6 L^p Spaces

L^p spaces are the most interesting examples of Banach spaces and play a salient role in modern analysis. In this chapter basic features of L^p spaces are studied; in particular, their dual spaces are identified. Special attention is directed towards $L^p(\Omega)$ where Ω is an open set in $\mathbb{R}^n$, for example, convolution and maximal function operators in L^p, are treated. An important class of function spaces, which is related to L^p spaces and was first introduced by S.L. Sobolev in his study of equations of mathematical physics, is briefly introduced in the last section of the chapter. Further study of this class of spaces is taken up in Chapter 7 by applying the method of Fourier integrals.

Some useful inequalities for functions in L^p spaces are collected in the first section for later reference. The second section on signed and complex measures is primarily preliminary in nature for this chapter, but it also has its own merit of interest, as is shown by the Riesz representation theorem in the concluding part of the section.

6.1 Some inequalities

Some inequalities which appear frequently in studies related to L^p spaces are collected here for later reference.

6.1.1 Markov inequality

Let $f \in L^p(\Omega, \Sigma, \mu)$, $1 \le p < \infty$, then

$$\mu(\{|f| \ge \lambda\}) \le \lambda^{-p}\|f\|_p^p, \tag{6.1}$$

for all $\lambda > 0$.

The inequality (6.1), called the **Markov inequality**, follows readily from the sequence of inequalities,

$$\lambda^p \mu(\{|f| \ge \lambda) \le \int_{\{|f|\ge\lambda\}} |f|^p d\mu \le \|f\|_p^p.$$

Remark Since $\lim_{\lambda\to\infty}\mu\{|f|\geq\lambda\}=0$ by (6.1), it follows from Exercise 2.5.9 (iii) that $\lim_{\lambda\to\infty}\int_{\{|f|\geq\lambda\}}|f|^p d\mu=0$, and hence

$$\lim_{\lambda\to\infty}\lambda^p\mu(\{|f|\geq\lambda\})=0. \tag{6.2}$$

6.1.2 Chebyshev inequality

Let $f\in L^2(\Omega,\Sigma,P)$, where (Ω,Σ,P) is a probability space, then the following **Chebyshev inequality** is a special case of (6.1):

$$P(\{|f-E(f)|\geq\lambda)\leq\lambda^{-2}\,\mathrm{Var}(f), \tag{6.3}$$

where $E(f)=\int_\Omega f dP$ and $\mathrm{Var}(f)=\int_\Omega|f-E(f)|^2 dP$.

Remark A measurable function f on a probability space is called a random variable. If $\int_\Omega f dP$ exists, it is called the expectation of the random variable f and is denoted by $E(f)$; if $E(f)$ is finite, $\int_\Omega|f-E(f)|^2 dP$ is called the variance of f and is denoted by $\mathrm{Var}(f)$. The significance of Chebyshev inequality in probability theory will become clear when the concept of independence is introduced in Chapter 7.

6.1.3 Jensen inequality

Suppose that φ is a convex function defined on $\mathbb{R}$, and f is an integrable function on a probability space (Ω,Σ,P), then

$$\varphi(E(f))\leq E(\varphi\circ f). \tag{6.4}$$

This inequality is referred to as the **Jensen inequality**. For the verification of (6.4), let us put $x=E(f)$ and choose $m\in[\varphi'_-(x),\varphi'_+(x)]$. By Proposition 5.4.1 (iv),

$$\varphi(x)+m(y-x)\leq\varphi(y)$$

for all $y\in\mathbb{R}$, and hence,

$$\varphi(x)+m(f(\omega)-x)\leq\varphi(f(\omega)) \tag{6.5}$$

for all $\omega\in\Omega$. It follows from (6.5) that

$$\{\varphi\circ f\}^-\leq|\varphi(x)|+|m||f|+|mx|,$$

and therefore $\{\varphi\circ f\}^-$ is integrable; consequently, $\int_\Omega\varphi\circ f dP$ exists. We can then integrate both sides of (6.5) over Ω to obtain

$$\varphi(x)+m(E(f)-x)\leq E(\varphi\circ f),$$

which reduces to (6.4), because $x=E(f)$. Thus the Jensen inequality is verified.

Remark If (Ω, Σ, μ) is a finite measure space and f is integrable on (Ω, Σ, μ), then the Jensen inequality leads to

$$\varphi\left(\frac{1}{\mu(\Omega)}\int_\Omega f d\mu\right) \leq \frac{1}{\mu(\Omega)}\int_\Omega \varphi \circ f d\mu. \tag{6.6}$$

In particular, $\left|\frac{1}{\mu(\Omega)}\int_\Omega f d\mu\right|^p \leq \frac{1}{\mu(\Omega)}\int |f|^p d\mu$ for $1 \leq p < \infty$.

6.1.4 Extended Hölder inequality

Suppose that $f_1, \dots, f_n$, $n \geq 3$, are measurable functions on a measure space (Ω, Σ, μ) and let $p_1 \geq 1, p_2 \geq 1, \dots, p_n \geq 1$ be extended real numbers such that $\sum_{i=1}^n p_i^{-1} = 1$, then the following **extended Hölder inequality** holds:

$$\int_\Omega \left|\prod_{i=1}^n f_i\right| d\mu \leq \prod_{i=1}^n \|f_i\|_{p_i}. \tag{6.7}$$

To see that (6.7) holds, it is sufficient to consider the case where $n = 3$; then (6.7) follows inductively. So consider the case where $n = 3$ and let $p^{-1} = \frac{1}{p_1} + \frac{1}{p_2}$. Since p and p_3 are conjugate exponents, by the Hölder inequality, we have

$$\int_\Omega |f_1 f_2 f_3| d\mu \leq \|f_1 f_2\|_p \cdot \|f_3\|_{p_3}. \tag{6.8}$$

Then put $p' = \frac{p_1}{p}$, $q' = \frac{p_2}{p}$ and apply the Hölder inequality, to obtain

$$\begin{aligned}\|f_1 f_2\|_p^p &= \int_\Omega |f_1|^p |f_2|^p d\mu \leq \left(\int_\Omega |f_1|^{pp'} d\mu\right)^{1/p'} \left(\int_\Omega |f_2|^{pq'} d\mu\right)^{1/q'} \\ &= \|f_1\|_{p_1}^p \|f_2\|_{p_2}^p,\end{aligned}$$

or $\|f_1 f_2\|_p \leq \|f_1\|_{p_1} \cdot \|f_2\|_{p_2}$. This last inequality and (6.8) imply that (6.7) holds when $n = 3$.

Exercise 6.1.1 Suppose that Ω is a measurable subset of $\mathbb{R}^n$ with $\lambda^n(\Omega) > 0$, and f is a measurable function on Ω. Show that $f \in L^p(\Omega)$, $1 \leq p < \infty$, if and only if for every $\varepsilon > 0$, there is a closed set $F \subset \Omega$ and a bounded continuous function g in $L^p(\mathbb{R}^n)$, such that $\lambda^n(\Omega \backslash F) < \varepsilon$, $f = g$ on F, and $\|f - g\|_p < \varepsilon$. (Hint: cf. (6.2) and Theorem 4.1.3.)

Exercise 6.1.2 Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is an orthonormal system in $L^2(\Omega, \Sigma, \mu)$. Show that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\left(\left\{\frac{\left|\sum_{k=1}^n f_k\right|}{n} \geq \varepsilon\right\}\right) = 0.$$

Exercise 6.1.3 Suppose that $\{f_n\}$ is a sequence in $L^p(\Omega, \Sigma, \mu)$, $1 \leq p < \infty$, which converges in $L^p(\Omega, \Sigma, \mu)$ to f. Show that $\{f_n\}$ has a subsequence which converges a.e. to f. (Hint: there are positive integers $n_1 < n_2 < \cdots < n_k < \cdots$ such that $\mu(\{|f_{n_k} - f| \geq \frac{1}{k}\}) \leq \frac{1}{k^2}$ for each $k \in \mathbb{N}$.)

Exercise 6.1.4 Suppose that $\sum_{n=1}^{\infty} \alpha_n = 1$, where $\alpha_n \geq 0$ for each n. Show that if $\{\beta_n\}$ is a sequence of real numbers such that $\sum_{n=1}^{\infty} \alpha_n |\beta_n| < \infty$, then

$$\left|\sum_{n=1}^{\infty} \alpha_n \beta_n\right|^p \leq \sum_{n=1}^{\infty} \alpha_n |\beta_n|^p$$

for $1 \leq p < \infty$.

6.2 Signed and complex measures

So far the integration is taken with respect to a measure on a measurable space (Ω, Σ), where a measure is understood to be a nonnegative σ-additive set function defined on Σ. But there naturally appear set functions which may take negative values, such as electric charges, and integration with respect to such set functions is a useful construct, such as the potential of the electric charge distribution. Our purpose in this section is firstly to generalize the concept of measure to cover situations when negative values might be assumed, and then to consider complex measures. In order to do this, we extend the concept of sum for systems of real numbers in Section 1.1 to systems which may contain ∞ or $-\infty$. This can be done naturally as follows. Let $\{c_\alpha\}_{\alpha \in I}$ be a system of extended real numbers; by considering $\{c_\alpha\}_{\alpha \in I}$ as a function on I, we say that the sum of $\{c_\alpha\}$ exists if its integral with respect to the counting measure on I exists. This integral is called the sum of $\{c_\alpha\}$ and is denoted by $\sum_{\alpha \in I} c_\alpha$, or $\sum_\alpha c_\alpha$ if I is clearly implied (cf. Examples 2.3.1 and 2.3.3). Note that $\{c_\alpha\}$ is summable if and only if $\sum_\alpha c_\alpha$ exists and is finite.

Let (Ω, Σ) be a measurable space; a set function $\sigma : \Sigma \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is called a **signed measure** on (Ω, Σ) if

(i) $\sigma(\emptyset) = 0$;

(ii) if $\{A_n\} \subset \Sigma$ is a disjoint sequence, then the sum $\sum_n \sigma(A_n)$ exists and

$$\sigma\left(\bigcup_n A_n\right) = \sum_n \sigma(A_n). \tag{6.9}$$

We remark first that if $\sigma(\bigcup_{n=1}^{\infty} A_n)$ is finite, then $\sum_n \sigma(A_n)$ on The right-hand side of (6.9) can be written as $\sum_{n=1}^{\infty} \sigma(A_n)$ which necessarily converges absolutely, because $\bigcup_{n=1}^{\infty} A_n$ does not depend on the order of $A_1, A_2, \ldots$. Secondly, we call attention to the fact that condition (ii) in the above definition forces σ to satisfy condition (iii);

(iii) The signed measure σ does not take both ∞ and $-\infty$ as its value.

In fact, if $\sigma(A) = -\infty$, $\sigma(B) = \infty$ for A, B in Σ, then,

$$\sigma(A \cup B) = \sigma(A \cap B) + \sigma(A \cap B^c) + \sigma(B \cap A^c)$$

does not make sense, because $-\infty$ and ∞ both appear on the right-hand side in all possible situations, as can easily be seen.

For definiteness, we shall assume that in the sequel, condition (iii)$'$ holds;

(iii)$'$ $\sigma(A) > -\infty$ for all $A \in \Sigma$.

Under this assumption, if $\{A_n\}$ is a disjoint sequence in Σ with $\sigma(\bigcup_n A_n) = \infty$, then $\sum_{n=1}^{\infty} \sigma(A_n)$ diverges to ∞.

Measures on Σ are certainly signed measures; to distinguish them from general signed measures, we shall sometimes refer to them as **positive measures**. Accordingly, if $\sigma(A) \leq 0$ for all $A \in \Sigma$, σ is called a **negative measure**.

Example 6.2.1 Let (Ω, Σ, μ) be a measure space.

(i) Suppose that $A_1, \ldots, A_k$ are disjoint sets from Σ with $\mu(A_j) < \infty$, $j = 1, \ldots, k$ and let $\alpha_1, \ldots, \alpha_k$ be real numbers. Define σ on Σ by

$$\sigma(A) = \sum_{j=1}^{k} \alpha_j \mu(A \cap A_j), \quad A \in \Sigma.$$

The set function σ is obviously a signed measure.

(ii) Suppose that f is a measurable function with $\int_\Omega f^- d\mu < \infty$, then

$$\sigma(A) = \int_A f d\mu, \quad A \in \Sigma,$$

is a signed measure.

Remark Signed measure σ, defined in Example 6.2.1 (ii), is usually referred to as the **indefinite integral** of f; but when Ω is a metric space and $\mathcal{B}(\Omega) \subset \Sigma$, the indefinite integral of f is sometimes restricted to $\mathcal{B}(\Omega)$. This should not cause any confusion, because the definite meaning of an indefinite integral will be clear from the context (cf. Example 3.8.1).

Example 6.2.2 Consider the measurable space $(\mathbb{R}, \mathcal{B})$ where $\mathcal{B}$ is the σ-algebra of all Borel sets in $\mathbb{R}$. Suppose that we order the set of all rational numbers by $\gamma_1, \gamma_2, \ldots, \gamma_n, \ldots$, and define σ on $\mathcal{B}$ by

$$\sigma(B) = \sum_{\gamma_n \in B} (-1)^n \frac{1}{n^2}, \quad B \in \mathcal{B}.$$

Then σ is a signed measure which assumes only finite values.

Exercise 6.2.1 Verify the following statements. Let σ be a signed measure on (Ω, Σ). If $\{E_n\} \subset \Sigma$ and $E_n \nearrow$, then

$$\sigma\left(\lim_{n\to\infty} E_n\right) = \sigma\left(\bigcup_n E_n\right) = \lim_{n\to\infty} \sigma(E_n);$$

if, on the other hand, $E_n \searrow$ and $\sigma(E_n) < \infty$ for some n, then

$$\sigma\left(\lim_{n\to\infty} E_n\right) = \sigma\left(\bigcap_n E_n\right) = \lim_{n\to\infty} \sigma(E_n).$$

Exercise 6.2.2 Show that if $|\sigma(E)| < \infty$, then $|\sigma(F)| < \infty$ for $F \subset E$.

We currently show that any signed measure is the difference of two positive measures, one of which is a finite measure.

In the following discussion, a fixed signed measure σ on a measurable space (Ω, Σ) is considered.

A set $E \in \Sigma$ is said to be **positive (negative)** if $\sigma(A \cap E) \geq 0$ (≤ 0) for all $A \in \Sigma$. Obviously, any measurable subset of a positive (negative) set is positive (negative). The empty set $\emptyset$ is both positive and negative. Certainly, if $A_1, A_2, \ldots, A_n, \ldots$ are positive (negative), then so is $\bigcup_n A_n$.

The family of all positive sets will be denoted by $\mathcal{P}_\sigma$, and that of all negative sets by $\mathcal{N}_\sigma$.

Lemma 6.2.1 *Let* $\beta = \inf_{E\in\mathcal{N}_\sigma} \sigma(E)$*; then* $-\infty < \beta \leq 0$ *and there is* $B \in \mathcal{N}_\sigma$ *such that* $\sigma(B) = \beta$.

Proof There is a sequence $\{B_n\}$ in $\mathcal{N}_\sigma$ such that

$$\beta = \lim_{n\to\infty} \sigma(B_n).$$

Take $B = \bigcup_n B_n$, then $B \in \mathcal{N}_\sigma$, and for each k,

$$\sigma(B) = \sigma(B_k) + \sigma(B \backslash B_k) \leq \sigma(B_k),$$

hence $\sigma(B) \leq \lim_{k\to\infty} \sigma(B_k) = \beta$. But $\sigma(B) \geq \beta$, so $\sigma(B) = \beta$. Since $\sigma(B) > -\infty$, we have $-\infty < \beta \leq 0$. ■

Theorem 6.2.1 (Hahn decomposition theorem) *There are disjoint sets A and B in* Σ *such that*

(i) $A \cup B = \Omega$;

(ii) $A \in \mathcal{P}_\sigma$ *and* $B \in \mathcal{N}_\sigma$.

Proof Let β and B be as in Lemma 6.2.1, and take $A = \Omega\backslash B$. It remains to show that $A \in \mathcal{P}_\sigma$. Suppose the contrary. Then there is a measurable set $E_0 \subset A$ such that $\sigma(E_0) < 0$. Naturally E_0 is not negative, because otherwise $B \cup E_0$ would be negative and $\sigma(B \cup E_0) = \sigma(B) + \sigma(E_0) < \beta$, contrary to the choice of β. Let k_1 be the smallest positive integer such that E_0 contains a measurable set E_1 with $\sigma(E_1) \geq \frac{1}{k_1}$.

Now, since $\sigma(E_0 \backslash E_1) = \sigma(E_0) - \sigma(E_1) \leq \sigma(E_0) - \frac{1}{k_1} < 0$, we can repeat the above argument with E_0 replaced by $E_0 \backslash E_1$. So, let k_2 be the smallest positive integer such that $E_0 \backslash E_1$ contains a measurable set E_2 with $\sigma(E_2) \geq \frac{1}{k_2}$. Continue in this fashion; we obtain a sequence of mutually disjoint measurable sets $E_1, E_2, \ldots, E_n, \ldots$ in E_0 and a sequence $k_1, k_2, \ldots, k_n, \ldots$ of positive integers such that for each $n \geq 2$, k_n is the smallest positive integer such that $E_0 \backslash (E_1 \cup \cdots \cup E_{n-1})$ contains a measurable set E_n with $\sigma(E_n) \geq \frac{1}{k_n}$. Since $\bigcup_{n=1}^{\infty} E_n \subset E_0$ and $|\sigma(E_0)| < \infty$, $|\sigma(\bigcup_{n=1}^{\infty} E_n)| < \infty$ (see Exercise 6.2.2), and hence,

$$\sum_{n=1}^{\infty} \frac{1}{k_n} \leq \sum_{n=1}^{\infty} \sigma(E_n) = \sigma\left(\bigcup_{n=1}^{\infty} E_n\right).$$

Thus $\sum_{n=1}^{\infty} \frac{1}{k_n}$ is a convergent series, and as a consequence,

$$\lim_{n \to \infty} \frac{1}{k_n} = 0. \tag{6.10}$$

Let $F_0 = E_0 \backslash \bigcup_{n=1}^{\infty} E_n$, then $\sigma(F_0) = \sigma(E_0) - \sum_{n=1}^{\infty} \sigma(E_n) \leq \sigma(E_0) < 0$. Consider a measurable set $F \subset F_0$; we claim that $\sigma(F) \leq 0$. If $\sigma(F) > 0$, then $\sigma(F) > \frac{1}{n_0}$ for some positive integer n_0; but (6.10) implies that $n_0 < k_n$ for sufficiently large n, thus contradicting the choice of k_n for such n's, because $F \subset E_0 \backslash \bigcup_{k=1}^{n-1} E_k$ for all n. Thus, $\sigma(F) \leq 0$ and consequently F_0 is a negative set. But then $F_0 \cup B$ is negative and $\sigma(F_0 \cup B) < \beta$, contrary to the choice of β. The contradiction proves the theorem. ■

The pair (A, B) in the statement of Theorem 6.2.1 is called a **Hahn decomposition** of Ω relative to the signed measure σ, or simply a **σ-decomposition** of Ω. In general, Hahn decomposition is not unique.

Exercise 6.2.3 Let σ be the signed measure of Example 6.2.2. Find two Hahn decompositions of $\mathbb{R}$ relative to σ.

Lemma 6.2.2 shows a close relation between any two Hahn decompositions of Ω relative to a signed measure σ.

Lemma 6.2.2 *Let (A_1, B_1) and (A_2, B_2) be Hahn decompositions of Ω relative to the signed measure σ; then for any $E \in \Sigma$ the following relations hold:*

$$\sigma(E \cap A_1) = \sigma(E \cap A_2); \quad \sigma(E \cap B_1) = \sigma(E \cap B_2).$$

Proof Since $A_1 \backslash A_2$ is positive, $\sigma(E \cap (A_1 \backslash A_2)) \geq 0$; on the other hand $E \cap (A_1 \backslash A_2) \subset B_2$ implies that $\sigma(E \cap (A_1 \backslash A_2)) \leq 0$. Hence $\sigma(E \cap (A_1 \backslash A_2)) = 0$; similarly, $\sigma(E \cap (A_2 \backslash A_1)) = 0$. Now,

$$\begin{aligned}
\sigma(E \cap A_1) &= \sigma(E \cap A_1) + \sigma(E \cap (A_2 \backslash A_1)) \\
&= \sigma(E \cap (A_1 \cup A_2)) = \sigma(E \cap A_2) + \sigma(E \cap (A_1 \backslash A_2)) \\
&= \sigma(E \cap A_2).
\end{aligned}$$

Similarly, $\sigma(E \cap B_1) = \sigma(E \cap B_2)$. ∎

For a Hahn decomposition (A, B) of Ω relative to σ, define for $E \in \Sigma$,

$$\sigma^+(E) = \sigma(E \cap A); \;\; \sigma^-(E) = -\sigma(E \cap B); \;\; \text{and } |\sigma|(E) = \sigma^+(E) + \sigma^-(E).$$

Obviously, σ^+, σ^-, and $|\sigma|$ are positive measures on Σ and are independent of the chosen Hahn decomposition (A, B), by Lemma 6.2.2. The measure $|\sigma|$ is called the **total variational measure** of σ, while σ^+ and σ^- are called respectively the **positive variational measure** and the **negative variational measure** of σ. Observe that $|\sigma(E)| \leq |\sigma|(E)$ for $E \in \Sigma$. Theorem 6.2.2 speaks for itself.

Theorem 6.2.2 *The measure σ^- is a finite positive measure and $\sigma = \sigma^+ - \sigma^-$. Furthermore, if σ is finite or σ-finite then so are σ^+ and $|\sigma|$.*

The decomposition $\sigma = \sigma^+ - \sigma^-$ is called the **Jordan decomposition** *of* σ.

Integrals and indefinite integrals of functions w.r.t. a signed measure σ are only defined for functions f in $L^1(\Omega, \Sigma, |\sigma|)$ by

$$\int_\Omega f d\sigma := \int_\Omega f d\sigma^+ - \int_\Omega f d\sigma^-;$$
$$\int_E f d\sigma := \int_E f d\sigma^+ - \int_E f d\sigma^-, \quad E \in \Sigma.$$

In the above definitions, f could be a complex-valued function.

Exercise 6.2.4 Show that for $E \in \Sigma$:

(i) $\sigma^+(E) = \max_{B\in\Sigma} \sigma(B \cap E)$;

(ii) $\sigma^-(E) = -\min_{B\in\Sigma} \sigma(B \cap E)$; and

(iii) $|\sigma|(E) = \sup\{\sum_{n=1}^{\infty} |\sigma(E_n)|\}$, where the supremum is taken over all decompositions of E into countable disjoint measurable sets $E_1, E_2, \ldots$.

Exercise 6.2.5 If σ is a finite signed measure, then

$$|\sigma|(E) = \sup \left| \int_E f d\sigma \right|,$$

where the supremum is taken over all measurable functions f with $|f| \leq 1$.

Exercise 6.2.6 Let σ be the signed measure in Example 6.2.1 (ii). Show that for $E \in \Sigma$, we have

$$\sigma^+(E) = \int_E f^+ d\mu; \quad \sigma^-(E) = \int_E f^- d\mu; \quad \text{and } |\sigma|(E) = \int_E |f| d\mu.$$

Also find a Hahn decomposition of Ω relative to σ.

Exercise 6.2.7 Let σ be a signed measure and $\sigma = \sigma_1 - \sigma_2$, where σ_1 and σ_2 are positive measures with σ_2 a finite measure. Show that there is a positive finite measure μ on Σ such that $\sigma_1 = \sigma^+ + \mu$ and $\sigma_2 = \sigma^- + \mu$. (Hint: use Exercise 6.2.4)

Remark The conclusion of Exercise 6.2.7 means that the Jordan decomposition of a signed measure σ is **the minimal decomposition** of σ into the difference of two positive measures. For the corresponding fact concerning decomposition of functions of bounded variation into the difference of two monotone increasing functions, see the paragraph following Theorem 4.4.1.

Now let μ be a positive measure on (Ω, Σ). A signed measure σ on Σ is said to be **μ-absolutely continuous** if $\sigma(A) = 0$ whenever $A \in \Sigma$ and $\mu(A) = 0$. It is easily verified that σ is μ-absolutely continuous if and only if σ^+, σ^- are μ-absolutely continuous; thus, σ is μ-absolutely continuous if and only if $|\sigma|$ is μ-absolutely continuous.

Theorem 6.2.3 *If σ is a finite signed measure, then σ is μ-absolutely continuous if and only if for any given $\varepsilon > 0$ there is $\delta > 0$ such that if $A \in \Sigma$ with $\mu(A) < \delta$, then $|\sigma|(A) < \varepsilon$.*

Proof Sufficiency is obvious.

Necessity: Suppose the contrary. Then for some $\varepsilon > 0$ and for any $n \in \mathbb{N}$, there is $A_n \in \Sigma$ such that $\mu(A_n) < 2^{-n}$ and $|\sigma|(A_n) \geq \varepsilon$. Let $A = \limsup_{n\to\infty} A_n$, then for each n,

$$\mu(A) = \mu\left(\lim_{n\to\infty} \bigcup_{k\geq n} A_k\right) \leq \mu\left(\bigcup_{k\geq n} A_k\right) < \sum_{k\geq n} 2^{-k};$$

letting $n \to \infty$, we then have $\mu(A) = 0$. But,

$$|\sigma|(A) = \lim_{n\to\infty} |\sigma|\left(\bigcup_{k\geq n} A_k\right) \geq \limsup_{n\to\infty} |\sigma|(A_n) \geq \varepsilon,$$

which contradicts the fact that $|\sigma|$ is μ-absolutely continuous. ■

Theorem 6.2.4 (Radon–Nikodym) *If (Ω, Σ, μ) is a σ-finite measure space and σ is a σ-finite μ-absolutely continuous signed measure on (Ω, Σ), then there is a unique measurable function f such that $\int_\Omega f^- d\mu < \infty$, and*

$$\sigma(A) = \int_A f d\mu, \quad A \in \Sigma.$$

Proof We know that σ^+ is σ-finite and σ^- is finite on Σ. By Exercise 5.7.1, there is $f_2 \in L^1(\Omega, \Sigma, \mu)$ such that $f_2 \geq 0$, and

$$\sigma^-(A) = \int_A f_2 d\mu, \quad A \in \Sigma;$$

and there is a measurable function f_1 with $f_1 \geq 0$ such that

$$\sigma^+(A) = \int_A f_1 d\mu, \quad A \in \Sigma.$$

Let $f = f_1 - f_2$, then

$$\sigma(A) = \int_A f d\mu, \quad A \in \Sigma.$$

One can verify (cf. Exercise 6.2.6) that $f^- = f_2$ a.e., hence $\int_\Omega f^- d\mu < \infty$. That f is unique is left as an exercise. ■

Exercise 6.2.8 Show that the function f in Theorem 6.2.4 is unique.

Now complex measures are introduced. Fix a measurable space (Ω, Σ); a set function $\sigma : \Sigma \to \mathbb{C}$ is called a **complex measure** if (i) $\sigma(\emptyset) = 0$; and (ii) $\sigma(\bigcup_n A_n) = \sum_{n=1}^\infty \sigma(A_n)$ for every disjoint sequence $\{A_n\}$ in Σ. Observe that in (ii) the convergence of $\sum_{n=1}^\infty \sigma(A_n)$ does not depend on how the sequence $\{A_n\}$ is ordered, hence for any disjoint sequence $\{A_n\} \subset \Sigma$, $\sum_{n=1}^\infty |\sigma(A_n)| < \infty$. We take a hint from Exercise 6.2.4 (iii) to define the total variational measure $|\sigma|$ of a complex measure by

$$|\sigma|(E) = \sup \left\{ \sum_{n=1}^\infty |\sigma(E_n)| \right\}$$

for $E \in \Sigma$, where the supremum is taken over all decompositions of E into countable disjoint measurable sets $E_1, E_2, \ldots$. When σ is a signed or complex measure on $\mathcal{B}(X)$, where X is a metric space, it is called a **Radon (Riesz) measure** if $|\sigma|^*$ is a Radon (Riesz) measure on X. Recall that $|\sigma|^*$ is the measure on X constructed from $|\sigma|$ by Method I.

Exercise 6.2.9 Show that the family of all complex Riesz measures on $\mathcal{B}(X)$ is a complex vector space.

For $A \in \Sigma$, let us put $\sigma_r(A) = \operatorname{Re}\sigma(A)$ and $\sigma_i(A) = \operatorname{Im}\sigma(A)$; then σ_r and σ_i are finite signed measures on Σ. If f is a complex-valued $|\sigma|$-integrable function on Ω, the **σ-integral** of f is defined by

$$\int_X f d\sigma := \int_X f d\sigma_r + i \int_X f d\sigma_i.$$

Suppose now that μ is a positive measure on Σ. A complex measure σ on Σ is **μ-absolutely continuous**, if $A \in \Sigma$ and $\mu(A) = 0$ implies $\sigma(A) = 0$. Obviously, σ is μ-absolutely continuous if and only if both σ_r and σ_i are μ-absolutely continuous.

Exercise 6.2.10 A complex measure σ on Σ is μ-absolutely continuous if and only if for any $\varepsilon > 0$, there is $\delta > 0$ such that if $A \in \Sigma$ with $\mu(A) < \delta$, then $|\sigma(A)| < \varepsilon$.

By applying Theorem 6.2.4 to σ_r and σ_i we obtain Theorem 6.2.5:

Theorem 6.2.5 *If (Ω, Σ, μ) is a σ-finite measure space and σ is a μ-absolutely continuous complex measure on Σ, then there is a unique μ-integrable function f on Ω such that*

$$\sigma(A) = \int_A f d\mu, \quad A \in \Sigma.$$

Henceforth, both Theorem 6.2.4 and Theorem 6.2.5 are to be referred to as the Radon–Nikodym theorem. We note in passing that the family of complex measures on Σ includes all finite signed measures on Σ.

As an application of the notion of signed (complex) measure, we present in the final part of this section the Riesz representation theorem for linear functions on $C_0(X)$; the space of all continuous functions vanishing at infinity on the locally compact metric space X. A function f on X is said to be **vanishing at infinity** if for any $\varepsilon > 0$ there is a compact set K such that $|f(x)| < \varepsilon$ for $x \in K^c$. The space $C_0(X)$ is a real or complex vector space, depending on whether the functions in question are real or complex-valued. Equipped with the norm defined by

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

for $f \in C_0(X)$, $C_0(X)$ is a normed vector space; clearly, $\|f\| = \max_{x \in X} |f(x)|$. The norm so defined on $C_0(X)$ is usually referred to as the **uniform norm**; and unless otherwise specified, $C_0(X)$ is equipped with this norm. For definiteness, we assume that functions in $C_0(X)$ are real-valued and hence $C_0(X)$ is a real vector space.

Exercise 6.2.11

(i) Show that $C_0(X)$ is a Banach space.

(ii) Show that if $f \in C_0(X)$, then both f^+ and f^- are in $C_0(X)$.

If ℓ is a positive linear functional on $C_0(X)$, it is, a fortiori, positive on $C_c(X)$; the measure μ constructed in Section 3.10 for ℓ considered as restricted to $C_c(X)$ is also referred to as the **measure for ℓ**. As we know in Section 3.10, μ is the unique Riesz measure on X such that

$$\ell(f) = \int_X f d\mu$$

for all $f \in C_c(X)$.

Lemma 6.2.3 *Suppose that ℓ is a bounded positive linear functional on $C_0(X)$ and μ is the measure for ℓ; then $\ell(f) = \int_X f d\mu$ for $f \in C_0(X)$ and $\|\ell\| = \mu(X)$.*

Proof Since ℓ is bounded, μ is a finite measure (cf. Exercise 3.10.1).

For $f \in C_0(X)$ and $\varepsilon > 0$, there is a compact set K in X such that $|f(x)| < \varepsilon$ for $x \in K^c$. By Corollary 1.10.1, there is $g \in U_c(X)$ satisfying $g = 1$ on K. Put $h = fg$, then $h \in C_c(X)$, and

$$\begin{aligned}\ell(f) &= \ell(f-h) + \ell(h) = \ell(f-h) + \int_X h d\mu \\ &= \ell(f-h) + \int_X f d\mu + \int_X (h-f) d\mu;\end{aligned}$$

hence,

$$\left| \ell(f) - \int_X f d\mu \right| \leq \|\ell\|\varepsilon + \varepsilon\mu(X),$$

because $\|f-h\| = \|f(1-g)\| \leq \sup_{x \in K^c} |f(x)| \leq \varepsilon$. By letting $\varepsilon \searrow 0$, we obtain

$$\ell(f) = \int_X f d\mu.$$

Now, if $f \in C_0(X)$ with $\|f\| = 1$, then

$$|\ell(f)| = \left| \int_X f d\mu \right| \leq \int_X |f| d\mu \leq \mu(X),$$

and consequently $\|\ell\| \leq \mu(X)$. On the other hand, for a compact set K in X, there is a function $f \in U_c(X)$ such that $f = 1$ on K (again by Corollary 1.10.1); then,

$$\mu(K) \leq \int_X f d\mu = \ell(f) \leq \|\ell\|,$$

from which $\mu(X) \leq \|\ell\|$ follows by the inner regularity of μ. Thus $\|\ell\| = \mu(X)$. ■

Suppose now that $\ell \in C_0(X)^*$; we shall decompose ℓ as a difference of two bounded positive linear functionals on $C_0(X)$ as follows.

Denote by $C_0(X)^+$ the family $\{f \in C_0(X) : f \geq 0\}$ and define a functional ℓ^+ on $C_0(X)^+$ by

$$\ell^+(f) = \sup\{\ell(g) : g \in C_0(X)^+ \text{ and } g \leq f\}$$

for $f \in C_0(X)^+$; since $\ell^+(f) \geq \ell(0) = 0$ and

$$\ell(g) \leq \|\ell\| \cdot \|g\| \leq \|\ell\| \cdot \|f\| < \infty$$

for $g \in C_0(X)^+$ satisfying $g \leq f$, ℓ^+ is nonnegative and $\ell^+(f) \leq \|\ell\| \cdot \|f\| < \infty$. Note that ℓ^+ is **positively homogeneous** on $C_0(X)^+$ in the sense that for $f \in C_0(X)^+$ and nonnegative number α, $\ell^+(\alpha f) = \alpha\ell^+(f)$.

Lemma 6.2.4 *The functional ℓ^+ is additive, i.e. if f and g are in $C_0(X)^+$, then $\ell^+(f+g) = \ell^+(f) + \ell^+(g)$.*

Proof Let u, v in $C_0(X)^+$ be such that $u \leq f$ and $v \leq g$, then $0 \leq u + v \leq f + g$, and hence,

$$\ell^+(f+g) \geq \ell(u+v) = \ell(u) + \ell(v),$$

from which it follows that

$$\ell^+(f+g) \geq \ell^+(f) + \ell^+(g).$$

On the other hand, if $u \in C_0(X)^+$ with $u \leq f + g$, by putting $u_1 = u \wedge f$ and $u_2 = u - u_1$, one verifies easily that

$$u = u_1 + u_2,\ u_1 \leq f,\ \text{and}\ u_2 \leq g;$$

and thus,

$$\ell(u) = \ell(u_1) + \ell(u_2) \leq \ell^+(f) + \ell^+(g),$$

implying that $\ell^+(f+g) \leq \ell^+(f) + \ell^+(g)$. ■

Now, extend ℓ^+ to $C_0(X)$ by defining

$$\ell^+(f) = \ell^+(f^+) - \ell^+(f^-)$$

for $f \in C_0(X)$. For $f \in C_0(X)$, note that both f^+ and f^- are in $C_0(X)^+$ (cf. Exercise 6.2.11 (ii)) and observe that if $f = g - h$, with g and h being in $C_0(X)^+$, then $g = f^+ + u$ and $h = f^- + u$ for some $u \in C_0(X)^+$, and hence,

$$\ell^+(f) = \ell^+(g) - \ell^+(h).$$

Therefore if f and g are in $C_0(X)$, we have

$$\begin{aligned}\ell^+(f+g) &= \ell^+(f^+ + g^+) - \ell^+(f^- + g^-)\\ &= \ell^+(f^+) + \ell^+(g^+) - \ell^+(f^-) - \ell^+(g^-)\\ &= \ell^+(f) + \ell^+(g),\end{aligned}$$

i.e. ℓ^+ is additive on $C_0(X)$. Obviously,

$$\ell^+(\alpha f) = \alpha \ell^+(f),$$

for $f \in C_0(X)$ and $\alpha \in \mathbb{R}$. Thus ℓ^+ is a positive linear functional on $C_0(X)$. Since

$$|\ell^+(f)| \leq \ell^+(f^+) + \ell^+(f^-) \leq \|\ell\|(\|f^+\| + \|f^-\|) \leq 2\|\ell\| \cdot \|f\|,$$

ℓ^+ is a bounded positive linear functional on $C_0(X)$.

If we let $\ell^- = \ell^+ - \ell$, then $\ell^- \in C_0(X)^*$ and $\ell = \ell^+ - \ell^-$. Since for $f \in C_0(X)^+$ we have $\ell^-(f) = \ell^+(f) - \ell(f) \geq 0$, ℓ^- is a bounded positive linear functional on $C_0(X)$. Thus, every $\ell \in C_0(X)^*$ can be decomposed as the difference $\ell^+ - \ell^-$ of two bounded positive linear functionals on $C_0(X)$. Let μ_+ and μ_- be respectively the measure for ℓ^+ and ℓ^-, and for $B \in \mathcal{B}(X)$ put $\mu(B) = \mu_+(B) - \mu_-(B)$, then μ is a finite signed measure on $\mathcal{B}(X)$ and

$$\ell(f) = \int_X f d\mu, \quad f \in C_0(X). \tag{6.11}$$

Denote as before the total variational measure of μ on $\mathcal{B}(X)$ by $|\mu|$, and let $|\mu|^*$ be the measure on X constructed from $|\mu|$ by Method I. We know from Corollary 3.4.1 that $|\mu|^*$ is the unique Borel regular measure extending $|\mu|$, and, since $|\mu|^*$ is finite, it is a Radon measure. We shall see presently that $|\mu|^*$ is a Riesz measure. For this purpose, set for the moment $\nu = \mu_+ + \mu_-$, then ν is a Riesz measure on X and $|\mu|^* \leq \nu$. Given that $B \in \mathcal{B}(X)$ and $\varepsilon > 0$, by outer regularity of ν and Proposition 3.10.1 there are $K \in \mathcal{K}$ and $G \in \mathcal{G}$ with $K \subset B \subset G$ such that $\nu(G \backslash K) < \varepsilon$ and, a fortiori, $|\mu|^*(G\backslash K) < \varepsilon$; consequently, $|\mu|^*(G) - \varepsilon < |\mu|^*(B) < |\mu|^*(K) + \varepsilon$, which in turn implies that

$$|\mu|^*(B) = \sup\{|\mu|^*(K) : K \in \mathcal{K}, K \subset B\} \tag{6.12}$$

and

$$|\mu|^*(B) = \inf\{|\mu|^*(G) : G \in \mathcal{G}, B \subset G\}.$$

Now for any $S \subset X$, there is $B \in \mathcal{B}(X)$ such that $B \supset S$ and $|\mu|^*(S) = |\mu|^*(B) = \inf\{|\mu|^*(G) : G \in \mathcal{G}, B \subset G\} \geq \inf\{|\mu|^*(G) : G \in \mathcal{G}, S \subset G\} \geq |\mu|^*(S)$; thus,

$$|\mu|^*(S) = \inf\{|\mu|^*(G) : G \in \mathcal{G}, S \subset G\},$$

i.e. $|\mu|^*$ is outer regular. Note that (6.12) implies in particular that $|\mu|^*$ is inner regular; hence $|\mu|^*$ is a Riesz measure on X and μ is a Riesz measure on $\mathcal{B}(X)$. This last fact and (6.11) prove the following Lemma 6.2.5.

Lemma 6.2.5 *For $\ell \in C_0(X)^*$ there is a finite Riesz measure μ on $\mathcal{B}(X)$ such that (6.11) holds.*

Lemma 6.2.6 *Suppose that μ is a finite Riesz measure on $\mathcal{B}(X)$. Define a linear functional ℓ on $C_0(X)$ by*

$$\ell(f) = \int_X f d\mu, \quad f \in C_0(X).$$

Then, $\ell \in C_0(X)^*$ and $\|\ell\| = |\mu|(X)$.

Proof For $f \in C_0(X)$,

$$|\ell(f)| = \left|\int_X f d\mu^+ - \int_X f d\mu^-\right| \leq \int_X |f| d\mu^+ + \int_X |f| d\mu^-$$
$$= \int_X |f| d|\mu| \leq \|f\| |\mu|(X),$$

hence, $\ell \in C_0(X)^*$ and $\|\ell\| \leq |\mu|(X)$.

Let (A, B) be a Hahn decomposition of X w.r.t. μ. Since $|\mu|^*$ is a finite Riesz measure on X, by Proposition 3.10.1, there are K_1 and K_2 in $\mathcal{K}$ with $K_1 \subset A$ and $K_2 \subset B$ such that $|\mu|^*(X \backslash (K_1 \cup K_2)) < \varepsilon$. Take a continuous function g on X such that $-1 \leq g \leq 1$, $g = 1$ on K_1 and $g = -1$ on K_2 according to Corollary 1.8.1, and a function $h \in U_c(X)$ such that $h = 1$ on $K_1 \cup K_2$ according to Corollary 1.10.1, and let $f = gh$; then, $f \in C_c(X)$, $-1 \leq f \leq 1$, $f = 1$ on K_1, and $f = -1$ on K_2. Now,

$$\begin{aligned}
\ell(f) = \int_X f d\mu &= \mu(K_1) - \mu(K_2) + \int_{X \backslash (K_1 \cup K_2)} f d\mu \\
&= |\mu|(K_1) + |\mu|(K_2) + \int_{X \backslash (K_1 \cup K_2)} f d\mu \\
&\geq |\mu|(K_1 \cup K_2) - \int_{X \backslash (K_1 \cup K_2)} |f| d|\mu| \\
&\geq |\mu|(X) - 2|\mu|(X \backslash (K_1 \cup K_2)) \\
&\geq |\mu|(X) - 2\varepsilon,
\end{aligned}$$

from which, since $\|f\| = 1$, it follows that $\|\ell\| \geq |\mu|(X) - 2\varepsilon$ and hence $\|\ell\| \geq |\mu|(X)$. Thus, $\ell \in C_0(X)^*$ and $\|\ell\| = |\mu|(X)$, because we already know that $\|\ell\| \leq |\mu|(X)$. ■

Theorem 6.2.6 (Riesz representation theorem) *For $\ell \in C_0(X)^*$ there is a unique finite Riesz measure μ on $\mathcal{B}(X)$ such that*

$$\ell(f) = \int_X f d\mu \tag{6.13}$$

for $f \in C_0(f)$.

Proof The existence of Riesz measure μ on $\mathcal{B}(X)$ such that (6.13) holds follows from Lemma 6.2.5. Suppose that μ_1 and μ_2 are Riesz measures on $\mathcal{B}(X)$ such that (6.13) holds, with μ replaced by either μ_1 or μ_2. Then $\mu_1 - \mu_2$ is a Riesz measure on $\mathcal{B}(X)$ (cf. Exercise 6.2.9) such that

$$\int_X f d(\mu_1 - \mu_2) = 0$$

for all $f \in C_0(X)$; it follows then from Lemma 6.2.6 that $|\mu_1 - \mu_2|(X) = 0$, i.e. $|\mu_1 - \mu_2|$ is a zero measure on $\mathcal{B}(X)$. But, for $B \in \mathcal{B}(X)$, $0 = |\mu_1 - \mu_2|(B) \geq |\mu_1(B) - \mu_2(B)|$ implies that $\mu_1(B) = \mu_2(B)$. Thus the uniqueness of μ is proved. ∎

Example 6.2.3 Let ℓ be a bounded linear functional on the real space $C[0, 1]$. Then there is a BV function g on $[a, b]$ such that g is right-continuous except at 0, and

$$\ell(f) = \int_0^1 f dg, \quad f \in C[0, 1].$$

Actually, let μ be the unique Riesz measure on $\mathcal{B}([0, 1])$ such that $\int_0^1 f d\mu = \ell(f)$ for $f \in C[0, 1]$, and let $g(0) = 0$ and $g(t) = \mu([0, t])$, $t \in [0, 1]$; then g is right-continuous except at 0. Consider any partition $0 = t_0 < t_1 < \cdots < t_n = 1$; we have $\sum_{k=1}^n |g(t_k) - g(t_{k-1})| = \sum_{k=1}^n |\mu((t_{k-1}, t_k])| + |\mu(\{0\})| \leq |\mu|([0, 1])$. Therefore g is a BV function. Clearly, $\ell(f) = \int_0^1 f dg$ for $f \in C[0, 1]$.

In the above discussion we assume that $C_0(X)$ is formed from real-valued functions; a brief account will now be given of the case when $C_0(X)$ consists of complex-valued functions. Recall that a complex-valued function f can be expressed as $\mathrm{Re} f + i \,\mathrm{Im} f$, where $\mathrm{Re} f$ and $\mathrm{Im} f$ are respectively the real and imaginary parts of f. Suppose that $\ell \in C_0(X)^*$, then

$$\ell(f) = \ell_r(f) + i\ell_i(f),$$

where $\ell_r(f) = \mathrm{Re}\{\ell(f)\}$ and $\ell_i(f) = \mathrm{Im}\{\ell(f)\}$; ℓ_r and ℓ_i are bounded linear functionals on $C_0(X)$ considered as a real vector space; in particular, they are bounded linear functionals on the real vector space of all real-valued functions in $C_0(X)$. By Theorem 6.2.6 there is a unique pair (μ_r, μ_i) of finite Riesz signed measures on $\mathcal{B}(X)$ such that

$$\ell(f) = \int_X f d\mu_r + i \int_X f d\mu_i$$

for real-valued functions f in $C_0(X)$. Let us put $\mu(B) = \mu_r(B) + i\mu_i(B)$ for $B \in \mathcal{B}(X)$; then μ is a complex Riesz measure on $\mathcal{B}(X)$, and for $f \in C_0(X)$ we have

$$\begin{aligned}
\ell(f) &= \ell(\mathrm{Re} f + i \,\mathrm{Im} f) = \ell(\mathrm{Re} f) + i\ell(\mathrm{Im} f) \\
&= \ell_r(\mathrm{Re} f) + i\ell_i(\mathrm{Re} f) + i\{\ell_r(\mathrm{Im} f) + i\ell_i(\mathrm{Im} f)\} \\
&= \int_X \mathrm{Re} f d\mu_r + i \int_X \mathrm{Re} f d\mu_i + i \int_X \mathrm{Im} f d\mu_r - \int_X \mathrm{Im} f d\mu_i \\
&= \int_X \mathrm{Re} f d\mu + i \int_X \mathrm{Im} f d\mu = \int_X f d\mu.
\end{aligned}$$

We leave it as an exercise to show the uniqueness of the Riesz measure μ on $\mathcal{B}(X)$ such that $\ell(f) = \int_X f d\mu$ for $f \in C_0(X)$, as well as the fact that $\|\ell\| = |\mu|(X)$. Hence, Theorem 6.2.6 also holds when the functions in $C_0(X)$ are complex-valued.

Exercise 6.2.12 When $C_0(X)$ consists of complex-valued functions and $\ell \in C_0(X)^*$, show that there is a unique Riesz measure on X such that

$$\ell(f) = \int_X f d\mu, \quad f \in C_0(X).$$

Furthermore, show that for such a measure, $\|\ell\| = |\mu|(X)$.

6.3 Linear functionals on L^p

Let p and q be conjugate exponents i.e. $p, q \geq 1$ and $p^{-1} + q^{-1} = 1$. We shall consider a fixed measure space (Ω, Σ, μ) throughout this section, therefore measurability of sets or functions is in reference to this measure space and the measure of a set A means $\mu(A)$ with $A \in \Sigma$. The space $L^p(\Omega, \Sigma, \mu)$ will be simply denoted by L^p for $p \geq 1$, and L^p-norm of f will be denoted by $\|f\|_p$.

Our purpose in this section is to identify $(L^p)^*$ with L^q in a sense to be specified later when μ is σ-finite and $p < \infty$.

For $g \in L^q$, define a linear functional ℓ_g on L^p by

$$\ell_g(f) = \int fgd\mu, \quad f \in L^p.$$

It follows from the Hölder inequality that ℓ_g is a bounded linear functional on L^p and its norm $\|\ell_g\| \leq \|g\|_q$.

We shall actually show that $\|\ell_g\| = \|g\|_q$ if $q < \infty$; and that this equality holds for all $q \geq 1$ if (Ω, Σ, μ) is σ-finite. This means that we may consider L^q as isometrically and isomorphically embedded in $(L^p)^*$ in either case, because the map $g \mapsto l_g$ is a linear map from L^q into $(L^p)^*$.

Lemma 6.3.1 *If $q < \infty$ and $g \in L^q$, then $\|g\|_q = \|\ell_g\|$.*

Proof We may assume that $g \neq 0$ on a set of positive measure, and let

$$f = \frac{|g|^{q-1}\overline{\operatorname{sgn} g}}{\|g\|_q^{q-1}},$$

where $\operatorname{sgn} g(x) = 0$ if $g(x) = 0$, and $= g(x)/|g(x)|$ if $g(x) \neq 0$. One sees easily that $\operatorname{sgn} g$ is a measurable function and $f \in L^p$ with $\|f\|_p = 1$. Now,

$$\|\ell_g\| \geq \left| \int fgd\mu \right| = \|g\|_q^{-(q-1)} \int |g|^q d\mu = \|g\|_q,$$

This, together with $\|\ell_g\| \leq \|g\|_q$, shows that $\|\ell_g\| = \|g\|_q$. ∎

Corollary 6.3.1 *If (Ω, Σ, μ) is σ-finite and $g \in L^q$, then $\|\ell_g\| = \|g\|_q$.*

Proof We need only to prove that $\|\ell_g\| \geq \|g\|_\infty$ for $g \in L^\infty$. For this purpose we may assume that $\|g\|_\infty > 0$ and for a given $0 < \varepsilon < \|g\|_\infty$, let $A = \{|g| \geq \|g\|_\infty - \varepsilon\}$. From the definition of $\|g\|_\infty$, $\mu(A) > 0$. Since μ is σ-finite, there is an increasing sequence $\{\Omega_n\} \subset \Sigma$ such that $\mu(\Omega_n) < \infty$ for each n and $\lim_{n\to\infty} \Omega_n = \Omega$. Then, $\mu(A) = \lim_{n\to\infty} \mu(A \cap \Omega_n)$ implies that $\mu(A \cap \Omega_n) > 0$ if n is large enough, say $n \geq n_0$; let $B = A \cap \Omega_{n_0}$, then $0 < \mu(B) < \infty$. Choose $f = \frac{1}{\mu(B)} I_B \overline{\operatorname{sgn} g}$, then $f \in L^1$ and $\|f\|_1 = 1$. Now,

$$\|\ell_g\| \geq \left|\int fg d\mu\right| = \frac{1}{\mu(B)} \int_B |g| d\mu \geq \|g\|_\infty - \varepsilon,$$

from which we infer that $\|\ell_g\| \geq \|g\|_\infty$ by letting $\varepsilon \searrow 0$. ■

For the statement of the next lemma (6.3.2), given a measurable function g which is finite a.e. on Ω, we denote by $S_p(g)$ the family of all those functions f such that $\|f\|_p = 1$ and fg is integrable.

Exercise 6.3.1 Suppose that (Ω, Σ, μ) is σ-finite and g is a measurable function which is finite a.e. on Ω. Show that $S_p(g)$ is nonempty.

Lemma 6.3.2 *Suppose that (Ω, Σ, μ) is a σ-finite measure space and g is measurable and finite almost everywhere. Then,*

$$\|g\|_q = \sup\left\{\left|\int fg d\mu\right| : f \in S_p(g)\right\}.$$

Proof From the Hölder inequality, $\|g\|_q \geq \sup\{|\int fg d\mu| : f \in S_p(g)\}$, it remains to show the converse inequality. For this purpose we may assume that $g \neq 0$ on a set of positive measure.

Let the sequence $\{\Omega_n\} \subset \Sigma$ be as in the proof of Corollary 6.3.1.

Step 1. Suppose that $q < \infty$. For each $n \in \mathbb{N}$, let $A_n = \{x \in \Omega : |g(x)| \leq n\} \cap \Omega_n$. $\{A_n\}$ is an increasing sequence in Σ such that $\mu(\Omega \setminus \bigcup_n A_n) = 0$. If we let $g_n = g I_{A_n}$, then g_n is bounded and $\neq 0$ on a set of positive measure when n is sufficiently large, say $n \geq n_0$. Define, for $n \geq n_0$,

$$f_n = \frac{|g_n|^{q-1} \overline{\operatorname{sgn} g_n}}{\|g_n\|_q^{q-1}}.$$

One can verify easily that $\|f_n\|_p = 1$. Since $f_n g = \|g_n\|_q^{1-q} |g_n|^q$, $f_n g$ is integrable and therefore $\{f_n\}_{n \geq n_0} \subset S_p(g)$. Now for $n \geq n_0$, using $f_n g = \|g_n\|_q^{1-q} |g_n|^q$, we have

$$\begin{aligned}
\|g_n\|_q^q &= \int |g_n|^q d\mu \\
&= \|g_n\|_q^{q-1} \int f_n g d\mu \leq \|g_n\|_q^{q-1} \sup\left\{\left|\int fg d\mu\right| : f \in S_p(g)\right\},
\end{aligned}$$

from which it follows that $\|g_n\|_q \leq \sup\left\{\left|\int fgd\mu\right| : f \in S_p(g)\right\}$. But $\mu(\Omega \setminus \bigcup_n A_n) = 0$ implies that $|g_n|$ increases to $|g|$ a.e. on Ω, hence, on letting $n \to \infty$, we obtain $\|g\|_q \leq \sup\{|\int fgd\mu| : f \in S_p(g)\}$. Thus, $\|g\|_q = \sup\{|\int fgd\mu| : f \in S_p(g)\}$ if $q < \infty$.

Step 2. Suppose that $q = \infty$, i.e. $p = 1$. Put

$$\gamma = \sup\left\{\left|\int fgd\mu\right| : f \in S_1(g)\right\}.$$

We may assume that $\gamma < \infty$.

Given that $\varepsilon > 0$, let $A = \{x \in \Omega : |g(x)| \geq \gamma + \varepsilon\}$. We claim that $\mu(A) = 0$; otherwise, let $B_n = A \cap \Omega_n \cap \{|g| \leq n\}$, then $0 < \mu(B_n) < \infty$ if $n \geq n_0$ for some $n_0 \in \mathbb{N}$. Put $B = B_{n_0}$ and let $f = \mu(B)^{-1} I_B \overline{\operatorname{sgn} g}$. Then, $\|f\|_1 = 1$ and $\int fgd\mu = \mu(B)^{-1} \int_B |g| d\mu \leq n_0$, thus $f \in S_1(g)$; but $\int fgd\mu = \mu(B)^{-1} \int_B |g| d\mu \geq \gamma + \varepsilon$, which contradicts the definition of γ. Hence, $\mu(A) = 0$ and consequently $\|g\|_\infty \leq \gamma + \varepsilon$. Let $\varepsilon \searrow 0$; we have $\|g\|_\infty \leq \gamma$. ∎

It is worthwhile noting that the proof of Lemma 6.3.2 actually shows that $\|g\|_q = \sup\{\mathrm{Re} \int_\Omega fgd\mu : f \in S_p(g)\}$, and that if $g \geq 0$, $\|g\|_q = \sup\{\int_\Omega fgd\mu : f \in S_p(g)$ and $f \geq 0\}$.

The following **integral version of the Minkowski inequality** follows from Lemma 6.3.2 with this note.

Corollary 6.3.2 *Suppose that $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ are σ-finite complete measure spaces and $f \geq 0$ is $\overline{\Sigma_1 \otimes \Sigma_2}$-measurable on $\Omega_1 \times \Omega_2$. Then for $1 \leq p < \infty$, the following inequality holds:*

$$\left\{\int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)d\mu_2(y)\right)^p d\mu_1(x)\right\}^{\frac{1}{p}} \leq \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^p d\mu_1(x)\right)^{\frac{1}{p}} d\mu_2(y). \tag{6.14}$$

Proof Put $F(x) = \int_{\Omega_2} f(x,y)d\mu_2(y), x \in \Omega_1$. F is measurable using the Fubini theorem.

Step 1. Suppose that $F(x) < \infty$ for μ_1-a.e. x. Let $h \geq 0$ be in $S_q(F) \subset L^q(\Omega_1, \Sigma_1, \mu_1)$, then

$$\begin{aligned}
\int_{\Omega_1} hFd\mu_1 &= \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)d\mu_2(y)\right) h(x)d\mu_1(x) \\
&= \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)h(x)d\mu_1(x)\right) d\mu_2(y) \\
&\leq \|h\|_q \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^p d\mu_1(x)\right)^{1/p} d\mu_2(y) \\
&= \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^p d\mu_1(x)\right)^{1/p} d\mu_2(y).
\end{aligned}$$

By Lemma 6.3.2, with p replaced by q, together with the note that follows it, we conclude that $\|F\|_p \le \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^p d\mu(x)\right)^{1/p} d\mu_2(y)$, i.e. (6.14) holds.

Step 2. Now suppose that $A = \{F = \infty\}$ has positive measure. Since μ_1 is σ-finite, there is a measurable set $A_0 \subset A$ such that $0 < \mu_1(A_0) < \infty$. Let $h = \mu_1(A_0)^{-\frac{1}{q}} I_{A_0}$ or I_{A_0} according to whether $q < \infty$ or $q = \infty$, then proceed as in Step 1; we have

$$\infty = \int_{\Omega_1} hFd\mu_1 \le \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^p d\mu_1(x)\right)^{1/p} d\mu_2(y).$$

Consequently (6.14) holds, because right-hand side of (6.14) is ∞. ■

Now we come to the main theorem of this section.

Theorem 6.3.1 *If (Ω, Σ, μ) is σ-finite and $1 \le p < \infty$, then $L^q = (L^p)^*$, through the map*

$$g \mapsto \ell_g, \quad g \in L^q.$$

Proof We already know that $L^q \subset (L^p)^*$ through the map $g \mapsto \ell_g$, by Corollary 6.3.1; it remains to show that for $\ell \in (L^p)^*$, there is a unique $g \in L^q$ such that $\ell = \ell_g$.

Step 1. Suppose that $\mu(\Omega) < \infty$.

For $A \in \Sigma$, let $\nu(A) = \ell(I_A)$. Since ℓ is linear, ν is an additive set function on Σ. Now suppose that $\{A_n\}_{n=1}^{\infty} \subset \Sigma$ is disjoint, then

$$\nu\left(\bigcup_n A_n\right) = \nu\left(\bigcup_{n=1}^{N} A_n\right) + \nu\left(\bigcup_{n=N+1}^{\infty} A_n\right),$$

hence, by putting $B_N = \bigcup_{n=N+1}^{\infty} A_n$, we have

$$\begin{aligned}\left|\nu\left(\bigcup_n A_n\right) - \sum_{n=1}^{N} \nu(A_n)\right| &\le |\nu(B_N)| \\ &\le \|\ell\| \|I_{B_N}\|_p \\ &= \|\ell\|[\mu(B_N)]^{1/p} \to 0\end{aligned}$$

as $N \to \infty$, because $B_N \downarrow \phi$ and $\mu(\Omega) < \infty$; consequently, $\nu(\bigcup_n A_n) = \sum_{n=1}^{\infty} \nu(A_n)$. Thus ν is a complex measure on Σ. Since ν is μ-absolutely continuous, from the Radon–Nikodym theorem, there is $g \in L^1$ such that $\nu(A) = \int_A gd\mu$, or

$$\ell(f) = \int fgd\mu \tag{6.15}$$

for simple functions f.

Suppose now that $f \in S_p(g)$. Choose a sequence $\{f_n\}$ of simple functions such that $f_n \to f$ pointwise and $|f_n| \le |f|$. Then, $|f_n g| \le |fg|$, by LDCT and (6.15),

$$\left|\int fg d\mu\right| = \lim_{n\to\infty}\left|\int f_n g d\mu\right| = \lim_{n\to\infty}|\ell(f_n)| \le \|\ell\|.$$

It then follows from Lemma 6.3.2 that $g \in L^q$ and $\|g\|_q \le \|\ell\|$.

Now let $f \in L^p$ and choose a sequence $\{\varphi_n\}$ of simple functions such that $\varphi_n \to f$ pointwise and $|\varphi_n| \le |f|$, then $\varphi_n \to f$ in L^p and by (6.15),

$$\int fg d\mu = \lim_{n\to\infty}\int \varphi_n g d\mu = \lim_{n\to\infty}\ell(\varphi_n) = \ell(f),$$

this means that $\ell = \ell_g$ and $\|\ell\| = \|\ell_g\| = \|g\|_q$.

Step 2. Suppose that (Ω, Σ, μ) is σ-finite.

Let $\{\Omega_n\} \subset \Sigma$ be as in the proof of Corollary 6.3.1. By Step 1, for each n, there is $g_n \in L^q$ with $\{g_n \neq 0\} \subset \Omega_n$ such that

$$\ell(f) = \int fg_n d\mu \tag{6.16}$$

for $f \in L^p$ with $\{f \neq 0\} \subset \Omega_n$. Define g on Ω by $g(x) = g_1(x)$ if $x \in \Omega_1$, and $g(x) = g_n(x)$ if $x \in \Omega_n \backslash \Omega_{n-1}$ for $n \ge 2$. Then, since $g_n(x) = g_{n-1}(x)$ for a.e. x in Ω_{n-1} when $n \ge 2$, $g(x) = g_n(x)$ for a.e. $x \in \Omega_n$.

Now let $f \in S_p(g)$, then $|fg_n| \le |fg|$ and $fg_n \to fg$ a.e., hence by (6.16),

$$\left|\int fg d\mu\right| = \lim_{n\to\infty}\left|\int fg_n d\mu\right| = \lim_{n\to\infty}|\ell(fI_{\Omega_n})| \le \|\ell\|.$$

From Lemma 6.3.2, $g \in L^q$ and hence for $f \in L^p$,

$$\int fg d\mu = \lim_{n\to\infty}\int fI_{\Omega_n} g_n d\mu = \lim_{n\to\infty}\ell(fI_{\Omega_n}) = \ell(f),$$

where the last equality comes from the obvious fact that $fI_{\Omega_n} \to f$ in L^p. Then $\ell = \ell_g$, and $\|\ell\| = \|g\|_q$. That g is uniquely determined is obvious.

■

Exercise 6.3.2 shows that Theorem 6.3.1 may not hold true when $p = \infty$.

Exercise 6.3.2 Consider $L^\infty[0,1]$ and let $x_0 \in [0,1]$. Show that there is $\ell \in L^\infty[0,1]^*$ with $\|\ell\| = 1$ such that $\ell(f) = f(x_0)$ for $f \in C[0,1]$. For this ℓ show that there is no $g \in L^1[0,1]$ such that $\ell(f) = \int_{[0,1]} fg d\lambda$ for all $f \in L^\infty[0,1]$.

Exercise 6.3.3 Suppose that (Ω, Σ, μ) is σ-finite and $1 < p < \infty$. Show that L^p is reflexive.

Exercise 6.3.4 Let D be a measurable set in $\mathbb{R}^n$ with positive measure. Show that every bounded sequence in $L^p(D)$, $1 < p < \infty$, has a subsequence which converges weakly. (Hint: cf. Exercise 5.10.5.)

6.4 Modular distribution function and Hardy–Littlewood maximal function

Suppose that f is a finite a.e. measurable function on a measure space (Ω, Σ, μ). Define a function $\lambda_f : (0, \infty) \to [0, \infty]$ by

$$\lambda_f(\alpha) = \mu(\{|f| > \alpha\}). \tag{6.17}$$

Then the function λ_f enjoys the following properties:

(1) λ_f is monotone decreasing and right continuous.
(2) If $|f| \leq |g|$, then $\lambda_f \leq \lambda_g$.
(3) If $|f_n| \nearrow |f|$, then $\lambda_{f_n} \nearrow \lambda_f$.
(4) If $f = g + h$, then $\lambda_f(\alpha + \beta) \leq \lambda_g(\alpha) + \lambda_h(\beta)$ for $\alpha, \beta > 0$.

Properties (1)–(3) follow directly from the definition, while (4) is a consequence of the fact that $\{|f| > \alpha + \beta\} \subset \{|g| > \alpha\} \cup \{|h| > \beta\}$.

The function λ_f is usually called the distribution function of f; but the distribution function of a measurable function is defined differently in Section 4.3, in agreement with the distribution function of a random variable in probability theory; we shall instead call λ_f the **modular distribution function** of f.

If $\lambda_f(\alpha) < \infty\ \forall \alpha > 0$, then λ_f generates a negative Radon measure ν on $(0, \infty)$ such that

$$\nu((a, b]) = \lambda_f(b) - \lambda_f(a), \quad 0 < a < b < \infty;$$

actually, ν is the negative of the Radon measure generated by the monotone increasing function $-\lambda_f$. We shall call ν the **Lebesque–Stieltjes measure** generated by λ_f. If φ is a Borel function on $(0, \infty)$ such that $\int_0^\infty \varphi d\nu = \int_{(0,\infty)} \varphi d\nu$ exists, then $\int_0^\infty \varphi d\nu$ will be denoted by $\int_0^\infty \varphi d\lambda_f$ or $\int_0^\infty \varphi(\alpha) d\lambda_f(\alpha)$ in this section, and called the **Lebesque–Stieltjes integral** of φ w.r.t. λ_f.

Lemma 6.4.1 *Suppose that $\lambda_f(\alpha) < \infty$ for all $\alpha > 0$ and let φ be a nonnegative Borel function on $(0, \infty)$, then*

$$\int_\Omega \varphi \circ |f| d\mu = -\int_0^\infty \varphi(\alpha) d\lambda_f(\alpha). \tag{6.18}$$

Proof We have

$$\nu((a,b]) = \lambda_f(b) - \lambda_f(a) = -\mu(\{a < |f| \le b\}) = -\mu(|f|^{-1}(a,b]),$$

from which it follows that

$$\nu(B) = -\mu(|f|^{-1}B)$$

for Borel set B in $(\frac{1}{k}, k]$ (by the $(\pi\text{-}\lambda)$ theorem), and therefore for Borel set B in $(0,\infty)$. This means that (6.18) holds if φ is the indicator function of Borel set B in $(0,\infty)$, and consequently, if φ is a nonnegative simple Borel function on $(0,\infty)$. For a general nonnegative Borel function on $(0,\infty)$, (6.18) follows then by approximating φ pointwise by an increasing sequence of nonnegative simple Borel functions on $(0,\infty)$. ■

Exercise 6.4.1 Give the detail of the first part of the proof of Lemma 6.4.1 where the $(\pi\text{-}\lambda)$ theorem is applied.

A measurable function f on (Ω, Σ, μ) is called a **weak L^p function**; $0 < p < \infty$, if there is $0 \le A < \infty$ depending only on f and p such that

$$\mu(\{|f| > \alpha\}) \le \frac{A^p}{\alpha^p}, \quad \alpha > 0. \tag{6.19}$$

One sees readily that f is a weak L^p function if and only if $\sup_{\alpha>0} \alpha^p \mu\{|f| > \alpha\} < \infty$.

Exercise 6.4.2 Show that if $|f|^p$, $0 < p < \infty$, is integrable, then f is a weak L^p function.

Theorem 6.4.1 *Suppose that f is a weak L^p function, $1 \le p < \infty$, then we have*

$$\int_\Omega |f|^p d\mu = -\int_0^\infty \alpha^p d\lambda_f(\alpha) = p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha. \tag{6.20}$$

Proof Since f is a weak L^p function, $1 \le p < \infty$, $\lambda_f(\alpha) < \infty$ for all $\alpha > 0$, hence the first equality in (6.20) follows from Lemma 6.4.1 by taking $\phi(\alpha) = \alpha^p$. It remains to show that

$$\int_\Omega |f|^p d\mu = p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha.$$

We observe first that the set

$$E := \{(x,\alpha) : x \in \Omega,\ 0 < \alpha < |f(x)|\}$$

is in $\Sigma \otimes \mathcal{B}$ (cf. Exercise 4.8.3). Since I_E is $\Sigma \otimes \mathcal{B}$-measurable, from Tonelli's theorem we have

$$\begin{aligned}
&\int_{\Omega\times(0,\infty)} pI_E(x,\alpha)\alpha^{p-1}d(\mu\times\lambda)(x,\alpha) \\
&= \int_\Omega \left(\int_0^{|f(x)|} p\alpha^{p-1}d\lambda(\alpha)\right) d\mu(x) \\
&= \int_\Omega |f(x)|^p d\mu(x);
\end{aligned}$$

but we also have

$$\begin{aligned}
&\int_{\Omega\times(0,\infty)} pI_E(x,\alpha)\alpha^{p-1}d(\mu\times\lambda)(x,\alpha) \\
&= p\int_0^\infty \alpha^{p-1}\left(\int_{\{|f|>\alpha\}} d\mu\right) d\lambda(\alpha) \\
&= p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha.
\end{aligned}$$

Hence $\int_\Omega |f|^p d\mu = p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha$. ■

The Hardy–Littlewood maximal function will now be introduced. Let f be a locally integrable function on $\mathbb{R}^n$; the **Hardy–Littlewood maximal function** of f, denoted Mf, is defined in terms of $|f|$ as follows:

$$Mf(x) = \sup_{r>0} \frac{1}{\lambda^n(B_r(x))}\int_{B_r(x)} |f(y)|dy, \tag{6.21}$$

where $Mf(x)$ could be infinite for some $x \in \mathbb{R}^n$. Since, for each $r > 0$, the function

$$x \mapsto \frac{1}{\lambda^n(B_r(x))}\int_{B_r(x)} |f(y)|dy$$

is continuous, $\{Mf > \alpha\}$ is open for $\alpha \in \mathbb{R}$. Hence Mf is a Borel function and is therefore measurable. We shall from now on simply call Mf the **maximal function** of f.

Theorem 6.4.2 *For $f \in L^1(\mathbb{R}^n)$, Mf is a weak L^1 function. Actually there is $A > 0$, depending only on n, such that*

$$\lambda^n(\{Mf > \alpha\}) \le A\|f\|_1\alpha^{-1} \tag{6.22}$$

for $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$.

Proof For $\alpha > 0$, put $E_\alpha = \{Mf > \alpha\}$. For $x \in E_\alpha$, there is a ball $B(x)$ centered at x such that

$$\int_{B(x)} |f(y)|dy > \alpha\lambda^n(B(x)). \tag{6.23}$$

Since $\lambda^n(B(x)) < \alpha^{-1}\|f\|_1$ by (6.23), $\mathcal{C} := \{B(x) : x \in E_\alpha\}$ is an admissible collection of balls. By Theorem 4.6.1, there is a disjoint sequence $\{B_k\}$ of balls from $\mathcal{C}$ such that $\bigcup \mathcal{C} \subset \bigcup_k \widehat{B}_k$, where $\widehat{B}_k$ is concentric with B_k and has a radius five times that of B_k. Then from (6.23),

$$\begin{aligned}\lambda^n(E_\alpha) &\leq \lambda^n\left(\bigcup \mathcal{C}\right) \leq \sum_k \lambda^n(\widehat{B}_k) = 5^n \sum_k \lambda^n(B_k) \\ &< 5^n\alpha^{-1}\sum_k \int_{B_k} |f(y)|dy = 5^n\alpha^{-1}\int_{\bigcup B_k} |f(y)|dy \\ &\leq 5^n\|f\|_1\alpha^{-1},\end{aligned}$$

from which we complete the proof by taking $A = 5^n$. ■

Exercise 6.4.3 For $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, show that Mf is a weak L^p function. (Hint: use Jensen's inequality to show that $Mf(x) \leq \{M|f|^p(x)\}^{1/p}$ for $x \in \mathbb{R}^n$.)

We note at this point that although Mf is a weak L^1 function, it can never be integrable except for the extreme case $f = 0$ a.e. To see this, suppose that $f \neq 0$ on a set of positive measure; then $\int_{B_R(0)} |f|d\lambda^n = c > 0$ for some $R > 0$ and hence if $|x| \geq R$, $B := B_{2|x|}(x)$ contains $B_R(0)$, from which

$$Mf(x) \geq \frac{1}{\lambda^n(B)}\int_B |f(y)|dy \geq 2^{-n}|x|^{-n}b_n^{-1}c = c_0|x|^{-n}$$

follows, where b_n is the measure of the unit ball in $\mathbb{R}^n$; thus by integrating Mf over $\mathbb{R}^n$ using polar coordinates (cf. Theorem 4.11.1), we conclude that $\int Mfd\lambda^n = \infty$. However, as the following theorem shows, $Mf \in L^p$ if $f \in L^p$ and the map $f \mapsto Mf$ is a bounded map from L^p into L^p when $p > 1$.

Theorem 6.4.3 *If $1 < p \leq \infty$, there is $A_p > 0$ such that for $f \in L^p(\mathbb{R}^n)$ we have*

$$\|Mf\|_p \leq A_p\|f\|_p.$$

Proof When $p = \infty$, this is obvious with $A_\infty = 1$. Consider now $1 < p < \infty$. For a fixed $\alpha > 0$, define f_1 by

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| \geq \frac{\alpha}{2}; \\ 0 & \text{otherwise,} \end{cases}$$

then, $f_1 \in L^1(\mathbb{R}^n)$ (see Exercise 6.4.4) and $|f(x)| \leq |f_1(x)| + \frac{\alpha}{2}$; hence $Mf \leq Mf_1 + \frac{\alpha}{2}$, which implies $\{Mf > \alpha\} \subset \{Mf_1 > \frac{\alpha}{2}\}$ and consequently by Theorem 6.4.2 (note that A can be taken to be 5^n),

$$\lambda_{Mf}(\alpha) \leq \frac{2 \cdot 5^n}{\alpha}\|f_1\|_1 = \frac{2 \cdot 5^n}{\alpha}\int_{\{|f| \geq \frac{\alpha}{2}\}} |f(x)|dx.$$

Now by (6.20),

$$\begin{aligned}
\|Mf\|_p^p &= p\int_0^\infty \alpha^{p-1}\lambda_{Mf}(\alpha)d\alpha \\
&\leq p\int_0^\infty \alpha^{p-1}\left(\frac{2 \cdot 5^n}{\alpha}\int_{\{|f| \geq \frac{\alpha}{2}\}} |f(x)|dx\right)d\alpha \\
&= 2 \cdot 5^n \cdot p\int_{\mathbb{R}^n} |f(x)|\int_0^{2|f(x)|} \alpha^{p-2}d\alpha dx \\
&= \frac{2 \cdot 5^n \cdot p}{p-1}\int_{\mathbb{R}^n} 2^{p-1}|f(x)|^p dx \\
&= 2^p \cdot 5^n \frac{p}{p-1}\|f\|_p^p = A_p^p\|f\|_p^p,
\end{aligned}$$

where $A_p = 2\left(\frac{5^n p}{p-1}\right)^{1/p}$. ■

Exercise 6.4.4 Show that the function f_1 defined at the beginning of the proof of Theorem 6.4.3 is integrable.

As an application of maximal function, a direct proof of Theorem 4.6.4 will now be given using Theorem 6.4.2 together with the Markov inequality (6.1). An application of Theorem 6.4.3 to the study of Sobolev space is presented in Section 6.6. Actually, we shall prove that if f is a locally integrable function on an open set $\Omega \subset \mathbb{R}^n$, then

$$\lim_{r \to 0} \frac{1}{\lambda^n(B_r(x))}\int_{B_r(x)} |f(y) - f(x)|dy = 0 \tag{6.24}$$

for a.e. $x \in \Omega$, and leave the proof for the general statement as an exercise. Because of the local nature of (6.24), we may assume that f is an integrable function on $\mathbb{R}^n$. Put $\theta(f, x) = \limsup_{r \to 0} \frac{1}{\lambda^n(B_r(x))}\int_{B_r(x)} |f(y) - f(x)|dy$; our aim is to show that $\theta(f, x) = 0$ for a.e. x in $\mathbb{R}^n$, or, equivalently, to show that $\lambda^n(\{\theta(f, \cdot) > \alpha\}) = 0$ for every $\alpha > 0$. Now, given that $\varepsilon > 0$, there is a continuous function g on $\mathbb{R}^n$ such that $\|f - g\|_1 < \varepsilon$ (cf. Exercise 6.1.1), then,

$$\theta(f, x) = \theta(f - g + g, x) \leq \theta(f - g, x) + \theta(g, x) = \theta(f - g, x),$$

because $\theta(g,x) = 0$; but $\theta(f-g,x) \le M(f-g)(x) + |f(x) - g(x)|$ and consequently,

$$\{\theta(f,\cdot) > \alpha\} \subset \{\theta(f-g,\cdot) > \alpha\} \subset \left\{M(f-g) > \frac{\alpha}{2}\right\} \cup \left\{|f-g| > \frac{\alpha}{2}\right\}.$$

Hence,

$$\lambda^n(\{\theta(f,\cdot) > \alpha\}) \le (A\|f-g\|_1 + \|f-g\|_1)\frac{2}{\alpha} \le 2(A+1)\frac{\varepsilon}{\alpha};$$

by letting $\varepsilon \to 0$, we have $\lambda^n(\{\theta(f,\cdot) > \lambda\}) = 0$. Thus, $\theta(f,x) = 0$ for a.e. x in $\mathbb{R}^n$ and (6.24) is established.

Exercise 6.4.5 Show that $\lim_{B\to x} \frac{1}{\lambda^n(B)} \int_B |f(y) - f(x)|dy = 0$ follows from (6.24). (Hint: if $x \in B$, then $B \subset B_{2r}(x)$, where r is the radius of B.)

6.5 Convolution

The operation of taking convolution was used in Section 4.9 when introducing the Friederichs mollifier for the purpose of smoothing functions. An account of general features of convolution for functions on $\mathbb{R}^n$ will be given in this section; its connection with the Fourier integral will be seen in Chapter 7. Referring to Exercises 1.6.6 and 1.6.7, we note in passing that convolution can be introduced for functions on groups with a measure invariant under translations w.r.t. the group operation and is often proved to be a useful operation.

We first state Proposition 4.8.2 as a lemma for later reference.

Lemma 6.5.1 *Let f be a measurable function on $\mathbb{R}^n$, then $F(x,y) := f(x-y)$, x,y in $\mathbb{R}^n$, is a measurable function on $\mathbb{R}^{2n} = \mathbb{R}^n \oplus \mathbb{R}^n$.*

Let f and g be measurable functions on $\mathbb{R}^n$. The **convolution** of f and g is the function $f * g$ defined for all those x for which the following integral exists and is finite:

$$f * g(x) = \int f(x-y)g(y)dy.$$

Exercise 6.5.1

(i) Show that if $f * g(x)$ exists and is finite, then $g * f(x)$ exists and is finite, and $g * f(x) = f * g(x)$.
(ii) Show that if $f * g$ exists and is finite for a.e. x, then $f * g$ is measurable. (Hint: apply Lemma 6.5.1 and the Fubini theorem.)

Exercise 6.5.2 Suppose that $[a,b]$ and $[c,d]$ are finite closed intervals of equal length. Find $I_{[a,b]} * I_{[c,d]}$; in particular, show that $I_{[-\frac{\alpha}{2},\frac{\alpha}{2}]} * I_{[-\frac{\alpha}{2},\frac{\alpha}{2}]}(x) = \alpha(1 - \frac{|x|}{\alpha})^+, \alpha > 0$.

Theorem 6.5.1 (Young inequality) *Suppose that $f \in L^p$, $p \geq 1$, and $g \in L^1$, then $f * g$ exists and is finite a.e., and*

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

Proof The case where $p = \infty$ is obvious. We consider the case where $1 \leq p < \infty$. Let $h(x, y) = f(x - y)g(y)$; h is measurable by Lemma 6.5.1. Using the integral version of the Minkowski inequality (Corollary 6.3.2), we have

$$\begin{aligned} \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)| dy \right)^p dx \right)^{\frac{1}{p}} &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|^p dx \right)^{\frac{1}{p}} dy \\ &= \|f\|_p \|g\|_1; \end{aligned}$$

then, $\left(\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x-y)g(y)dy \right|^p dx \right)^{\frac{1}{p}} \leq \|f\|_p \|g\|_1$, and consequently $f * g$ exists and is finite a.e., and $\|f * g\|_p \leq \|f\|_p \|g\|_1$. ■

Example 6.5.1 We give here another proof of Theorem 5.6.1 without recourse to the integral version of the Minkowski inequality. Since

$$\begin{aligned} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)| dy \right) dx &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)| dx \right) dy \\ &= \|f\|_p^p \int_{\mathbb{R}^n} |g(y)| dy = \|f\|_p^p \|g\|_1, \end{aligned}$$

therefore, $\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)| dy < \infty$ for a.e. x, and hence

$$\int_{\mathbb{R}^n} |f(x-y)||g(y)| dy \leq \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)| dy \right)^{\frac{1}{p}} \cdot \|g\|_1^{\frac{1}{q}} < \infty \text{ for a.e. } x,$$

which implies that $f * g$ exists and is finite a.e., and

$$\begin{aligned} \|f * g\|_p^p &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)||g(y)| dy \right)^p dx \\ &\leq \|g\|_1^{\frac{p}{q}} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)| dy \right) dx \\ &= \|g\|_1^{\frac{p}{q}} \|f\|_p^p \|g\|_1 = \|f\|_p^p \|g\|_1^p, \end{aligned}$$

or

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

Lemma 6.5.2 *For $f \in L^p$, $1 \leq p < \infty$, and $y \in \mathbb{R}^n$, let $f^y(x) = f(x - y)$. Then,* $\lim_{y \to 0} \|f^y - f\|_p = 0$.

Proof Given $\varepsilon > 0$, there is a continuous function g with compact support such that $\|f - g\|_p < \frac{\varepsilon}{3}$, by Proposition 4.6.1. Then,

$$\begin{aligned}\|f^y - f\|_p &= \|f^y - g^y + g^y - g + g - f\|_p \\ &\leq \|f^y - g^y\|_p + \|g - f\|_p + \|g^y - g\|_p \\ &< \frac{2}{3}\varepsilon + \|g^y - g\|_p,\end{aligned}$$

but since g is continuous with compact support, $\|g^y - g\|_p < \frac{\varepsilon}{3}$ when $|y|$ is small. Thus, $\|f^y - f\|_p < \varepsilon$ when $|y|$ is small. ■

We shall denote by $C_0(\mathbb{R}^n)$ the space of all those continuous functions f on $\mathbb{R}^n$ with the property that for any $\varepsilon > 0$, there is $R > 0$ such that $|f(x)| < \varepsilon$ whenever $|x| > R$. Functions in $C_0(\mathbb{R}^n)$ are functions **vanishing at infinity**, introduced in Section 6.2. Clearly, $C_c(\mathbb{R}^n) \subset C_0(\mathbb{R}^n)$.

Theorem 6.5.2 *If p and q are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g(x)$ exists and is finite for all x, and $f * g$ is bounded and uniformly continuous on $\mathbb{R}^n$. Furthermore, $\|f * g\|_\infty \leq \|f\|_p \|g\|_q$; and if $1 < p < \infty$, then $f * g \in C_0(\mathbb{R}^n)$.*

Proof From the Hölder inequality, $|f * g(x)| \leq \|f\|_p \|g\|_q$ for all x, hence $f * g(x)$ exists and is finite for all x, and $\|f * g\|_\infty \leq \|f\|_p \|g\|_q$.

To show that $f * g$ is uniformly continuous, we may assume that $1 \leq p < \infty$ (otherwise interchange p and q). Now,

$$|f * g(x - y) - f * g(x)| = |(f^y - f) * g(x)| \leq \|f^y - f\|_p \|g\|_q,$$

hence $f * g$ is uniformly continuous on $\mathbb{R}^n$, by Lemma 6.5.2.

Finally, suppose that $1 < p < \infty$ (then $1 < q < \infty$). Choose sequences $\{f_k\}$, $\{g_k\}$ in $C_c(\mathbb{R}^n)$ so that $\|f_k - f\|_p \to 0$ and $\|g_k - g\|_q \to 0$ as $k \to \infty$; this is possible by Proposition 4.6.1. Then, $\{f_k * g_k\}$ is a sequence of continuous functions with compact support, and

$$\begin{aligned}\sup_{x \in \mathbb{R}^n} |f_k * g_k(x) - f * g(x)| &= \sup_{x \in \mathbb{R}^n} |f_k * (g_k - g)(x) + (f_k - f) * g(x)| \\ &\leq \|f_k\|_p \|g_k - g\|_q + \|f_k - f\|_p \|g\|_q \to 0\end{aligned}$$

as $k \to \infty$, because $\{f_k\}$, being a convergent sequence in L^p, is bounded in L^p.

Now given $\varepsilon > 0$, from what we have just shown choose k_0 large enough so that $\sup_{x \in \mathbb{R}^n} |f_{k_0} * g_{k_0}(x) - f * g(x)| < \varepsilon$, and then choose $R > 0$ such that $f_{k_0} * g_{k_0}(x) = 0$ when $|x| > R$; thus $|f * g(x)| < \varepsilon$, when $|x| > R$. This shows that $f * g \in C_0(\mathbb{R}^n)$. ■

Remark Theorem 6.5.2 is an example showing the smoothing effect of convolution.

Exercise 6.5.3 Show that for f, g, and h in L^1, $(f * g) * h = f * (g * h)$.

Example 6.5.2 The Friederich mollifier $\{J_\varepsilon\}_{\varepsilon>0}$ constructed from a mollifying function φ introduced in Section 4.9 can be expressed as

$$J_\varepsilon f(x) = f * \varphi_\varepsilon(x), \quad x \in \mathbb{R}^n,$$

for $f \in L^{\text{loc}}(\mathbb{R}^n)$. By Proposition 4.9.2 and Theorem 6.5.2, $J_\varepsilon f \in C^\infty(\mathbb{R}^n) \cap C_0(\mathbb{R}^n)$ if $f \in L^p$, $1 < p < \infty$.

Exercise 6.5.4 Show that there is no $u \in L^1$ such that $u * f = f$ for all $f \in L^1$. (Hint: if there is such a u, then $u * \varphi_\varepsilon = \varphi_\varepsilon$ for all $\varepsilon > 0$, where φ is a mollifying function.)

Example 6.5.3 Suppose that f, g are in L^1 and $f \in C^1(\mathbb{R}^n)$ with bounded partial derivatives. Since $f \in C^1(\mathbb{R}^n)$ with bounded partial derivatives, f is uniformly continuous; consequently if $f * g(x)$ exists and is finite, then $f * g(x')$ exists and is finite if $|x' - x| < \delta$, where $\delta > 0$ is chosen so that $|f(z) - f(z')| < 1$ if $|z - z'| < \delta$. This, together with the known fact that $f * g$ exists and is finite a.e., shows that $f * g$ exists and is finite everywhere and is uniformly continuous on $\mathbb{R}^n$. Now for any x, y in $\mathbb{R}^n$, $\frac{|f(x)-f(y)|}{|x-y|} \leq M$ for a fixed $M > 0$, because partial derivatives of f are bounded. We can then apply LDCT to infer that

$$\frac{\partial}{\partial x_j} f * g(x) = \frac{\partial f}{\partial x_j} * g(x), \quad x \in \mathbb{R}^n, \ j = 1, \ldots, n.$$

But from Theorem 6.5.2, $\frac{\partial f}{\partial x_j} * g$ is bounded and continuous. Hence, $f * g \in C^1(\mathbb{R}^n)$ and its partial derivatives are bounded.

By the Young inequality (Theorem 6.5.1), L^1 is closed under the binary operation of convolution, which is associative (cf. Exercise 6.5.3) and clearly distributive w.r.t. the addition of elements in L^1. Thus with the introduction of the binary operation $*$ into L^1, L^1 becomes a commutative algebra; it is an example of the so-called Banach algebras, in that it is a Banach space which is also an algebra that satisfies the inequality $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ for f, g in L^1. Because of the conclusion of Exercise 6.5.4, there exists no identity element in L^1 w.r.t. the multiplication operation $*$. However, if φ is a mollifying function (cf. Example 6.5.2), $\lim_{\varepsilon\to 0} \varphi_\varepsilon * f = f$ in L^1, by Theorem 4.9.2; such a family $\{\varphi_\varepsilon\}_{\varepsilon>0}$ is called an **approximate identity** for L^1. Just as we construct the approximate identity $\{\varphi_\varepsilon\}_{\varepsilon>0}$ from a mollifying function φ, starting from an integrable function h on $\mathbb{R}^n$ with $\int h d\lambda^n = 1$, we define for each $t > 0$ a function h_t by

$$h_t(x) = t^{-n} h\left(\frac{x}{t}\right), \quad x \in \mathbb{R}^n,$$

then, $\int h_t d\lambda^n = 1$. We shall see that $\{h_t\}_{t>0}$ is an approximate identity for L^1.

Lemma 6.5.3 *For $\varepsilon > 0$ and $\delta > 0$, there is $t_0 > 0$ such that*

$$\int_{|y|\geq\delta} |h_t(y)|dy < \varepsilon,$$

whenever $0 < t \leq t_0$.

Proof Since $h \in L^1$, there is $R > 0$ such that $\int_{|y|\geq R} |h(y)|dy < \varepsilon$. Then,

$$\int_{|y|\geq\delta} |h_t(y)|dy = \int_{|y|\geq\frac{\delta}{t}} |h(y)|dy < \varepsilon$$

if $\frac{\delta}{t} \geq R$. We choose $t_0 = \frac{\delta}{R}$ to complete the proof. ■

Theorem 6.5.3 *$\{h_t\}_{t>0}$ is an approximate identity for L^1, i.e.*

$$\lim_{t\to 0} \|h_t * f - f\|_1 = 0, \quad f \in L^1.$$

Proof For $f \in L^1$ and $\varepsilon > 0$, there is $\delta > 0$ such that

$$\int_{\mathbb{R}^n} |f(x-y) - f(x)|dx < \frac{\varepsilon}{2\|h\|_1} \tag{6.25}$$

if $|y| < \delta$. Since we may assume that $\|f\|_1 > 0$, there is $t_0 > 0$ such that

$$\int_{|y|\geq\delta} |h_t(y)|dy < \frac{\varepsilon}{4\|f\|_1} \tag{6.26}$$

whenever $0 < t \leq t_0$, by Lemma 6.5.3. Now,

$$\begin{aligned}
&\int_{\mathbb{R}^n} |h_t * f - f| d\lambda^n \\
&= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \{f(x-y) - f(x)\} h_t(y)dy \right| dx \\
&\leq \int_{\mathbb{R}^n} |h_t(y)| \int_{\mathbb{R}^n} |f(x-y) - f(x)|dxdy \\
&\leq \int_{|y|<\delta} |h_t(y)| \int_{\mathbb{R}^n} |f(x-y) - f(x)|dxdy + 2\|f\|_1 \int_{|y|\geq\delta} |h_t(y)|dy \\
&< \frac{\varepsilon}{2\|h\|_1} \int_{|y|<\delta} |h_t(y)|dy + \frac{\varepsilon}{2} \leq \varepsilon
\end{aligned}$$

if $0 < t \leq t_0$ by (6.25) and (6.26). ■

We know from Theorem 6.5.2 that $f * h_t$ is a bounded and uniformly continuous function for each $t > 0$ if $f \in L^\infty$; we show now, as a supplement to Theorem 6.5.3, that $f * h_t$ converges to f uniformly on every compact set of $\mathbb{R}^n$ as $t \to 0$ if $f \in L^\infty \cap C(\mathbb{R}^n)$.

Theorem 6.5.4 *If f is a bounded continuous function on $\mathbb{R}^n$, then $\lim_{t\to 0} f * h_t = f$ uniformly on every compact set K of $\mathbb{R}^n$.*

Proof For a compact set K in $\mathbb{R}^n$ and $\varepsilon > 0$, there is $\delta > 0$ such that $|f(x-y) - f(x)| < \frac{\varepsilon}{2\|h\|_1}$ whenever $x \in K$ and $|y| < \delta$. Then by Lemma 6.5.3, there is $t_0 > 0$ such that $\int_{|y|\geq\delta} |h_t(y)|dy < \frac{\varepsilon}{4\|f\|_\infty}$ if $0 < t \leq t_0$. Now for $x \in K$ and $0 < t \leq t_0$, we have

$$
\begin{aligned}
|h_t * f(x) - f(x)| &= \left| \int_{\mathbb{R}^n} h_t(y)\{f(x-y) - f(x)\}dy \right| \\
&\leq \int_{\mathbb{R}^n} |h_t(y)||f(x-y) - f(x)|dy \\
&= \int_{|y|<\delta} |h_t(y)||f(x-y) - f(x)|dy + \int_{|y|\geq\delta} |h_t(y)||f(x-y) - f(x)|dy \\
&< \frac{\varepsilon}{2\|h\|_1} \int_{|y|<\delta} |h_t(y)|dy + 2\|f\|_\infty \int_{|y|\geq\delta} |h_t(y)|dy \\
&< \frac{\varepsilon}{2} + 2\|f\|_\infty \frac{\varepsilon}{4\|f\|_\infty} = \varepsilon,
\end{aligned}
$$

which means that $h_t * f(x) \to f(x)$ uniformly for $x \in K$ as $t \to 0$. ■

Exercise 6.5.5 Let $p(x) = \frac{1}{\pi}\frac{1}{1+x^2}$ and write $p_t(x) = \frac{t}{\pi}\frac{1}{t^2+x^2}$ as $p(x,t)$ for $x \in \mathbb{R}$ and $t > 0$. The function $(x,t) \mapsto p(x,t)$ on $\mathbb{R} \times (0,\infty)$ is called the **Poisson kernel**.

(i) For $f \in L^1(\mathbb{R})$, let

$$\Pi(x,t) = p_t * f = \int_{\mathbb{R}} p(x-y,t)f(y)dy, \quad (x,t) \in \mathbb{R} \times (0,\infty).$$

Show that

$$
\begin{aligned}
\frac{\partial^2 \Pi}{\partial x^2}(x,t) &= \int_{\mathbb{R}} \frac{\partial^2 p}{\partial x^2}(x-y,t)f(y)dy; \\
\frac{\partial^2 \Pi}{\partial t^2}(x,t) &= \int_{\mathbb{R}} \frac{\partial^2 p}{\partial t^2}(x-y,t)f(y)dy.
\end{aligned}
$$

(Hint: $\frac{\partial p}{\partial x}(x,t)$, $\frac{\partial^2 p}{\partial x^2}(x,t)$, $\frac{\partial p}{\partial t}(x,t)$, $\frac{\partial^2 p}{\partial t^2}(x,t)$ are bounded on $\mathbb{R} \times (t_0,\infty)$ for any $t_0 > 0$.)

(ii) Let f and Π be as in (i). Show that Π is harmonic on $\mathbb{R} \times (0,\infty)$. Furthermore, if f is bounded and continuous, show that Π can be extended continuously to $\mathbb{R} \times [0,\infty)$ and that $\Pi(x,0) = f(x)$ for $x \in \mathbb{R}$.

6.6 The Sobolev space $W^{k,p}(\Omega)$

A brief account of Sobolev spaces, which are fundamental in modern theory of partial differential equations and calculus of variations, will now be given.

A locally integrable function u defined on an open set $\Omega \subset \mathbb{R}^n$ is said to be **weakly differentiable up to order** k on Ω, k being a positive integer, if for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n \le k$ there is a locally integrable function g_α on Ω, such that

$$\int_\Omega u\partial^\alpha \varphi d\lambda^n = (-1)^{|\alpha|} \int_\Omega g_\alpha \varphi d\lambda^n, \tag{6.27}$$

for all $\varphi \in C_c^\infty(\Omega)$. Observe that g_α is uniquely determined by u in the sense that any two such functions are equivalent. We therefore denote g_α by u_α . Note that $u_0 = u_{(0,\ldots,0)} = u$. Clearly, functions u in $C^k(\Omega)$ are weakly differentiable up to order k on Ω with $u_\alpha = \partial^\alpha u$. For $p \ge 1$, let $W^{k,p}(\Omega)$ be the equivalence class of all such functions u in $L^p(\Omega)$ which is weakly differentiable up to order k on Ω such that $u_\alpha \in L^p(\Omega)$ for all α with $|\alpha| \le k$. $W^{k,p}(\Omega)$ is a vector space with the usual definition of addition and multiplication by scalar. On $W^{k,p}(\Omega)$ a norm $\|\cdot\|_{k,p}$ is defined by

$$\begin{aligned} \|u\|_{k,p} &= \left(\sum_{|\alpha|\le k} \|u_\alpha\|_p^p \right)^{\frac{1}{p}} \quad \text{if } p < \infty; \\ &= \sum_{|\alpha|\le k} \|u_\alpha\|_\infty \quad \text{if } p = \infty. \end{aligned} \tag{6.28}$$

To see that $\|u\|_{k,p}$ is actually a norm, we need only verify that triangle inequality holds when $1 \le p < \infty$: $\|u+v\|_{k,p} \le (\sum_{|\alpha|\le k}\{\|u_\alpha\|_p + \|v_\alpha\|_p\}^p)^{\frac{1}{p}} \le (\sum_{|\alpha|\le k} \|u_\alpha\|_p^p)^{\frac{1}{p}} + (\sum_{|\alpha|\le k} \|v_\alpha\|_p^p)^{\frac{1}{p}} = \|u\|_{k,p} + \|v\|_{k,p}$, where we have used the Minkowski inequality for $l^p(S)$ with S a finite set. Of course, there are equivalent norms for $W^{k,p}(\Omega)$; for example, we may also define $\|u\|_{k,p}$ as $\sum_{|\alpha|\le k} \|u_\alpha\|_p$. We prefer the norm defined in (6.28), because when $p = 2$, the norm comes from an inner product on $W^{k,2}(\Omega)$, defined by

$$(u,v)_k = \sum_{|\alpha|\le k} \int_\Omega u_\alpha \bar{v}_\alpha d\lambda^n. \tag{6.29}$$

If u is weakly differentiable to certain order, u_α's are called generalized partial derivatives of u, and often u_α is denoted by $\partial^\alpha u$ or $\frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$; many notations related to smooth functions are also borrowed to be applied to weakly differentiable functions, for example, if u is weakly differentiable to first order, ∇u is used to denote $(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n})$ and is called the **generalized gradient** of u.

In what follows in this section, p and q are conjugate exponents.

Theorem 6.6.1 *$W^{k,p}(\Omega)$ is a Banach space.*

Proof Let $\{u^{(j)}\}$ be a Cauchy sequence in $W^{k,p}(\Omega)$. For each α with $|\alpha| \le k$, $\{u_\alpha^{(j)}\}$ is a Cauchy sequence in $L^p(\Omega)$, hence, $\lim_{j\to\infty} \|u_\alpha^{(j)} - g_\alpha\|_p = 0$ for some $g_\alpha \in L^p(\Omega)$. If we put $u = g_0$, we shall show that $u \in W^{k,p}(\Omega)$ and $\lim_{j\to\infty} \|u^{(j)} - u\|_{k,p} = 0$. For any given $\varphi \in C_c^\infty(\Omega)$,

$$\begin{aligned}
&\left|\int_\Omega u\partial^\alpha \varphi d\lambda^n - (-1)^{|\alpha|}\int_\Omega g_\alpha \varphi d\lambda^n\right| \\
&= \left|\int_\Omega (u - u^{(j)})\partial^\alpha \varphi d\lambda^n + (-1)^{|\alpha|}\int_\Omega (u_\alpha^{(j)} - g_\alpha)\varphi d\lambda^n\right| \\
&\le \|u - u^{(j)}\|_p \|\partial^\alpha \varphi\|_q + \|g_\alpha - u_\alpha^{(j)}\|_p \|\varphi\|_q,
\end{aligned}$$

from which by letting $j \to \infty$, we have

$$\int_\Omega u\partial^\alpha \varphi d\lambda^n = (-1)^\alpha \int_\Omega g_\alpha \varphi d\lambda^n,$$

and hence u is weakly differentiable up to order k with $u_\alpha = g_\alpha$. Thus $u \in W^{k,p}(\Omega)$. That $\lim_{j\to\infty} \|u - u^{(j)}\|_{k,p} = 0$ follows from $\lim_{j\to\infty} \|u_\alpha - u_\alpha^{(j)}\|_p = 0$, for each α with $|\alpha| \le k$. ■

Theorem 6.6.1 implies in particular that $W^{k,2}(\Omega)$ is a Hilbert space with inner product defined by (6.29).

Exercise 6.6.1 A locally integrable function u defined on an open set Ω in $\mathbb{R}^n$ is in $W^{k,p}(\Omega)$, $p > 1$, if and only if for each multi-index α with $|\alpha| \le k$, there is a constant $C_\alpha > 0$ such that

$$\left|\int_\Omega u\partial^\alpha \varphi d\lambda^n\right| \le C_\alpha \|\varphi\|_q$$

for all $\varphi \in C_c^\infty(\Omega)$, where p, q are conjugate exponents.

Exercise 6.6.2 Let $\{J_\varepsilon\}_{\varepsilon>0}$ be a Friederich mollifier and suppose that u is weakly differentiable up to order k on an open set $\Omega \subset \mathbb{R}^n$. Show that for any multi-index α with $|\alpha| \le k$, we have

$$\partial^\alpha (J_\varepsilon u)(x) = J_\varepsilon u_\alpha(x), \quad x \in \Omega_\varepsilon,$$

where $\Omega_\varepsilon = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > \varepsilon\}$.

Exercise 6.6.3 Let $u \in W^{k,p}(\Omega)$, $1 \le p < \infty$. Show that there is a sequence $\{v_j\} \subset C^\infty(\mathbb{R}^n)$ such that for every $\varepsilon > 0$, $v_j \in W^{k,p}(\Omega_\varepsilon)$ when j is large and $v_j \to u$ in $W^{k,p}(\Omega_\varepsilon)$. Note that $v_j \in W^{k,p}(\Omega_\varepsilon)$ implicitly implies that the restriction of v_j to Ω_ε is also denoted by v_j.

Exercise 6.6.4 Let I be an open interval in $\mathbb{R}$. Show that a locally integrable function f on I is in $W^{1,1}(I)$ if and only if it is equivalent to a function g which is absolutely continuous on every finite closed interval in I and $g' \in L^1(I)$.

Theorem 6.6.2 *Suppose that $u \in W^{1,1}(\mathbb{R}^n)$, then*

$$u(x) = \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot\nabla u(\xi)}{|x-\xi|^n}\,d\xi$$

for a.e. x in $\mathbb{R}^n$, where $b_n = \lambda^n(B_1(0))$ and $\nabla u = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right)$.

Proof We know from Exercises 6.1.3 and 6.6.3 that there is a sequence $\{u_j\}$ in $C^\infty(\mathbb{R}^n) \cap W^{1,1}(\mathbb{R}^n)$ such that $\lim_{j\to\infty} \|u_j - u\|_{1,1} = 0$ and $u_j(x) \to u(x)$ for a.e. x in $\mathbb{R}^n$. Apply Corollary 4.11.1 to each u_j; we have

$$u_j(x) = \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot\nabla u_j(\xi)}{|x-\xi|^n}\,d\xi, \quad x \in \mathbb{R}^n. \tag{6.30}$$

Fix $R > 0$. Let $\Omega = B_{R+1}(0)$, $D = B_R(0)$, and put

$$g_j(x) = \int_\Omega \frac{|\nabla u_j(\xi) - \nabla u(\xi)|}{|x-\xi|^{n-1}}\,d\xi, \quad x \in D.$$

By Theorem 4.11.2, $\|g_j\|_1 \to 0$ as $j \to \infty$; hence, $\{g_j\}$ has a subsequence $\{g_{j'}\}$ such that $g_{j'}(x) \to 0$ as $j' \to \infty$ for a.e. x in D, by Exercise 6.1.3. Now,

$$u_{j'}(x) = \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot(\nabla u_{j'}(\xi) - \nabla u(\xi))}{|x-\xi|^n}\,d\xi + \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot\nabla u(\xi)}{|x-\xi|^n}\,d\xi;$$

if we show that $\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot(\nabla u_{j'}(\xi)-\nabla u(\xi))}{|x-\xi|^n}\,d\xi \to 0$ for a.e. x in D as $j' \to \infty$, then $u(x) = \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot\nabla(\xi)}{|x-\xi|^n}\,d\xi$ for a.e. x in D. But, for $x \in D$, we have

$$\begin{aligned}
\left|\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot(\nabla u_{j'}(\xi) - \nabla u(\xi))}{|x-\xi|^n}\,d\xi\right| &\le \int_{\mathbb{R}^n} \frac{|\nabla u_{j'}(\xi) - \nabla u(\xi)|}{|x-\xi|^{n-1}}\,d\xi \\
&= g_{j'}(x) + \int_{\mathbb{R}^n\setminus\Omega} \frac{|\nabla u_{j'}(\xi) - \nabla u(\xi)|}{|x-\xi|^{n-1}}\,d\xi \\
&\le g_{j'}(x) + \int_{\mathbb{R}^n} |\nabla u_{j'}(\xi) - \nabla u(\xi)|\,d\xi \to 0
\end{aligned}$$

as $j' \to \infty$ for those x where $g_{j'}(x) \to 0$. Thus $u(x) = \frac{1}{nb_n}\int_{\mathbb{R}^n} \frac{(x-\xi)\cdot\nabla u(\xi)}{|x-\xi|^n}\,d\xi$ for a.e. x in D. Since $R > 0$ is arbitrary, the theorem is proved. ■

The closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$ is denoted by $\overset{\circ}{W}{}^{k,p}(\Omega)$; functions in $\overset{\circ}{W}{}^{k,p}(\Omega)$ are said to **vanish on $\partial\Omega$ in a generalized sense**.

Exercise 6.6.5 Suppose that Ω is bounded and let $u \in \overset{\circ}{W}{}^{k,\infty}(\Omega)$. Show that u is equivalent to a function $v \in C^k(\Omega)$ which can be continuously extended to be zero on $\partial\Omega$, together with all its partial derivatives up to order k.

Exercise 6.6.6 Show that if $u \in W^{k,p}(\Omega)$, then

$$\int_\Omega u\partial^\alpha v d\lambda^n = (-1)^{|\alpha|} \int_\Omega u_\alpha v d\lambda^n$$

for all $v \in \overset{\circ}{W}{}^{k,q}(\Omega)$ if $|\alpha| \leq k$.

Exercise 6.6.7 Let g be in $C^\infty(\mathbb{R}^n)$ satisfying $0 \leq g \leq 1$, $g = 0$ outside $B_2(0)$, and $g = 1$ on $B_1(0)$. For $j \in \mathbb{N}$, let g_j be the function defined on $\mathbb{R}^n$ by

$$g_j(x) = g(j^{-1}x), \quad x \in \mathbb{R}^n.$$

(i) Suppose that $u \in C^k(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$, $1 \leq p < \infty$. Show that $\lim_{j\to\infty} \|g_j u - u\|_{k,p} = 0$.

(ii) Show that $\overset{\circ}{W}{}^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$ if $1 \leq p < \infty$.

Theorem 6.6.3 (Poincaré) *If Ω is a bounded open set in $\mathbb{R}^n$, then on $\overset{\circ}{W}{}^{k,p}(\Omega)$ the norm $\|\cdot\|_{k,p}$ is equivalent to the norm $|\cdot|_{k,p}$, defined for $u \in \overset{\circ}{W}{}^{k,p}(\Omega)$ by*

$$|u|_{k,p} = \left(\sum_{|\alpha|=k} \|u_\alpha\|_p^p\right)^{1/p}, \quad p < \infty;$$
$$|u|_{k,\infty} = \sum_{|\alpha|=k} \|u_\alpha\|_\infty.$$

Proof We prove the theorem for $k = 1$ and $p < \infty$; the proof for the general case will be clear from the proof of this particular case.

For $u \in \overset{\circ}{W}{}^{1,p}(\Omega)$, we are going to show that there is $C > 0$, independent of u, such that $\|u\|_{1,p} \leq C|u|_{1,p}$. From the definition of $\overset{\circ}{W}{}^{1,p}(\Omega)$, we may assume that $u \in C_c^\infty(\Omega)$. By letting $u = 0$ outside Ω, we may further assume that $u \in C_c^\infty(I)$, where I is an open oriented cube containing Ω and with side-width $= l$. Express I as $I = I_1 \times \hat{I}_1$, where $I_1 = (a, b) \subset \mathbb{R}$ and $\hat{I}_1 \subset \mathbb{R}^{n-1}$; then for $x \in I$, x can be expressed as $(x_1, \hat{x}_1)$ with $x_1 \in (a, b)$ and $\hat{x}_1 \in \hat{I}_1$. Now, $u(x) = \int_a^{x_1} \frac{\partial u}{\partial x_1}(t, \hat{x}_1)dt$ implies that $|u(x)|^p \leq (x_1 - a)^{p/q} \int_a^b |\frac{\partial u}{\partial x_1}(t, \hat{x}_1)|^p dt$ and hence,

$$\|u\|_p^p \le (b-a)^{p/q}(b-a)\int_I \left|\frac{\partial u}{\partial x_1}(x)\right|^p dx = (b-a)^p \left\|\frac{\partial u}{\partial x_1}\right\|_p^p$$
$$\le (b-a)^p \sum_{j=1}^n \left\|\frac{\partial u}{\partial x_j}\right\|_p^p,$$

from which it follows that

$$\|u\|_{1,p}^p \le \{1+(b-a)^p\}|u|_{1,p}^p;$$

therefore $\|u\|_{1,p} \le C|u|_{1,p}$, where $C = \{1+(b-a)^p\}^{1/p}$. Then,

$$|u|_{1,p} \le \|u\|_{1,p} \le C|u|_{1,p},$$

implying that $\|\cdot\|_{1,p}$ and $|\cdot|_{1,p}$ are equivalent. ∎

Remark Since $|u|_{k,p} \le \|u\|_{k,p}$ for $u \in \overset{\circ}{W}{}^{k,p}(\Omega)$, Theorem 6.6.3 is equivalent to the statement that there is $C > 0$ such that

$$\|u\|_{k,p} \le C|u|_{k,p} \tag{6.31}$$

for all $u \in \overset{\circ}{W}{}^{k,p}(\Omega)$. Inequality (6.31) is called the **Poincaré inequality**; and Theorem 6.6.3 is usually referred to as the Poincaré inequality.

The following lemma is a generalization of Example 4.11.2.

Lemma 6.6.1 *Let $u \in W^{1,p}(\mathbb{R}^n)$, $1 \le p < \infty$, then,*

$$\int_{B_R(x)} \frac{|u(\xi)-u(x)|}{|\xi-x|}\,d\xi \le M|\nabla u|(x)$$

for x in $\mathbb{R}^n$, where $M|\nabla u|$ is the maximal function of ∇u.

Proof Fix a Friederichs mollifier $\{J_\varepsilon\}_{\varepsilon>0}$, and let $u_\varepsilon = J_\varepsilon u$ (cf. Section 4.9), then $\lim_{\varepsilon\to 0}\|J_\varepsilon u - u\|_{1,p} = 0$, by Exercise 6.6.2 and Theorem 4.9.2; hence $u_\varepsilon \to u$, $\nabla u_\varepsilon \to \nabla u$ in $L^p(\mathbb{R}^n)$. Fix $x \in \mathbb{R}^n$ and $R > 0$, in terms of polar coordinates of $y-x$; we have

$$\int_{B_R(x)} |u_\varepsilon(y)-u(y)|dy = \int_0^R \rho^{n-1}\int_{S^{n-1}} |u_k(\rho,\theta)-u(\rho,\theta)|d\sigma(\theta)d\rho$$
$$= \int_{S^{n-1}}\int_0^R \rho^{n-1}|u_k(\rho,\theta)-u(\rho,\theta)|d\rho d\sigma(\theta) \to 0$$

as $\varepsilon \searrow 0$. We infer then from Example 4.8.2 that there is a sequence $\varepsilon_k \searrow 0$ such that $\int_0^R \rho^{n-1}|u_{\varepsilon_k}(\rho,\theta)-u(\rho,\theta)|d\rho \to 0$ as $k\to\infty$ for σ-a.e. $\theta \in S^{n-1}$.

Then for any $0 < \delta < R$, $\int_\delta^R |u_{\varepsilon_k}(\rho,\theta) - u(\rho,\theta)|d\rho \to 0$ as $k \to \infty$. Similarly, $\int_\delta^R |\nabla u_{\varepsilon_k'}(\rho,\theta) - \nabla u(\rho,\theta)|d\rho \to 0$ as $k \to \infty$ for σ-a.e. $\theta \in S^{n-1}$. Since we may choose ε_k', a subsequence of ε_k, we conclude that there is a sequence $\{u_k\}$ in $C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$ such that for a.e. y in $B_R(x)\backslash B_\delta(x)$,

$$\int_\delta^R |u_k(x+t(y-x)) - u(x+t(y-x))|dt \to 0$$

and

$$\int_\delta^R |\nabla u_k(x+t(y-x)) - \nabla u(x+t(y-x))|dt \to 0$$

as $k \to \infty$. Therefore for a.e. y in $B_R(x)\backslash B_\delta(x)$, $u(x+t(y-x))$ is AC on $[\delta, R]$ and $\frac{d}{dt}u(x+t(y-x)) = \nabla u(x+t(y-x)) \cdot (y-x)$ for a.e. t on $[\delta, R]$. Then, as in Example 4.11.2,

$$\begin{aligned}\int_{B_R(x)\backslash B_\delta(x)} \frac{|u(\xi)-u(x)|}{|\xi - x|}d\xi &\leq \int_0^1 \frac{1}{t^n}\int_{B_{Rt}(x)\backslash B_{\delta(t)}(x)} |\nabla u(z)|dz\\ &\leq \int_0^1 \frac{1}{t^n}\int_{B_{Rt}(x)} |\nabla u(z)|dz\\ &\leq \lambda^n(B_R(x)) \cdot M|\nabla u|(x).\end{aligned}$$

We conclude the proof by letting $\delta \searrow 0$. ∎

Theorem 6.6.4 *There is a positive constant $\theta = \theta(n,p)$, $1 < p < \infty$ with the property that if $u \in W^{1,p}(\mathbb{R}^n)$, then for $\varepsilon > 0$ there is a closed set $F \subset \mathbb{R}^n$ such that $u|_F$, the restriction of u to F, is Lipschitz with Lipschitz constant $\mathrm{Lip}(u|_F)$, satisfying*

$$\mathrm{Lip}(u|_F)^p\lambda^n(\mathbb{R}^n\backslash F) < \theta(n,p)\varepsilon.$$

Proof For x, y in $\mathbb{R}^n$, put $q(x,y) = \frac{|u(y)-u(x)|}{|y-x|}$, then

$$\frac{1}{\sigma_R}\int_{B_R(x)} q(x,y)dy \leq M|\nabla u|(x), \tag{6.32}$$

from Lemma 6.6.1, where $\sigma_R = \lambda^n(B_R(x))$. For $x \in \mathbb{R}^n$ and $\lambda > 0$, let $W_R(x,\lambda) = \{y \in B_R(x) : q(x,y) \leq \lambda\}$; we have from (6.1) and (6.32),

$$\lambda^n(B_R(x)\backslash W_R(x,\lambda)) \leq \frac{1}{\lambda}\int_{B_R(x)} q(x,y)dy \leq \frac{\sigma_R}{\lambda}M|\nabla u|(x). \tag{6.33}$$

Now put $Z_\delta = \{x \in \mathbb{R}^n : M|\nabla u|(x) \le \delta\}$, and choose $k_0 > 1$ such that

$$\lambda^n(B_R(x) \cap B_R(y)) > \frac{2}{k_0}\sigma_R, \quad R = |x - y|. \tag{6.34}$$

Consider now x, y in Z_δ; we have from (6.33),

$$\lambda^n(B_R(z) \backslash W_R(z, k_0\delta)) \le \frac{M|\nabla u|(z)}{k_0\delta}\sigma_R \le \frac{1}{k_0}\sigma_R, \tag{6.35}$$

for $z = x$ or y and $R = |x - y|$. It follows from (6.34) and (6.35) that $W_R(x, k_0\delta) \cap W_R(y, k_0\delta) \neq \emptyset$; choose $z_0 \in W_R(x, k_0\delta) \cap W_R(y, k_0\delta)$, then

$$q(x, y) \le q(x, z_0) + q(y, z_0) \le 2k_0\delta. \tag{6.36}$$

Given that $\varepsilon > 0$, by (6.2) there is $\delta > 0$ such that $\delta^p\lambda^n(\{M|\nabla u| > \delta\}) < \varepsilon$. Choose then a closed set F in Z_δ with $\lambda^n(\mathbb{R}^n \backslash F) < 2\lambda^n(\{M|\nabla u| > \delta\})$. The restriction of u to F is a Lipschitz function with Lipschitz constant $\le 2k_0\delta$, by (6.36); therefore $(\frac{\mathrm{Lip}(u|_F)}{2k_0})^p\lambda^n(\mathbb{R}^n \backslash F) < 2\varepsilon$. We choose $\theta = \theta(n, p) = 2^{p+1}k_0^p$ to complete the proof. ■

Remark If $M(|u| + |\nabla u|)$ is substituted for $M|\nabla u|$ in Theorem 6.6.4, the closed set F can be chosen so that $\|u|_F\|_\infty + \mathrm{Lip}(u|_F) \le 2\,\mathrm{Lip}(u|_F)$; this observation, together with the known fact that $u|_F$ can be extended to a Lipschitz function v on $\mathbb{R}^n$ such that $\|v\|_\infty + \mathrm{Lip}(v) \le A(\|u|_F\|_\infty + \mathrm{Lip}(u|_F))$, where A is a constant depending only on n (cf. [St, Chapter VI]), shows that Theorem 6.6.4 can be formulated as follows. A function $u \in L^p(\mathbb{R}^n)$ is in $W^{1,p}(\mathbb{R}^n)$ if and only if for any given $\varepsilon > 0$ there is a Lipschitz function v on $\mathbb{R}^n$, and a closed set F such that $u = v$ on F, $\lambda^n(\mathbb{R}^n \backslash F) < \varepsilon$, and $\|u - v\|_{1,p} < \varepsilon$.

Besides, Theorem 6.6.4 also holds when $p = 1$, because in the last paragraph of the proof of the theorem, δ can be chosen so that $\delta\lambda^n(\{M|\nabla u| > \delta\}) < \varepsilon$ follows from the improved form of Theorem 6.4.2:

$$\lambda^n(\{Mf > \alpha\}) \le 2A\alpha^{-1}\int_{\{|f| > \frac{\alpha}{2}\}} |f|\, d\lambda^n,$$

of which we refer to [St, P.7].

Since $W^{k,2}(\Omega)$ is a Hilbert space, it will be denoted by $H^k(\Omega)$; accordingly, $\mathring{W}^{k,2}(\Omega)$ is denoted by $\mathring{H}^k(\Omega)$. By Exercise 6.6.7 (ii), $\mathring{H}^k(\mathbb{R}^n) = H^k(\mathbb{R}^n)$; $H^k(\mathbb{R}^n)$ is usually abbreviated to H^k. In Chapter 7, with the help of the Fourier integral, H^s will also be defined for fractional number s.