $\|f_n - f\|_p \to 0$ (see exercise 3.6), and hence

$$\begin{aligned} \|f_n g_n - fg\|_1 &\leq \|(f_n - f)g_n\|_1 + \|f(g_n - g)\|_1 \\ &\leq \|f_n - f\|_p \, \|g_n\|_q + \|f\|_p \, \|g_n - g\|_q \;\to\; 0 \quad \text{as } n \to \infty. \end{aligned}$$

Then, by continuity of G on $L^p(I)$,

$$\int_I fg = \lim_{n\to\infty} \int_I f_n g_n = \lim_{n\to\infty} G(f_n) = G(f).$$

Moreover, $\|G\| \leq \|g\|_q$, and so $\|g|_q = \|G\|$. Uniqueness of such g follows as in the case of bounded I. This completes the proof. $\square$

Corollary 6.4. *If $1 < p < \infty$, then $L^p(I)$ is reflexive.*

Proof. Let q be the conjugate exponent of p. Consider $i_p : L^p \to (L^q)^*$ defined by $i_p(f)(g) = \int fg$. The Riesz representation theorem tells us that i_p is an isometric isomorphism. Likewise, $i_q(g)(f) = \int gf$ defines an isometric isomorphism i_q from L^q to $(L^p)^*$. By Remark 5.2, its adjoint $i_q^* : (L^p)^{**} \to (L^q)^*$ is invertible and $(i_q^*)^{-1} = (i_q^{-1})^*$. The corollary follows by observing that the canonical embedding $\tau : L^p \to (L^p)^{**}$ is equal to $(i_q^*)^{-1} \circ i_p$. Verification for this relation is straightforward: given $f \in L^p$, $G \in (L^p)^*$,

$$\begin{aligned}\tau(f)(G) &= G(f) = (i_q \circ i_q^{-1} \circ G)(f) = \int i_q^{-1}(G) f \\ &= i_p(f)(i_q^{-1}(G)) = (i_q^{-1})^*(i_p(f))(G) = (i_q^*)^{-1} \circ i_p(f)(G).\end{aligned}$$

□

Exercises.

6.1. Let $I \subset \mathbb{R}$ be a bounded interval. Given $1 \le p < \infty$, $G \in L^p(I)^*$, $g \in L^1(I)$. Suppose $G(\chi_J) = \int_J g$ on any subinterval $J \subset I$. Show that $G(\chi_A) = \int_A g$ for any measurable set $A \subset I$.

6.2. Prove Lemma 6.2 for the case $p = 1$.

6.3. Let $I \subset \mathbb{R}$ be an interval. Suppose $g \in L^\infty(I)$. Show that for any $\varepsilon > 0$ there exists $f \in L^1(I)$, $\|f\|_1 \neq 0$, such that

$$\int_I fg \ge \|f\|_1(\|g\|_\infty - \varepsilon).$$

6.4. Let $I = [0, 1]$. Consider the linear functional G on $C(I)$ defined by $G(f) = f(1)$. Use this linear functional and the assumption in Exercise 5.4 to show that $L^1(I)$ is not isometrically isomorphic to $L^\infty(I)^*$.

6.5. Determine a representation for $(\ell^p)^*$, $1 \le p < \infty$.

7. Maximal Functions on L^p

We have seen that the Hardy-Littlewood maximal function plays a critical role in the theory of differentiation. It is defined for functions in $L^1_{\text{loc}}(\mathbb{R}^d)$, and that certainly includes $L^p(\mathbb{R}^d)$ for $1 \le p < \infty$. Here we denote the Hardy-Littlewood maximal function of f by f^*:

$$f^*(x) = \sup_{0<r<\infty} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f|.$$

Recall that for any $f \in L^1(\mathbb{R}^d)$, its maximal function f^* is not in $L^1(\mathbb{R}^d)$ unless $f = 0$ a.e.. In sharp contrast, for any $f \in L^p(\mathbb{R}^d)$ with $1 < p < \infty$, we always have $f^* \in L^p(\mathbb{R}^d)$. Furthermore, the mapping $M : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ defined by $M(f) = f^*$ is bounded (but nonlinear). This result, to which this section is devoted, has important applications in harmonic analysis and ergodic theory.

Definition 7.1. Let f be a measurable function on $E \subset \mathbb{R}^d$. The *distribution function* of f on E is defined by

$$w_{f,E}(\alpha) := m(\{x \in E : f(x) > \alpha\}).$$

Clearly w is monotone decreasing. If $m(E) < \infty$, then w is bounded, so it has bounded variation and its total variation is $m(E)$. We begin with an observation which links Lebesgue integral with improper Riemann integral:

Theorem 7.1. *Suppose $f \ge 0$ is measurable on $E \subset \mathbb{R}^d$ and its distribution function $w = w_{f,E}$ is bounded. Given any real-valued C^1 function ϕ on $(0, \infty)$, we have*

$$\int_E \phi \circ f = -\phi w\Big|_0^\infty + \int_0^\infty \phi'(\alpha) w(\alpha) d\alpha.$$

PROOF. Given $0 < a < b < \infty$. Let $E_{a,b} = \{a < f \le b\}$. We claim that

$$\int_{E_{a,b}} \phi \circ f \;=\; -\phi w\Big|_a^b + \int_a^b \phi'(\alpha) w(\alpha) d\alpha.$$

The theorem follows by observing $E = \bigcup_{k=1}^{\infty} E_{1/k,k}$ and letting $a \to 0+$, $b \to \infty$.

We first observe that $\phi \circ f$ is measurable (since ϕ is C^1), so the Lebesgue integral on the left side makes sense. Also, $\phi' w$ has a countable number of discontinuities, so the Riemann integral on the right side also makes sense.

Since ϕ is of bounded variation on $[a, b]$, it can be written as the difference of two bounded increasing functions, so there is no loss of generality by assuming that ϕ is increasing. Then, given any partition $\mathcal{P} = \{\alpha_0 < \alpha_1 < \cdots < \alpha_n\}$ of $[a, b]$,

$$\begin{aligned}
\int_{E_{a,b}} \phi \circ f &= \sum_{k=1}^{n} \int_{E_{\alpha_{k-1},\alpha_k}} \phi \circ f \\
&\ge \sum_{k=1}^{n} \phi(\alpha_{k-1}) m(E_{\alpha_{k-1},\alpha_k}) \\
&= \sum_{k=1}^{n} \phi(\alpha_{k-1}) \left(w(\alpha_{k-1}) - w(\alpha_k)\right) \\
&= \phi(\alpha_0) w(\alpha_0) - \phi(\alpha_{n-1}) w(\alpha_n) + \sum_{k=1}^{n-1} w(\alpha_k) \left(\phi(\alpha_k) - \phi(\alpha_{k-1})\right) \\
&= \phi(a) w(a) - \phi(\alpha_{n-1}) w(b) + \sum_{k=1}^{n-1} w(\alpha_k) \phi'(\xi_k) \left(\alpha_k - \alpha_{k-1}\right)
\end{aligned}$$

for some $\xi_k \in [\alpha_{k-1}, \alpha_k]$. Let $|\mathcal{P}| \to 0$, we find a lower bound for $\int_{E_{a,b}} \phi \circ f$:

$$\phi(a) w(a) - \phi(b) w(b) + \int_a^b w(\alpha) \phi'(\alpha) d\alpha.$$

Similarly,

$$\begin{aligned}
\int_{E_{a,b}} \phi \circ f &\le \sum_{k=1}^{n} \phi(\alpha_k) m(E_{\alpha_{k-1},\alpha_k}) \\
&= \sum_{k=1}^{n} \phi(\alpha_k) \left(w(\alpha_{k-1}) - w(\alpha_k)\right) \\
&= \phi(\alpha_1) w(\alpha_0) - \phi(\alpha_n) w(\alpha_n) + \sum_{k=1}^{n-1} w(\alpha_{k-1}) \left(\phi(\alpha_k) - \phi(\alpha_{k-1})\right) \\
&= \phi(\alpha_1) w(a) - \phi(b) w(b) + \sum_{k=1}^{n-1} w(\alpha_{k-1}) \phi'(\eta_k) \left(\alpha_k - \alpha_{k-1}\right)
\end{aligned}$$

for some $\eta_k \in [\alpha_{k-1}, \alpha_k]$. Let $|\mathcal{P}| \to 0$, then the lower bound for $\int_{E_{a,b}} \phi \circ f$ we found above becomes an upper bound, so they must be equal. $\square$

THEOREM 7.2. *Suppose $f \in L^p(E)$, $1 < p < \infty$. Then*

$$\int_E |f|^p \;=\; p\int_0^\infty \alpha^{p-1} w_{|f|,E}(\alpha)d\alpha.$$

PROOF. The distribution function $w_{|f|,E}$ of $|f|$ on E is bounded because (as the proof of Chebyshev's inequality)

$$\alpha^p w_{|f|,E}(\alpha) \;=\; \alpha^p m\left(\{|f| > \alpha\}\right) \;\le\; \int_{|f|>\alpha} |f|^p \;<\; \infty. \tag{9}$$

Replace the function f in Theorem 7.1 by $|f|$ and let $\phi(s) = s^p$. Denote $w_{|f|,E}$ by w for convenience. Then the corollary follows once we have proved

$$\lim_{\alpha\to 0+} \alpha^{p-1} w(\alpha) \;=\; 0 \;=\; \lim_{\alpha\to\infty} \alpha^{p-1} w(\alpha).$$

The second identity follows trivially from (9). We leave the first identity as exercise. $\square$

THEOREM 7.3. *If $f \in L^p(\mathbb{R}^d)$, $1 < p \le \infty$, then $f^* \in L^p(\mathbb{R}^d)$ and*

$$\|f^*\|_p \le C\|f\|_p$$

for some constant C which depends only on d and p.

PROOF. The case $p = \infty$ is obvious since $\|f^*\|_\infty \le \|f\|_\infty$. Assume $1 < p < \infty$. Denote the distribution function of f^* on $\mathbb{R}^d$ by w.

Given $\alpha > 0$. Let $g = f \cdot \chi_{\{|f| \ge \frac{\alpha}{2}\}}$. Then

$$\begin{aligned} |f(x)| &\le |g(x)| + \frac{\alpha}{2}. \\ f^*(x) &\le \sup_{0<r<\infty} \frac{1}{m(B_r)} \int_{B_r(x)} |g(y)| + \frac{\alpha}{2} \;=\; g^*(x) + \frac{\alpha}{2}. \end{aligned}$$

Therefor, $\{f^* > \alpha\} \subset \{g^* > \frac{\alpha}{2}\}$ and hence

$$w(\alpha) \le m(\{g^* > \frac{\alpha}{2}\}) \le \frac{2c}{\alpha}\|g\|_1 = \frac{2c}{\alpha}\int_{\{|f|\ge\frac{\alpha}{2}\}} |f| \quad \forall\, \alpha > 0.$$

Note that $g \in L^1$ because $E = \{|f| \ge \frac{\alpha}{2}\}$ has finite measure and $L^p(E) \subset L^1(E)$.

$$\begin{aligned} \int_{\mathbb{R}^d} (f^*)^p &= p\int_0^\infty \alpha^{p-1} w(\alpha)d\alpha \\ &\le 2cp\int_0^\infty \int_{\{|f|\ge\frac{\alpha}{2}\}} \alpha^{p-2}|f(x)|dx\,d\alpha \\ &= 2cp\int_{\mathbb{R}^d}\int_0^{2|f(x)|} \alpha^{p-2}|f(x)|d\alpha\,dx \quad \text{(by Tonelli's Theorem)} \\ &= \frac{2cp}{p-1}\int_{\mathbb{R}^d} |f(x)|\alpha^{p-1}\Big|_{\alpha=0}^{\alpha=2|f(x)|}dx \\ &= \frac{2^p cp}{p-1}\int_{\mathbb{R}^d} |f(x)|^p dx \end{aligned}$$

This completes the proof by taking $C = 2(\frac{cp}{p-1})^{\frac{1}{p}}$. $\square$