Chapter 4

Convolutions and Approximations to the Identity

In this Chapter we introduce an important operation on functions called convolution. This operation and its properties play an important role in proving, among other things, that smooth functions of compact support are dense in $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. This is one of the main results of this chapter. Here we shall also study approximations to the identity. Of particular interest is the case of the Hardy-Littlewood maximal function and its applications, especially the Calderón–Zygmund decomposition theorem. This will be discussed in Chapter 5.

4.1 Minkowski's Integral Inequality

We begin with a theorem which usefulness will become clear shortly.

Theorem 4.1 (Minkowski's Integral Inequality)**.** *Suppose* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \nu)$ *be two* σ*–finite measure spaces. Let* $1 \leq p < \infty$ *and suppose* $F(x, y)$ *is measurable with respect to* $\mathcal{A} \times \mathcal{B}$*. Then*

$$\left\| \int_X F(x, y)\, d\mu(x) \right\|_{L^p(d\nu)} \leq \int_X \|F(x, y)\|_{L^p(d\nu)}\, d\mu(x). \tag{4.1}$$

Remark 4.1. Recall that if f_1, f_2 are two functions then the classical Minkowski inequality gives

$$\|f_1 + f_2\|_p \leq \|f_1\|_p + \|f_2\|_p \tag{4.2}$$

and in general

$$\left\| \sum_{j=1}^{\infty} f_j \right\|_p \leq \sum_{j=1}^{\infty} \|f_j\|_p. \tag{4.3}$$

If we let μ be the counting measure on the positive integers $\mathbb{N}$ we can write (4.3) as

$$\sum_{j=1}^{\infty} f_j = \int_{\mathbb{N}} f_j d\mu(j)$$

Thus (4.3) is a special case of (4.1). However, as we will see in a moment, the proof of (4.1) is exactly the same as the proof of (4.2). An alternative proof of this, which we leave to the reader as an exercise, is also possible based on Theorem 2.17 of Chapter 2.

Proof. Assume, as we may, that $F(x, y) \geq 0$ and, by approximating with simple functions if necessary, that both sides of the inequality are finite. The case $p = 1$ follows directly from Fubini's theorem and so we take $1 < p < \infty$ and let $q = p/(p-1)$ be its conjugate exponent. Set

$$G(y) = \left(\int_X F(x, y)\, d\mu(x) \right)^{p-1}.$$

Then

$$\begin{aligned} \|G\|_{L^q(d\nu)} &= \left\| \left(\int_X F(x, t)\, d\mu(x) \right)^{p-1} \right\|_{L^q(d\nu)} \\ &= \left(\int_Y \left(\int_X F(x, y)\, d\mu(x) \right)^p d\nu(y) \right)^{1/q} \\ &= \left\| \int_X F(x, y)\, d\mu(x) \right\|_{L^p(d\nu)}^{p-1}. \end{aligned}$$

By Fubini:

$$\begin{aligned} &\left\| \int_X F(x, y)\, d\mu(x) \right\|_{L^p(d\nu)}^p \\ &= \int_Y G(y) \int_X F(x, y)\, d\mu(x)\, d\nu(y) \end{aligned}$$

$$
\begin{aligned}
&= \int_X \int_Y G(y)F(x,y)\,d\nu(y)\,d\mu(x) \\
&\le \int_X \|G(y)\|_{L^q(d\nu)} \|F(x,y)\|_{L^p(d\nu)}\,d\mu(x) \\
&= \|G\|_{L^q(d\nu)} \int_X \|F(x,y)\|_{L^p(d\nu)}\,d\mu(x).
\end{aligned}
$$

Dividing by $\|G\|_{L^q(d\nu)}$ proves (4.1). □

Exercise 4.1.1.
Let $1 < p < \infty$ and $q = p/(p-1)$ be its conjugate exponent. Let $K(x, y)$, $x, y \in (0, \infty)$ be nonnegative and homogeneous of degree -1. That is, K has the property that for all $\lambda > 0$, $K(\lambda x, \lambda y) = \lambda^{-1} K(x, y)$. Suppose that

$$\int_0^\infty K(x,1)x^{-1/p}\,dx = \int_0^\infty K(1,y)y^{-1/q}\,dy = a$$

and denote the $L^p(0, \infty)$ norm of a function f by $\|f\|_p$. Prove that

$$\left\| \int_0^\infty K(x,y)f(x)\,dx \right\|_{L^p(dy)} \le a\|f\|_p$$

and that

$$\left| \int_0^\infty \int_0^\infty K(x,y)f(x)g(y)\,dx\,dy \right| \le a\|f\|_p \|g\|_q.$$

Exercise 4.1.2.
Let a_j and b_k be two sequences of nonnegative numbers. Prove that

$$\sum_{j=1}^\infty \sum_{k=1}^\infty \frac{a_j b_k}{j+k} \le \frac{\pi}{\sin(\pi/p)} \Big(\sum_{j=1}^\infty a_j^p \Big)^{1/p} \Big(\sum_{k=1}^\infty b_k^q \Big)^{1/q}.$$

Exercise 4.1.3.
Hardy's Inequality. Let f be a nonnegative function on $[0, \infty)$, $1 < p < \infty$ and $0 < r < \infty$. Prove that

$$\left(\int_0^\infty \left(\int_0^x f(y)\,dy \right)^p \frac{dx}{x^{r+1}} \right)^{1/p} \le \frac{p}{r} \left(\int_0^\infty (yf(y))^p \frac{dy}{y^{r+1}} \right)^{1/p}. \tag{4.4}$$

Exercise 4.1.4.
Suppose $r = p - 1$ in Exercise 4.1.3. In this case we can write the inequality (4.4) as

$$\left(\int_0^\infty \Big(\frac{1}{x}\int_0^x f(y)\,dy\Big)^p\,dx\right)^{1/p} \leq \frac{p}{p-1}\left(\int_0^\infty \big(f(y)\big)^p\,dy\right)^{1/p}.$$

Prove that the constant $p/(p-1)$ cannot be improved.

Exercise 4.1.5.
Let $\{a_k\}$ be a sequence of nonnegative numbers. Prove that

$$\left(\sum_{N=1}^\infty \Big(\frac{1}{N}\sum_{j=1}^N a_j\Big)^p\right)^{1/p} \leq \frac{p}{p-1}\Big(\sum_{j=1}^\infty a_j^p\Big)^{1/p}.$$

Exercise 4.1.6.
Let f be a differentiable function of compact support in $(0,\infty]$. Prove that

$$\int_0^\infty \frac{|f(y)|^2}{y^2}\,dy \leq 4\int_0^\infty |f'(y)|^2\,dy.$$

Exercise 4.1.7.
Let $u(x,y)$ be a differentiable function in the upper half space $\mathbb{R}^2_+ = \{(x,y) : x \in \mathbb{R}, y > 0\}$ of compact support. Prove that

$$\int_{\mathbb{R}^2_+} \frac{|u(x,y)|^2}{y^2}\,dx\,dy \leq 4\int_{\mathbb{R}^2_+} |\nabla u(x,y)|^2\,dx\,dy,$$

where $\nabla u(x,y)$ denotes the gradient of the function u at the point (x,y).

Exercise 4.1.8. (i) Let D be a simply connected domain in the plane. Use Exercise 4.1.7 and the Koebe 1/4-theorem to conclude that for any smooth function $u(x,y)$ with compact support in D,

$$\int_D \frac{|u(x,y)|^2}{d_D^2(x,y)}\,dx\,dy \leq 16\int_D |\nabla u(x,y)|^2\,dx\,dy,$$

where $d_D(x,y)$ represents the distance from the point $(x,y) \in D$ to D^c. That is, the distance from the point (x,y) to the boundary of D.

(ii) Use this inequality to conclude that if we define the inner radius of the domain D to be $R_D = \sup\{d_D(x,y) : (x,y) \in D\}$ and the quantity

$$\lambda_D = \inf\left(\frac{\int_D |\nabla u(x,y)|^2\,dx\,dy}{\int_D |u(x,y)|^2\,dx\,dy}\right),$$

where the infimum is taken over all smooth functions u of compact support in D, then

$$\lambda_D \geq \frac{1}{16R_D^2}.$$

Remark 4.2. The quantity λ_D defined above is the lowest eigenvalue of the Laplacian with Dirichlet boundary conditions in D. The inequality proves that if the "drum" D produces a low tone then it must contain a large disk. The converse is also true and very easy to prove. That is, if it contains a large disk then it produces a low tone. The reader interested in this connection can see [BC].

4.2 The Convolution Operator

Theorem 4.2. *Let $g \in L^1(\mathbb{R}^n)$, $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$. Define*

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy.$$

Then

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

Remark 4.3. $f * g$ is called the convolution of f and g. Notice that by a simple change of variables, $f * g = g * f$.

Proof. The case $p = \infty$ is trivial and the case $p = 1$ follows directly from Fubini's theorem. However, before we apply Fubini's theorem we must verify that the function $F(x,y) = f(x-y)g(y)$ is measurable in $\mathbb{R}^n \times \mathbb{R}^n$. We leave this as an exercise to the reader. Assume now that $1 < p < \infty$. Since $g \in L^1(\mathbb{R}^n)$, the measure $d\mu(y) = |g(y)|\,dy$ is finite and applying the Minkowski integral inequality we obtain

$$\|f * g\|_p \leq \left\| \int_{\mathbb{R}^n} |f(x-y)|\ |g(y)|\,dy \right\|_{L^p(dx)}$$

$$
\begin{aligned}
&\le \int_{\mathbb{R}^n} \|f(x-y)\|_{L^p(dx)} |g(y)|\, dy \\
&= \|f\|_{L^p(dx)} \|g\|_{L^1(dx)},
\end{aligned}
$$

which proves the Theorem for all p. □

The above theorem shows that the convolution of two functions inherits the integrability properties of the better function. The next main result shows that this is also the case as far as smoothness is concerned. This will permit us to show that the space of infinitely differentiable functions with compact support is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$. We shall use the following notation for the rest of these notes. We denote by $C_0(\mathbb{R}^n)$ the space of continuous functions of compact support in $\mathbb{R}^n$ and by $C_0^m(\mathbb{R}^n)$, $m = 1, 2 \ldots,$ the space of functions of compact support which also have m continuous derivatives each with compact support. We will denote functions which are m times differentiable (not necessarily of compact support), by $C^m(\mathbb{R}^n)$. Any polynomial is clearly in $C^\infty(\mathbb{R}^n)$ but not in any $L^p(\mathbb{R}^n)$. Examples of functions in $C^\infty(\mathbb{R}^n)$ which also have good integrability properties are: $e^{-|x|}$, $e^{-|x|^2}$ and

$$
\frac{1}{(1+|x|^2)^{(n+1)/2}} \tag{4.5}
$$

These functions will play an important role later in these notes. To construct C^∞–functions of compact support, let $\psi(t) = e^{-1/t}$ for $t < 0$ and 0 for $t \ge 0$. It follows from L'Hospital's rule that the function $\varphi(x) = \psi(|x|^2 - 1)$ is in $C_0^\infty(\mathbb{R}^n)$.

Proposition 4.3. *The space $C_0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$. That is, given $f \in L^p(\mathbb{R}^n)$ and $\varepsilon > 0$ there is a $g \in C_0(\mathbb{R}^n)$ such that*

$$
\|f - g\|_p < \varepsilon
$$

Remark 4.4. This proposition is proved in an elementary measure theory course. We observe that because of the density of the simple functions in $L^p(\mathbb{R}^n)$, it reduces to proving the result when f is a simple function and this can be further reduced to the case when f is the characteristic functions of sets of finite measure. By the regularity of the Lebesgue measure this in turn reduces to the case of characteristic functions of open sets.

Lemma 4.4. *The translation operator $y \to f(x-y)$ is continuous from $\mathbb{R}^n$ to $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$. That is, if $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, then*

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

Proof. Let $\varepsilon > 0$ be given. By Proposition 4.3 there is a function $g \in C_0(\mathbb{R}^n)$ such that

$$\|g(x) - f(x)\|_p < \varepsilon/3. \tag{4.6}$$

Thus

$$\begin{aligned} &\|f(x+h) - f(x)\|_p \\ &\quad \le \|f(x+h) - g(x+h)\|_p \\ &\qquad +\|g(x+h) - g(x)\|_p \\ &\qquad +\|g(x) - f(x)\|_p \\ &\quad < \tfrac{2}{3}\varepsilon + \|g(x+h) - g(x)\|_p. \end{aligned} \tag{4.7}$$

Let $B = B(0,r)$ be the ball centered at the origin and of radius r with r such that the support of the function g is contained in B. Since g has compact support, g is in fact uniformly continuous. Thus given $\eta > 0$ there is a $\delta > 0$ such that

$$|g(x+h) - g(x)| \le \eta, \tag{4.8}$$

for all $|h| < \delta$. We may suppose that δ is small enough so that for $|h| < \delta$, the function $g(x+h) - g(x)$ vanishes off the set $\widetilde{B} = B(0, 2r)$. Then

$$\begin{aligned} &\int_{\mathbb{R}^n} |g(x+h) - g(x)|^p \, dx \\ &\quad \le \int_{\widetilde{B}} |g(x+h) - g(x)|^p \, dx \\ &\quad \le \eta^p m(\widetilde{B}). \end{aligned} \tag{4.9}$$

If we take η such that $\eta^p m(\widetilde{B}) < (\varepsilon/3)^p$ and substitute the estimate from (4.9) into (4.7) we obtain the desired result. □

Theorem 4.5. *Let $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$ and $\varphi \in C_0^m(\mathbb{R}^n)$, $m \ge 1$. Then $f * \varphi \in L^p(\mathbb{R}^n) \cap C^m(\mathbb{R}^n)$.*

Proof. By Theorem 4.1, $f * \varphi \in L^p(\mathbb{R}^n)$ and all we need to do is show that $f * \varphi \in C^m$. We begin by proving that

(i) $f * \varphi$ is continuous. For this, apply Hölder's inequality to obtain,

$$\begin{aligned}
&|(f*\varphi)(x+h)-(f*\varphi)(x)| \\
&\quad = \ |(\varphi*f)(x+h)-(\varphi*f)(x)| \\
&\quad \le \ \int_{\mathbb{R}^n} |\varphi(x+h-y)f(y)-\varphi(x-y)f(y)|\, dy \\
&\quad \le \ \int_{\mathbb{R}^n} |f(y)\,[\varphi(x+h-y)-\varphi(x-y)]|\, dy \\
&\quad \le \ \|f\|_p \|\varphi(x+h-\cdot)-\varphi(x-\cdot)\|_{L^q(dy)}. \qquad (4.10)
\end{aligned}$$

If $1 < p$ then $q < \infty$ and the continuity of $f * \varphi$ follows from Lemma 4.5. If $p = 1$ then $q = \infty$ and we get as in (4.10),

$$|(f*\varphi)(x+h)-(f*\varphi)(x)| \le \|f\|_1 \sup_y |\varphi(y-h)-\varphi(y)|\,. \qquad (4.11)$$

By the uniform continuity of φ the right hand side of (4.11) goes to zero as $|h| \to 0$. This proves (i). Notice that in proving (i) we only used the fact that $\varphi \in C_0^1(\mathbb{R}^n)$.

(ii) $f * \varphi$ is differentiable and

$$\frac{\partial}{\partial x_j}(f*\varphi) = f * \frac{\partial \varphi}{\partial x_j}, \text{ for } j = 1, \ldots n. \qquad (4.12)$$

To prove (ii), let e_j be the jth coordinate vector in $\mathbb{R}^n$. Fix the point $x \in \mathbb{R}^n$ and consider the function

$$F_{x,t,j}(y) = \frac{\varphi(x+te_j-y)-\varphi(x-y)}{t} - \frac{\partial \varphi}{\partial x_j}(x-y)$$

with $t > 0$. Then

$$\frac{f*\varphi(x+te_j)-f*\varphi(x)}{t} - \left(f*\tfrac{\partial\varphi}{\partial x_j}\right)(x) = \int_{\mathbb{R}^n} f(y)F_{x,t,j}(y)\,dy. \quad (4.13)$$

Since the function φ has compact support, is continuous and differentiable, we have that as $t \to 0$, $F_{x,t,j}(y) \to 0$, uniformly in y. Furthermore, for each fixed x, the function $F_{x,t,j}(y)$ is uniformly bounded in y and of compact support. By the dominated convergence theorem the right side of (4.13) goes to zero as t goes to zero and we have proved (ii).

Now repeat the proofs of (i) and (ii) with $\frac{\partial \varphi}{\partial x_j}$ in place of φ to complete the proof. □

Exercise 4.2.1.
Prove that if f and g both have compact support so does $f * g$.

We end this section by stating another important inequality for convolutions. The proof follows directly from the Riesz–Thorin interpolation theorem given in Section 5.4 of Chapter 5 below.

Theorem 4.6 (Young's Inequality)**.** *Let $r \geq 1$ satisfy*

$$\frac{1}{r} = 1 + \frac{1}{q} - \frac{1}{p}$$

for $1 \leq p \leq q \leq \infty$ and suppose $f \in L^p(\mathbb{R}^n)$ and $g \in L^r(\mathbb{R}^n)$. Then $f * g \in L^q(\mathbb{R}^n)$ and

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

4.3 Approximations to the Identity

For any function φ defined on $\mathbb{R}^n$ and $\varepsilon > 0$, we define the dilation operator $\tau_\varepsilon \varphi(x) = \varphi(\varepsilon x)$ and the function $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$.

Lemma 4.7. *Let $\varphi \geq 0$, and $\varphi \in L^1(\mathbb{R}^n)$. Then for any $\varepsilon > 0$,*

$$\int_{\mathbb{R}^n} \varphi_\varepsilon(x)\, dx = \int_{\mathbb{R}^n} \varphi(x)\, dx \tag{4.14}$$

and

$$\lim_{\varepsilon \to 0} \int\limits_{\{|x|>\delta\}} \varphi_\varepsilon(x)\, dx = 0, \text{ for any } \delta > 0. \tag{4.15}$$

Furthermore, if we set $Tf(x) = \varphi * f(x)$ we have

$$(\tau_{\varepsilon^{-1}} T \tau_\varepsilon)\, f(x) = \varphi_\varepsilon * f(x). \tag{4.16}$$

Proof. (4.14) follows from the obvious change of variables. For (4.15) observe that by the dominated convergence theorem,

$$\int\limits_{\{|x|>k\}} \varphi(x)\, dx \to 0,$$

as $k \to \infty$. Thus

$$\int\limits_{\{|x|>\delta\}} \varphi_\varepsilon(x)\,dx = \frac{1}{\varepsilon^n} \int\limits_{\{|x|>\delta\}} \varphi(x/\varepsilon)\,dx = \int\limits_{\{|x|>\delta/\varepsilon\}} \varphi(x)\,dx \to 0$$

as $\varepsilon \to 0$.

For (4.16) simply observe that the left hand side is

$$\int_{\mathbb{R}^n} \varphi(\varepsilon^{-1}x - y) f(\varepsilon y)\,dy$$

and change variables. □

Theorem 4.8. *Let $\varphi \geq 0$ with $\int_{\mathbb{R}^n} \varphi(y)\,dy = 1$.*

(i) Suppose $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$. Then

$$\lim_{\varepsilon \to 0} \|f * \varphi_\varepsilon - f\|_p = 0. \tag{4.17}$$

(ii) Suppose $f \in L^\infty(\mathbb{R}^n)$. Then for every x which is a point of continuity of f,

$$\lim_{\varepsilon \to 0} f * \varphi_\varepsilon(x) = f(x). \tag{4.18}$$

Proof. By our assumption on φ and (4.14), $\varphi_\varepsilon\,dy$ is a probability measure. Writing

$$|f * \varphi_\varepsilon(x) - f(x)| = \left| \int_{\mathbb{R}^n} (f(x-y) - f(x))\,\varphi_\varepsilon(y)\,dy \right|$$

we obtain by the Minkowski integral inequality that

$$\begin{aligned} \|f * \varphi_\varepsilon - f\|_p &\leq \left\| \int_{\mathbb{R}^n} |f(x-y) - f(x)| \varphi_\varepsilon(y)\,dy \right\|_p \\ &\leq \int_{\mathbb{R}^n} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\,dy. \end{aligned} \tag{4.19}$$

Set

$$I = \int\limits_{\{|y|\leq\delta\}} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\,dy$$

and
$$II = \int_{\{|y|>\delta\}} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\, dy$$

By the continuity of the translations in $L^p(\mathbb{R}^n)$, Lemma 4.4, given $\eta > 0$ there is a δ such that
$$\|f(x-y) - f(x)\|_{L^p(dx)} < \eta$$
for all $|y| < \delta$. Thus with such a δ,
$$I < \eta \int_{\{|y|<\delta\}} \varphi_\varepsilon(y)\, dy \le \eta.$$
From the fact that
$$\|f(x-y) - f(x)\|_{L^p(dx)} \le 2\|f\|_p$$
it follows that
$$II \le 2\|f\|_p \int_{\{|y|>\delta\}} \varphi_\varepsilon(y)\, dy,$$
which goes to zero as ε goes to zero. This and (4.19) completes the proof of (4.17).

For (4.18) we begin exactly in the same way and obtain
$$|f * \varphi_\varepsilon(x) - f(x)| \le \int_{\{|y|<\delta\}} |f(x-y) - f(x)| \varphi_\varepsilon(y)\, dy$$
$$+ \int_{\{|y|\ge\delta\}} |f(x-y) - f(x)| \varphi_\varepsilon(y)\, dy = I + II.$$

II is handled exactly as in (i). For I, suppose f is continuous at x. There is a $\delta > 0$ (depending on x) such that
$$|f(x-y) - f(x)| \le \eta, \text{ if } |y| < \delta.$$
Thus
$$I < \eta \int_{\{|y|<\delta\}} \varphi_\varepsilon(y)\, dy \le \eta,$$
which proves (4.18) and hence the Theorem. $\square$

In subsequent chapters we will make repeated applications of the following corollary.

Corollary 4.9. *The space of infinitely differentiable functions with compact support, $C_0^\infty(\mathbb{R}^n)$, is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$. That is, given an $f \in L^p(\mathbb{R}^n)$ and and $\eta > 0$ there is a $\varphi \in C_0^\infty(\mathbb{R}^n)$ such that*

$$\|f - \varphi\|_p < \eta.$$

Proof. Fix a nonnegative $\varphi \in C_0^\infty(\mathbb{R}^n)$ with integral 1. Let $f \in L^p(\mathbb{R}^n)$ and $\eta > 0$. Let r large enough so that

$$\int_{\{|y|>r\}} |f(y)|^p \, dy < \eta^p \tag{4.20}$$

and set

$$f^\varepsilon(x) = (f\chi_{B(0,r)}) * \varphi_\varepsilon(x),$$

where $B(0,r)$. By Theorem 4.5 and Exercise 4.2.1, $f^\varepsilon \in C_0^\infty(\mathbb{R}^n)$. By the Minkowski inequality and (4.20),

$$\begin{aligned} \|f - f^\varepsilon\|_p &\le \left\|f\chi_{B(0,r)} - f^\varepsilon\right\|_p + \left\|f(1 - \chi_{B(0,r)})\right\|_p \\ &\le \ \left\|f\chi_{B(0,r)} - f^\varepsilon\right\|_p + \eta = \left\|f\chi_{B(0,r)} - \varphi_\varepsilon * (f\chi_{B(0,r)})\right\|_p + \eta. \end{aligned}$$

By Theorem 4.8, the first term on the right hand side of (4.21) goes to zero as $\varepsilon \to 0$. Since $\eta > 0$ was arbitrary, this proves the corollary. $\square$

It follows from Theorem 4.8 that given $f \in L^p(\mathbb{R}^n)$ there is a sequence ε_j converging to zero such that

$$\lim_{\varepsilon_j \to 0} f * \varphi_{\varepsilon_j}(x) = f(x) \text{ a.e.}$$

The following stronger result will be proved in the next Chapter where we will study the *Hardy–Littlewood Maximal Functions*, one of the most fundamental operators in real analysis.

Theorem 4.10. *With φ as in Theorem 4.8 and $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, we have*

$$\lim_{\varepsilon \to 0} f * \varphi_\varepsilon(x) = f(x) \ \textit{a.e.}$$

Chapter 5

The Hardy–Littlewood Maximal Function

In this chapter we study the Hardy-Littlewood maximal function. This is one of the most fundamental operators in real analysis with many applications in different areas. The Hardy-Littlewood maximal function will play an important role for the rest of these notes. We will first prove its L^p–boundedness and then apply this to prove some classical differentiation theorems. The Hardy-Littlewood maximal function is a maximal operator obtained by convolution, as in Chapter 4, with the characteristic function of the unit ball normalized to have integral one. It is remarkable that all other such convolutions can be controlled by this one. This is presented in Theorem 5.4 below; Theorem 4.10 in Chapter 4 immediately follows from this. The Calderón–Zygmund decomposition, another result of fundamental importance in real analysis, will be proved in Section 5.2. As an illustration of the usefulness of the Calderón–Zygmund decomposition, we will discuss in Section 5.3 some of its applications to BMO. This is the John–Nirenberg space of functions with *bounded mean oscillation*. The Marcinkiewicz and Riesz–Thorin interpolation theorems are presented in Section 5.4. These interpolation theorems will be important in proving the L^p–boundedness of singular integrals in Chapter 7.

5.1 The L^p–inequalities

Let us denote by $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ the space of measurable functions on $\mathbb{R}^n$ which are integrable on any bounded subset of $\mathbb{R}^n$. Of fundamental importance for the rest of these notes is the Hardy-Littlewood maximal operator (also referred to as the maximal function) defined for any $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ by

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy, \tag{5.1}$$

where as before $|B(x,r)|$ denotes the volume of the ball $B(x,r)$. Often we will write this simply as $\gamma_n r^n$ where γ_n is the volume of the unit ball. This operator arises naturally in studying the following question on differentiation: *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, for what x's is it true that*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy = f(x)? \tag{5.2}$$

The Lebesgue Differentiation Theorem asserts that (5.2) holds for almost every x. Notice that if we take

$$\varphi = \frac{1}{|B(0,1)|} \chi_{B(0,1)} \tag{5.3}$$

we have

$$\varphi_r * f(x) = \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy$$

and

$$Mf(x) = \sup_{r>0} \big(\varphi_r * |f|\big)(x);$$

hence the connection to approximations to the identity and to Theorem 4.10 of Chapter 4. However, it is a remarkable fact that all other approximations to the identity can be controlled by this one, as we shall see in Theorem 5.4 below.

We begin with the basic boundedness properties of the maximal operator. First, let us observe that if $f \not\equiv 0$, then $Mf \notin L^1(\mathbb{R}^n)$. We leave it as an exercise to the reader to check that if $f \not\equiv 0$ then $Mf(x) > C_f^1/|x|^n$ for all

$|x| > C_f^2$, where the constant C_f^1 and C_f^2 depend on f. Thus by integrating in polar coordinates (exercise 3.2.7 in page 50) we have

$$\int_{\mathbb{R}^n} Mf(x)\,dx \geq C_f^1 \int_{\mathbb{R}^n\setminus B(0,C_f^2)} \frac{dx}{|x|^n} = \sigma(S^{n-1}) \int_{C_f^2}^{\infty} \frac{r^{n-1}}{r^n}\,dr = \infty. \tag{5.4}$$

We next introduce some standard terminology which we shall use in these notes. Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be two measure spaces. For $1 \leq p \leq \infty$ denote their corresponding L^p–spaces by $L^p(\mu)$ and $L^p(\nu)$, respectively, with corresponding norms $\|f\|_{L^p(\mu)}$ and $\|f\|_{L^p(\nu)}$. We recall that for any $0 < p < \infty$ and any measure space (X, μ), the weak–$L^p(\mu)$ space is the collection of all measurable functions with the property that

$$\sup_{\lambda>0} \Big\{ \lambda^p \mu\{x \in X : |f(x)| > \lambda\} \Big\} < \infty.$$

For $p = \infty$, weak–L^∞ is just the same as L^∞. A mapping $T : L^p(\mu) \to L^q(\nu)$ is said to be of strong–type (p, q) if there is a constant $A = A_{p,q}$ independent of f such that

$$\|Tf\|_{L^q(\nu)} \leq A\|f\|_{L^p(\mu)}.$$

The mapping T is said to be weak–type (p, q) if there is a constant A_1 independent of f such that

$$\nu\{y \in Y : |Tf(y)| > \alpha\} \leq \left(\frac{A_1}{\alpha} \|f\|_{L^p(\mu)} \right)^q,$$

for $q < \infty$. If $q = \infty$, weak–type (p, q) is just the same as strong–type (p, q)

Theorem 5.1. *The Hardy–Littlewood maximal operator M is weak–type $(1, 1)$ and strong–type (p, p) for $1 < p \leq \infty$. In fact,*

(i)

$$m\{x \in \mathbb{R}^n \colon Mf(x) > \alpha\} \leq \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(y)|\,dy$$

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

(ii)

$$\|Mf\|_p \leq C_{n,p}\|f\|_p,$$

$1 < p \leq \infty$, $f \in L^p(\mathbb{R}^n)$. *Furthermore,*

$$C_n = 5^n \tag{5.5}$$

and

$$C_{n,p} = 2\left(\frac{5^n p}{p-1}\right)^{1/p}. \tag{5.6}$$

Proof. Let $E_\alpha = \{x \in \mathbb{R}^n \colon Mf(x) > \alpha\}$. By Exercise 5.1.1, this set is open hence measurable. For each $x \in E_\alpha$ there exists a ball centered at x and radius r_x, which we denote by $B_{r_x}(x)$, such that

$$\int_{B_{r_x}(x)} |f(y)|\, dy > \alpha |B_{r_x}(x)|.$$

Also, $|B_{r_x}(x)| \leq \frac{1}{\alpha}\|f\|_1$ and hence these balls have bounded diameter. Clearly we have

$$E_\alpha \subset \bigcup_{x \in E_\alpha} B_{r_x}(x).$$

We may therefore, by Theorem 1.4 in Chapter 1, pick a disjoint sequence $B_1, B_2, \ldots$ of these balls such that $|E_\alpha| \leq 5^n \sum_j |B_j|$. By the way the balls were chosen, we know that

$$\sum_j |B_j| \leq \frac{1}{\alpha} \sum_j \int_{B_j} |f(y)|\, dy = \frac{1}{\alpha} \int_{\cup B_j} |f(y)|\, dy \leq \frac{1}{\alpha} \int_{\mathbb{R}^n} |f(y)|\, dy,$$

proving (i) with $C_n = 5^n$.

For (ii), when $p = \infty$, the result is trivial and in fact $\|Mf\|_\infty \leq \|f\|_\infty$. Assume therefore that $1 < p < \infty$. Let $\alpha > 0$ and define

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

Then

$$|f(x)| \leq |f_1(x)| + \alpha/2$$

from which we conclude that

$$Mf(x) \leq Mf_1(x) + \alpha/2.$$

Since

$$\int_{\mathbb{R}^n} |f_1(y)|\, dy = \int_{\{|f(x)| \geq \alpha/2\}} |f(x)|\, dx,$$

we see that $f_1 \in L^1(\mathbb{R}^n)$ and we have by (i) that

$$m\{x \in \mathbb{R}^n \colon Mf(x) > \alpha\} \le m\{Mf_1(x) > \alpha/2\} \le \tfrac{2\cdot 5^n}{\alpha} \int_{\mathbb{R}^n} |f_1(y)|\, dy.$$

This together with Exercise 3.2.3 on page 49, and Fubini's theorem imply

$$\begin{aligned}
\int_{\mathbb{R}^n} Mf(x)^p\, dx &= p \int_0^\infty \alpha^{p-1} m\{Mf > \alpha\}\, d\alpha \\
&\le p \int_0^\infty \alpha^{p-1} \frac{2\cdot 5^n}{\alpha} \int\limits_{\{|f| \ge \alpha/2\}} |f(x)|\, dx\, d\alpha \\
&= 2\cdot 5^n p \int_{\mathbb{R}^n} |f(x)| \int_0^{2|f(x)|} \alpha^{p-2}\, d\alpha\, dx \\
&= 2^p \cdot \frac{5^n p}{p-1} \int_{\mathbb{R}^n} |f(x)|^p\, dx.
\end{aligned}$$

Thus we obtain

$$\|Mf\|_p \le 2 \left(\frac{5^n p}{p-1}\right)^{1/p} \|f\|_p, \quad 1 < p < \infty,$$

which completes the proof of the theorem. □

The following corollary of the above proof is often useful.

Corollary 5.2. *For all $\alpha > 0$ the following inequality holds:*

$$m\{x \in \mathbb{R}^n : Mf(x) > \alpha\} \le \frac{C_n}{\alpha} \int\limits_{\{x \in \mathbb{R}^n : |f(x)| > \alpha/2\}} |f(x)|\, dx$$

with the constant C_n depending only on n.

Remark 5.1. Note that for fixed n, $C_{n,p} \to \infty$, as $p \to 1$ and $C_{n,p} \to 1$, as $p \to \infty$, reflecting the fact that the operator is bounded on $L^\infty(\mathbb{R}^n)$ but not on $L^1(\mathbb{R}^n)$. Also, as far as the asymptotic behavior in p is concerned, the constants in (5.5) and (5.6) are already best possible. However, this is not the case with respect to the dimension n. It has been proved by Stein and Strönberg [St-St] that the constant in (ii) can be taken to be independent of the dimension, and that the constant in (i) can be taken to be $Cn \log n$ with

C independent of n. Whether the constant in (i) can be taken independent of the dimension remains an interesting and challenging open problem. For other interesting problems concerning the dependence on n in other clasical operators in analysis, see Stein [St2].

Corollary 5.3 (Lebesgue Differentiation Theorem)**.** *Let $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$. Then*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int\limits_{B(x,r)} f(y)\,dy = f(x) \ \textit{a.e.}$$

Proof. Let us first deal with the case $p = 1$. Let

$$\omega(f)(x) = \left| \limsup_{r\to 0} f * \varphi_r(x) - \liminf_{r\to 0} f * \varphi_r(x) \right|,$$

where

$$\varphi = \frac{1}{|B(0,1)|} \chi_{B(0,1)}$$

By Theorem 4.8, Chapter 4, $\|f * \varphi_r - f\|_1 \to 0$ as $r \to 0$ and if $h \in C_0^\infty(\mathbb{R}^n)$, $h * \varphi_r(x) \to h(x)$ uniformly in x. Thus for such an h we have $\omega(h)(x) = 0$ for all x. Given $f \in L^1(\mathbb{R}^n)$ and $\eta > 0$, choose an $h \in C_0^\infty(\mathbb{R}^n)$ such that

$$\|f - h\|_1 < \eta.$$

With this we arrive at

$$\omega(f)(x) \leq \omega(f - h)(x) \leq 2M(f - h)(x).$$

Fix $\varepsilon > 0$. By (i) of Theorem 5.1 we have that

$$\begin{aligned} |\{\omega(f)(x) > \varepsilon\}| &\leq |\{\omega(f-h) > \varepsilon\}| \\ &\leq |\{M(f-h) > \varepsilon/2\}| \\ &\leq \frac{C_n}{\varepsilon} \int_{\mathbb{R}^n} |f - h|\,dx \leq \frac{C_n \eta}{\varepsilon}. \end{aligned}$$

Since this is true for any $\eta > 0$ we obtain that

$$|\{\omega(f)(x) > \varepsilon\}| = 0.$$

Since $\varepsilon > 0$ was arbitrary, this can only happen if $\omega(f)(x) = 0$ for almost every x, which proves that $\lim_{r\to 0} \varphi_r * f$ exists almost everywhere. By the convergence of $\varphi_r * f$ to f in L^1, there exists a sequence $r_k \downarrow 0$ such that $\lim_{r_k \to 0} \varphi_{r_k} * f \to f$ almost everywhere. Thus

$$\lim_{r \to 0} \varphi_r * f = f \text{ a.e.}$$

and we have proved the case $p = 1$. The case $1 < p < \infty$ is exactly the same except that we use Chebychev's inequality and (ii) of Theorem 5.1. We now present the case $p = \infty$ which is a little different due to the fact that we cannot approximate functions in L^∞ by functions in $C_0^\infty(\mathbb{R}^n)$. Suppose $f \in L^\infty$. We shall show that $f * \varphi_r(x) \to f(x)$ for almost every $x \in B$ for any fixed ball B. Let B_1 be another ball satisfying $B \subset B_1$ and with $\delta =$ distance$(B, B_1^c) > 0$. Set

$$f_1(x) = \begin{cases} f(x) & \text{if } x \in B_1 \\ 0 & \text{if } x \notin B_1 \end{cases}$$

Then $f_1 \in L^1$ and by what we have already proved, $f_1 * \varphi_r(x) \to f_1(x)$, as $r \to 0$, for almost every x. In particular,

$$f_1 * \varphi_r(x) \to f_1(x) = f(x), \text{ a.e. in } B,$$

as $r \to 0$. Set $f_2(x) = f(x) - f_1(x)$. We claim that

$$f_2 * \varphi_r(x) \to 0, \quad x \in B,$$

as $r \to 0$. To see this observe that for $x \in B$,

$$\begin{aligned} |f_2 * \varphi_r(x)| &= \left| \int_{\mathbb{R}^n} f_2(x-y)\varphi_r(y)\, dy \right| \\ &\leq \int_{\{|y| \geq \delta > 0\}} |f_2(x-y)||\varphi_r(y)|\, dy \\ &\leq \|f\|_{L^\infty} \int_{\{|y| \geq \delta\}} |\varphi_r(y)|\, dy \\ &= \|f\|_{L^\infty} \int_{\{|y| \geq \delta/r\}} |\varphi(y)|\, dy \end{aligned}$$

and this last quantity goes to zero as $r \to 0$. □

As we mentioned earlier, the above result for the characteristic (indicator) function of the ball implies the more general result stated in Theorem 3.2 of Chapter IV. To do this we need the following

Theorem 5.4. *Suppose $\varphi \in L^1(\mathbb{R}^n)$ and that the least decreasing radial majorant of φ is in $L^1(\mathbb{R}^n)$. That is, define*

$$\psi(x) = \sup_{|y| \geq |x|} |\varphi(y)|$$

and suppose that

$$\int_{\mathbb{R}^n} \psi(x)\, dx = A < \infty.$$

Let $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$. Then

$$\sup_{\varepsilon > 0} |f * \varphi_\varepsilon(x)| \leq AMf(x).$$

Proof. By bringing absolute values inside the integral and majorizing φ by ψ, it is enough to prove

$$|f * \psi_\varepsilon(x)| \leq AMf(x) \tag{5.7}$$

for $f \geq 0$. Since (5.7) is translation invariant with respect to f and dilation invariant with respect to ψ, we only need to show that

$$|f * \psi_1(0)| \leq AMf(0). \tag{5.8}$$

That is, once (5.8) is proved we apply it with f replaced by $f_x(y) = f(x - y)$ and ψ replaced by ψ_ε which has the same A. To prove (5.8), assume our function ψ is a simple function of the form

$$\psi(x) = \sum_{k=1}^{m} c_k \chi_{B(0,r_k)}(x) \tag{5.9}$$

where the numbers c_k are positive. With this we see that

$$A = \sum_{k=1}^{m} c_k |B(0, r_k)|.$$

Then

$$
\begin{aligned}
|f * \psi_1(0)| = \left| \int_{\mathbb{R}^n} f(x)\psi(x)\,dx \right| &\le \sum_{k=1}^{m} c_k \int_{B(0,r_k)} |f(x)|\,dx \\
= \quad \sum_{k=1}^{m} c_k |B(0,r_k)| \frac{1}{|B(0,r_k)|} & \int_{B(0,r_k)} |f(x)|\,dx \le AMf(0). \qquad (5.10)
\end{aligned}
$$

The case of general ψ is accomplished by Exercise 5.1.9 below. Indeed, given the exercise we have by the monotone convergence theorem that

$$
\begin{aligned}
\left| \int_{\mathbb{R}^n} f(x)\psi(x)\,dx \right| &\le \int_{\mathbb{R}^n} |f(x)|\psi(x)\,dx \\
&= \lim_{j\to\infty} \int_{\mathbb{R}^n} |f(x)|\psi_j(x)\,dx \le AMf(0),
\end{aligned}
$$

and this proves the theorem. □

An alternative proof of Theorem 5.4 goes as follows. Set $E = \{(y,t) : \psi(y) > t > 0\}$. Note that

$$
\psi(y) = \int_0^\infty \chi_E(y,t)\,dt
$$

By Fubini's Theorem and the fact that for each $t > 0$ the set $\{y \in \mathbb{R}^n : \psi(y) > t\}$ is a ball centered at the origin of radius r_t, we have

$$
\begin{aligned}
|f * \psi(x)| &\le \int_{\mathbb{R}^n} |f(x-y)| \int_0^\infty \chi_E(y,t)\,dt\,dy \\
&= \int_0^\infty \int_{\{y\in\mathbb{R}^n : \psi(y) > t\}} |f(x-y)|\,dy\,dt \\
&= \int_0^\infty \left(\frac{1}{|B_{r_t}|} \int_{B_{r_t}} |f(x-y)|\,dy \right) |B_{r_t}|\,dt \\
&\le \|\psi\|_1 Mf(x) = AMf(x),
\end{aligned}
$$

which proves the theorem.

The following two corollaries are direct from Theorems 5.1, 5.4 and the proof of Corollary 5.3. Observe that Corollary 5.5 below is in fact more general than Theorem 4.10 of Chapter 4.

Corollary 5.5. *Let* φ *and* A *be as in Theorem 5.4. Set*

$$f^*(x) = \sup_{\varepsilon>0} |\varphi_\varepsilon * f(x)| .$$

Then

$$m\{x \in \mathbb{R}^n : f^*(x) \geq \alpha\} \leq \frac{AC_n}{\alpha}\|f\|_1$$

and

$$\|f^*\|_p \leq AC_{p,n}\|f\|_p,$$

for $1 < p \leq \infty$*. The constants* C_n *and* $C_{p,n}$ *are the same as those in Theorem 1.1.*

Corollary 5.6. *Let* φ *be as in Theorem 5.4 with* $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$*.*

(i) If $f \in L^p(\mathbb{R}^n)$ *,* $1 \leq p \leq \infty$*, then*

$$\lim_{\varepsilon\to 0} f * \varphi_\varepsilon(x) = f(x) \text{ a.e.}$$

(ii) If $f \in L^p(\mathbb{R}^n)$ *,* $1 \leq p < \infty$*, then*

$$\lim_{\varepsilon\to 0} \|f * \varphi_\varepsilon - f\|_p = 0.$$

Exercise 5.1.1.
Prove that the maximal function is lower semi-continuous. That is, for all $\lambda > 0$, $\{x \in \mathbb{R}^n : Mf(x) > \lambda\}$ is open.

Exercise 5.1.2.
Prove that

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\,dy = f(x) \text{ a.e.}$$

for $f \in L^1_{loc}(\mathbb{R}^n)$.

Exercise 5.1.3.
Suppose the function f is supported in the ball B with $f \in L^1(B)$ and $Mf \in L^1(B)$. Prove that

$$\int_{\{x\in\mathbb{R}^n : Mf(x)>\alpha\}} Mf(x)\,dx < \infty,$$

for any $\alpha > 0$.

Exercise 5.1.4.
Prove that in the definition of the maximal function the ball $B(x,r)$ can be replaced by a cube. That is if instead we define

$$Mf(x) = \sup_{Q} \frac{1}{|Q|} \int_{Q} |f(y)|\, dy,$$

where the sup is taken over all cubes Q centered at x, then M satisfies the conclusions of Theorem 5.1.

Exercise 5.1.5.
For $f \in L^1_{loc}(\mathbb{R}^n)$. For $1 \leq p < \infty$, define

$$M_p(f)(x) = \sup_{r>0} \left(\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|^p\, dy \right)^{1/p}.$$

Prove that M_p is weak–type (p,p) and strong–type (q,q) for $p < q \leq \infty$.

Exercise 5.1.6.
Let μ be a Borel measure on $\mathbb{R}^n$ satisfying the doubling property. That is, there is a constant c such that $\mu(B(x,2r)) \leq c\mu(B(x,r))$ for all r. Define

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu(y).$$

Prove that

(i) $\mu\{x \in \mathbb{R}^n \colon M_\mu f(x) > \alpha\} \leq \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(y)|\, d\mu(y)$, for $f \in L^1(\mu)$

(ii) $\|M_\mu f\|_{L^p(\mu)} \leq C_{n,p} \|f\|_{L^p(\mu)}$, $1 < p \leq \infty$, $f \in L^p(\mathbb{R}^n)$.

Exercise 5.1.7.
Kolmogorov's inequality. Let $f \in L^1(\mathbb{R}^n)$ and $0 < p < 1$. Prove that

$$\int_E (Mf(x))^p\, dx \leq C_p |E|^{1-p} \|f\|_1^p,$$

for all measurable $E \in \mathbb{R}^n$.

Exercise 5.1.8.
Let μ be a finite Borel measure on $\mathbb{R}^n$. We define its maximal function by

$$M\mu(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \mu(B(x,r)).$$