Chapter 6
Smoothness and Function Spaces

In this chapter we study differentiability and smoothness of functions. There are several ways to interpret smoothness and numerous ways to describe it and quantify it. A fundamental fact is that smoothness can be measured and fine-tuned using the Fourier transform, and this point of view is of great importance. In fact, the investigation of the subject is based on this point. It is not surprising, therefore, that Littlewood–Paley theory plays a crucial and deep role in this study.

Certain spaces of functions are introduced to serve the purpose of measuring smoothness. The main function spaces we study are Lipschitz, Sobolev, and Hardy spaces, although the latter measure smoothness within the realm of rough distributions. Hardy spaces also serve as a substitute for L^p when $p < 1$. We also take a quick look at Besov–Lipschitz and Triebel–Lizorkin spaces, which provide an appropriate framework that unifies the scope and breadth of the subject. One of the main achievements of this chapter is the characterization of these spaces using Littlewood–Paley theory. Another major accomplishment of this chapter is the atomic characterization of these function spaces. This is obtained from the Littlewood–Paley characterization of these spaces in a single way for all of them.

Before one embarks on a study of function spaces, it is important to understand differentiability and smoothness in terms of the Fourier transform. This can be achieved using the Laplacian and the potential operators and is discussed in the first section.

6.1 Riesz Potentials, Bessel Potentials, and Fractional Integrals

Recall the Laplacian operator

$$\Delta = \partial_1^2 + \cdots + \partial_n^2 ,$$

which may act on functions or tempered distributions. The Fourier transform of a Schwartz function (or even a tempered distribution f) satisfies the following

L. Grafakos, *Modern Fourier Analysis*, DOI: 10.1007/978-0-387-09434-2_6,

identity:

$$-\widehat{\Delta(f)}(\xi) = 4\pi^2|\xi|^2\widehat{f}(\xi)\,.$$

Motivated by this identity, we replace the exponent 2 by a complex exponent z and we define $(-\Delta)^{z/2}$ as the operator given by the multiplication with the function $(2\pi|\xi|)^z$ on the Fourier transform. More precisely, for $z \in \mathbf{C}$ and Schwartz functions f we define

$$(-\Delta)^{z/2}(f)(x) = ((2\pi|\xi|)^z\widehat{f}(\xi))^\vee(x)\,. \tag{6.1.1}$$

Roughly speaking, the operator $(-\Delta)^{z/2}$ is acting as a derivative of order z if z is a positive integer. If z is a complex number with real part less than $-n$, then the function $|\xi|^z$ is not locally integrable on $\mathbf{R}^n$ and so (6.1.1) may not be well defined. For this reason, whenever we write (6.1.1), we assume that either $\operatorname{Re} z > -n$ or $\operatorname{Re} z \le -n$ and that $\widehat{f}$ vanishes to sufficiently high order at the origin so that the expression $|\xi|^z\widehat{f}(\xi)$ is locally integrable. Note that the family of operators $(-\Delta)^z$ satisfies the semigroup property

$$(-\Delta)^z(-\Delta)^w = (-\Delta)^{z+w}\,, \qquad \text{for all } z,w \in \mathbf{C},$$

when acting on spaces of suitable functions.

The operator $(-\Delta)^{z/2}$ is given by convolution with the inverse Fourier transform of $(2\pi)^z|\xi|^z$. Theorem 2.4.6 gives that this inverse Fourier transform is equal to

$$(2\pi)^z(|\xi|^z)^\vee(x) = (2\pi)^z\frac{\pi^{-\frac{z}{2}}}{\pi^{\frac{z+n}{2}}}\frac{\Gamma(\frac{n+z}{2})}{\Gamma(\frac{-z}{2})}|x|^{-z-n}\,. \tag{6.1.2}$$

The expression in (6.1.2) is in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$ only when $-\operatorname{Re} z - n > -n$, that is when $\operatorname{Re} z < 0$. In general, (6.1.2) is a distribution. Thus only in the range $-n < \operatorname{Re} z < 0$ are both the function $|\xi|^z$ and its inverse Fourier transform locally integrable functions.

6.1.1 Riesz Potentials

When z is a negative real number, the operation $f \mapsto (-\Delta)^{z/2}(f)$ is not really "differentiating" f, but "integrating" it instead. For this reason, we introduce a slightly different notation in this case by replacing z by $-s$.

Definition 6.1.1. Let s be a complex number with $\operatorname{Re} s > 0$. The *Riesz potential* of order s is the operator

$$I_s = (-\Delta)^{-s/2}.$$

Using identity (6.1.2), we see that I_s is actually given in the form

$$I_s(f)(x) = 2^{-s}\pi^{-\frac{n}{2}}\frac{\Gamma(\frac{n-s}{2})}{\Gamma(\frac{s}{2})}\int_{\mathbf{R}^n} f(x-y)|y|^{-n+s}\,dy\,,$$

and the integral is convergent if f is a function in the Schwartz class.

We begin with a simple, yet interesting, remark concerning the homogeneity of the operator I_s.

Remark 6.1.2. Suppose that for s real we had an estimate

$$\big\|I_s f\big\|_{L^q(\mathbf{R}^n)} \le C(p,q,n,s)\big\|f\big\|_{L^p(\mathbf{R}^n)} \tag{6.1.3}$$

for some positive indices p,q and all $f \in L^p(\mathbf{R}^n)$. Then p and q must be related by

$$\frac{1}{p} - \frac{1}{q} = \frac{s}{n}. \tag{6.1.4}$$

This follows by applying (6.1.3) to the dilation $\delta^a(f)(x) = f(ax)$ of the function f, $a > 0$, in lieu of f, for some fixed f, say $f(x) = e^{-|x|^2}$. Indeed, replacing f by $\delta^a(f)$ in (6.1.3) and carrying out some algebraic manipulations using the identity $I_s(\delta^a(f)) = a^{-s}\delta^a(I_s(f))$, we obtain

$$a^{-\frac{n}{q}-s}\big\|I_s(f)\big\|_{L^q(\mathbf{R}^n)} \le C(p,q,n,s)a^{-\frac{n}{p}}\big\|f\big\|_{L^p(\mathbf{R}^n)}. \tag{6.1.5}$$

Suppose now that $\frac{1}{p} > \frac{1}{q} + \frac{s}{n}$. Then we can write (6.1.5) as

$$\big\|I_s(f)\big\|_{L^q(\mathbf{R}^n)} \le C(p,q,n,s)a^{\frac{n}{q}-\frac{n}{p}+s}\big\|f\big\|_{L^p(\mathbf{R}^n)} \tag{6.1.6}$$

and let $a \to \infty$ to obtain that $I_s(f) = 0$, a contradiction. Similarly, if $\frac{1}{p} < \frac{1}{q} + \frac{s}{n}$, we could write (6.1.5) as

$$a^{-\frac{n}{q}+\frac{n}{p}-s}\big\|I_s(f)\big\|_{L^q(\mathbf{R}^n)} \le C(p,q,n,s)\big\|f\big\|_{L^p(\mathbf{R}^n)} \tag{6.1.7}$$

and let $a \to 0$ to obtain that $\big\|f\big\|_{L^p} = \infty$, again a contradiction. It follows that (6.1.4) must necessarily hold.

We conclude that the homogeneity (or dilation structure) of an operator dictates a relationship on the indices p and q for which it (may) map L^p to L^q.

As we saw in Remark 6.1.2, if the Riesz potentials map L^p to L^q for some p,q, then we must have $q > p$. Such operators that improve the integrability of a function are called *smoothing*. The importance of the Riesz potentials lies in the fact that they are indeed smoothing operators. This is the essence of the *Hardy–Littlewood–Sobolev theorem on fractional integration*, which we now formulate and prove.

Theorem 6.1.3. *Let s be a real number with $0 < s < n$ and let $1 \le p < q < \infty$ satisfy (6.1.4). Then there exist constants $C(n,s,p) < \infty$ such that for all f in $L^p(\mathbf{R}^n)$ we have*

$$\big\|I_s(f)\big\|_{L^q} \le C(n,s,p)\big\|f\big\|_{L^p}$$

when $p > 1$, and also $\big\|I_s(f)\big\|_{L^{q,\infty}} \le C(n,s)\big\|f\big\|_{L^1}$ when $p = 1$.

We note that the $L^p \to L^{q,\infty}$ estimate in Theorem 6.1.3 is a consequence of Theorem 1.2.13, for the kernel $|x|^{-n+s}$ of I_s lies in the space $L^{r,\infty}$ when $r = \frac{n}{n-s}$, and (1.2.15) is satisfied for this r. Applying Theorem 1.4.19, we obtain the required conclusion. Nevertheless, for the sake of the exposition, we choose to give another self-contained proof of Theorem 6.1.3.

Proof. We begin by observing that the function $I_s(f)$ is well defined whenever f is bounded and has some decay at infinity. This makes the operator I_s well defined on a dense subclass of all the L^p spaces with $p < \infty$. Second, we may assume that $f \geq 0$, since $|I_s(f)| \leq I_s(|f|)$.

Under these assumptions we write the convolution

$$\int_{\mathbf{R}^n} f(x-y)|y|^{s-n}\,dy = J_1(f)(x) + J_2(f)(x),$$

where, in the spirit of interpolation, J_1 and J_2 are defined by

$$\begin{aligned} J_1(f)(x) &= \int_{|y|<R} f(x-y)|y|^{s-n}\,dy, \\ J_2(f)(x) &= \int_{|y|\geq R} f(x-y)|y|^{s-n}\,dy, \end{aligned}$$

for some R to be determined later. Observe that J_1 is given by convolution with the function $|y|^{-n+s}\chi_{|y|<R}(y)$, which is radial, integrable, and symmetrically decreasing about the origin. It follows from Theorem 2.1.10 that

$$J_1(f)(x) \leq M(f)(x) \int_{|y|<R} |y|^{-n+s}\,dy = \frac{\omega_{n-1}}{s} R^s M(f)(x), \tag{6.1.8}$$

where M is the Hardy–Littlewood maximal function. Now Hölder's inequality gives that

$$\begin{aligned} |J_2(f)(x)| &\leq \left(\int_{|y|\geq R} (|y|^{-n+s})^{p'}\,dy \right)^{\frac{1}{p'}} \|f\|_{L^p(\mathbf{R}^n)} \\ &= \left(\frac{q\omega_{n-1}}{p'n} \right)^{\frac{1}{p'}} R^{-\frac{n}{q}} \|f\|_{L^p(\mathbf{R}^n)}, \end{aligned} \tag{6.1.9}$$

and note that this estimate is also valid when $p = 1$ (in which case $q = \frac{n}{n-s}$), provided the $L^{p'}$ norm is interpreted as the L^∞ norm and the constant $\left(\frac{q\omega_{n-1}}{p'n}\right)^{\frac{1}{p'}}$ is replaced by 1. Combining (6.1.8) and (6.1.9), we obtain that

$$I_s(f)(x) \leq C'_{n,s,p} \left(R^s M(f)(x) + R^{-\frac{n}{q}} \|f\|_{L^p} \right) \tag{6.1.10}$$

for all $R > 0$. A constant multiple of the quantity

$$R = \|f\|_{L^p}^{\frac{p}{n}} \big(M(f)(x)\big)^{-\frac{p}{n}}$$

minimizes the expression on the right in (6.1.10). This choice of R yields the estimate

$$I_s(f)(x) \le C_{n,s,p} M(f)(x)^{\frac{p}{q}} \big\|f\big\|_{L^p}^{1-\frac{p}{q}}. \tag{6.1.11}$$

The required inequality for $p>1$ follows by raising to the power q, integrating over $\mathbf{R}^n$, and using the boundedness of the Hardy–Littlewood maximal operator M on $L^p(\mathbf{R}^n)$. The case $p=1$, $q=\frac{n}{n-s}$ also follows from (6.1.11) by the weak type $(1,1)$ property of M. Indeed,

$$\begin{aligned}
\big|\{C_{n,s,1}M(f)^{\frac{n-s}{n}}\big\|f\big\|_{L^1}^{\frac{s}{n}} > \lambda\}\big| &= \Big|\Big\{M(f) > \Big(\frac{\lambda}{C_{n,s,1}\big\|f\big\|_{L^1}^{\frac{s}{n}}}\Big)^{\frac{n}{n-s}}\Big\}\Big| \\
&\le 3^n\Big(\frac{C_{n,s,1}\big\|f\big\|_{L^1}^{\frac{s}{n}}}{\lambda}\Big)^{\frac{n}{n-s}}\big\|f\big\|_{L^1} \\
&= C(n,s)\Big(\frac{\big\|f\big\|_{L^1}}{\lambda}\Big)^{\frac{n}{n-s}}.
\end{aligned}$$

We now give an alternative proof of the case $p=1$ that corresponds to $q=\frac{n}{n-s}$. Without loss of generality we may assume that $f\ge 0$ has L^1 norm 1. Once this case is proved, the general case follows by scaling. Observe that

$$\int_{\mathbf{R}^n} f(x-y)|y|^{s-n}\,dy \le \sum_{j\in\mathbf{Z}} 2^{(j-1)(s-n)}\int_{|y|\le 2^j} f(x-y)\,dy. \tag{6.1.12}$$

Let $E_\lambda = \{x:\ I_s(f)(x)>\lambda\}$. Then

$$\begin{aligned}
|E_\lambda| &\le \frac{1}{\lambda}\int_{E_\lambda} I_s(f)(x)\,dx \\
&= \frac{1}{\lambda}\int_{E_\lambda}\int_{\mathbf{R}^n}|y|^{s-n}f(x-y)\,dy\,dx \\
&\le \frac{1}{\lambda}\int_{E_\lambda}\sum_{j\in\mathbf{Z}}2^{(j-1)(s-n)}\int_{|y|\le 2^j}f(x-y)\,dy\,dx \\
&= \frac{1}{\lambda}\sum_{j\in\mathbf{Z}}2^{(j-1)(s-n)}\int_{E_\lambda}\int_{|y|\le 2^j}f(x-y)\,dy\,dx \\
&\le \frac{1}{\lambda}\sum_{j\in\mathbf{Z}}2^{(j-1)(s-n)}\min(|E_\lambda|,v_n2^{jn}) \\
&= \frac{1}{\lambda}\sum_{2^j>|E_\lambda|^{\frac{1}{n}}}2^{(j-1)(s-n)}|E_\lambda| + \frac{v_n}{\lambda}\sum_{2^j\le|E_\lambda|^{\frac{1}{n}}}2^{(j-1)(s-n)}2^{jn} \\
&\le \frac{C}{\lambda}\big(|E_\lambda|^{\frac{s-n}{n}}|E_\lambda|+|E_\lambda|^{\frac{s}{n}}\big) \\
&= \frac{2C}{\lambda}|E_\lambda|^{\frac{s}{n}}.
\end{aligned} \tag{6.1.13}$$

It follows that $|E_\lambda|^{\frac{n-s}{n}} \le \frac{2C}{\lambda}$, which implies the weak type $(1, \frac{n}{n-s})$ estimate for I_s. Here C is a constant that depends on n and s. $\square$

6.1.2 Bessel Potentials

While the behavior of the kernels $|x|^{-n+s}$ as $|x| \to 0$ is well suited to their smoothing properties, their decay as $|x| \to \infty$ gets worse as s increases. We can slightly adjust the Riesz potentials so that we maintain their essential behavior near zero but achieve exponential decay at infinity. The simplest way to achieve this is by replacing the "nonnegative" operator $-\Delta$ by the "strictly positive" operator $I - \Delta$. Here the terms nonnegative and strictly positive, as one may have surmised, refer to the Fourier multipliers of these operators.

Definition 6.1.4. Let s be a complex number with $0 < \operatorname{Re} s < \infty$. The *Bessel potential* of order s is the operator

$$\mathscr{J}_s = (I - \Delta)^{-s/2},$$

whose action on functions is given by

$$\mathscr{J}_s(f) = \big(\widehat{f}\,\widehat{G_s}\big)^{\vee} = f * G_s\,,$$

where

$$G_s(x) = \big((1 + 4\pi^2|\xi|^2)^{-s/2}\big)^{\vee}(x)\,.$$

Let us see why this adjustment yields exponential decay for G_s at infinity.

Proposition 6.1.5. *Let $s > 0$. Then G_s is a smooth function on $\mathbf{R}^n \setminus \{0\}$ that satisfies $G_s(x) > 0$ for all $x \in \mathbf{R}^n$. Moreover, there exist positive finite constants $C(s,n), c(s,n), C_{s,n}$ such that*

$$G_s(x) \le C(s,n) e^{-\frac{|x|}{2}}, \qquad \textit{when } |x| \ge 2, \tag{6.1.14}$$

and such that

$$\frac{1}{c(s,n)} \le \frac{G_s(x)}{H_s(x)} \le c(s,n)\,, \qquad \textit{when } |x| \le 2,$$

where H_s is equal to

$$H_s(x) = \begin{cases} |x|^{s-n} + 1 + O(|x|^{s-n+2}) & \textit{for } 0 < s < n, \\ \log\frac{2}{|x|} + 1 + O(|x|^2) & \textit{for } s = n, \\ 1 + O(|x|^{s-n}) & \textit{for } s > n, \end{cases}$$

and $O(t)$ is a function with the property $|O(t)| \le C_{s,n}|t|$ for $0 \le t \le 4$.

Proof. For $A, s > 0$ we have the gamma function identity

$$A^{-\frac{s}{2}} = \frac{1}{\Gamma(\frac{s}{2})} \int_0^\infty e^{-tA} t^{\frac{s}{2}} \frac{dt}{t},$$

which we use to obtain

$$(1 + 4\pi^2|\xi|^2)^{-\frac{s}{2}} = \frac{1}{\Gamma(\frac{s}{2})} \int_0^\infty e^{-t} e^{-\pi|2\sqrt{\pi t}\,\xi|^2} t^{\frac{s}{2}} \frac{dt}{t}.$$

Note that the previous integral converges at both ends. Now take the inverse Fourier transform in ξ and use the fact that the function $e^{-\pi|\xi|^2}$ is equal to its Fourier transform (Example 2.2.9) to obtain

$$G_s(x) = \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_0^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}.$$

This proves that $G_s(x) > 0$ for all $x \in \mathbf{R}^n$ and that G_s is smooth on $\mathbf{R}^n \setminus \{0\}$. Now suppose $|x| \geq 2$. Then $t + \frac{|x|^2}{4t} \geq t + \frac{1}{t}$ and also $t + \frac{|x|^2}{4t} \geq |x|$. This implies that

$$-t - \frac{|x|^2}{4t} \leq -\frac{t}{2} - \frac{1}{2t} - \frac{|x|}{2},$$

from which it follows that when $|x| \geq 2$,

$$|G_s(x)| \leq \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \left(\int_0^\infty e^{-\frac{t}{2}} e^{-\frac{1}{2t}} t^{\frac{s-n}{2}} \frac{dt}{t} \right) e^{-\frac{|x|}{2}} = C_{s,n} e^{-\frac{|x|}{2}}.$$

This proves (6.1.14).

Suppose now that $|x| \leq 2$. Write $G_s(x) = G_s^1(x) + G_s^2(x) + G_s^3(x)$, where

$$\begin{aligned} G_s^1(x) &= \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_0^{|x|^2} e^{-t'} e^{-\frac{|x|^2}{4t'}} (t')^{\frac{s-n}{2}} \frac{dt'}{t'} \\ &= |x|^{s-n} \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_0^1 e^{-t|x|^2} e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \\ G_s^2(x) &= \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_{|x|^2}^4 e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}, \\ G_s^3(x) &= \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_4^\infty e^{-t} e^{-\frac{|x|^2}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t}. \end{aligned}$$

In G_s^1 we have $e^{-t|x|^2} = 1 + O(t|x|^2)$, since $t|x|^2 \leq 4$; thus we can write

$$\begin{aligned} G_s^1(x) &= |x|^{s-n} \frac{(2\sqrt{\pi})^{-n}}{\Gamma(\frac{s}{2})} \int_0^1 e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} \frac{dt}{t} + \frac{O(|x|^{s-n+2})}{\Gamma(\frac{s}{2})} \int_0^1 e^{-\frac{1}{4t}} t^{\frac{s-n}{2}} dt \\ &= c_{s,n}^1 |x|^{s-n} + O(|x|^{s-n+2}) \qquad \text{as } |x| \to 0. \end{aligned}$$

Since $0 \le \frac{|x|^2}{4t} \le \frac{1}{4}$ and $0 \le t \le 4$ in G_s^2, we have $e^{-\frac{17}{4}} \le e^{-t-\frac{|x|^2}{4t}} \le 1$; thus as $|x| \to 0$ we obtain

$$G_s^2(x) \approx \int_{|x|^2}^4 t^{\frac{s-n}{2}} \frac{dt}{t} = \begin{cases} \frac{2}{n-s}|x|^{s-n} - \frac{2^{s-n+1}}{n-s} & \text{for } s < n, \\ 2\log\frac{2}{|x|} & \text{for } s = n, \\ \frac{1}{s-n}2^{s-n+1} - \frac{2}{s-n}|x|^{s-n} & \text{for } s > n. \end{cases}$$

Finally, we have $e^{-\frac{1}{4}} \le e^{-\frac{|x|^2}{4t}} \le 1$ in G_s^3, which yields that $G_s^3(x)$ is bounded above and below by fixed positive constants. Combining the estimates for $G_s^1(x)$, $G_s^2(x)$, and $G_s^3(x)$, we obtain the required conclusion. □

We end this section with a result analogous to that of Theorem 6.1.3 for the operator $\mathscr{J}_s$.

Corollary 6.1.6. *(a) For all $0 < s < \infty$, the operator $\mathscr{J}_s$ maps $L^r(\mathbf{R}^n)$ to itself with norm 1 for all $1 \le r \le \infty$.*
(b) Let $0 < s < n$ and $1 \le p < q < \infty$ satisfy (6.1.4). Then there exist constants $C_{p,q,n,s} < \infty$ such that for all f in $L^p(\mathbf{R}^n)$ with $p > 1$ we have

$$\big\|\mathscr{J}_s(f)\big\|_{L^q} \le C_{p,q,n,s}\big\|f\big\|_{L^p}$$

and also $\big\|\mathscr{J}_s(f)\big\|_{L^{q,\infty}} \le C_{1,q,n,s}\big\|f\big\|_{L^1}$ when $p = 1$.

Proof. (a) Since $\widehat{G_s}(0) = 1$ and $G_s > 0$, it follows that G_s has L^1 norm 1. The operator $\mathscr{J}_s$ is given by convolution with the positive function G_s, which has L^1 norm 1; thus it maps $L^r(\mathbf{R}^n)$ to itself with norm 1 for all $1 \le r \le \infty$ (see Exercise 1.2.9).
(b) In the special case $0 < s < n$ we have that the kernel G_s of $\mathscr{J}_s$ satisfies

$$G_s(x) \approx \begin{cases} |x|^{-n+s} & \text{when } |x| \le 2, \\ e^{-\frac{|x|}{2}} & \text{when } |x| \ge 2. \end{cases}$$

Then we can write

$$\begin{aligned} \mathscr{J}_s(f)(x) &\le C_{n,s}\left[\int_{|y|\le 2} |f(x-y)|\,|y|^{-n+s}\,dy + \int_{|y|\ge 2} |f(x-y)|\,e^{-\frac{|y|}{2}}\,dy\right] \\ &\le C_{n,s}\left[I_s(|f|)(x) + \int_{\mathbf{R}^n} |f(x-y)|\,e^{-\frac{|y|}{2}}\,dy\right]. \end{aligned}$$

We now use that the function $y \mapsto e^{-|y|/2}$ is in L^r for all $r < \infty$, Theorem 1.2.12 (Young's inequality), and Theorem 6.1.3 to complete the proof of the corollary. □

Exercises

6.1.1. (a) Let $0 < s,t < \infty$ be such that $s+t < n$. Show that $I_s I_t = I_{s+t}$.
(b) Prove the operator identities

$$I_s(-\Delta)^z = (-\Delta)^z I_s = I_{s-2z} = (-\Delta)^{z-\frac{s}{2}}$$

whenever $\operatorname{Re} s > 2\operatorname{Re} z$.
(c) Prove that for all $z \in \mathbf{C}$ we have

$$\big\langle (-\Delta)^z(f) \,|\, (-\Delta)^{-z}(g) \big\rangle = \big\langle f \,|\, g \big\rangle$$

whenever the Fourier transforms of f and g vanish to sufficiently high order at the origin.
(d) Given $\operatorname{Re} s > 0$, find an $\alpha \in \mathbf{C}$ such that the identity

$$\big\langle I_s(f) \,|\, f \big\rangle = \big\| (-\Delta)^{\alpha}(f) \big\|_{L^2}^2$$

is valid for all functions f as in part (c).

6.1.2. Use Exercise 2.2.14 to prove that for $-\infty < \alpha < n/2 < \beta < \infty$ we have

$$\big\| f \big\|_{L^\infty(\mathbf{R}^n)} \le C \big\| \Delta^{\alpha/2}(f) \big\|_{L^2(\mathbf{R}^n)}^{\frac{\beta-n/2}{\beta-\alpha}} \big\| \Delta^{\beta/2}(f) \big\|_{L^2(\mathbf{R}^n)}^{\frac{n/2-\alpha}{\beta-\alpha}},$$

where C depends only on α, n, β.

6.1.3. Show that when $0 < s < n$ we have

$$\sup_{\|f\|_{L^1(\mathbf{R}^n)}=1} \big\| I_s(f) \big\|_{L^{\frac{n}{n-s}}(\mathbf{R}^n)} = \sup_{\|f\|_{L^1(\mathbf{R}^n)}=1} \big\| \mathscr{J}_s(f) \big\|_{L^{\frac{n}{n-s}}(\mathbf{R}^n)} = \infty .$$

Thus I_s and $\mathscr{J}_s$ are not of strong type $(1, \frac{n}{n-s})$.
$\big[$*Hint:* Consider an approximate identity.$\big]$

6.1.4. Let $0 < s < n$. Consider the function $h(x) = |x|^{-s}(\log \frac{1}{|x|})^{-\frac{s}{n}(1+\delta)}$ for $|x| \le 1/e$ and zero otherwise. Prove that when $0 < \delta < \frac{n-s}{s}$ we have $h \in L^{\frac{n}{s}}(\mathbf{R}^n)$ but that $\lim_{x\to 0} I_s(h)(x) = \infty$. Conclude that I_s does not map $L^{\frac{n}{s}}(\mathbf{R}^n)$ to $L^\infty(\mathbf{R}^n)$.

6.1.5. For $1 \le p \le \infty$ and $0 < s < \infty$ define the *Bessel potential space* $\mathscr{L}_s^p(\mathbf{R}^n)$ as the space of all functions $f \in L^p(\mathbf{R}^n)$ for which there exists another function f_0 in $L^p(\mathbf{R}^n)$ such that $\mathscr{J}_s(f_0) = f$. Define a norm on these spaces by setting $\big\| f \big\|_{\mathscr{L}_s^p} = \big\| f_0 \big\|_{L^p}$. Prove the following properties of these spaces:
(a) $\big\| f \big\|_{L^p} \le \big\| f \big\|_{\mathscr{L}_s^p}$; hence $\mathscr{L}_s^p(\mathbf{R}^n)$ is a subspace of $L^p(\mathbf{R}^n)$.
(b) For all $0 < t, s < \infty$ we have $G_s * G_t = G_{s+t}$ and thus

$$\mathscr{L}_s^p(\mathbf{R}^n) * \mathscr{L}_t^q(\mathbf{R}^n) \subseteq \mathscr{L}_{s+t}^r(\mathbf{R}^n),$$

where $1 \le p,q,r \le \infty$ and $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}+1$.
(c) The sequence of norms $\|f\|_{\mathscr{L}_s^p}$ increases, and therefore the spaces $\mathscr{L}_s^p(\mathbf{R}^n)$ decrease as s increases.
(d) The map $\mathscr{I}_t$ is an isomorphism from the space $\mathscr{L}_s^p(\mathbf{R}^n)$ onto $\mathscr{L}_{s+t}^p(\mathbf{R}^n)$.
[Note: Note that the Bessel potential space $\mathscr{L}_s^p(\mathbf{R}^n)$ coincides with the Sobolev space $L_s^p(\mathbf{R}^n)$, introduced in Section 6.2.]

6.1.6. For $0 \le s < n$ define the *fractional maximal function*

$$M^s(f)(x) = \sup_{t>0} \frac{1}{(v_n t^n)^{\frac{n-s}{n}}} \int_{|y|\le t} |f(x-y)|\,dy,$$

where v_n is the volume of the unit ball in $\mathbf{R}^n$.
(a) Show that for some constant C we have

$$M^s(f) \le C I_s(f)$$

for all $f \ge 0$ and conclude that M^s maps L^p to L^q whenever I_s does.
(b) (*Adams [1]*) Let $s>0$, $1<p<\frac{n}{s}$, $1 \le q \le \infty$ be such that $\frac{1}{r}=\frac{1}{p}-\frac{s}{n}+\frac{sp}{nq}$. Show that there is a constant $C>0$ (depending on the previous parameters) such that for all positive functions f we have

$$\big\|I_s(f)\big\|_{L^r} \le C\big\|M^{n/p}(f)\big\|_{L^q}^{\frac{sp}{n}} \big\|f\big\|_{L^p}^{1-\frac{sp}{n}}.$$

[*Hint:* For $f \ne 0$, write $I_s(f) = I_1 + I_2$, where

$$I_1 = \int_{|x-y|\le\delta} f(y)\,|y|^{s-n}\,dy, \qquad I_2 = \int_{|x-y|>\delta} f(y)\,|y|^{s-n}\,dy.$$

Show that $I_1 \le C\delta^s M^0(f)$ and that $I_2(f) \le C\delta^{s-\frac{n}{p}} M^{n/p}(f)$. Optimize over $\delta>0$ to obtain

$$I_s(f) \le C M^{n/p}(f)^{\frac{sp}{n}} M^0(f)^{1-\frac{sp}{n}},$$

from which the required conclusion follows easily.]

6.1.7. Suppose that a function K defined on $\mathbf{R}^n$ satisfies $|K(y)| \le C(1+|y|)^{-s+n-\varepsilon}$, where $0<s<n$ and $0<C,\varepsilon<\infty$. Prove that the maximal operator

$$\sup_{t>0} t^{-n+s}\left|\int_{\mathbf{R}^n} f(x-y)K(y/t)\,dy\right|$$

maps $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$ whenever I_s maps $L^p(\mathbf{R}^n)$ to $L^q(\mathbf{R}^n)$.
[*Hint:* Control this operator by the maximal function M^s of Exercise 6.1.6.]

6.1.8. Let $0<s<n$. Use the following steps to obtain a simpler proof of Theorem 6.1.3 based on more delicate interpolation.
(a) Prove that $\|I_s(\chi_E)\|_{L^\infty} \le |E|^{\frac{s}{n}}$ for any set E of finite measure.
(b) For any two sets E and F of finite measure show that

$$\int_F |I_s(\chi_E)(x)|\,dx \le |E|\,|F|^{\frac{s}{n}}\,.$$

(c) Use Exercise 1.1.12 to obtain that

$$\big\|I_s(\chi_E)\big\|_{L^{\frac{n}{n-s},\infty}} \le C_{ns}|E|\,.$$

(d) Use parts (a), (c), and Theorem 1.4.19 to obtain another proof of Theorem 6.1.3.
[*Hint:* Parts (a) and (b): Use that when $\lambda > 0$, the integral $\int_E |y|^{-\lambda}\,dy$ becomes largest when E is a ball centered at the origin equimeasurable to E.]

6.1.9. (*Welland [329]*) Let $0 < \alpha < n$ and suppose $0 < \varepsilon < \min(\alpha, n-\alpha)$. Show that there exists a constant depending only on α, ε, and n such that for all compactly supported bounded functions f we have

$$|I_\alpha(f)| \le C\sqrt{M^{\alpha-\varepsilon}(f)M^{\alpha+\varepsilon}(f)}\,,$$

where $M^\beta(f)$ is the fractional maximal function of Exercise 6.1.6.
[*Hint:* Write

$$|I_\alpha(f)| \le \int_{|x-y|<s} \frac{|f(y)|\,dy}{|x-y|^{n-\alpha}} + \int_{|x-y|\ge s} \frac{|f(y)|\,dy}{|x-y|^{n-\alpha}}$$

and split each integral into a sum of integrals over annuli centered at x to obtain the estimate

$$|I_\alpha(f)| \le C\big(s^\varepsilon M^{\alpha-\varepsilon}(f) + s^{-\varepsilon}M^{\alpha+\varepsilon}(f)\big)\,.$$

Then optimize over s.]

6.1.10. Show that the *discrete fractional integral operator*

$$\{a_j\}_{j\in\mathbf{Z}^n} \to \Big\{\sum_{k\in\mathbf{Z}^n} \frac{a_k}{(|j-k|+1)^{n-\alpha}}\Big\}_{j\in\mathbf{Z}^n}$$

maps $\ell^s(\mathbf{Z}^n)$ to $\ell^t(\mathbf{Z}^n)$ when $0 < \alpha < n$, $1 < s < t$, and

$$\frac{1}{s} - \frac{1}{t} = \frac{\alpha}{n}\,.$$

6.1.11. Show that the bilinear operator

$$B_\alpha(f,g)(x) = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} f(y)g(z)(|x-y|+|x-z|)^{-2n+\alpha}\,dy\,dz$$

maps $L^p(\mathbf{R}^n)\times L^q(\mathbf{R}^n)$ to $L^r(\mathbf{R}^n)$ when $1 < p,q < \infty$ and

$$\frac{1}{p} + \frac{1}{q} = \frac{\alpha}{n} + \frac{1}{r}\,.$$

[*Hint:* Control $B_\alpha(f,g)$ by the product of two fractional integrals.]

6.1.12. (*Grafakos and Kalton [148]/Kenig and Stein [189]*) (a) Prove that the bilinear operator

$$S(f,g)(x) = \int_{|t|\le 1} |f(x+t)g(x-t)|\,dt$$

maps $L^1(\mathbf{R}^n) \times L^1(\mathbf{R}^n)$ to $L^{\frac{1}{2}}(\mathbf{R}^n)$.
(b) For $0 < \alpha < n$ prove that the *bilinear fractional integral operator*

$$I_\alpha(f,g)(x) = \int_{\mathbf{R}^n} f(x+t)g(x-t)|t|^{-n+\alpha}\,dt$$

maps $L^1(\mathbf{R}^n) \times L^1(\mathbf{R}^n)$ to $L^{\frac{n}{2n-\alpha},\infty}(\mathbf{R}^n)$.
$\big[$*Hint:* Part (a): Write $f = \sum_{k\in\mathbf{Z}^n} f_k$, where each f_k is supported in the cube $k+[0,1]^n$ and similarly for g. Observe that the resulting double sum reduces to a single sum and use that $(\sum_j a_j)^{1/2} \le \sum_j a_j^{1/2}$ for $a_j \ge 0$. Part (b): Use part (a) and adjust the argument in (6.1.13) to a bilinear setting.$\big]$

6.2 Sobolev Spaces

In this section we study a quantitative way of measuring smoothness of functions. Sobolev spaces serve exactly this purpose. They measure the smoothness of a given function in terms of the integrability of its derivatives. We begin with the classical definition of Sobolev spaces.

Definition 6.2.1. Let k be a nonnegative integer and let $1 < p < \infty$. The *Sobolev space* $L^p_k(\mathbf{R}^n)$ is defined as the space of functions f in $L^p(\mathbf{R}^n)$ all of whose distributional derivatives $\partial^\alpha f$ are also in $L^p(\mathbf{R}^n)$ for all multi-indices α that satisfy $|\alpha| \le k$. This space is normed by the expression

$$\big\|f\big\|_{L^p_k} = \sum_{|\alpha|\le k} \big\|\partial^\alpha f\big\|_{L^p}, \tag{6.2.1}$$

where $\partial^{(0,\dots,0)} f = f$.

Sobolev spaces measure smoothness of functions. The index k indicates the "degree" of smoothness of a given function in L^p_k. As k increases the functions become smoother. Equivalently, these spaces form a decreasing sequence

$$L^p \supset L^p_1 \supset L^p_2 \supset L^p_3 \supset \cdots,$$

meaning that each $L^p_{k+1}(\mathbf{R}^n)$ is a subspace of $L^p_k(\mathbf{R}^n)$. This property, which coincides with our intuition of smoothness, is a consequence of the definition of the Sobolev norms.

We next observe that the space $L^p_k(\mathbf{R}^n)$ is complete. Indeed, if f_j is a Cauchy sequence in the norm given by (6.2.1), then $\{\partial^\alpha f_j\}_j$ are Cauchy sequences for all

$|\alpha| \le k$. By the completeness of L^p, there exist functions f_α such that $\partial^\alpha f_j \to f_\alpha$ in L^p. This implies that for all φ in the Schwartz class we have

$$(-1)^{|\alpha|} \int_{\mathbf{R}^n} f_j \, (\partial^\alpha \varphi) \, dx = \int_{\mathbf{R}^n} (\partial^\alpha f_j) \, \varphi \, dx \to \int_{\mathbf{R}^n} f_\alpha \, \varphi \, dx.$$

Since the first expression converges to

$$(-1)^{|\alpha|} \int_{\mathbf{R}^n} f_0 \, (\partial^\alpha \varphi) \, dx \, ,$$

it follows that the distributional derivative $\partial^\alpha f_0$ is f_α. This implies that $f_j \to f_0$ in $L^p_k(\mathbf{R}^n)$ and proves the completeness of this space.

Our goal in this section is to investigate relations between these spaces and the Riesz and Bessel potentials discussed in the previous section and to obtain a Littlewood–Paley characterization of them. Before we embark on this study, we note that we can extend the definition of Sobolev spaces to the case in which the index k is not necessarily an integer. In fact, we extend the definition of the spaces $L^p_k(\mathbf{R}^n)$ to the case in which the number k is real.

6.2.1 *Definition and Basic Properties of General Sobolev Spaces*

Definition 6.2.2. Let s be a real number and let $1 < p < \infty$. The *inhomogeneous Sobolev space* $L^p_s(\mathbf{R}^n)$ is defined as the space of all tempered distributions u in $\mathscr{S}'(\mathbf{R}^n)$ with the property that

$$((1+|\xi|^2)^{\frac{s}{2}} \widehat{u})^\vee \tag{6.2.2}$$

is an element of $L^p(\mathbf{R}^n)$. For such distributions u we define

$$\big\|u\big\|_{L^p_s} = \big\|((1+|\cdot|^2)^{\frac{s}{2}} \, \widehat{u})^\vee \big\|_{L^p(\mathbf{R}^n)} \, .$$

Note that the function $(1+|\xi|^2)^{\frac{s}{2}}$ is $\mathscr{C}^\infty$ and has at most polynomial growth at infinity. Since $\widehat{u} \in \mathscr{S}'(\mathbf{R}^n)$, the product in (6.2.2) is well defined.

Several observations are in order. First, we note that when $s = 0$, $L^p_s = L^p$. It is natural to ask whether elements of L^p_s are always L^p functions. We show that this is the case when $s \ge 0$ but not when $s < 0$. We also show that the space L^p_s coincides with the space L^p_k given in Definition 6.2.1 when $s = k$ and k is an integer.

To prove that elements of L^p_s are indeed L^p functions when $s \ge 0$, we simply note that if $f_s = ((1+|\xi|^2)^{s/2} \widehat{f})^\vee$, then

$$f = \big(\widehat{f_s}(\xi) \widehat{G_s}(\xi/2\pi)\big)^\vee = f_s * (2\pi)^n \, G_s(2\pi(\cdot)) \, ,$$

where G_s is given in Definition 6.1.4. Thus a certain dilation of f can be expressed as the Bessel potential of itself; hence Corollary 6.1.6 yields that

$$c^{-1}\big\|f\big\|_{L^p} \le \big\|f_s\big\|_{L^p} = \big\|f\big\|_{L^p_s},$$

for some constant c.

We now prove that if $s = k$ is a nonnegative integer and $1 < p < \infty$, then the norm of the space L^p_k as given in Definition 6.2.1 is comparable to that in Definition 6.2.2. Suppose that $f \in L^p_k$ according to Definition 6.2.2. Then for all $|\alpha| \le k$ we have

$$\partial^\alpha f = c_\alpha(\widehat{f}(\xi)\xi^\alpha)^\vee = c_\alpha\bigg(\widehat{f}(\xi)(1+|\xi|^2)^{\frac{k}{2}}\frac{\xi^\alpha}{(1+|\xi|^2)^{\frac{k}{2}}}\bigg)^\vee. \tag{6.2.3}$$

Theorem 5.2.7 gives that the function

$$\frac{\xi^\alpha}{(1+|\xi|^2)^{k/2}}$$

is an L^p multiplier. Since by assumption $\big(\widehat{f}(\xi)(1+|\xi|^2)^{\frac{k}{2}}\big)^\vee$ is in $L^p(\mathbf{R}^n)$, it follows from (6.2.3) that $\partial^\alpha f$ is in L^p and also that

$$\sum_{|\alpha|\le k}\big\|\partial^\alpha f\big\|_{L^p} \le C_{p,n,k}\big\|((1+|\cdot|^2)^{\frac{k}{2}}\widehat{f})^\vee\big\|_{L^p}.$$

Conversely, suppose that $f \in L^p_k$ according to Definition 6.2.1; then

$$(1+\xi_1^2+\cdots+\xi_n^2)^{\frac{k}{2}} = \sum_{|\alpha|\le k}\frac{k!}{\alpha_1!\cdots\alpha_n!(k-|\alpha|)!}\,\xi^\alpha\,\frac{\xi^\alpha}{(1+|\xi|^2)^{\frac{k}{2}}}.$$

As we have already observed, the functions $m_\alpha(\xi) = \xi^\alpha(1+|\xi|^2)^{-\frac{k}{2}}$ are L^p multipliers whenever $|\alpha| \le k$. Since

$$\big((1+|\xi|^2)^{\frac{k}{2}}\widehat{f}\big)^\vee = \sum_{|\alpha|\le k}c_{\alpha,k}\big(m_\alpha(\xi)\xi^\alpha\widehat{f}\big)^\vee = \sum_{|\alpha|\le k}c'_{\alpha,k}\big(m_\alpha(\xi)\widehat{\partial^\alpha f}\big)^\vee,$$

it follows that

$$\big\|(\widehat{f}(\xi)(1+|\xi|^2)^{\frac{k}{2}})^\vee\big\|_{L^p} \le C_{p,n,k}\sum_{|\gamma|\le k}\big\|(\widehat{f}(\xi)\xi^\gamma)^\vee\big\|_{L^p}.$$

Example 6.2.3. Every Schwartz function lies in $L^p_s(\mathbf{R}^n)$ for s real. Sobolev spaces with negative indices s can indeed contain tempered distributions that are not locally integrable functions. For example, Dirac mass at the origin δ_0 is an element of $L^p_{-s}(\mathbf{R}^n)$ for all $s > n/p'$. Indeed, when $0 < s < n$, Proposition 6.1.5 gives that G_s [i.e., the inverse Fourier transform of $(1+|\xi|^2)^{-\frac{s}{2}}$] is integrable to the power p as long as $(s-n)p > -n$ (i.e., $s > n/p'$). When $s \geq n$, G_s is integrable to any positive power.

We now continue with the *Sobolev embedding theorem.*

Theorem 6.2.4. *(a) Let $0 < s < \frac{n}{p}$ and $1 < p < \infty$. Then the Sobolev space $L^p_s(\mathbf{R}^n)$ continuously embeds in $L^q(\mathbf{R}^n)$ when*

$$\frac{1}{p} - \frac{1}{q} = \frac{s}{n}.$$

(b) Let $0 < s = \frac{n}{p}$ and $1 < p < \infty$. Then $L^p_s(\mathbf{R}^n)$ continuously embeds in $L^q(\mathbf{R}^n)$ for any $\frac{n}{s} < q < \infty$.
(c) Let $\frac{n}{p} < s < \infty$ and $1 < p < \infty$. Then every element of $L^p_s(\mathbf{R}^n)$ can be modified on a set of measure zero so that the resulting function is bounded and uniformly continuous.

Proof. (a) If $f \in L^p_s$, then $f_s(x) = ((1+|\xi|^2)^{\frac{s}{2}} \widehat{f}\,)^\vee(x)$ is in $L^p(\mathbf{R}^n)$. Thus

$$f(x) = ((1+|\xi|^2)^{-\frac{s}{2}} \widehat{f_s}\,)^\vee(x);$$

hence $f = G_s * f_s$. Since $s < n$, Proposition 6.1.5 gives that

$$|G_s(x)| \leq C_{s,n} |x|^{s-n}$$

for all $x \in \mathbf{R}^n$. This implies that $|f| = |G_s * f_s| \leq C_{s,n} I_s(|f_s|)$. Theorem 6.1.3 now yields the required conclusion

$$\big\|f\big\|_{L^q} \leq C'_{s,n} \big\|I_s(|f_s|)\big\|_{L^q} \leq C''_{s,n} \big\|f\big\|_{L^p_s}.$$

(b) Given any $\frac{n}{s} < q < \infty$ we can find $t > 1$ such that

$$1 + \frac{1}{q} = \frac{s}{n} + \frac{1}{t} = \frac{1}{p} + \frac{1}{t}.$$

Then $1 < \frac{s}{n} + \frac{1}{t}$, which implies that $(-n+s)t > -n$. Thus the function $|x|^{-n+s}\chi_{|x|\leq 2}$ is integrable to the tth power, which implies that G_s is in L^t. Since $f = G_s * f_s$, Young's inequality gives that

$$\big\|f\big\|_{L^q(\mathbf{R}^n)} \leq \big\|f_s\big\|_{L^p(\mathbf{R}^n)} \big\|G_s\big\|_{L^t(\mathbf{R}^n)} = C_{n,s} \big\|f\big\|_{L^p_{n/p}}.$$

(c) As before, $f = G_s * f_s$. If $s \geq n$, then Proposition 6.1.5 gives that the function G_s is in $L^{p'}(\mathbf{R}^n)$. Now if $n > s$, then $G_s(x)$ looks like $|x|^{-n+s}$ near zero. This function is integrable to the power p' near the origin if and only if $s > n/p$, which is what we are assuming. Thus f is given as the convolution of an L^p function and an $L^{p'}$ function, and hence it is bounded and can be identified with a uniformly continuous function (cf. Exercise 1.2.3). □

We now introduce the homogeneous Sobolev spaces $\dot{L}^p_s$. The main difference with the inhomogeneous spaces L^p_s is that elements of $\dot{L}^p_s$ may not themselves be elements of L^p. Another difference is that elements of homogeneous Sobolev spaces are not tempered distributions but equivalence classes of tempered distributions.

We would expect the homogeneous Sobolev space $\dot{L}^p_s$ to be the space of all distributions u in $\mathscr{S}'(\mathbf{R}^n)$ for which the expression

$$(|\xi|^s \widehat{u})^\vee \tag{6.2.4}$$

is an L^p function. Since the function $|\xi|^s$ is not (always) smooth at the origin, some care is needed in defining the product in (6.2.4). The idea is that when u lies in $\mathscr{S}'/\mathscr{P}$, then the value of $\widehat{u}$ at the origin is irrelevant, since we may add to $\widehat{u}$ a distribution supported at the origin and obtain another element of the equivalence class of u (Proposition 2.4.1). It is because of this irrelevance that we are allowed to multiply $\widehat{u}$ by a function that may be nonsmooth at the origin (and which has polynomial growth at infinity).

To do this, we fix a smooth function $\eta(\xi)$ on $\mathbf{R}^n$ that is equal to 1 when $|\xi| \geq 2$ and vanishes when $|\xi| \leq 1$. Then for $s \in \mathbf{R}$, $u \in \mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$, and $\varphi \in \mathscr{S}(\mathbf{R}^n)$ we define

$$\big\langle |\xi|^s \widehat{u}, \varphi \big\rangle = \lim_{\varepsilon \to 0} \big\langle \widehat{u}, \eta(\tfrac{\xi}{\varepsilon}) |\xi|^s \varphi(\xi) \big\rangle ,$$

provided that the last limit exists. Note that this defines $|\xi|^s \widehat{u}$ as another element of $\mathscr{S}'/\mathscr{P}$, and this definition is independent of the function η, as follows easily from (2.3.23).

Definition 6.2.5. Let s be a real number and let $1 < p < \infty$. The *homogeneous Sobolev space* $\dot{L}^p_s(\mathbf{R}^n)$ is defined as the space of all tempered distributions modulo polynomials u in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ for which the expression

$$(|\xi|^s \widehat{u})^\vee$$

exists and is an $L^p(\mathbf{R}^n)$ function. For distributions u in $\dot{L}^p_s(\mathbf{R}^n)$ we define

$$\big\|u\big\|_{\dot{L}^p_s} = \big\|(|\cdot|^s \widehat{u})^\vee\big\|_{L^p(\mathbf{R}^n)} . \tag{6.2.5}$$

As noted earlier, to avoid working with equivalence classes of functions, we identify two distributions in $\dot{L}^p_s(\mathbf{R}^n)$ whose difference is a polynomial. In view of this identification, the quantity in (6.2.5) is a norm.

6.2.2 *Littlewood–Paley Characterization of Inhomogeneous Sobolev Spaces*

We now present the first main result of this section, the characterization of the inhomogeneous Sobolev spaces using Littlewood–Paley theory.

For the purposes of the next theorem we need the following setup. We fix a radial Schwartz function Ψ on $\mathbf{R}^n$ whose Fourier transform is nonnegative, supported in the annulus $1-\frac{1}{7} \le |\xi| \le 2$, equal to 1 on the smaller annulus $1 \le |\xi| \le 2-\frac{2}{7}$, and satisfies $\widehat{\Psi}(\xi)+\widehat{\Psi}(\xi/2)=1$ on the annulus $1 \le |\xi| \le 4-\frac{4}{7}$. This function has the property

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1 \tag{6.2.6}$$

for all $\xi \neq 0$. We define the associated Littlewood–Paley operators Δ_j given by multiplication on the Fourier transform side by the function $\widehat{\Psi}(2^{-j}\xi)$, that is,

$$\Delta_j(f) = \Delta_j^{\Psi}(f) = \Psi_{2^{-j}} * f. \tag{6.2.7}$$

Notice that the support properties of the Δ_j's yield the simple identity

$$\Delta_j = \big(\Delta_{j-1} + \Delta_j + \Delta_{j+1}\big)\Delta_j$$

for all $j \in \mathbf{Z}$. We also define a Schwartz function Φ so that

$$\widehat{\Phi}(\xi) = \begin{cases} \sum_{j\le 0} \widehat{\Psi}(2^{-j}\xi) & \text{when } \xi \neq 0, \\ 1 & \text{when } \xi = 0. \end{cases} \tag{6.2.8}$$

Note that $\widehat{\Phi}(\xi)$ is equal to 1 for $|\xi| \le 2-\frac{2}{7}$, vanishes when $|\xi| \ge 2$, and satisfies

$$\widehat{\Phi}(\xi) + \sum_{j=1}^{\infty} \widehat{\Psi}(2^{-j}\xi) = 1 \tag{6.2.9}$$

for all ξ in $\mathbf{R}^n$. We now introduce an operator S_0 by setting

$$S_0(f) = \Phi * f. \tag{6.2.10}$$

Identity (6.2.9) yields the operator identity

$$S_0 + \sum_{j=1}^{\infty} \Delta_j = I,$$

in which the series converges in $\mathscr{S}'(\mathbf{R}^n)$; see Exercise 2.3.12. (Note that $S_0(f)$ and $\Delta_j(f)$ are well defined functions when f is a tempered distribution.)

Having introduced the relevant background, we are now ready to state and prove the following result.

Theorem 6.2.6. *Let Φ, Ψ satisfy (6.2.6) and (6.2.8) and let Δ_j, S_0 be as in (6.2.7) and (6.2.10). Fix $s \in \mathbf{R}$ and all $1 < p < \infty$. Then there exists a constant C_1 that depends only on n, s, p, Φ, and Ψ such that for all $f \in L^p_s$ we have*

$$\big\|S_0(f)\big\|_{L^p} + \Big\|\Big(\sum_{j=1}^{\infty} (2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \le C_1 \big\|f\big\|_{L^p_s}. \tag{6.2.11}$$

Conversely, there exists a constant C_2 that depends on the parameters n,s,p, Φ, and Ψ such that every tempered distribution f that satisfies

$$\big\|S_0(f)\big\|_{L^p}+\Big\|\Big(\sum_{j=1}^{\infty}(2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p}<\infty$$

is an element of the Sobolev space L^p_s with norm

$$\big\|f\big\|_{L^p_s}\le C_2\Big(\big\|S_0(f)\big\|_{L^p}+\Big\|\Big(\sum_{j=1}^{\infty}(2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p}\Big). \tag{6.2.12}$$

Proof. We denote by C a generic constant that depends on the parameters n,s,p,Φ, and Ψ and that may vary in different occurrences. For a given tempered distribution f we define another tempered distribution f_s by setting

$$f_s=\big((1+|\cdot|^2)^{\frac{s}{2}}\widehat{f}\big)^{\vee},$$

so that we have $\big\|f\big\|_{L^p_s}=\big\|f_s\big\|_{L^p}$ if $f\in L^p_s$.

We first assume that the expression on the right in (6.2.12) is finite and we show that the tempered distribution f lies in the space L^p_s by controlling the L^p norm of f_s by a multiple of this expression. We begin by writing

$$f_s=\big(\widehat{\Phi}\widehat{f_s}\big)^{\vee}+\big((1-\widehat{\Phi})\widehat{f_s}\big)^{\vee},$$

and we plan to show that both quantities on the right are in L^p. Pick a smooth function with compact support η_0 that is equal to 1 on the support of $\widehat{\Phi}$. It is a simple fact that for all $s\in\mathbf{R}$ the function $(1+|\xi|^2)^{\frac{s}{2}}\eta_0(\xi)$ is in $\mathscr{M}_p(\mathbf{R}^n)$ (i.e., it is an L^p Fourier multiplier). Since

$$\big(\widehat{\Phi}\widehat{f_s}\big)^{\vee}(x)=\big\{\big((1+|\xi|^2)^{\frac{s}{2}}\eta_0(\xi)\big)\widehat{S_0(f)}(\xi)\big\}^{\vee}(x), \tag{6.2.13}$$

we have the estimate

$$\big\|\big(\widehat{\Phi}\widehat{f_s}\big)^{\vee}\big\|_{L^p}\le C\|S_0(f)\|_{L^p}. \tag{6.2.14}$$

We now introduce a smooth function η_∞ that vanishes in a neighborhood of the origin and is equal to 1 on the support of $1-\widehat{\Phi}$. Using Theorem 5.2.7, we can easily see that the function

$$\frac{(1+|\xi|^2)^{\frac{s}{2}}}{|\xi|^s}\eta_\infty(\xi)$$

is in $\mathscr{M}_p(\mathbf{R}^n)$ (with constant depending on n, p, η_∞, and s). Since

$$\big((1+|\xi|^2)^{\frac{s}{2}}(1-\widehat{\Phi}(\xi))\widehat{f}\big)^{\vee}(x)=\Big(\frac{(1+|\xi|^2)^{\frac{s}{2}}\eta_\infty(\xi)}{|\xi|^s}\,|\xi|^s(1-\widehat{\Phi}(\xi))\widehat{f}\Big)^{\vee}(x),$$

we obtain the estimate

$$\left\| \left((1 - \widehat{\Phi}) \widehat{f_s} \right)^{\vee} \right\|_{L^p} \le C \left\| f_\infty \right\|_{L^p}, \tag{6.2.15}$$

where f_∞ is another tempered distribution defined via

$$f_\infty = \left(|\xi|^s (1 - \widehat{\Phi}(\xi)) \widehat{f} \right)^{\vee}.$$

We are going to show that the quantity $\|f_\infty\|_{L^p}$ is finite using Littlewood–Paley theory. To achieve this, we introduce a smooth bump $\widehat{\zeta}$ supported in the annulus $\frac{1}{2} \le |\xi| \le 4$ and equal to 1 on the support of $\widehat{\Psi}$. Then we define $\widehat{\theta}(\xi) = |\xi|^s \widehat{\zeta}(\xi)$ and we introduce Littlewood–Paley operators

$$\Delta_j^{\theta}(g) = g * \theta_{2^{-j}},$$

where $\theta_{2^{-j}}(t) = 2^{jn}\theta(2^j t)$. Recalling that

$$1 - \widehat{\Phi}(\xi) = \sum_{k \ge 1} \widehat{\Psi}(2^{-k}\xi),$$

we obtain that

$$\widehat{f_\infty} = \sum_{j=1}^{\infty} |\xi|^s \widehat{\Psi}(2^{-j}\xi) \widehat{\zeta}(2^{-j}\xi) \widehat{f} = \sum_{j=1}^{\infty} 2^{js} \widehat{\Psi}(2^{-j}\xi) \widehat{\theta}(2^{-j}\xi) \widehat{f}$$

and hence

$$f_\infty = \sum_{j=1}^{\infty} \Delta_j^{\theta}(2^{js}\Delta_j(f)).$$

Using estimate (5.1.20), we obtain

$$\left\| f_\infty \right\|_{L^p} \le C \left\| \Big(\sum_{j=1}^{\infty} |2^{js}\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \right\|_{L^p} < \infty. \tag{6.2.16}$$

Combining (6.2.14), (6.2.15), and (6.2.16), we deduce the estimate in (6.2.12). (Incidentally, this argument shows that f_∞ is a function.)

To obtain the converse inequality (6.2.11) we essentially have to reverse our steps. Here we assume that $f \in L^p_s$ and we show the validity of (6.2.11). First, we have the estimate

$$\left\| S_0(f) \right\|_{L^p} \le C \left\| f_s \right\|_{L^p} = C \left\| f \right\|_{L^p_s}, \tag{6.2.17}$$

since we can obtain the Fourier transform of $S_0(f) = \Phi * f$ by multiplying $\widehat{f_s}$ by the L^p Fourier multiplier $(1 + |\xi|^2)^{-\frac{s}{2}} \widehat{\Phi}(\xi)$. Second, setting $\widehat{\sigma}(\xi) = |\xi|^{-s}\widehat{\Psi}(\xi)$ and letting Δ_j^{σ} be the Littlewood–Paley operator associated with the bump $\widehat{\sigma}(2^{-j}\xi)$, we have

$$2^{js}\widehat{\Psi}(2^{-j}\xi)\widehat{f} = \widehat{\sigma}(2^{-j}\xi)|\xi|^s\widehat{f} = \widehat{\sigma}(2^{-j}\xi)|\xi|^s(1 - \widehat{\Phi}(\xi))\widehat{f},$$

when $j \ge 2$ [since $\widehat{\Phi}$ vanishes on the support of $\widehat{\sigma}(2^{-j}\xi)$ when $j \ge 2$]. This yields the operator identity

$$2^{js}\Delta_j(f) = \Delta_j^{\sigma}(f_\infty)\,. \tag{6.2.18}$$

Using identity (6.2.18) we obtain

$$\Big\|\Big(\sum_{j=2}^{\infty}|2^{js}\Delta_j(f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} = \Big\|\Big(\sum_{j=2}^{\infty}|\Delta_j^{\sigma}(f_\infty)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \le C\big\|f_\infty\big\|_{L^p}\,, \tag{6.2.19}$$

where the last inequality follows by Theorem 5.1.2. Notice that

$$f_\infty = \big(|\xi|^s(1-\widehat{\Phi}(\xi))\,\widehat{f}\,\big)^{\vee} = \Big(\frac{|\xi|^s(1-\widehat{\Phi}(\xi))}{(1+|\xi|^2)^{\frac{s}{2}}}\,\widehat{f_s}\Big)^{\vee},$$

and since the function $|\xi|^s(1-\widehat{\Phi}(\xi))(1+|\xi|^2)^{-\frac{s}{2}}$ is in $\mathscr{M}_p(\mathbf{R}^n)$ by Theorem 5.2.7, it follows that

$$\big\|f_\infty\big\|_{L^p} \le C\big\|f_s\big\|_{L^p} = C\big\|f\big\|_{L^p_s}\,,$$

which combined with (6.2.19) yields

$$\Big\|\Big(\sum_{j=2}^{\infty}|2^{js}\Delta_j(f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \le C\big\|f\big\|_{L^p_s}\,. \tag{6.2.20}$$

Finally, we have

$$2^s\Delta_1(f) = 2^s\big(\widehat{\Psi}(\tfrac{1}{2}\xi)(1+|\xi|^2)^{-\frac{s}{2}}(1+|\xi|^2)^{\frac{s}{2}}\,\widehat{f}\,\big)^{\vee} = 2^s\big(\widehat{\Psi}(\tfrac{1}{2}\xi)(1+|\xi|^2)^{-\frac{s}{2}}\,\widehat{f_s}\,\big)^{\vee},$$

and since the function $\widehat{\Psi}(\frac{1}{2}\xi)(1+|\xi|^2)^{-\frac{s}{2}}$ is smooth with compact support and thus in $\mathscr{M}_p$, it follows that

$$\big\|2^s\Delta_1(f)\big\|_{L^p} \le C\big\|f_s\big\|_{L^p} = C\big\|f\big\|_{L^p_s}\,. \tag{6.2.21}$$

Combining estimates (6.2.17), (6.2.20), and (6.2.21), we conclude the proof of (6.2.11). □

6.2.3 Littlewood–Paley Characterization of Homogeneous Sobolev Spaces

We now state and prove the homogeneous version of the previous theorem.

Theorem 6.2.7. *Let Ψ satisfy (6.2.6) and let Δ_j be the Littlewood–Paley operator associated with Ψ. Let $s \in \mathbf{R}$ and $1 < p < \infty$. Then there exists a constant C_1 that depends only on n, s, p, and Ψ such that for all $f \in \dot{L}^p_s(\mathbf{R}^n)$ we have*

$$\Big\|\Big(\sum_{j\in\mathbf{Z}}(2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \le C_1\big\|f\big\|_{\dot{L}^p_s}\,. \tag{6.2.22}$$

Conversely, there exists a constant C_2 that depends on the parameters n, s, p, and Ψ such that every element f of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ that satisfies

$$\Big\|\Big(\sum_{j\in\mathbf{Z}}(2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p}<\infty$$

lies in the homogeneous Sobolev space $\dot{L}^p_s$ and we have

$$\big\|f\big\|_{\dot{L}^p_s}\le C_2\Big\|\Big(\sum_{j\in\mathbf{Z}}(2^{js}|\Delta_j(f)|)^2\Big)^{\frac{1}{2}}\Big\|_{L^p}. \tag{6.2.23}$$

Proof. The proof of the theorem is similar but a bit simpler than that of Theorem 6.2.6. To obtain (6.2.22) we start with $f\in\dot{L}^p_s$ and we note that

$$2^{js}\Delta_j(f)=2^{js}\big(|\xi|^s|\xi|^{-s}\widehat{\Psi}(2^{-j}\xi)\widehat{f}\,\big)^\vee=\big(\widehat{\sigma}(2^{-j}\xi)\widehat{f_s}\,\big)^\vee=\Delta_j^\sigma(f_s),$$

where $\widehat{\sigma}(\xi)=\widehat{\Psi}(\xi)|\xi|^{-s}$ and Δ_j^σ is the Littlewood–Paley operator given on the Fourier transform side by multiplication with the function $\widehat{\sigma}(2^{-j}\xi)$. We have

$$\Big\|\Big(\sum_{j\in\mathbf{Z}}|2^{js}\Delta_j(f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}=\Big\|\Big(\sum_{j\in\mathbf{Z}}|\Delta_j^\sigma(f_s)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}\le C\big\|f_s\big\|_{L^p}=C\big\|f\big\|_{\dot{L}^p_s},$$

where the last inequality follows from Theorem 5.1.2. This proves (6.2.22).

Next we show that if the expression on the right in (6.2.23) is finite, then the distribution f in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ must lie the in the homogeneous Sobolev space $\dot{L}^p_s$ with norm controlled by a multiple of this expression.

Define Littlewood–Paley operators Δ_j^η given by convolution with $\eta_{2^{-j}}$, where $\widehat{\eta}$ is a smooth bump supported in the annulus $\frac{4}{5}\le|\xi|\le 2$ that satisfies

$$\sum_{k\in\mathbf{Z}}\widehat{\eta}(2^{-k}\xi)=1,\qquad \xi\neq 0, \tag{6.2.24}$$

or, in operator form,

$$\sum_{k\in\mathbf{Z}}\Delta_k^\eta=I,$$

where the convergence is in the sense of $\mathscr{S}'/\mathscr{P}$ in view of Exercise 2.3.12. We introduce another family of Littlewood–Paley operators Δ_j^θ given by convolution with $\theta_{2^{-j}}$, where $\widehat{\theta}(\xi)=\widehat{\eta}(\xi)|\xi|^s$. Given $f\in\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$, we set $f_s=\big(|\xi|^s\widehat{f}\,\big)^\vee$, which is also an element of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$. In view of (6.2.24) we can use the reverse estimate (5.1.8) in Theorem 5.1.2 to obtain for some polynomial Q,

$$\|f\|_{\dot{L}^p_s}=\|f_s-Q\|_{L^p}\le C\Big\|\Big(\sum_{j\in\mathbf{Z}}|\Delta_j^\eta(f_s)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}=C\Big\|\Big(\sum_{j\in\mathbf{Z}}|2^{js}\Delta_j^\theta(f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}.$$

Recalling the definition of Δ_j (see the discussion before the statement of Theorem 6.2.6), we notice that the function

$$\widehat{\Psi}(\tfrac{1}{2}\xi)+\widehat{\Psi}(\xi)+\widehat{\Psi}(2\xi)$$

is equal to 1 on the support of $\widehat{\theta}$ (which is the same as the support of η). It follows that

$$\Delta_j^\theta = \big(\Delta_{j-1}+\Delta_j+\Delta_{j+1}\big)\Delta_j^\theta\,.$$

We therefore have the estimate

$$\Big\|\Big(\sum_{j\in\mathbf{Z}}|2^{js}\Delta_j^\theta(f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}\le\sum_{r=-1}^{1}\Big\|\Big(\sum_{j\in\mathbf{Z}}|\Delta_j^\theta\Delta_{j+r}(2^{js}f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p},$$

and applying Proposition 5.1.4, we can control the last expression $\big($and thus $\|f\|_{\dot L^p_s}\big)$ by a constant multiple of

$$\Big\|\Big(\sum_{j\in\mathbf{Z}}|\Delta_j(2^{js}f)|^2\Big)^{\frac{1}{2}}\Big\|_{L^p}.$$

This proves that the homogeneous Sobolev norm of f is controlled by a multiple of the expression in (6.2.23). In particular, the distribution f lies in the homogeneous Sobolev space $\dot L^p_s$. This ends the proof of the converse direction and completes the proof of the theorem. □

Exercises

6.2.1. Show that the spaces $\dot L^p_s$ and L^p_s are complete and that the latter are decreasing as s increases.

6.2.2. (a) Let $1<p<\infty$ and $s\in\mathbf{Z}^+$. Suppose that $f\in L^p_s(\mathbf{R}^n)$ and that φ is in $\mathscr{S}(\mathbf{R}^n)$. Prove that φf is also an element of $L^p_s(\mathbf{R}^n)$.
(b) Let v be a function whose Fourier transform is a bounded compactly supported function. Prove that if f is in $L^2_s(\mathbf{R}^n)$, then so is vf.

6.2.3. Let $s>0$ and α a fixed multi-index. Find the set of p in $(1,\infty)$ such that the distribution $\partial^\alpha\delta_0$ belongs to L^p_{-s}.

6.2.4. Let I be the identity operator, I_1 the Riesz potential of order 1, and R_j the usual Riesz transform. Prove that

$$I=\sum_{j=1}^{n} I_1 R_j\partial_j\,,$$

and use this identity to obtain Theorem 6.2.4 when $s=1$.
$\big[$*Hint:* Take the Fourier transform.$\big]$

6.2.5. Let f be in L^p_s for some $1 < p < \infty$. Prove that $\partial^\alpha f$ is in $L^p_{s-|\alpha|}$.

6.2.6. Prove that for all $\mathscr{C}^1$ functions f that are supported in a ball B we have

$$|f(x)| \le \frac{1}{\omega_{n-1}} \int_B |\nabla f(y)||x-y|^{-n+1}\, dy,$$

where $\omega_{n-1} = |\mathbf{S}^{n-1}|$. For such functions obtain the local Sobolev inequality

$$\big\|f\big\|_{L^q(B)} \le C_{q,r,n}\big\|\nabla f\big\|_{L^p(B)},$$

where $1 < p < q < \infty$ and $1/p = 1/q + 1/n$.
$\big[$*Hint:* Start from $f(x) = \int_0^\infty \nabla f(x - t\theta)\cdot\theta\, dt$ and integrate over $\theta \in \mathbf{S}^{n-1}$.$\big]$

6.2.7. Show that there is a constant C such that for all $\mathscr{C}^1$ functions f that are supported in a ball B we have

$$\frac{1}{|B'|}\int_{B'} |f(x) - f(z)|\, dz \le C \int_B |\nabla f(y)||x-y|^{-n+1}\, dy$$

for all B' balls contained in B and all $x \in B'$.
$\big[$*Hint:* Start with $f(z) - f(x) = \int_0^1 \nabla f(x + t(z-x))\cdot(z-x)\, dt$.$\big]$

6.2.8. Let $1 < p < \infty$ and $s > 0$. Show that

$$f \in L^p_s \iff f \in L^p \quad \text{and} \quad f \in \dot{L}^p_s.$$

Conclude that $\dot{L}^p_s \cap L^p = L^p_s$ and obtain an estimate for the corresponding norms.
$\big[$*Hint:* If f is in $\dot{L}^p_s \cap L^p$ use Theorem 5.2.7 to obtain that $\big\|f\big\|_{L^p_s}$ is controlled by a multiple of the L^p norm of $(\widehat{f}(\xi)(1+|\xi|^s))^\vee$. Use the same theorem to show that $\big\|f\big\|_{\dot{L}^p_s} \le C\big\|f\big\|_{L^p_s}$.$\big]$

6.2.9. (*Gagliardo [139]/Nirenberg [249]*) Prove that all Schwartz functions on $\mathbf{R}^n$ satisfy the estimate

$$\big\|f\big\|_{L^q} \le \prod_{j=1}^n \big\|\partial_j f\big\|_{L^1}^{1/n},$$

where $1/q + 1/n = 1$.
$\big[$*Hint:* Use induction beginning with the case $n = 1$. Assuming that the inequality is valid for $n-1$, set $I_j(x_1) = \int_{\mathbf{R}^{n-1}} |\partial_j f(x_1, x')|\, dx'$ for $j = 2, \dots, n$, where $x = (x_1, x') \in \mathbf{R}\times\mathbf{R}^{n-1}$ and $I_1(x') = \int_{\mathbf{R}^1} |\partial_1 f(x_1, x')|\, dx_1$. Apply the induction hypothesis to obtain

$$\big\|f(x_1, \cdot)\big\|_{L^{q'}} \le \prod_{j=2}^n I_j(x_1)^{1/(n-1)}$$

and use that $|f|^q \le I_1(x')^{1/(n-1)}|f|$ and Hölder's inequality to calculate $\big\|f\big\|_{L^q}$.$\big]$

6.2.10. Let $f \in L^2_1(\mathbf{R}^n)$. Prove that there is a constant $c_n > 0$ such that

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \frac{|f(x+t)+f(x-t)-2f(x)|^2}{|t|^{n+2}}\,dx\,dt = c_n \int_{\mathbf{R}^n}\sum_{j=1}^n |\partial_j f(x)|^2\,dx.$$

6.2.11. (*Christ [61]*) Let $0 \le \beta < \infty$ and let

$$C_0 = \int_{\mathbf{R}^n} |\widehat{g}(\xi)|^2 (1+|\xi|)^n \big(\log(2+|\xi|)\big)^{-\beta}\,d\xi\,.$$

(a) Prove that there is a constant $C(n,\beta,C_0)$ such that for every $q > 2$ we have

$$\|g\|_{L^q(\mathbf{R}^n)} \le C(n,\beta,C_0)\,q^{\frac{\beta+1}{2}}\,.$$

(b) Conclude that for any compact subset K of $\mathbf{R}^n$ we have

$$\int_K e^{|g(x)|^\gamma}\,dx < \infty$$

whenever $\gamma < \frac{2}{\beta+1}$.
[*Hint:* Part (a): For $q > 2$ control $\|g\|_{L^q(\mathbf{R}^n)}$ by $\|\widehat{g}\|_{L^{q'}(\mathbf{R}^n)}$ and apply Hölder's inequality with exponents $\frac{2}{q'}$ and $\frac{2(q-1)}{q-2}$. Part (b): Expand the exponential in a Taylor series.]

6.2.12. Suppose that $m \in L^2_s(\mathbf{R}^n)$ for some $s > \frac{n}{2}$ and let $\lambda > 0$. Define the operator T_λ by setting $\widehat{T_\lambda(f)}(\xi) = m(\lambda\xi)\widehat{f}(\xi)$. Show that there exists a constant $C = C(n,s)$ such that for all f and $u \ge 0$ and $\lambda > 0$ we have

$$\int_{\mathbf{R}^n} |T_\lambda(f)(x)|^2\,u(x)\,dx \le C\int_{\mathbf{R}^n} |f(x)|^2\,M(u)(x)\,dx\,.$$

6.3 Lipschitz Spaces

The classical definition says that a function f on $\mathbf{R}^n$ is Lipschitz (or Hölder) continuous of order $\gamma > 0$ if there is constant $C < \infty$ such that for all $x,y \in \mathbf{R}^n$ we have

$$|f(x+y)-f(x)| \le C|y|^\gamma. \tag{6.3.1}$$

It turns out that only constant functions satisfy (6.3.1) when $\gamma > 1$, and the corresponding definition needs to be suitably adjusted in this case. This is discussed in this section. The key point is that any function f that satisfies (6.3.1) possesses a certain amount of smoothness "measured" by the quantity γ. The Lipschitz norm of a function is introduced to serve this purpose, that is, to precisely quantify and exactly measure this smoothness. In this section we formalize these concepts and we explore connections they have with the orthogonality considerations of the previous chapter. The main achievement of this section is a characterization of Lipschitz spaces using Littlewood–Paley theory.

6.3.1 Introduction to Lipschitz Spaces

Definition 6.3.1. Let $0<\gamma<1$. A function f on $\mathbf{R}^n$ is said to be *Lipschitz of order* γ if it is bounded and satisfies (6.3.1) for some $C<\infty$. In this case we let

$$\big\|f\big\|_{\Lambda_\gamma(\mathbf{R}^n)} = \big\|f\big\|_{L^\infty} + \sup_{x\in\mathbf{R}^n}\sup_{h\in\mathbf{R}^n\setminus\{0\}} \frac{|f(x+h)-f(x)|}{|h|^\gamma}$$

and we set

$$\Lambda_\gamma(\mathbf{R}^n) = \{f:\mathbf{R}^n\to\mathbf{C} \text{ continuous}:\ \big\|f\big\|_{\Lambda_\gamma(\mathbf{R}^n)}<\infty\}.$$

Note that functions in $\Lambda_\gamma(\mathbf{R}^n)$ are automatically continuous when $\gamma<1$, so we did not need to make this part of the definition. We call $\Lambda_\gamma(\mathbf{R}^n)$ the *inhomogeneous Lipschitz space* of order γ. For reasons of uniformity we also set

$$\Lambda_0(\mathbf{R}^n) = L^\infty(\mathbf{R}^n)\cap C(\mathbf{R}^n),$$

where $C(\mathbf{R}^n)$ is the space of all continuous functions on $\mathbf{R}^n$. See Exercise 6.3.2.

Example 6.3.2. The function $h(x)=\cos(x\cdot a)$ for some fixed $a\in\mathbf{R}^n$ is in Λ_γ for all $\gamma<1$. Simply notice that $|h(x)-h(y)|\le\min(2,|a|\,|x-y|)$.

We now extend this definition to indices $\gamma\ge 1$.

Definition 6.3.3. For $h\in\mathbf{R}^n$ define the *difference operator* D_h by setting

$$D_h(f)(x) = f(x+h)-f(x)$$

for a continuous function $f:\ \mathbf{R}^n\to\mathbf{C}$. We may check that

$$\begin{aligned} D_h^2(f)(x) &= D_h(D_hf)(x) = f(x+2h)-2f(x+h)+f(x),\\ D_h^3(f)(x) &= D_h(D_h^2f)(x) = f(x+3h)-3f(x+2h)+3f(x+h)-f(x)\,, \end{aligned}$$

and in general, that $D_h^{k+1}(f)=D_h^k(D_h(f))$ is given by

$$D_h^{k+1}(f)(x) = \sum_{s=0}^{k+1}(-1)^{k+1-s}\binom{k+1}{s} f(x+sh) \tag{6.3.2}$$

for a nonnegative integer k. See Exercise 6.3.3. For $\gamma>0$ define

$$\big\|f\big\|_{\Lambda_\gamma} = \big\|f\big\|_{L^\infty} + \sup_{x\in\mathbf{R}^n}\sup_{h\in\mathbf{R}^n\setminus\{0\}} \frac{|D_h^{[\gamma]+1}(f)(x)|}{|h|^\gamma},$$

where $[\gamma]$ denotes the integer part of γ, and set

$$\Lambda_\gamma = \{f:\mathbf{R}^n \to \mathbf{C} \text{ continuous}:\ \big\|f\big\|_{\Lambda_\gamma} < \infty\}.$$

We call $\Lambda_\gamma(\mathbf{R}^n)$ the inhomogeneous *Lipschitz space* of order $\gamma \in \mathbf{R}^+$.

For a tempered distribution u we also define another distribution $D_h^k(u)$ via the identity

$$\big\langle D_h^k(u),\varphi\big\rangle = \big\langle u, D_{-h}^k(\varphi)\big\rangle$$

for all φ in the Schwartz class.

We now define the homogeneous Lipschitz spaces. We adhere to the usual convention of using a dot on a space to indicate its homogeneous nature.

Definition 6.3.4. For $\gamma > 0$ we define

$$\big\|f\big\|_{\dot\Lambda_\gamma} = \sup_{x\in\mathbf{R}^n}\sup_{h\in\mathbf{R}^n\setminus\{0\}} \frac{|D_h^{[\gamma]+1}(f)(x)|}{|h|^\gamma}$$

and we also let $\dot\Lambda_\gamma$ be the space of all continuous functions f on $\mathbf{R}^n$ that satisfy $\big\|f\big\|_{\dot\Lambda_\gamma} < \infty$. We call $\dot\Lambda_\gamma$ the *homogeneous Lipschitz space* of order γ. We note that elements of $\dot\Lambda_\gamma$ have at most polynomial growth at infinity and thus they are elements of $\mathscr{S}'(\mathbf{R}^n)$.

A few observations are in order here. Constant functions f satisfy $D_h(f)(x) = 0$ for all $h,x \in \mathbf{R}^n$, and therefore the homogeneous quantity $\big\|\cdot\big\|_{\dot\Lambda_\gamma}$ is insensitive to constants. Similarly the expressions $D_h^{k+1}(f)$ and $\big\|f\big\|_{\dot\Lambda_\gamma}$ do not recognize polynomials of degree up to k. Moreover, polynomials are the only continuous functions with this property; see Exercise 6.3.1. This means that the quantity $\big\|f\big\|_{\dot\Lambda_\gamma}$ is not a norm but only a seminorm. To make it a norm, we need to consider functions modulo polynomials, as we did in the case of homogeneous Sobolev spaces. For this reason we think of $\dot\Lambda_\gamma$ as a subspace of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$.

We make use of the following proposition concerning properties of the difference operators D_h^k.

Proposition 6.3.5. *Let f be a $\mathscr{C}^m$ function on $\mathbf{R}^n$ for some $m \in \mathbf{Z}^+$. Then for all $h = (h_1,\dots,h_n)$ and $x \in \mathbf{R}^n$ the following identity holds:*

$$D_h(f)(x) = \int_0^1 \sum_{j=1}^n h_j\,(\partial_j f)(x+sh)\,ds. \tag{6.3.3}$$

More generally, we have that

$$D_h^m(f)(x) = \int_{[0,1]^m} \sum_{\substack{1\le j_\ell \le n\\ 1\le \ell\le m}} h_{j_1}\cdots h_{j_m}(\partial_{j_1}\cdots\partial_{j_m} f)(x+(s_1+\cdots+s_m)h)\,ds_1\cdots ds_m. \quad (6.3.4)$$

Proof. Identity (6.3.3) is a consequence of the fundamental theorem of calculus applied to the function $t\mapsto f((1-t)x+t(x+h))$ on $[0,1]$, while identity (6.3.4) follows by induction. □

6.3.2 Littlewood–Paley Characterization of Homogeneous Lipschitz Spaces

We now characterize the homogeneous Lipschitz spaces using the Littlewood–Paley operators Δ_j. As in the previous section, we fix a radial Schwartz function Ψ whose Fourier transform is nonnegative, supported in the annulus $1-\frac{1}{7}\le|\xi|\le 2$, is equal to one on the annulus $1\le|\xi|\le 2-\frac{2}{7}$, and that satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1 \quad (6.3.5)$$

for all $\xi\neq 0$. The Littlewood–Paley operators $\Delta_j=\Delta_j^\Psi$ associated with Ψ are given by multiplication on the Fourier transform side by the smooth bump $\widehat{\Psi}(2^{-j}\xi)$.

Theorem 6.3.6. *Let Δ_j be as above and $\gamma>0$. Then there is a constant $C=C(n,\gamma)$ such that for every f in $\dot{\Lambda}_\gamma$ we have the estimate*

$$\sup_{j\in\mathbf{Z}} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} \le C\big\|f\big\|_{\dot{\Lambda}_\gamma}. \quad (6.3.6)$$

Conversely, every element f of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ that satisfies

$$\sup_{j\in\mathbf{Z}} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} < \infty \quad (6.3.7)$$

is an element of $\dot{\Lambda}_\gamma$ with norm

$$\big\|f\big\|_{\dot{\Lambda}_\gamma} \le C' \sup_{j\in\mathbf{Z}} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} \quad (6.3.8)$$

for some constant $C'=C'(n,\gamma)$.

Note that condition (6.3.7) remains invariant if a polynomial is added to the function f; this is consistent with the analogous property of the mapping $f\mapsto\big\|f\big\|_{\dot{\Lambda}_\gamma}$.

Proof. We begin with the proof of (6.3.8). Let $k=[\gamma]$ be the integer part of γ. Let us pick a Schwartz function η on $\mathbf{R}^n$ whose Fourier transform is nonnegative, supported in the annulus $\frac{4}{5}\le|\xi|\le 2$, and that satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\eta}(2^{-j}\xi)^2 = 1 \tag{6.3.9}$$

for all $\xi \neq 0$. Associated with η, we define the Littlewood–Paley operators Δ_j^η given by multiplication on the Fourier transform side by the smooth bump $\widehat{\eta}(2^{-j}\xi)$. With Ψ as in (6.2.6) we set

$$\widehat{\Theta}(\xi) = \widehat{\Psi}(\tfrac{1}{2}\xi) + \widehat{\Psi}(\xi) + \widehat{\Psi}(2\xi),$$

and we denote by $\Delta_j^\Theta = \Delta_{j-1} + \Delta_j + \Delta_{j+1}$ the Littlewood–Paley operator given by multiplication on the Fourier transform side by the smooth bump $\widehat{\Theta}(2^{-j}\xi)$.

The fact that the previous function is equal to 1 on the support of $\widehat{\eta}$ together with the functional identity (6.3.9) yields the operator identity

$$I = \sum_{j\in\mathbf{Z}} (\Delta_j^\eta)^2 = \sum_{j\in\mathbf{Z}} \Delta_j^\Theta \Delta_j^\eta \Delta_j^\eta,$$

with convergence in the sense of the space $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$. Since convolution is a linear operation, we have $D_h^{k+1}(F*G) = F*D_h^{k+1}(G)$, from which we deduce

$$\begin{aligned} D_h^{k+1}(f) &= \sum_{j\in\mathbf{Z}} \Delta_j^\Theta(f) * D_h^{k+1}(\eta_{2^{-j}}) * \eta_{2^{-j}} \\ &= \sum_{j\in\mathbf{Z}} D_h^{k+1}(\Delta_j^\Theta(f)) * (\eta*\eta)_{2^{-j}} \end{aligned} \tag{6.3.10}$$

for all tempered distributions f. The convergence of the series in (6.3.10) is in the sense of $\mathscr{S}'/\mathscr{P}$ in view of Exercise 5.2.2. The convergence of the series in (6.3.10) in the L^∞ norm is a consequence of condition (6.3.7) and is contained in the following argument.

Using (6.3.2), we easily obtain the estimate

$$\big\|D_h^{k+1}(\Delta_j^\Theta(f)) * (\eta*\eta)_{2^{-j}}\big\|_{L^\infty} \le 2^{k+1}\big\|\eta*\eta\big\|_{L^1}\big\|\Delta_j^\Theta(f)\big\|_{L^\infty}. \tag{6.3.11}$$

We first integrate over $(s_1,\dots,s_{k+1}) \in [0,1]^{k+1}$ the identity

$$\begin{aligned} &\sum_{r_1=1}^n \cdots \sum_{r_{k+1}=1}^n h_{r_1}\cdots h_{r_{k+1}} (\partial_{r_1}\cdots\partial_{r_{k+1}}\eta_{2^{-j}})(x+(s_1+\cdots+s_{k+1})h) \\ &\quad = 2^{j(k+1)} \sum_{r_1=1}^n \cdots \sum_{r_{k+1}=1}^n h_{r_1}\cdots h_{r_{k+1}} (\partial_{r_1}\dots\partial_{r_{k+1}}\eta)_{2^{-j}}(x+(s_1+\cdots+s_{k+1})h). \end{aligned}$$

We then use (6.3.4) with $m = k+1$, and we integrate over $x \in \mathbf{R}^n$ to obtain

$$\big\|D_h^{k+1}(\eta_{2^{-j}})\big\|_{L^1} \le 2^{j(k+1)}|h|^{k+1} \sum_{r_1=1}^n \cdots \sum_{r_{k+1}=1}^n \big\|\partial_{r_1}\cdots\partial_{r_{k+1}}\eta\big\|_{L^1}.$$

We deduce the validity of the estimate

$$\begin{aligned}\big\|\Delta_j^{\Theta}(f)*D_h^{k+1}(\eta_{2^{-j}})*\eta_{2^{-j}}\big\|_{L^\infty}&\\ \le\ \big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}\big\|D_h^{k+1}(\eta_{2^{-j}})*\eta_{2^{-j}}\big\|_{L^1}&\\ \le\ \big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}|2^jh|^{k+1}c_k\sum_{|\alpha|\le k+1}\big\|\partial^\alpha\eta\big\|_{L^1}\big\|\eta\big\|_{L^1}.&\end{aligned}\tag{6.3.12}$$

Combining (6.3.11) and (6.3.12), we obtain

$$\begin{aligned}\big\|\Delta_j^{\Theta}(f)*D_h^{k+1}(\eta_{2^{-j}})*\eta_{2^{-j}}\big\|_{L^\infty}&\\ \le C_{\eta,n,k}\big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}\min\big(1,|2^jh|^{k+1}\big).&\end{aligned}\tag{6.3.13}$$

We insert estimate (6.3.13) in (6.3.10) to deduce

$$\frac{\big\|D_h^{k+1}(f)\big\|_{L^\infty}}{|h|^\gamma}\le\frac{C'}{|h|^\gamma}\sum_{j\in\mathbf{Z}}2^{j\gamma}\big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}\min\big(2^{-j\gamma},2^{j(k+1-\gamma)}|h|^{k+1}\big),$$

from which it follows that

$$\begin{aligned}\big\|f\big\|_{\dot\Lambda_\gamma}\ &\le\ \sup_{h\in\mathbf{R}^n\setminus\{0\}}\frac{C'}{|h|^\gamma}\sum_{j\in\mathbf{Z}}2^{j\gamma}\big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}\min\big(2^{-j\gamma},2^{j(k+1-\gamma)}|h|^{k+1}\big)\\ &\le\ C'\sup_{j\in\mathbf{Z}}2^{j\gamma}\big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty}\sup_{h\neq0}\sum_{j\in\mathbf{Z}}\min\big(|h|^{-\gamma}2^{-j\gamma},2^{j(k+1-\gamma)}|h|^{k+1-\gamma}\big)\\ &\le\ C'\sup_{j\in\mathbf{Z}}2^{j\gamma}\big\|\Delta_j^{\Theta}(f)\big\|_{L^\infty},\end{aligned}$$

since the last numerical series converges ($\gamma<k+1=[\gamma]+1$). This proves (6.3.8) with the difference that instead of Δ_j we have Δ_j^{Θ} on the right. The passage to Δ_j is a trivial matter, since $\Delta_j^{\Theta}=\Delta_{j-1}+\Delta_j+\Delta_{j+1}$.

Having established (6.3.8), we now turn to the proof of (6.3.6). We first consider the case $0<\gamma<1$, which is very simple. Since each Δ_j is given by convolution with a function with mean value zero, we may write

$$\begin{aligned}\Delta_j(f)(x)\ &=\ \int_{\mathbf{R}^n}f(x-y)\Psi_{2^{-j}}(y)\,dy\\ &=\ \int_{\mathbf{R}^n}(f(x-y)-f(x))\Psi_{2^{-j}}(y)\,dy\\ &=\ 2^{-j\gamma}\int_{\mathbf{R}^n}\frac{D_{-y}(f)(x)}{|y|^\gamma}|2^jy|^\gamma2^{jn}\Psi(2^jy)\,dy,\end{aligned}$$

and the previous expression is easily seen to be controlled by a constant multiple of $2^{-j\gamma}\big\|f\big\|_{\dot\Lambda_\gamma}$. This proves (6.3.6) when $0<\gamma<1$. In the case $\gamma\ge1$ we have to work a bit harder.

As before, set $k=[\gamma]$. Notice that for Schwartz functions g we have the identity

$$D_h^{k+1}(g) = \big(\widehat{g}(\xi)\,(e^{2\pi i\xi\cdot h}-1)^{k+1}\big)^{\vee}.$$

To express $\Delta_j(g)$ in terms of $D_h^{k+1}(g)$, we need to introduce the function

$$\xi \mapsto \widehat{\Psi}(2^{-j}\xi)\,(e^{2\pi i\xi\cdot h}-1)^{-(k+1)}.$$

But as the support of $\widehat{\Psi}(2^{-j}\xi)$ may intersect the set of all ξ for which $\xi\cdot h$ is an integer, the previous function is not well defined. To deal with this problem, we pick a finite family of unit vectors $\{u_r\}_r$ so that the annulus $\frac{1}{2}\le|\xi|\le 2$ is covered by the union of sets

$$U_r = \big\{\xi\in\mathbf{R}^n:\ \tfrac{1}{2}\le|\xi|\le 2,\quad \tfrac{1}{4}\le|\xi\cdot u_r|\le 2\big\}.$$

Then we write $\widehat{\Psi}$ as a finite sum of smooth functions $\widehat{\Psi^{(r)}}$, where each $\widehat{\Psi^{(r)}}$ is supported in U_r. Setting

$$h_r = \frac{1}{8}2^{-j}u_r\,,$$

we note that

$$\begin{aligned}\Psi^{(r)}_{2^{-j}} * f \;&=\; \big(\widehat{\Psi^{(r)}}(2^{-j}\xi)\,(e^{2\pi i\xi\cdot h_r}-1)^{-(k+1)}(e^{2\pi i\xi\cdot h_r}-1)^{k+1}\widehat{f}(\xi)\big)^{\vee}\\ &=\; \big(\widehat{\Psi^{(r)}}(2^{-j}\xi)\,(e^{2\pi i 2^{-j}\xi\cdot\frac{1}{8}u_r}-1)^{-(k+1)}\widehat{D^{k+1}_{h_r}(f)}(\xi)\big)^{\vee}\end{aligned} \tag{6.3.14}$$

and observe that the exponential is never equal to 1, since

$$2^{-j}\xi\in U_r \implies \tfrac{1}{32}\le|2^{-j}\xi\cdot\tfrac{1}{8}u_r|\le\tfrac{1}{4}\,.$$

Since the function $\widehat{\zeta^{(r)}}=\widehat{\Psi^{(r)}}(\xi)\,(e^{2\pi i\xi\cdot\frac{1}{8}u_r}-1)^{-(k+1)}$ is well defined and smooth with compact support, it follows that

$$\Psi^{(r)}_{2^{-j}} * f = (\zeta^{(r)})_{2^{-j}} * D^{k+1}_{2^{-j}\frac{1}{8}u_r}(f)\,,$$

which implies that

$$\begin{aligned}\big\|\Psi^{(r)}_{2^{-j}} * f\big\|_{L^\infty} \;&\le\; \big\|(\zeta^{(r)})_{2^{-j}}\big\|_{L^1}\big\|D^{k+1}_{2^{-j}\frac{1}{8}u_r}(f)\big\|_{L^\infty}\\ &\le\; \big\|\zeta^{(r)}\big\|_{L^1}\big\|f\big\|_{\dot{\Lambda}_\gamma}2^{-j\gamma}.\end{aligned}$$

Summing over the finite number of r, we obtain the estimate

$$\big\|\Delta_j(f)\big\|_{L^\infty}\le C\big\|f\big\|_{\dot{\Lambda}_\gamma}2^{-j\gamma},$$

which concludes the proof of the theorem. □

6.3.3 *Littlewood–Paley Characterization of Inhomogeneous Lipschitz Spaces*

We have seen that quantities involving the Littlewood–Paley operators Δ_j characterize homogeneous Lipschitz spaces. We now address the same question for inhomogeneous spaces.

As in the Littlewood–Paley characterization of inhomogeneous Sobolev spaces, we need to treat the contribution of the frequencies near zero separately. We recall the Schwartz function Φ introduced in Section 6.2.2:

$$\widehat{\Phi}(\xi) = \begin{cases} \sum_{j\le 0} \widehat{\Psi}(2^{-j}\xi) & \text{when } \xi \neq 0, \\ 1 & \text{when } \xi = 0. \end{cases} \tag{6.3.15}$$

Note that $\widehat{\Phi}(\xi)$ is equal to 1 for $|\xi| \le 2 - \frac{2}{7}$ and vanishes when $|\xi| \ge 2$. We also recall the operator $S_0(f) = \Phi * f$. One should not be surprised to find out that a result analogous to that in Theorem 6.2.6 is valid for Lipschitz spaces as well.

Theorem 6.3.7. *Let Ψ and Δ_j be as in the Theorem 6.3.6, Φ as in (6.3.15), and $\gamma > 0$. Then there is a constant $C = C(n,\gamma)$ such that for every f in Λ_γ we have the estimate*

$$\big\|S_0(f)\big\|_{L^\infty} + \sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} \le C\big\|f\big\|_{\Lambda_\gamma}. \tag{6.3.16}$$

Conversely, every tempered distribution f that satisfies

$$\big\|S_0(f)\big\|_{L^\infty} + \sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} < \infty \tag{6.3.17}$$

can be identified with an element of Λ_γ. Moreover, there is a constant $C' = C'(n,\gamma)$ such that for all f that satisfy (6.3.17) we have

$$\big\|f\big\|_{\Lambda_\gamma} \le C'\Big(\big\|S_0(f)\big\|_{L^\infty} + \sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty}\Big). \tag{6.3.18}$$

Proof. The proof of (6.3.16) is immediate, since we trivially have

$$\big\|S_0(f)\big\|_{L^\infty} = \big\|f * \Phi\big\|_{L^\infty} \le \big\|\Phi\big\|_{L^1}\big\|f\big\|_{L^\infty} \le C\big\|f\big\|_{\Lambda_\gamma}$$

and also

$$\sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} \le C\big\|f\big\|_{\dot\Lambda_\gamma} \le C\big\|f\big\|_{\Lambda_\gamma}$$

by the previous theorem.

Therefore, the main part of the argument is contained in the proof of the converse estimate (6.3.18). Here we introduce Schwartz functions ζ,η so that

$$\widehat{\zeta}(\xi)^2 + \sum_{j=1}^{\infty} \widehat{\eta}(2^{-j}\xi)^2 = 1$$

and such that $\widehat{\eta}$ is supported in the annulus $\frac{4}{5} \le |\xi| \le 2$ and $\widehat{\zeta}$ is supported in the ball $|\xi| \le 1$. We associate Littlewood–Paley operators Δ_j^{η} given by convolution with the functions $\eta_{2^{-j}}$ and we also let $\Delta_j^{\Theta} = \Delta_{j-1} + \Delta_j + \Delta_{j+1}$. Note that $\widehat{\boldsymbol{\Phi}}$ is equal to one on the support of $\widehat{\zeta}$. Moreover, $\Delta_j^{\Theta}\Delta_j^{\eta} = \Delta_j^{\eta}$; hence for tempered distributions f we have the identity

$$f = \zeta * \zeta * \boldsymbol{\Phi} * f + \sum_{j=1}^{\infty} \eta_{2^{-j}} * \eta_{2^{-j}} * \Delta_j^{\Theta}(f), \tag{6.3.19}$$

where the series converges in $\mathscr{S}'(\mathbf{R}^n)$. With $k = [\gamma]$ we write

$$\frac{D_h^{k+1}(f)}{|h|^{\gamma}} = \zeta * \frac{D_h^{k+1}(\zeta)}{|h|^{\gamma}} * \boldsymbol{\Phi} * f + \sum_{j=1}^{\infty} \eta_{2^{-j}} * \frac{D_h^{k+1}(\eta_{2^{-j}})}{|h|^{\gamma}} * \Delta_j^{\Theta}(f), \tag{6.3.20}$$

and we use Proposition 6.3.5 to estimate the L^{∞} norm of the term $\zeta * \frac{D_h^{k+1}(\zeta)}{|h|^{\gamma}} * \boldsymbol{\Phi} * f$ in the previous sum as follows:

$$\begin{aligned} \Big\|\zeta * \tfrac{D_h^{k+1}(\zeta)}{|h|^{\gamma}} * \boldsymbol{\Phi} * f\Big\|_{L^{\infty}} &\le \Big\|\tfrac{D_h^{k+1}(\zeta)}{|h|^{\gamma}}\Big\|_{L^{\infty}} \big\|\zeta * \boldsymbol{\Phi} * f\big\|_{L^1} \\ &\le C \min\big(\tfrac{1}{|h|^{\gamma}}, \tfrac{|h|^{k+1}}{|h|^{\gamma}}\big) \big\|\boldsymbol{\Phi} * f\big\|_{L^{\infty}} \\ &\le C\big\|\boldsymbol{\Phi} * f\big\|_{L^{\infty}}. \end{aligned} \tag{6.3.21}$$

The corresponding L^{∞} estimates for $\Delta_j^{\Theta}(f) * \eta_{2^{-j}} * D_h^{k+1}(\eta_{2^{-j}})$ have already been obtained in (6.3.13). Indeed, we obtained

$$\big\|D_h^{k+1}(\eta_{2^{-j}}) * \eta_{2^{-j}} * \Delta_j^{\Theta}(f)\big\|_{L^{\infty}} \le C_{\eta,n,k}\big\|\Delta_j^{\Theta}(f)\big\|_{L^{\infty}} \min\big(1, |2^j h|^{k+1}\big),$$

from which it follows that

$$\begin{aligned} &\Big\|\sum_{j=1}^{\infty} \eta_{2^{-j}} * \frac{D_h^{k+1}(\eta_{2^{-j}})}{|h|^{\gamma}} * \Delta_j^{\Theta}(f)\Big\|_{L^{\infty}} \\ &\quad\le C\Big(\sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j^{\Theta}(f)\big\|_{L^{\infty}}\Big) \sum_{j=1}^{\infty} 2^{-j\gamma}|h|^{-\gamma} \min\big(1, |2^j h|^{k+1}\big) \\ &\quad\le C\Big(\sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^{\infty}}\Big) \sum_{j=1}^{\infty} \min\big(|2^j h|^{-\gamma}, |2^j h|^{k+1-\gamma}\big) \\ &\quad\le C \sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^{\infty}}, \end{aligned} \tag{6.3.22}$$

where the last series is easily seen to converge uniformly in $h \in \mathbf{R}^n$, since $k+1 = [\gamma]+1 > \gamma$. We now combine identity (6.3.20) with estimates (6.3.21) and (6.3.22) to obtain that the expression on the right in (6.3.19) has a bounded L^{∞} norm. This implies that f can be identified with a bounded function that satisfies (6.3.18). □

Next, we obtain consequences of the Littlewood–Paley characterization of Lipschitz spaces. In the following corollary we identify Λ_0 with L^∞.

Corollary 6.3.8. *For* $0 \le \gamma \le \delta < \infty$ *there is a constant* $C_{n,\gamma,\delta} < \infty$ *such that for all* $f \in \Lambda_\delta(\mathbf{R}^n)$ *we have*

$$\big\|f\big\|_{\Lambda_\gamma} \le C_{n,\gamma,\delta}\big\|f\big\|_{\Lambda_\delta}.$$

In other words, the space $\Lambda_\delta(\mathbf{R}^n)$ *can be identified with a subspace of* $\Lambda_\gamma(\mathbf{R}^n)$.

Proof. If $0 < \gamma \le \delta$ and $j \ge 0$, then we must have $2^{j\gamma} \le 2^{j\delta}$ and thus

$$\sup_{j\ge 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty} \le \sup_{j\ge 1} 2^{j\delta}\big\|\Delta_j(f)\big\|_{L^\infty}.$$

Adding $\big\|S_0(f)\big\|_{L^\infty}$ and using Theorem 6.3.7, we obtain the required conclusion. The case $\gamma = 0$ is trivial. □

Remark 6.3.9. We proved estimates (6.3.18) and (6.3.8) using the Littlewood–Paley operators Δ_j constructed by a fixed choice of the function Ψ; Φ also depended on Ψ. It should be noted that the specific choice of the functions Ψ and Φ was unimportant in those estimates. In particular, if we know (6.3.18) and (6.3.8) for some choice of Littlewood–Paley operators $\widetilde{\Delta}_j$ and some Schwartz function $\widetilde{\Phi}$ whose Fourier transform is supported in a neighborhood of the origin, then (6.3.18) and (6.3.8) would also hold for our fixed choice of Δ_j and Φ. This situation is illustrated in the next corollary.

Corollary 6.3.10. *Let* $\gamma > 0$ *and let* α *be a multi-index with* $|\alpha| < \gamma$. *If* $f \in \Lambda_\gamma$, *then the distributional derivative* $\partial^\alpha f$ *(of* f*) lies in* $\Lambda_{\gamma-|\alpha|}$. *Likewise, if* $f \in \dot{\Lambda}_\gamma$, *then* $\partial^\alpha f \in \dot{\Lambda}_{\gamma-|\alpha|}$. *Precisely, we have the norm estimates*

$$\big\|\partial^\alpha f\big\|_{\Lambda_{\gamma-|\alpha|}} \le C_{\gamma,\alpha}\big\|f\big\|_{\Lambda_\gamma}, \tag{6.3.23}$$

$$\big\|\partial^\alpha f\big\|_{\dot{\Lambda}_{\gamma-|\alpha|}} \le C_{\gamma,\alpha}\big\|f\big\|_{\dot{\Lambda}_\gamma}. \tag{6.3.24}$$

In particular, elements of Λ_γ *and* $\dot{\Lambda}_\gamma$ *are in* $\mathscr{C}^\alpha$ *for all* $|\alpha| < \gamma$.

Proof. Let α be a multi-index with $|\alpha| < \gamma$. We denote by $\Delta_j^{\partial^\alpha\Psi}$ the Littlewood–Paley operator associated with the bump $(\partial^\alpha\Psi)_{2^{-j}}$. It is straightforward to check that the identity

$$\Delta_j(\partial^\alpha f) = 2^{j|\alpha|}\Delta_j^{\partial^\alpha\Psi}(f)$$

is valid for any tempered distribution f. Using the support properties of Ψ, we obtain

$$2^{j(\gamma-|\alpha|)}\Delta_j(\partial^\alpha f) = 2^{j\gamma}\Delta_j^{\partial^\alpha\Psi}(\Delta_{j-1}+\Delta_j+\Delta_{j+1})(f), \tag{6.3.25}$$

and from this it easily follows that

$$\sup_{j\in\mathbf{Z}} 2^{j(\gamma-|\alpha|)}\big\|\Delta_j(\partial^\alpha f)\big\|_{L^\infty} \le (2^\gamma+2)\,\big\|\partial^\alpha\Psi\big\|_{L^1}\sup_{j\in\mathbf{Z}} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty}$$

and also that

$$\sup_{j\geq 1} 2^{j(\gamma-|\alpha|)}\big\|\Delta_j(\partial^\alpha f)\big\|_{L^\infty} \leq (2^\gamma+2)\,\big\|\partial^\alpha \Psi\big\|_{L^1} \sup_{j\geq 1} 2^{j\gamma}\big\|\Delta_j(f)\big\|_{L^\infty}. \tag{6.3.26}$$

Using Theorem 6.3.6, we deduce that if $f\in\dot{\Lambda}_\gamma$, then $\partial^\alpha f\in\dot{\Lambda}_{\gamma-|\alpha|}$, and we also obtain (6.3.24). To derive the inhomogeneous version, we note that

$$S_0(\partial^\alpha f) = \Phi*(\partial^\alpha f) = (\partial^\alpha\Phi * f) = \big(\partial^\alpha\Phi*(\Phi+\Psi_{2^{-1}})*f\big),$$

since the function $\widehat{\Phi}+\widehat{\Psi_{2^{-1}}}$ is equal to 1 on the support of $\widehat{\partial^\alpha\Phi}$. Taking L^∞ norms, we obtain

$$\begin{aligned}\big\|S_0(\partial^\alpha f)\big\|_{L^\infty} &\leq \big\|\partial^\alpha\Phi\big\|_{L^1}\big(\big\|\Phi*f\big\|_{L^\infty}+\big\|\Psi_{2^{-1}}*f\big\|_{L^\infty}\big)\\ &\leq \big\|\partial^\alpha\Phi\big\|_{L^1}\Big(\big\|S_0(f)\big\|_{L^\infty}+\sup_{j\geq 1}\big\|\Delta_j(f)\big\|_{L^\infty}\Big),\end{aligned}$$

which, combined with (6.3.26), yields $\big\|\partial^\alpha f\big\|_{\Lambda_{\gamma-|\alpha|}} \leq C_{\gamma,\alpha}\big\|f\big\|_{\Lambda_\gamma}$. □

Exercises

6.3.1. Fix $k\in\mathbf{Z}^+$. Show that

$$D_h^k(f)(x)=0$$

for all x,h in $\mathbf{R}^n$ if and only if f is a polynomial of degree at most $k-1$.
[*Hint:* One direction may be proved by direct verification. For the converse direction, show that $\widehat{f}$ is supported at the origin and use Proposition 2.4.1.]

6.3.2. (a) Extend Definition 6.3.1 to the case $\gamma=0$ and show that for all continuous functions f we have

$$\big\|f\big\|_{L^\infty} \leq \big\|f\big\|_{\Lambda_0} \leq 3\big\|f\big\|_{L^\infty};$$

hence the space $\Lambda_0(\mathbf{R}^n)$ can be identified with $L^\infty(\mathbf{R}^n)\cap C(\mathbf{R}^n)$.
(b) Given a measurable function f on $\mathbf{R}^n$ we define

$$\big\|f\big\|_{\dot{L}^\infty} = \inf\big\{\big\|f+c\big\|_{L^\infty}:\ c\in\mathbf{C}\big\}.$$

Let $\dot{L}^\infty(\mathbf{R}^n)$ be the space of equivalent classes of bounded functions whose difference is a constant, equipped with this norm. Show that for all continuous functions f on $\mathbf{R}^n$ we have

$$\big\|f\big\|_{\dot{L}^\infty} \leq \sup_{x,h\in\mathbf{R}^n}|f(x+h)-f(x)| \leq 2\big\|f\big\|_{\dot{L}^\infty}.$$

In other words, $\dot{\Lambda}_0(\mathbf{R}^n)$ can be identified with $\dot{L}^\infty(\mathbf{R}^n)\cap C(\mathbf{R}^n)$.

6.3.3. (a) For a continuous function f prove the identity

$$D_h^{k+1}(f)(x) = \sum_{s=0}^{k+1}(-1)^{k+1-s}\binom{k+1}{s} f(x+sh)$$

for all $x,h \in \mathbf{R}^n$ and $k \in \mathbf{Z}^+ \cup \{0\}$.
(b) Prove that $D_h^k D_h^l = D_h^{k+l}$ for all $k,l \in \mathbf{Z}^+ \cup \{0\}$.

6.3.4. For $x \in \mathbf{R}$ let

$$f(x) = \sum_{k=1}^{\infty} 2^{-k} e^{2\pi i 2^k x}.$$

(a) Prove that $f \in \Lambda_\gamma(\mathbf{R})$ for all $0 < \gamma < 1$.
(b) Prove that there is an $A < \infty$ such that

$$\sup_{x,t\neq 0} |f(x+t) + f(x-t) - 2f(x)|\,|t|^{-1} \le A\,;$$

thus $f \in \Lambda_1(\mathbf{R})$.
(c) Show, however, that for all $x \in [0,1]$ we have

$$\sup_{0<|t|<1} |f(x+t) - f(x)|\,|t|^{-1} = \infty;$$

thus f is nowhere differentiable.
[*Hint:* Part (c): Use that $f(x)$ is 1-periodic and thus

$$\int_0^1 |f(x+t) - f(x)|^2\,dx = \sum_{k=1}^{\infty} 2^{-2k} |e^{2\pi i 2^k t} - 1|^2.$$

Observe that when $2^k|t| \le \frac{1}{2}$ we have $|e^{2\pi i 2^k t} - 1| \ge 2^{k+2}|t|$.]

6.3.5. For $0 < a,b < \infty$ and $x \in \mathbf{R}$ let

$$g_{ab}(x) = \sum_{k=1}^{\infty} 2^{-ak} e^{2\pi i 2^{bk} x}.$$

Show that g_{ab} lies in $\Lambda_{\frac{a}{b}}(\mathbf{R})$.
[*Hint:* Use the estimate $|D_h^L(e^{2\pi i 2^{bk} x})| \le C \min\big(1, (2^{bk}|h|)^L\big)$ with $L = [a/b] + 1$ and split the sum into two parts.]

6.3.6. Let $\gamma > 0$ and let $k = [\gamma]$.
(a) Use Exercise 6.3.3(b) to prove that if $|D_h^k(f)(x)| \le C|h|^\gamma$ for all $x,h \in \mathbf{R}^n$, then $|D_h^{k+l}(f)(x)| \le C2^l |h|^\gamma$ for all $l \ge 1$.
(b) Conversely, assuming that for some $l \ge 1$ we have

$$\sup_{x,h \in \mathbf{R}^n} \frac{|D_h^{k+l}(f)(x)|}{|h|^\gamma} < \infty,$$

show that $f \in \dot{\Lambda}_\gamma$.
$\big[$*Hint:* Part (b): Use (6.3.14) but replace $k+1$ by $k+l$.$\big]$

6.3.7. Let Ψ and Δ_j be as in Theorem 6.3.7. Define a continuous operator Q_t by setting

$$Q_t(f) = f * \Psi_t, \qquad \Psi_t(x) = t^{-n}\Psi(t^{-1}x).$$

Show that all tempered distributions f satisfy

$$\sup_{t>0} t^{-\gamma} \big\|Q_t(f)\big\|_{L^\infty} \approx \sup_{j\in\mathbf{Z}} 2^{j\gamma} \big\|\Delta_j(f)\big\|_{L^\infty}$$

with the interpretation that if either term is finite, then it controls the other term by a constant multiple of itself.
$\big[$*Hint:* Observe that $Q_t = Q_t(\Delta_{j-2}+\Delta_{j-1}+\Delta_j+\Delta_{j+1})$ when $2^{-j} \le t \le 2^{1-j}$.$\big]$

6.3.8. (a) Let $0 \le \gamma < 1$ and suppose that $\partial_j f \in \dot{\Lambda}_\gamma$ for all $1 \le j \le n$. Show that for some constant C we have

$$\big\|f\big\|_{\dot{\Lambda}_{\gamma+1}} \le C \sum_{j=1}^{n} \big\|\partial_j f\big\|_{\dot{\Lambda}_\gamma}$$

and conclude that $f \in \dot{\Lambda}_{\gamma+1}$.
(b) Let $\gamma \ge 0$. If we have $\partial^\alpha f \in \dot{\Lambda}_\gamma$ for all multi-indices α with $|\alpha| = r$, then there is an estimate

$$\big\|f\big\|_{\dot{\Lambda}_{\gamma+r}} \le C_\gamma \sum_{|\alpha|=r} \big\|\partial^\alpha f\big\|_{\dot{\Lambda}_\gamma},$$

and thus $f \in \dot{\Lambda}_{\gamma+r}$.
(c) Use Corollary 6.3.10 to obtain that the estimates in both (a) and (b) can be reversed.
$\big[$*Hint:* Part (a): Write

$$D_h^2(f)(x) = \int_0^1 \sum_{j=1}^{n} \big[\partial_j f(x+th+2h) - \partial_j f(x+th+h)\big]\, h_j\, dt\,.$$

Part (b): Use induction.$\big]$

6.3.9. Introduce a difference operator

$$\mathscr{D}^\beta(f)(x) = \left[\int_{\mathbf{R}^n} \frac{|D_y^{[\beta]+1}(f)(x)|^2}{|y|^{n+2\beta}}\, dy\right]^{\frac{1}{2}},$$

where $\beta > 0$. Show that for some constant $c_0(n,\beta)$ we have

$$\big\|\mathscr{D}^\beta(f)\big\|_{L^2(\mathbf{R}^n)}^2 = c_0(n,\beta) \int_{\mathbf{R}^n} |\widehat{f}(\xi)|^2 |\xi|^{2\beta}\, d\xi$$

for all functions $f \in \dot{L}^2_\beta(\mathbf{R}^n)$.

6.4 Hardy Spaces

Having been able to characterize L^p spaces, Sobolev spaces, and Lipschitz spaces using Littlewood–Paley theory, it should not come as a surprise that the theory can be used to characterize other spaces as well. This is the case with the Hardy spaces $H^p(\mathbf{R}^n)$, which form a family of spaces with some remarkable properties in which the integrability index p can go all the way down to zero.

There exists an abundance of equivalent characterizations for Hardy spaces, of which only a few representative ones are discussed in this section. A reader interested in going through the material quickly may define the Hardy space H^p as the space of all tempered distributions f modulo polynomials for which

$$\|f\|_{H^p} = \Big\| \Big(\sum_{j\in\mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} < \infty \tag{6.4.1}$$

whenever $0 < p \le 1$. An atomic decomposition for Hardy spaces can be obtained from this definition (see Section 6.6), and once this is in hand, the analysis of these spaces is significantly simplified. For historical reasons, however, we choose to define Hardy spaces using a more classical approach, and as a result, we have to go through a considerable amount of work to obtain the characterization alluded to in (6.4.1).

6.4.1 *Definition of Hardy Spaces*

To give the definition of Hardy spaces on $\mathbf{R}^n$, we need some background. We say that a tempered distribution v is *bounded* if $\varphi * v \in L^\infty(\mathbf{R}^n)$ whenever φ is in $\mathscr{S}(\mathbf{R}^n)$. We observe that if v is a bounded tempered distribution and $h \in L^1(\mathbf{R}^n)$, then the convolution $h * v$ can be defined as a distribution via the convergent integral

$$\langle h * v, \varphi \rangle = \langle \widetilde{\varphi} * v, \widetilde{h} \rangle = \int_{\mathbf{R}^n} (\widetilde{\varphi} * v)(x) \widetilde{h}(x)\, dx,$$

where φ is a Schwartz function, and as usual, we set $\widetilde{\varphi}(x) = \varphi(-x)$.

Let us recall the Poisson kernel P introduced in (2.1.13):

$$P(x) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \frac{1}{(1+|x|^2)^{\frac{n+1}{2}}} . \tag{6.4.2}$$

For $t > 0$, let $P_t(x) = t^{-n} P(t^{-1}x)$. If v is a bounded tempered distribution, then $P_t * v$ is a well defined distribution, since P_t is in L^1. We claim that $P_t * v$ can be identified with a well defined bounded function. To see this, write $1 = \widehat{\varphi}(\xi) + \eta(\xi)$, where $\widehat{\varphi}$ has compact support and η is a smooth function that vanishes in a neighborhood of the origin. Then the function ψ defined by $\widehat{\psi}(\xi) = e^{-2\pi|\xi|}\eta(\xi)$ is in the Schwartz class, and one has that

$$\widehat{P_t}(\xi) = e^{-2\pi t|\xi|} = e^{-2\pi t|\xi|}\widehat{\varphi}(t\xi) + \widehat{\psi}(t\xi)$$

is a sum of a compactly supported function and a Schwartz function. Then

$$P_t * v = P_t * (\varphi_t * v) + \psi_t * v,$$

but $\varphi_t * v$ and $\psi_t * v$ are bounded functions, since φ_t and ψ_t are in the Schwartz class. The last identity proves that $P_t * v$ is a bounded function.

Before we define Hardy spaces we introduce some notation.

Definition 6.4.1. Let $a,b > 0$. Let Φ be a Schwartz function and let f be a tempered distribution on $\mathbf{R}^n$. We define the *smooth maximal function of f with respect to Φ* as

$$M(f;\Phi)(x) = \sup_{t>0} |(\Phi_t * f)(x)|.$$

We define the *nontangential maximal function (with aperture a) of f with respect to Φ* as

$$M_a^*(f;\Phi)(x) = \sup_{t>0} \sup_{\substack{y\in\mathbf{R}^n \\ |y-x|\le at}} |(\Phi_t * f)(y)|.$$

We also define the *auxiliary maximal function*

$$M_b^{**}(f;\Phi)(x) = \sup_{t>0} \sup_{y\in\mathbf{R}^n} \frac{|(\Phi_t * f)(x-y)|}{(1+t^{-1}|y|)^b},$$

and we observe that

$$M(f;\Phi) \le M_a^*(f;\Phi) \le (1+a)^b M_b^{**}(f;\Phi) \tag{6.4.3}$$

for all $a,b > 0$. We note that if Φ is merely integrable, for example, if Φ is the Poisson kernel, the maximal functions $M(f;\Phi)$, $M_a^*(f;\Phi)$, and $M_b^{**}(f;\Phi)$ are well defined only for bounded tempered distributions f on $\mathbf{R}^n$.

For a fixed positive integer N and a Schwartz function φ we define the quantity

$$\mathfrak{N}_N(\varphi) = \int_{\mathbf{R}^n} (1+|x|)^N \sum_{|\alpha|\le N+1} |\partial^\alpha \varphi(x)|\,dx. \tag{6.4.4}$$

We now define

$$\mathscr{F}_N = \Big\{\varphi \in \mathscr{S}(\mathbf{R}^n):\ \mathfrak{N}_N(\varphi) \le 1\Big\}, \tag{6.4.5}$$

and we also define the *grand maximal function of f (with respect to N)* as

$$\mathscr{M}_N(f)(x) = \sup_{\varphi\in\mathscr{F}_N} M_1^*(f;\varphi)(x).$$

Having introduced a variety of smooth maximal operators useful in the development of the theory, we proceed with the definition of Hardy spaces.

Definition 6.4.2. Let f be a bounded tempered distribution on $\mathbf{R}^n$ and let $0 < p < \infty$. We say that f lies in the *Hardy space* $H^p(\mathbf{R}^n)$ if the *Poisson maximal function*

$$M(f;P)(x) = \sup_{t>0} |(P_t * f)(x)| \tag{6.4.6}$$

is in $L^p(\mathbf{R}^n)$. If this is the case, we set

$$\big\|f\big\|_{H^p} = \big\|M(f;P)\big\|_{L^p}.$$

At this point we don't know whether these spaces coincide with any other known spaces for some values of p. In the next theorem we show that this is the case when $1 < p < \infty$.

Theorem 6.4.3. *(a) Let* $1 < p < \infty$. *Then every bounded tempered distribution* f *in* H^p *is an element of* L^p. *Moreover, there is a constant* $C_{n,p}$ *such that for all such* f *we have*

$$\big\|f\big\|_{L^p} \le \big\|f\big\|_{H^p} \le C_{n,p}\big\|f\big\|_{L^p},$$

and therefore $H^p(\mathbf{R}^n)$ *coincides with* $L^p(\mathbf{R}^n)$.
(b) When $p = 1$, *every element of* H^1 *is an integrable function. In other words,* $H^1(\mathbf{R}^n) \subseteq L^1(\mathbf{R}^n)$ *and for all* $f \in H^1$ *we have*

$$\big\|f\big\|_{L^1} \le \big\|f\big\|_{H^1}. \tag{6.4.7}$$

Proof. (a) Let $f \in H^p(\mathbf{R}^n)$. The set $\{P_t * f : \, t > 0\}$ lies in a multiple of the unit ball of L^p. By the Banach–Alaoglu–Bourbaki theorem there exists a sequence $t_j \to 0$ such that $P_{t_j} * f$ converges to some L^p function f_0 in the weak* topology of L^p. On the other hand, we see that $P_t * \varphi \to \varphi$ in $\mathscr{S}(\mathbf{R}^n)$ as $t \to 0$ for all φ in $\mathscr{S}(\mathbf{R}^n)$. Thus

$$P_t * f \to f \qquad \text{in } \mathscr{S}'(\mathbf{R}^n), \tag{6.4.8}$$

and it follows that the distribution f coincides with the L^p function f_0. Since the family $\{P_t\}_{t>0}$ is an approximate identity, Theorem 1.2.19 gives that

$$\big\|P_t * f - f\big\|_{L^p} \to 0 \qquad \text{as } t \to 0,$$

from which it follows that

$$\big\|f\big\|_{L^p} \le \big\|\sup_{t>0} |P_t * f|\big\|_{L^p} = \big\|f\big\|_{H^p}. \tag{6.4.9}$$

The converse inequality is a consequence of the fact that

$$\sup_{t>0} |P_t * f| \le M(f),$$

where M is the Hardy–Littlewood maximal operator. (See Corollary 2.1.12.)

(b) The case $p=1$ requires only a small modification of the case $p>1$. Embedding L^1 into the space of finite Borel measures $\mathscr{M}$ whose unit ball is weak* compact, we can extract a sequence $t_j \to 0$ such that $P_{t_j} * f$ converges to some measure μ in the topology of measures. In view of (6.4.8), it follows that the distribution f can be identified with the measure μ.

It remains to show that μ is absolutely continuous with respect to Lebesgue measure, which would imply that it coincides with some L^1 function. Let $|\mu|$ be the total variation of μ. We show that μ is absolutely continuous by showing that for all subsets E of $\mathbf{R}^n$ we have $|E|=0 \implies |\mu|(E)=0$. Given an $\varepsilon>0$, there exists a $\delta>0$ such that for any measurable subset F of $\mathbf{R}^n$ we have

$$|F|<\delta \implies \int_F \sup_{t>0} |P_t * f|\,dx < \varepsilon\,.$$

Given E with $|E|=0$, we can find an open set U such that $E\subseteq U$ and $|U|<\delta$. Then for any g continuous function supported in U we have

$$\begin{aligned}\left|\int_{\mathbf{R}^n} g\,d\mu\right| &= \lim_{j\to\infty}\left|\int_{\mathbf{R}^n} g(x)\,(P_{t_j}*f)(x)\,dx\right| \\ &\le \big\|g\big\|_{L^\infty}\int_U \sup_{t>0}|P_t*f|\,dx \\ &< \varepsilon\big\|g\big\|_{L^\infty}\,.\end{aligned}$$

But we have

$$|\mu(U)| = \sup\left\{\left|\int_{\mathbf{R}^n} g\,d\mu\right| :\ g \text{ continuous supported in } U \text{ with } \big\|g\big\|_{L^\infty}\le 1\right\},$$

which implies that $|\mu(U)|<\varepsilon$. Since ε was arbitrary, it follows that $|\mu|(E)=0$; hence μ is absolutely continuous with respect to Lebesgue measure. Finally, (6.4.7) is a consequence of (6.4.9), which is also valid for $p=1$. □

We may wonder whether H^1 coincides with L^1. We show in Theorem 6.7.4 that elements of H^1 have integral zero; thus H^1 is a proper subspace of L^1.

We now proceed to obtain some characterizations of these spaces.

6.4.2 Quasinorm Equivalence of Several Maximal Functions

It is a fact that all the maximal functions of the preceding subsection have comparable L^p quasinorms for all $0<p<\infty$. This is the essence of the following theorem.

Theorem 6.4.4. *Let $0<p<\infty$. Then the following statements are valid:*
(a) There exists a Schwartz function Φ with $\int_{\mathbf{R}^n}\Phi(x)\,dx\neq 0$ and a constant C_1 (which does not depend on any parameter) such that

$$\big\|M(f;\Phi)\big\|_{L^p}\le C_1\big\|f\big\|_{H^p} \tag{6.4.10}$$

for all bounded $f \in \mathscr{S}'(\mathbf{R}^n)$.
(b) For every $a > 0$ *and* Φ *in* $\mathscr{S}(\mathbf{R}^n)$ *there exists a constant* $C_2(n,p,a,\Phi)$ *such that*

$$\big\|M_a^*(f;\Phi)\big\|_{L^p} \le C_2(n,p,a,\Phi)\big\|M(f;\Phi)\big\|_{L^p} \tag{6.4.11}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$.
(c) For every $a > 0$, $b > n/p$, *and* Φ *in* $\mathscr{S}(\mathbf{R}^n)$ *there exists a constant* $C_3(n,p,a,b,\Phi)$ *such that*

$$\big\|M_b^{**}(f;\Phi)\big\|_{L^p} \le C_3(n,p,a,b,\Phi)\big\|M_a^*(f;\Phi)\big\|_{L^p} \tag{6.4.12}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$.
(d) For every $b > 0$ *and* Φ *in* $\mathscr{S}(\mathbf{R}^n)$ *with* $\int_{\mathbf{R}^n} \Phi(x)\,dx \neq 0$ *there exists a constant* $C_4(b,\Phi)$ *such that if* $N = [b]+1$ *we have*

$$\big\|\mathscr{M}_N(f)\big\|_{L^p} \le C_4(b,\Phi)\big\|M_b^{**}(f;\Phi)\big\|_{L^p} \tag{6.4.13}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$.
(e) For every positive integer N *there exists a constant* $C_5(n,N)$ *such that every tempered distribution* f *with* $\big\|\mathscr{M}_N(f)\big\|_{L^p} < \infty$ *is a bounded distribution and satisfies*

$$\big\|f\big\|_{H^p} \le C_5(n,N)\big\|\mathscr{M}_N(f)\big\|_{L^p}\,, \tag{6.4.14}$$

that is, it lies in the Hardy space H^p.

We conclude that for $f \in H^p(\mathbf{R}^n)$, the inequality in (6.4.14) can be reversed, and therefore for *any* Schwartz function Φ with $\int_{\mathbf{R}^n} \Phi(x)\,dx \neq 0$, we have

$$\big\|M_a^*(f;\Phi)\big\|_{L^p} \le C(a,n,p,\Phi)\big\|f\big\|_{H^p}\,.$$

Consequently, there exists $N \in \mathbf{Z}^+$ large enough such that for $f \in \mathscr{S}'(\mathbf{R}^n)$ we have

$$\big\|\mathscr{M}_N(f)\big\|_{L^p} \approx \big\|M_b^{**}(f;\Phi)\big\|_{L^p} \approx \big\|M_a^*(f;\Phi)\big\|_{L^p} \approx \big\|M(f;\Phi)\big\|_{L^p} \approx \big\|f\big\|_{H^p}$$

for all Schwartz functions Φ with $\int_{\mathbf{R}^n} \Phi(x)\,dx \neq 0$ and constants that depend only on Φ,a,b,n,p. This furnishes a variety of characterizations for Hardy spaces.

The proof of this theorem is based on the following lemma.

Lemma 6.4.5. *Let* $m \in \mathbf{Z}^+$ *and let* Φ *in* $\mathscr{S}(\mathbf{R}^n)$ *satisfy* $\int_{\mathbf{R}^n} \Phi(x)\,dx = 1$. *Then there exists a constant* $C_0(\Phi,m)$ *such that for any* Ψ *in* $\mathscr{S}(\mathbf{R}^n)$, *there exist Schwartz functions* $\Theta^{(s)}$, $0 \le s \le 1$, *with the properties*

$$\Psi(x) = \int_0^1 (\Theta^{(s)} * \Phi_s)(x)\,ds \tag{6.4.15}$$

and

$$\int_{\mathbf{R}^n} (1+|x|)^m |\Theta^{(s)}(x)|\,dx \le C_0(\Phi,m)\,s^m\,\mathfrak{N}_m(\Psi). \tag{6.4.16}$$

Proof. We start with a smooth function ζ supported in $[0,1]$ that satisfies

$$\begin{aligned} 0 \le \zeta(s) \le & \frac{2s^m}{m!} && \text{for all } 0 \le s \le 1\,, \\ \zeta(s) = & \frac{s^m}{m!} && \text{for all } 0 \le s \le \frac{1}{2}\,, \\ \frac{d^r\zeta}{dt^r}(1) = & \; 0 && \text{for all } 0 \le r \le m+1\,. \end{aligned}$$

We define

$$\Theta^{(s)} = \Xi^{(s)} - \frac{d^{m+1}\zeta(s)}{ds^{m+1}} \overbrace{\Phi_s * \cdots * \Phi_s}^{m+1 \text{ terms}} * \Psi\,, \tag{6.4.17}$$

where

$$\Xi^{(s)} = (-1)^{m+1}\zeta(s)\frac{\partial^{m+1}}{\partial s^{m+1}}\Big(\overbrace{\Phi_s * \cdots * \Phi_s}^{m+2 \text{ terms}}\Big) * \Psi\,,$$

and we claim that (6.4.15) holds for this choice of $\Theta^{(s)}$. To verify this assertion, we apply $m+1$ integration by parts to write

$$\begin{aligned} \int_0^1 \Theta^{(s)} * \Phi_s\, ds = \int_0^1 \Xi^{(s)} * \Phi_s\, ds + \frac{d^m\zeta(s)}{ds^m}(0) \lim_{s\to 0+} \big(\overbrace{\Phi * \cdots * \Phi}^{m+2 \text{ terms}}\big)_s * \Psi \\ - (-1)^{m+1} \int_0^1 \zeta(s)\, \frac{\partial^{m+1}}{\partial s^{m+1}}\Big(\overbrace{\Phi_s * \cdots * \Phi_s}^{m+2 \text{ terms}}\Big) * \Psi\, ds\,, \end{aligned}$$

noting that all the boundary terms vanish except for the one in the first integration by parts at $s = 0$. The first and the third terms in the previous expression on the right add up to zero, while the second term is equal to Ψ, since Φ has integral one, which implies that the family $\{(\Phi * \cdots * \Phi)_s\}_{s>0}$ is an approximate identity as $s \to 0$. Therefore, (6.4.15) holds.

We now prove estimate (6.4.16). Let Ω be the $(m+1)$-fold convolution of Φ. For the second term on the right in (6.4.17), we note that the $(m+1)$st derivative of $\zeta(s)$ vanishes on $\left[0, \frac{1}{2}\right]$, so that we may write

$$\begin{aligned} &\int_{\mathbf{R}^n} (1+|x|)^m \Big|\frac{d^{m+1}\zeta(s)}{ds^{m+1}}\Big|\, |\Omega_s * \Psi(x)|\, dx \\ &\le\; C_m\, \chi_{[\frac{1}{2},1]}(s) \int_{\mathbf{R}^n} (1+|x|)^m \Big[\int_{\mathbf{R}^n} \tfrac{1}{s^n}\big|\Omega(\tfrac{x-y}{s})\big|\, |\Psi(y)|\, dy\Big]\, dx \\ &\le\; C_m\, \chi_{[\frac{1}{2},1]}(s) \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} (1+|y+sx|)^m |\Omega(x)|\, |\Psi(y)|\, dy\, dx \\ &\le\; C_m\, \chi_{[\frac{1}{2},1]}(s) \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} (1+|sx|)^m |\Omega(x)|\, (1+|y|)^m |\Psi(y)|\, dy\, dx \\ &\le\; C_m\, \chi_{[\frac{1}{2},1]}(s) \Big(\int_{\mathbf{R}^n} (1+|x|)^m |\Omega(x)|\, dx\Big)\Big(\int_{\mathbf{R}^n} (1+|y|)^m |\Psi(y)|\, dy\Big) \\ &\le\; C_0'(\Phi, m)\, s^m\, \mathfrak{N}_m(\Psi)\,, \end{aligned}$$

since $\chi_{[\frac{1}{2},1]}(s) \le 2^m s^m$. To obtain a similar estimate for the first term on the right in (6.4.17), we argue as follows:

$$\begin{aligned}
&\int_{\mathbf{R}^n} (1+|x|)^m |\zeta(s)| \left| \frac{d^{m+1}(\Omega_s * \Psi)}{ds^{m+1}}(x) \right| dx \\
&= \int_{\mathbf{R}^n} (1+|x|)^m \, |\zeta(s)| \left| \frac{d^{m+1}}{ds^{m+1}} \int_{\mathbf{R}^n} \frac{1}{s^n} \Omega\Big(\frac{x-y}{s}\Big) \Psi(y)\, dy \right| dx \\
&= \int_{\mathbf{R}^n} (1+|x|)^m \, |\zeta(s)| \left| \int_{\mathbf{R}^n} \Omega(y) \frac{d^{m+1}\Psi(x-sy)}{ds^{m+1}}\, dy \right| dx \\
&\le C_m' \int_{\mathbf{R}^n} (1+|x|)^m \, |\zeta(s)| \int_{\mathbf{R}^n} |\Omega(y)| \Big[\sum_{|\alpha|\le m+1} |\partial^\alpha \Psi(x-sy)| \, |y|^{|\alpha|} \Big] \, dy\, dx \\
&\le C_m' |\zeta(s)| \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} (1+|x+sy|)^m |\Omega(y)| \sum_{|\alpha|\le m+1} |\partial^\alpha \Psi(x)| \, (1+|y|)^{m+1}\, dy\, dx \\
&\le C_m' |\zeta(s)| \int_{\mathbf{R}^n} (1+|y|)^{m+1} \, |\Omega(y)| \, (1+|y|)^m \, dy \int_{\mathbf{R}^n} (1+|x|)^m \sum_{|\alpha|\le m+1} |\partial^\alpha \Psi(x)| \, dx \\
&\le C_0''(\Phi, m) \, s^m \, \mathfrak{N}_m(\Psi) \, .
\end{aligned}$$

We now set $C_0(\Phi,m) = C_0'(\Phi,m) + C_0''(\Phi,m)$ to conclude the proof of (6.4.16). $\square$

Next, we discuss the proof of Theorem 6.4.4.

Proof. (a) We pick a continuous and integrable function $\psi(s)$ on the interval $[1,\infty)$ that decays faster than the reciprocal of any polynomial (i.e., $|\psi(s)| \le C_N s^{-N}$ for all $N > 0$) such that

$$\int_1^\infty s^k \, \psi(s)\, ds = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k = 1,2,3,\dots . \end{cases} \tag{6.4.18}$$

Such a function exists; in fact, we may take

$$\psi(s) = \frac{e}{\pi} \frac{1}{s} \operatorname{Im} \Big(e^{(\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2})(s-1)^{\frac{1}{4}}} \Big) . \tag{6.4.19}$$

See Exercise 6.4.4. We now define a function

$$\Phi(x) = \int_1^\infty \psi(s) P_s(x)\, ds \, , \tag{6.4.20}$$

where P_s is the Poisson kernel. The Fourier transform Φ is

$$\widehat{\Phi}(\xi) = \int_1^\infty \psi(s) \widehat{P_s}(\xi)\, ds = \int_1^\infty \psi(s) e^{-2\pi s|\xi|}\, ds$$

(cf. Exercise 2.2.11), which is easily seen to be rapidly decreasing as $|\xi| \to \infty$. The same is true for all the derivatives of $\widehat{\Phi}$. The function $\widehat{\Phi}$ is clearly smooth on $\mathbf{R}^n \setminus \{0\}$. Moreover,

$$\partial_j \widehat{\Phi}(\xi) = \sum_{k=0}^{L-1} (-2\pi)^{k+1} \frac{|\xi|^k}{k!} \frac{\xi_j}{|\xi|} \int_1^\infty s^{k+1} \psi(s)\, ds + O(|\xi|^L) = O(|\xi|^L)$$

as $|\xi| \to 0$, which implies that the distributional derivative $\partial_j \widehat{\Phi}$ is continuous at the origin. Since

$$\partial_\xi^\alpha (e^{-2\pi s|\xi|}) = s^{|\alpha|} p_\alpha(\xi) |\xi|^{-m_\alpha} e^{-2\pi s|\xi|}$$

for some $m_\alpha \in \mathbf{Z}^+$ and some polynomial p_α, choosing L sufficiently large gives that every derivative of $\widehat{\Phi}$ is also continuous at the origin. We conclude that the function $\widehat{\Phi}$ is in the Schwartz class, and thus so is Φ. It also follows from (6.4.18) and (6.4.20) that

$$\int_{\mathbf{R}^n} \Phi(x)\, dx = 1 \neq 0\,.$$

Finally, we have the estimate

$$\begin{aligned} M(f;\Phi)(x) &= \sup_{t>0} |(\Phi_t * f)(x)| \\ &= \sup_{t>0} \left| \int_1^\infty \psi(s)(f * P_{ts})(x)\, ds \right| \\ &\le \int_1^\infty |\psi(s)|\, ds M(f;P)(x)\,, \end{aligned}$$

and the required conclusion follows with $C_1 = \int_1^\infty |\psi(s)|\, ds$. Note that we actually obtained the stronger pointwise estimate

$$M(f;\Phi) \le C_1 M(f;P)$$

rather than (6.4.10).

(b) The control of the nontagential maximal function $M_a^*(\cdot\,;\Phi)$ in terms of the vertical maximal function $M(\cdot\,;\Phi)$ is the hardest and most technical part of the proof. For matters of exposition, we present the proof only in the case that $a = 1$ and we note that the case of general $a > 0$ presents only notational differences. We derive (6.4.11) as a consequence of the estimate

$$\left(\int_{\mathbf{R}^n} M_1^*(f;\Phi)^{\varepsilon,N}(x)^p dx \right)^{\frac{1}{p}} \le C_2(n,p,N,\Phi) \big\| M(f;\Phi) \big\|_{L^p}\,, \tag{6.4.21}$$

where N is a large enough integer depending on f, $0 < \varepsilon < 1$, and

$$M_1^*(f;\Phi)^{\varepsilon,N}(x) = \sup_{0<t<\frac{1}{\varepsilon}} \sup_{|y-x|\le t} \big|(\Phi_t * f)(y)\big| \left(\frac{t}{t+\varepsilon}\right)^N \frac{1}{(1+\varepsilon|y|)^N}\,.$$

Let us a fix an element f in $\mathscr{S}'(\mathbf{R}^n)$ such that $M(f;\Phi) \in L^p$. We first show that $M_1^*(f;\Phi)^{\varepsilon,N}$ lies in $L^p(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$. Indeed, using (2.3.22) (with $\alpha = 0$), we obtain the following estimate for some constants C_f, m, and l (depending on f):

$$
\begin{aligned}
|(\Phi_t * f)(y)| &\leq C_f \sum_{|\beta|\leq l} \sup_{z\in\mathbf{R}^n} (|y|^m + |z|^m)|(\partial^\beta \widetilde{\Phi}_t)(z)| \\
&\leq C_f (1+|y|^m) \sum_{|\beta|\leq l} \sup_{z\in\mathbf{R}^n} (1+|z|^m)|(\partial^\beta \Phi_t)(-z)| \\
&\leq C_f \frac{(1+|y|^m)}{\min(t^n,t^{n+l})} \sum_{|\beta|\leq l} \sup_{z\in\mathbf{R}^n} (1+|z|^m)|(\partial^\beta \Phi)(-z/t)| \\
&\leq C_f \frac{(1+|y|)^m}{\min(t^n,t^{n+l})}(1+t^m) \sum_{|\beta|\leq l} \sup_{z\in\mathbf{R}^n} (1+|z/t|^m)|(\partial^\beta \Phi)(-z/t)| \\
&\leq C(f,\Phi)(1+\varepsilon|y|)^m \varepsilon^{-m}(1+t^m)(t^{-n}+t^{-n-l}).
\end{aligned}
$$

Multiplying by $(\frac{t}{t+\varepsilon})^N(1+\varepsilon|y|)^{-N}$ for some $0<t<\frac{1}{\varepsilon}$ and $|y-x|<t$ yields

$$
\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N} \leq C(f,\Phi)\frac{\varepsilon^{-m}(1+\varepsilon^{-m})(\varepsilon^{n-N}+\varepsilon^{n+l-N})}{(1+\varepsilon|y|)^{N-m}},
$$

and using that $1+\varepsilon|y|\geq \frac{1}{2}(1+\varepsilon|x|)$, we obtain for some $C(f,\Phi,\varepsilon,n,l,m,N)<\infty$,

$$
M_1^*(f;\Phi)^{\varepsilon,N}(x) \leq \frac{C(f,\Phi,\varepsilon,n,l,m,N)}{(1+\varepsilon|x|)^{N-m}}.
$$

Taking $N>(m+n)/p$, we deduce that $M_1^*(f;\Phi)^{\varepsilon,N}$ lies in $L^p(\mathbf{R}^n)\cap L^\infty(\mathbf{R}^n)$.

We now introduce a parameter $L>0$ and functions

$$
U(f;\Phi)^{\varepsilon,N}(x) = \sup_{0<t<\frac{1}{\varepsilon}} \sup_{|y-x|<t} t\big|\nabla(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N}
$$

and

$$
V(f;\Phi)^{\varepsilon,N,L}(x) = \sup_{0<t<\frac{1}{\varepsilon}} \sup_{y\in\mathbf{R}^n} \big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N}\Big(\frac{t}{t+|x-y|}\Big)^L.
$$

We fix an integer $L>n/p$. We need the norm estimate

$$
\big\|V(f;\Phi)^{\varepsilon,N,L}\big\|_{L^p} \leq C_{n,p}\big\|M_1^*(f;\Phi)^{\varepsilon,N}\big\|_{L^p} \tag{6.4.22}
$$

and the pointwise estimate

$$
U(f;\Phi)^{\varepsilon,N} \leq A(\Phi,N,n,p)\,V(f;\Phi)^{\varepsilon,N,L}, \tag{6.4.23}
$$

where

$$
A(\Phi,N,n,p) = 2^L C_0(\partial_j\Phi;N+L)\,\mathfrak{N}_{N+L}(\partial_j\Phi).
$$

To prove (6.4.22) we observe that when $z\in B(y,t)\subseteq B(x,|x-y|+t)$ we have

$$\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N} \le M_1^*(f;\Phi)^{\varepsilon,N}(z),$$

from which it follows that for any $0 < q < \infty$ and $y \in \mathbf{R}^n$,

$$\begin{aligned}
&\big|(\Phi_t * f)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N} \\
&\qquad \le \bigg(\frac{1}{|B(y,t)|}\int_{B(y,t)} M_1^*(f;\Phi)^{\varepsilon,N}(z)^q\,dz\bigg)^{\frac{1}{q}} \\
&\qquad \le \bigg(\frac{|x-y|+t}{t}\bigg)^{\frac{n}{q}}\bigg(\frac{1}{|B(x,|x-y|+t)|}\int_{B(x,|x-y|+t)} M_1^*(f;\Phi)^{\varepsilon,N}(z)^q\,dz\bigg)^{\frac{1}{q}} \\
&\qquad \le \bigg(\frac{|x-y|+t}{t}\bigg)^{L} M\Big(\big[M_1^*(f;\Phi)^{\varepsilon,N}\big]^q\Big)^{\frac{1}{q}}(x),
\end{aligned}$$

where we used that $L > n/p$. We now take $0 < q < p$ and we use the boundedness of the Hardy–Littlewood maximal operator M on $L^{p/q}$ to obtain (6.4.22).

In proving (6.4.23), we may assume that Φ has integral 1; otherwise we can multiply Φ by a suitable constant to arrange for this to happen. We note that

$$t\big|\nabla(\Phi_t * f)\big| = \big|(\nabla\Phi)_t * f\big| \le \sqrt{n}\sum_{j=1}^n |(\partial_j\Phi)_t * f|,$$

and it suffices to work with each partial derivative $\partial_j\Phi$ of Φ. Using Lemma 6.4.5 we write

$$\partial_j\Phi = \int_0^1 \Theta^{(s)} * \Phi_s\,ds$$

for suitable Schwartz functions $\Theta^{(s)}$. Fix $x \in \mathbf{R}^n$, $t > 0$, and y with $|y-x| < t < 1/\varepsilon$. Then we have

$$\begin{aligned}
&\big|\big((\partial_j\Phi)_t * f\big)(y)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N} \\
&\qquad = \Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y|)^N}\bigg|\int_0^1 \big((\Theta^{(s)})_t * \Phi_{st} * f\big)(y)\,ds\bigg| \qquad (6.4.24)\\
&\qquad \le \Big(\frac{t}{t+\varepsilon}\Big)^N \int_0^1\int_{\mathbf{R}^n} t^{-n}\big|\Theta^{(s)}(t^{-1}z)\big|\,\frac{\big|\big(\Phi_{st} * f\big)(y-z)\big|}{(1+\varepsilon|y|)^N}\,dz\,ds.
\end{aligned}$$

Inserting the factor 1 written as

$$\bigg(\frac{ts}{ts+|x-(y-z)|}\bigg)^L\Big(\frac{ts}{ts+\varepsilon}\Big)^N\bigg(\frac{ts+|x-(y-z)|}{ts}\bigg)^L\Big(\frac{ts+\varepsilon}{ts}\Big)^N$$

in the preceding z-integral and using that

$$\frac{1}{(1+\varepsilon|y|)^N} \le \frac{(1+\varepsilon|z|)^N}{(1+\varepsilon|y-z|)^N}$$

and the fact that $|x-y|<t<1/\varepsilon$, we obtain the estimate

$$\begin{aligned}
&\Big(\frac{t}{t+\varepsilon}\Big)^N \int_0^1 \int_{\mathbf{R}^n} t^{-n}\big|\Theta^{(s)}(t^{-1}z)\big| \frac{\big|\big(\Phi_{st}*f\big)(y-z)\big|}{(1+\varepsilon|y|)^N}\,dz\,ds \\
&\le V(f;\Phi)^{\varepsilon,N,L}(x)\int_0^1 \int_{\mathbf{R}^n} (1+\varepsilon|z|)^N \Big(\frac{ts+|x-(y-z)|}{ts}\Big)^L t^{-n}\big|\Theta^{(s)}(t^{-1}z)\big|\,dz\,\frac{ds}{s^N} \\
&\le V(f;\Phi)^{\varepsilon,N,L}(x)\int_0^1 \int_{\mathbf{R}^n} s^{-L-N}(1+\varepsilon t|z|)^N(s+1+|z|)^L\big|\Theta^{(s)}(z)\big|\,dz\,ds \\
&\le 2^L C_0(\partial_j\Phi;N+L)\,\mathfrak{N}_{N+L}(\partial_j\Phi)\,V(f;\Phi)^{\varepsilon,N,L}(x)
\end{aligned}$$

in view of conclusion (6.4.16) of Lemma 6.4.5. Combining this estimate with (6.4.24), we deduce (6.4.23). Having established both (6.4.22) and (6.4.23), we conclude that

$$\big\|U(f;\Phi)^{\varepsilon,N}\big\|_{L^p} \le C_{n,p}A(\Phi,N,n,p)\big\|M_1^*(f;\Phi)^{\varepsilon,N}\big\|_{L^p}. \tag{6.4.25}$$

We now set

$$E_\varepsilon = \big\{x\in\mathbf{R}^n:\ U(f;\Phi)^{\varepsilon,N}(x)\le KM_1^*(f;\Phi)^{\varepsilon,N}(x)\big\}$$

for some constant K to be determined shortly. With $A=A(\Phi,N,n,p)$ we have

$$\begin{aligned}
\int_{(E_\varepsilon)^c}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^p\,dx &\le \frac{1}{K^p}\int_{(E_\varepsilon)^c}\big[U(f;\Phi)^{\varepsilon,N}(x)\big]^p\,dx \\
&\le \frac{1}{K^p}\int_{\mathbf{R}^n}\big[U(f;\Phi)^{\varepsilon,N}(x)\big]^p\,dx \\
&\le \frac{C_{n,p}^pA^p}{K^p}\int_{\mathbf{R}^n}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^p\,dx \\
&\le \frac{1}{2}\int_{\mathbf{R}^n}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^p\,dx,
\end{aligned} \tag{6.4.26}$$

provided we choose K such that $K^p=2C_{n,p}^pA^p$. Obviously $K=K(\Phi,N,n,p)$, i.e., it depends on all these variables, in particular on N, which depends on f.

It remains to estimate the contribution of the integral of $\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^p$ over the set E_ε. We claim that the following pointwise estimate is valid:

$$M_1^*(f;\Phi)^{\varepsilon,N}(x)\le C_{n,N,K}\,M\big(M(f;\Phi)^r\big)^{\frac{1}{q}}(x) \tag{6.4.27}$$

for any $x\in E_\varepsilon$ and $0<q<\infty$. Note that $C_{n,N,K}$ depends on K. To prove (6.4.27) we fix $x\in E_\varepsilon$ and we also fix y such that $|y-x|<t$.

By the definition of $M_1^*(f;\Phi)^{\varepsilon,N}(x)$ there exists a point $(y_0,t)\in\mathbf{R}_+^{n+1}$ such that $|x-y_0|<t<\frac{1}{\varepsilon}$ and

$$\big|(\Phi_t * f)(y_0)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|y_0|)^N} \geq \frac{1}{2} M_1^*(f;\Phi)^{\varepsilon,N}(x). \tag{6.4.28}$$

Also by the definitions of E_ε and $U(f;\Phi)^{\varepsilon,N}$, for any $x \in E_\varepsilon$ we have

$$t\big|\nabla(\Phi_t * f)(\xi)\big|\Big(\frac{t}{t+\varepsilon}\Big)^N \frac{1}{(1+\varepsilon|\xi|)^N} \leq K M_1^*(f;\Phi)^{\varepsilon,N}(x) \tag{6.4.29}$$

for all ξ satisfying $|\xi - x| < t < \frac{1}{\varepsilon}$. It follows from (6.4.28) and (6.4.29) that

$$t\big|\nabla(\Phi_t * f)(\xi)\big| \leq 2K\big|(\Phi_t * f)(y_0)\big| \left(\frac{1+\varepsilon|\xi|}{1+\varepsilon|y_0|}\right)^N \tag{6.4.30}$$

for all ξ satisfying $|\xi - x| < t < \frac{1}{\varepsilon}$. We let z be such that $|z-x| < t$. Applying the mean value theorem and using (6.4.30), we obtain, for some ξ between y_0 and z,

$$\begin{aligned} \big|(\Phi_t * f)(z) - (\Phi_t * f)(y_0)\big| &= \big|\nabla(\Phi_t * f)(\xi)\big|\,|z-y_0| \\ &\leq \frac{2K}{t}\big|(\Phi_t * f)(\xi)\big|\left(\frac{1+\varepsilon|\xi|}{1+\varepsilon|y_0|}\right)^N |z-y_0| \\ &\leq \frac{2^{N+1}K}{t}\big|(\Phi_t * f)(y_0)\big|\,|z-y_0| \\ &\leq \frac{1}{2}\big|(\Phi_t * f)(y_0)\big|, \end{aligned}$$

provided z also satisfies $|z-y_0| < 2^{-N-2}K^{-1}t$ in addition to $|z-x| < t$. Therefore, for z satisfying $|z-y_0| < 2^{-N-2}K^{-1}t$ and $|z-x| < t$ we have

$$\big|(\Phi_t * f)(z)\big| \geq \frac{1}{2}\big|(\Phi_t * f)(y_0)\big| \geq \frac{1}{4} M_1^*(f;\Phi)^{\varepsilon,N}(x),$$

where the last inequality uses (6.4.28). Thus we have

$$\begin{aligned} M\big(M(f;\Phi)^q\big)(x) &\geq \frac{1}{|B(x,t)|}\int_{B(x,t)} \big[M(f;\Phi)(w)\big]^q\,dw \\ &\geq \frac{1}{|B(x,t)|}\int_{B(x,t)\cap B(y_0,2^{-N-2}K^{-1}t)} \big[M(f;\Phi)(w)\big]^q\,dw \\ &\geq \frac{1}{|B(x,t)|}\int_{B(x,t)\cap B(y_0,2^{-N-2}K^{-1}t)} \frac{1}{4^q}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^q\,dw \\ &\geq \frac{|B(x,t)\cap B(y_0,2^{-N-2}K^{-1}t)|}{|B(x,t)|}\frac{1}{4^q}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^r \\ &\geq C_{n,N,K}4^{-q}\big[M_1^*(f;\Phi)^{\varepsilon,N}(x)\big]^q, \end{aligned}$$

where we used the simple geometric fact that if $|x-y_0| \leq t$ and $\delta > 0$, then

$$\frac{|B(x,t)\cap B(y_0,\delta t)|}{|B(x,t)|} \geq c_{n,\delta} > 0\,,$$

the minimum of this constant being obtained when $|x-y_0| = t$. See Figure 6.1.

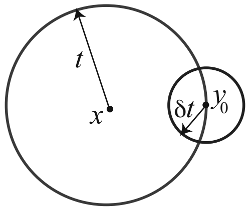

Fig. 6.1 The ball $B(y_0,\delta t)$ captures at least a fixed proportion of the ball $B(x,t)$.

This proves (6.4.27). Taking $r < p$ and applying the boundedness of the Hardy–Littlewood maximal operator yields

$$\int_{E_\varepsilon} \big[M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}(x)\big]^p\,dx \leq C'_{\boldsymbol{\Phi},N,K,n,p}\int_{\mathbf{R}^n} M(f;\boldsymbol{\Phi})(x)^p\,dx\,. \tag{6.4.31}$$

Combining this estimate with (6.4.26), we obtain

$$\int_{\mathbf{R}^n} \big[M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}\big]^p\,dx \leq C^p_{\boldsymbol{\Phi},N,K,n,p}\int_{\mathbf{R}^n} M(f;\boldsymbol{\Phi})^p\,dx + \frac{1}{2}\int_{\mathbf{R}^n} \big[M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}\big]^p\,dx\,,$$

and using the fact (obtained earlier) $\big\|M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}\big\|_{L^p} < \infty$, we obtain the required conclusion (6.4.11). This proves the inequality

$$\big\|M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}\big\|_{L^p} \leq 2^{1/p} C_{\boldsymbol{\Phi},N,K,n,p}\big\|M(f;\boldsymbol{\Phi})\big\|_{L^p}\,. \tag{6.4.32}$$

The previous constant depends on f but is independent of ε. Notice that

$$M_1^*(f;\boldsymbol{\Phi})^{\varepsilon,N}(x) \geq \frac{2^{-N}}{(1+\varepsilon|x|)^N}\sup_{0<t<1/\varepsilon}\Big(\frac{t}{t+\varepsilon}\Big)^N \sup_{|y-x|<t}\big|(\boldsymbol{\Phi}_t * f)(y)\big|$$

and that the preceding expression on the right increases to

$$2^{-N}M_1^*(f;\boldsymbol{\Phi})(x)$$

as $\varepsilon \downarrow 0$. Since the constant in (6.4.32) does not depend on ε, an application of the Lebesgue monotone convergence theorem yields

$$\big\|M_1^*(f;\boldsymbol{\Phi})\big\|_{L^p} \leq 2^{N+\frac{1}{p}} C_{\boldsymbol{\Phi},N,K,n,p}\big\|M(f;\boldsymbol{\Phi})\big\|_{L^p}\,. \tag{6.4.33}$$

The problem with this estimate is that the finite constant $2^N C_{\boldsymbol{\Phi},N,K,n,p}$ depends on N and thus on f. However, we have managed to show that under the assumption

$\big\|M(f;\Phi)\big\|_{L^p} < \infty$, one must necessarily have $\big\|M_1^*(f;\Phi)\big\|_{L^p} < \infty$. This is a significant observation that allows us now to repeat the preceding argument from the point where the functions $U(f;\phi)^{\varepsilon,N}$ and $V(f;\phi)^{\varepsilon,N,L}$ are introduced, setting $\varepsilon = N = 0$. Since the resulting constant no longer depends on the tempered distribution f, the required conclusion follows.

(c) As usual, $B(x,R)$ denotes a ball centered at x with radius R. It follows from the definition of $M_a^*(f;\Phi)$ that

$$|(\Phi_t * f)(y)| \le M_a^*(f;\Phi)(z) \qquad \text{if } z \in B(y,at)\,.$$

But the ball $B(y,at)$ is contained in the ball $B(x,|x-y|+at)$; hence it follows that

$$\begin{aligned} |(\Phi_t * f)(y)|^{\frac{n}{b}} &\le \frac{1}{|B(y,at)|}\int_{B(y,at)} M_a^*(f;\Phi)(z)^{\frac{n}{b}}\,dz \\ &\le \frac{1}{|B(y,at)|}\int_{B(x,|x-y|+at)} M_a^*(f;\Phi)(z)^{\frac{n}{b}}\,dz \\ &\le \left(\frac{|x-y|+at}{at}\right)^n M\big(M_a^*(f;\Phi)^{\frac{n}{b}}\big)(x) \\ &\le \max(1,a^{-n})\left(\frac{|x-y|}{t}+1\right)^n M\big(M_a^*(f;\Phi)^{\frac{n}{b}}\big)(x)\,, \end{aligned}$$

from which we conclude that for all $x \in \mathbf{R}^n$ we have

$$M_b^{**}(f;\Phi)(x) \le \max(1,a^{-n})\Big\{M\big(M_a^*(f;\Phi)^{\frac{n}{b}}\big)(x)\Big\}^{\frac{b}{n}}\,.$$

Raising to the power p and using the fact that $p > n/b$ and the boundedness of the Hardy–Littlewood maximal operator M on $L^{pb/n}$, we obtain the required conclusion (6.4.12).

(d) In proving (d) we may replace b by the integer $b_0 = [b]+1$. Let Φ be a Schwartz function with nonvanishing integral. Multiplying Φ by a constant, we can assume that Φ has integral equal to 1. Applying Lemma 6.4.5 with $m = b_0$, we write any function φ in $\mathscr{F}_N$ as

$$\varphi(y) = \int_0^1 (\Theta^{(s)} * \Phi_s)(y)\,ds$$

for some choice of Schwartz functions $\Theta^{(s)}$. Then we have

$$\varphi_t(y) = \int_0^1 ((\Theta^{(s)})_t * \Phi_{ts})(y)\,ds$$

for all $t > 0$. Fix $x \in \mathbf{R}^n$. Then for y in $B(x,t)$ we have

$$
\begin{aligned}
|(\varphi_t * f)(y)| \le\ & \int_0^1 \int_{\mathbf{R}^n} |(\Theta^{(s)})_t(z)|\,|(\Phi_{ts} * f)(y-z)|\,dz\,ds \\
\le\ & \int_0^1 \int_{\mathbf{R}^n} |(\Theta^{(s)})_t(z)|\,M^{**}_{b_0}(f;\Phi)(x)\left(\frac{|x-(y-z)|}{st}+1\right)^{b_0} dz\,ds \\
\le\ & \int_0^1 s^{-b_0} \int_{\mathbf{R}^n} |(\Theta^{(s)})_t(z)|\,M^{**}_{b_0}(f;\Phi)(x)\left(\frac{|x-y|}{t}+\frac{|z|}{t}+1\right)^{b_0} dz\,ds \\
\le\ & 2^{b_0} M^{**}_{b_0}(f;\Phi)(x) \int_0^1 s^{-b_0} \int_{\mathbf{R}^n} |\Theta^{(s)}(w)|\,\big(|w|+1\big)^{b_0}\,dw\,ds \\
\le\ & 2^{b_0} M^{**}_{b_0}(f;\Phi)(x) \int_0^1 s^{-b_0} C_0(\Phi,b_0)\, s^{b_0}\, \mathfrak{N}_{b_0}(\varphi)\,ds\,,
\end{aligned}
$$

where we applied conclusion (6.4.16) of Lemma 6.4.5. Setting $N = b_0 = [b]+1$, we obtain for y in $B(x,t)$ and $\varphi \in \mathscr{F}_N$,

$$
|(\varphi_t * f)(y)| \le 2^{b_0} C_0(\Phi,b_0)\, M^{**}_{b_0}(f;\Phi)(x)\,.
$$

Taking the supremum over all y in $B(x,t)$, over all $t>0$, and over all φ in $\mathscr{F}_N$, we obtain the pointwise estimate

$$
\mathscr{M}_N(f)(x) \le 2^{b_0} C_0(\Phi,b_0)\, M^{**}_{b_0}(f;\Phi)(x)\,, \qquad x \in \mathbf{R}^n,
$$

where $N = b_0+1$. This clearly yields (6.4.13) if we set $C_4 = 2^{b_0} C_0(\Phi,b_0)$.

(e) We fix an $f \in \mathscr{S}'(\mathbf{R}^n)$ that satisfies $\big\|\mathscr{M}_N(f)\big\|_{L^p} < \infty$ for some fixed positive integer N. To show that f is a bounded distribution, we fix a Schwartz function φ and we observe that for some positive constant $c = c_\varphi$, we have that $c\,\varphi$ is an element of $\mathscr{F}_N$ and thus $M_1^*(f;c\,\varphi) \le \mathscr{M}_N(f)$. Then

$$
\begin{aligned}
c^p\,|(\varphi * f)(x)|^p \ \le\ & \inf_{|y-x|\le 1} \sup_{|z-y|\le 1} |(c\,\varphi * f)(z)|^p \\
\le\ & \inf_{|y-x|\le 1} M_1^*(f;c\,\varphi)(y)^p \\
\le\ & \frac{1}{v_n} \int_{|y-x|\le 1} M_1^*(f;c\,\varphi)(y)^p\,dy \\
\le\ & \frac{1}{v_n} \int_{\mathbf{R}^n} M_1^*(f;c\,\varphi)(y)^p\,dy \\
\le\ & \frac{1}{v_n} \int_{\mathbf{R}^n} \mathscr{M}_N(f)(y)^p\,dy < \infty\,,
\end{aligned}
$$

which implies that $\varphi * f$ is a bounded function. We conclude that f is a bounded distribution. We now proceed to show that f is an element of H^p. We fix a smooth function with compact support θ such that

$$
\theta(x) = \begin{cases} 1 & \text{if} \quad |x| < 1, \\ 0 & \text{if} \quad |x| > 2. \end{cases}
$$

We observe that the identity

$$\begin{aligned} P(x) &= P(x)\theta(x)+\sum_{k=1}^{\infty}\big(\theta(2^{-k}x)P(x)-\theta(2^{-(k-1)}x)P(x)\big)\\ &= P(x)\theta(x)+\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\sum_{k=1}^{\infty}2^{-k}\bigg(\frac{\theta(\cdot)-\theta(2(\cdot))}{(2^{-2k}+|\cdot|^2)^{\frac{n+1}{2}}}\bigg)_{2^k}(x) \end{aligned}$$

is valid for all $x\in\mathbf{R}^n$. Setting

$$\Phi^{(k)}(x)=\big(\theta(x)-\theta(2x)\big)\frac{1}{(2^{-2k}+|x|^2)^{\frac{n+1}{2}}},$$

we note that for some fixed constant $c_0=c_0(n,N)$, the functions $c_0\,\theta\,P$ and $c_0\Phi^{(k)}$ lie in $\mathscr{F}_N$ uniformly in $k=1,2,3,\dots$. Combining this observation with the identity for $P(x)$ obtained earlier, we conclude that

$$\begin{aligned} \sup_{t>0}|P_t*f| &\le \sup_{t>0}|(\theta P)_t*f|+\frac{1}{c_0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\sup_{t>0}\sum_{k=1}^{\infty}2^{-k}\big|(c_0\Phi^{(k)})_{2^kt}*f\big|\\ &\le C_5(n,N)\mathscr{M}_N(f), \end{aligned}$$

which proves the required conclusion (6.4.14).

We observe that the last estimate also yields the stronger estimate

$$M_1^*(f;P)(x)=\sup_{t>0}\sup_{\substack{y\in\mathbf{R}^n\\ |y-x|\le at}}|(P_t*f)(y)|\le C_5(n,N)\mathscr{M}_N(f)(x). \tag{6.4.34}$$

It follows that the quasinorm $\big\|M_1^*(f;P)\big\|_{L^p(\mathbf{R}^n)}$ is also equivalent to $\big\|f\big\|_{H^p}$. This fact is very useful. □

Remark 6.4.6. To simplify the understanding of the equivalences just proved, a first-time reader may wish to define the H^p quasinorm of a distribution f as

$$\big\|f\big\|_{H^p}=\big\|M_1^*(f;P)\big\|_{L^p}$$

and then study only the implications (a) $\Longrightarrow$ (c), (c) $\Longrightarrow$ (d), (d) $\Longrightarrow$ (e), and (e) $\Longrightarrow$ (a) in the proof of Theorem 6.4.4. In this way one avoids passing through the statement in part (b). For many applications, the identification of $\big\|f\big\|_{H^p}$ with $\big\|M_1^*(f;\Phi)\big\|_{L^p}$ for some Schwartz function Φ (with nonvanishing integral) suffices.

We also remark that the proof of Theorem 6.4.4 yields

$$\big\|f\big\|_{H^p(\mathbf{R}^n)}\approx\big\|\mathscr{M}_N(f)\big\|_{L^p(\mathbf{R}^n)},$$

where $N=[\frac{n}{p}]+1$.

6.4.3 Consequences of the Characterizations of Hardy Spaces

In this subsection we look at a few consequences of Theorem 6.4.4. In many applications we need to be working with dense subspaces of H^p. It turns out that both $H^p \cap L^2$ and $H^p \cap L^1$ are dense in H^p.

Proposition 6.4.7. *Let $0 < p \le 1$ and let r satisfy $p \le r \le \infty$. Then $L^r \cap H^p$ is dense in H^p. Hence, $H^p \cap L^2$ and $H^p \cap L^1$ are dense in H^p.*

Proof. Let f be a distribution in $H^p(\mathbf{R}^n)$. Recall the Poisson kernel $P(x)$ and set $N = [\frac{n}{p}] + 1$. For any fixed $x \in \mathbf{R}^n$ and $t > 0$ we have

$$|(P_t * f)(x)| \le M_1^*(f;P)(y) \le C\mathscr{M}_N(f)(y) \tag{6.4.35}$$

for any $|y - x| \le t$. Indeed, the first estimate in (6.4.35) follows from the definition of $M_1^*(f;P)$, and the second estimate by (6.4.34). Raising (6.4.35) to the power p and averaging over the ball $B(x,t)$, we obtain

$$|(P_t * f)(x)|^p \le \frac{C^p}{v_n t^n} \int_{B(x,t)} \mathscr{M}_N(f)(y)^p \, dy \le \frac{C_1^p}{t^n} \|f\|_{H^p}^p .$$

It follows that the function $P_t * f$ is in $L^\infty(\mathbf{R}^n)$ with norm at most a constant multiple of $t^{-n/p}\|f\|_{H^p}$. Moreover, this function is also in $L^p(\mathbf{R}^n)$, since it is controlled by $M(f;P)$. Therefore, the functions $P_t * f$ lie in $L^r(\mathbf{R}^n)$ for all $r \le p \le \infty$. It remains to show that $P_t * f$ also lie in H^p and that $P_t * f \to f$ in H^p as $t \to 0$.

To see that $P_t * f$ lies in H^p, we use the semigroup formula $P_t * P_s = P_{t+s}$ for the Poisson kernel, which is a consequence of the fact that $\widehat{P_t}(\xi) = e^{-2\pi t|\xi|}$ by applying the Fourier transform. Therefore, for any $t > 0$ we have

$$\sup_{s>0} |P_s * P_t * f| = \sup_{s>0} |P_{s+t} * f| \le \sup_{s>0} |P_s * f| ,$$

which implies that

$$\big\|P_t * f\big\|_{H^p} \le \big\|f\big\|_{H^p}$$

for all $t > 0$. We now need to show that $P_t * f \to f$ in H^p as $t \to 0$. This will be a consequence of the Lebesgue dominated convergence theorem once we know that

$$\sup_{s>0} |(P_s * P_t * f - P_s * f)(x)| \to 0 \qquad \text{as} \quad t \to 0 \tag{6.4.36}$$

pointwise for all $x \in \mathbf{R}^n$ and also

$$\sup_{s>0} |P_s * P_t * f - P_s * f| \le 2 \sup_{s>0} |P_s * f| \in L^p(\mathbf{R}^n) . \tag{6.4.37}$$

Statement (6.4.37) is a trivial consequence of the Poisson semigroup formula. As far as (6.4.36) is concerned, we note that for all $x \in \mathbf{R}^n$ the function

$$s \mapsto |(P_s * P_t * f)(x) - (P_s * f)(x)| = |(P_{s+t} * f)(x) - (P_s * f)(x)|$$

is bounded by a constant multiple of $s^{-n/p}$ and therefore tends to zero as $s \to \infty$. Given any $\varepsilon > 0$, there exists an $M > 0$ such that for all $t > 0$ we have

$$\sup_{s>M} |(P_s * P_t * f - P_s * f)(x)| < \frac{\varepsilon}{2}. \tag{6.4.38}$$

Moreover, the function $t \mapsto \sup_{0\le s\le M} |(P_s * P_t * f - P_s * f)(x)|$ is continuous in t. Therefore, there exists a $t_0 > 0$ such that for $t < t_0$ we have

$$\sup_{0\le s\le M} |(P_s * P_t * f - P_s * f)(x)| < \frac{\varepsilon}{2}. \tag{6.4.39}$$

Combining (6.4.38) and (6.4.39) proves (6.4.36). □

Next we observe the following consequence of Theorem 6.4.4.

Corollary 6.4.8. *For any two Schwartz functions Φ and Θ with nonvanishing integral we have*

$$\big\| \sup_{t>0} |\Theta_t * f| \big\|_{L^p} \approx \big\| \sup_{t>0} |\Phi_t * f| \big\|_{L^p} \approx \big\| f \big\|_{H^p}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$, with constants depending only on n, p, Φ, and Θ.

Proof. See the discussion after Theorem 6.4.4. □

Next we define a *norm* on Schwartz functions relevant in the theory of Hardy spaces:

$$\mathfrak{N}_N(\varphi; x_0, R) = \int_{\mathbf{R}^n} \Big(1 + \Big|\frac{x - x_0}{R}\Big|\Big)^N \sum_{|\alpha|\le N+1} R^{|\alpha|} |\partial^\alpha \varphi(x)|\, dx.$$

Note that $\mathfrak{N}_N(\varphi; 0, 1) = \mathfrak{N}_N(\varphi)$.

Corollary 6.4.9. *(a) For any $0 < p \le 1$, any $f \in H^p(\mathbf{R}^n)$, and any $\varphi \in \mathscr{S}(\mathbf{R}^n)$ we have*

$$\big|\langle f, \varphi \rangle\big| \le \mathfrak{N}_N(\varphi) \inf_{|z|\le 1} \mathscr{M}_N(f)(z), \tag{6.4.40}$$

where $N = [\frac{n}{p}] + 1$. More generally, for any $x_0 \in \mathbf{R}^n$ and $R > 0$ we have

$$\big|\langle f, \varphi \rangle\big| \le \mathfrak{N}_N(\varphi; x_0, R) \inf_{|z-x_0|\le R} \mathscr{M}_N(f)(z). \tag{6.4.41}$$

(b) Let $0 < p \le 1$ and $p \le r \le \infty$. For any $f \in H^p$ we have the estimate

$$\big\| \varphi * f \big\|_{L^r} \le C(p,n) \mathfrak{N}_N(\varphi) \big\| f \big\|_{H^p},$$

where $N = [n/p] + 1$.

Proof. (a) Set $\psi(x) = \varphi(-Rx + x_0)$. It follows directly from Definition 6.4.1 that for any fixed z with $|z - x_0| \le R$ we have

$$\begin{aligned}\big|\langle f,\varphi\rangle\big| &= R^n|(f*\psi_R)(x_0)| \\ &\le \sup_{y:\,|y-z|\le R} R^n|(f*\psi_R)(y)| \\ &\le R^n\bigg[\int_{\mathbf{R}^n}(1+|w|)^N\sum_{|\alpha|\le N+1}|\partial^\alpha\psi(w)|\,dw\bigg]\mathscr{M}_N(f)(z),\end{aligned}$$

from which the second assertion in the corollary follows easily by the change of variables $x=-Rw+x_0$. Taking the infimum over all z with $|z-x_0|\le R$ yields the required conclusion.

(b) For any fixed $x\in\mathbf{R}^n$ and $t>0$ we have

$$|(\varphi*f)(x)|\le\mathfrak{N}_N(\varphi)M_1^*\Big(f;\frac{\varphi}{\mathfrak{N}_N(\varphi)}\Big)(y)\le\mathfrak{N}_N(\varphi)\mathscr{M}_N(f)(y) \tag{6.4.42}$$

for all y satisfying $|y-x|\le 1$. Hence

$$|(\varphi*f)(x)|^p\le\frac{\mathfrak{N}_N(\varphi)^p}{|B(x,1)|}\int_{B(x,1)}\mathscr{M}_N(f)^p(y)\,dy\le\mathfrak{N}_N(\varphi)^pC_{p,n}^p\big\|f\big\|_{H^p}^p.$$

This implies that $\big\|\varphi*f\big\|_{L^\infty}\le C_{p,n}\mathfrak{N}_N(\varphi)\big\|f\big\|_{H^p}$. Choosing $y=x$ in (6.4.42) and then taking L^p quasinorms yields a similar estimate for $\big\|\varphi*f\big\|_{L^p}$. By interpolation we deduce $\big\|\varphi*f\big\|_{L^r}\le\mathfrak{N}_N(\varphi)\big\|f\big\|_{H^p}$. □

Proposition 6.4.10. *Let* $0<p\le 1$. *Then the following statements are valid:*
(a) Convergence in H^p *implies convergence in* $\mathscr{S}'$.
(b) H^p *is a complete quasinormed metrizable space.*

Proof. Part (a) says that if a sequence f_j tends to f in $H^p(\mathbf{R}^n)$, then $f_j\to f$ in $\mathscr{S}'(\mathbf{R}^n)$. But this easily follows from the estimate

$$\big|\langle f,\varphi\rangle\big|\le C_\varphi\inf_{|z|\le 1}\mathscr{M}_N(f)(z)\le\frac{C_\varphi}{v_n}\int_{\mathbf{R}^n}\mathscr{M}_N(f)^p\,dz\le C_\varphi C_{n,p}\big\|f\big\|_{H^p}^p,$$

which is a direct consequence of (6.4.40) for all φ in $\mathscr{S}(\mathbf{R}^n)$. As before, here $N=[\frac{n}{p}]+1$.

To obtain the statement in (b), we first observe that the map $(f,g)\mapsto\big\|f-g\big\|_{H^p}^p$ is a metric on H^p that generates the same topology as the quasinorm $f\mapsto\big\|f\big\|_{H^p}$. To show that H^p is a complete space, it suffices to show that for any sequence of functions f_j that satisfies

$$\sum_j\int_{\mathbf{R}^n}\mathscr{M}_N(f_j)^p\,dx<\infty,$$

the series $\sum_j f_j$ converges in $H^p(\mathbf{R}^n)$. The partial sums of this series are Cauchy in $H^p(\mathbf{R}^n)$ and therefore are Cauchy in $\mathscr{S}'(\mathbf{R}^n)$ by part (a). It follows that the sequence $\sum_{-k}^k f_j$ converges to some tempered distribution f in $\mathscr{S}'(\mathbf{R}^n)$. Sublinearity gives

$$\int_{\mathbf{R}^n} \mathscr{M}_N(f)^p\,dx = \int_{\mathbf{R}^n} \mathscr{M}_N\Big(\sum_j f_j\Big)^p\,dx \le \sum_j \int_{\mathbf{R}^n} \mathscr{M}_N(f_j)^p\,dx < \infty,$$

which implies that $f \in H^p$. Finally,

$$\int_{\mathbf{R}^n} \mathscr{M}_N\Big(f - \sum_{j=-k}^{k} f_j\Big)^p\,dx \le \sum_{|j|\ge k+1} \int_{\mathbf{R}^n} \mathscr{M}_N(f_j)^p\,dx \to 0$$

as $k \to \infty$; thus the series converges in H^p. □

6.4.4 Vector-Valued H^p and Its Characterizations

We now obtain a vector-valued analogue of Theorem 6.4.4 crucial in the characterization of Hardy spaces using Littlewood–Paley theory. To state this analogue we need to extend the definitions of the maximal operators to sequences of distributions. Let $a,b > 0$ and let Φ be a Schwartz function on $\mathbf{R}^n$. In accordance with Definition 6.4.1, we give the following sequence of definitions.

Definition 6.4.11. For a sequence $\vec{f} = \{f_j\}_{j\in\mathbf{Z}}$ of tempered distributions on $\mathbf{R}^n$ we define the *smooth maximal function of* $\vec{f}$ *with respect to* Φ as

$$M(\vec{f};\Phi)(x) = \sup_{t>0} \big\|\{(\Phi_t * f_j)(x)\}_j\big\|_{\ell^2}.$$

We define the *nontangential maximal function (with aperture a) of f with respect to* Φ as

$$M_a^*(\vec{f};\Phi)(x) = \sup_{t>0} \sup_{\substack{y\in\mathbf{R}^n \\ |y-x|\le at}} \big\|\{(\Phi_t * f_j)(y)\}_j\big\|_{\ell^2}.$$

We also define the *auxiliary maximal function*

$$M_b^{**}(\vec{f};\Phi)(x) = \sup_{t>0} \sup_{y\in\mathbf{R}^n} \frac{\big\|\{(\Phi_t * f_j)(x-y)\}_j\big\|_{\ell^2}}{(1+t^{-1}|y|)^b}.$$

We note that if the function Φ is not assumed to be Schwartz but merely integrable, for example, if Φ is the Poisson kernel, the maximal functions $M(\vec{f};\Phi)$, $M_a^*(\vec{f};\Phi)$, and $M_b^{**}(\vec{f};\Phi)$ are well defined for sequences $\vec{f} = \{f_j\}_j$ whose terms are bounded tempered distributions on $\mathbf{R}^n$.

For a fixed positive integer N we define the *grand maximal function of* $\vec{f}$ *(with respect to N)* as

$$\mathscr{M}_N(\vec{f}) = \sup_{\varphi\in\mathscr{F}_N} M_1^*(\vec{f};\varphi), \tag{6.4.43}$$

where

$$\mathscr{F}_N = \Big\{\varphi \in \mathscr{S}(\mathbf{R}^n) : \mathfrak{N}_N(\varphi) \le 1\Big\}$$

is as defined in (6.4.5).

We note that as in the scalar case, we have the sequence of simple inequalities

$$M(\vec{f};\Phi) \le M_a^*(\vec{f};\Phi) \le (1+a)^b M_b^{**}(\vec{f};\Phi). \tag{6.4.44}$$

We now define the vector-valued Hardy space $H^p(\mathbf{R}^n,\ell^2)$.

Definition 6.4.12. Let $\vec{f} = \{f_j\}_j$ be a sequence of bounded tempered distributions on $\mathbf{R}^n$ and let $0 < p < \infty$. We say that $\vec{f}$ lies in the vector-valued Hardy space $H^p(\mathbf{R}^n,\ell^2)$ if the *Poisson maximal function*

$$M(\vec{f};P)(x) = \sup_{t>0} \big\| \{(P_t * f_j)(x)\}_j \big\|_{\ell^2}$$

lies in $L^p(\mathbf{R}^n)$. If this is the case, we set

$$\big\|\vec{f}\,\big\|_{H^p(\mathbf{R}^n,\ell^2)} = \big\|M(\vec{f};P)\big\|_{L^p(\mathbf{R}^n)} = \Big\| \sup_{\varepsilon>0} \Big(\sum_j |f_j * P_\varepsilon|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p(\mathbf{R}^n)}.$$

The next theorem provides a vector-valued analogue of Theorem 6.4.4.

Theorem 6.4.13. *Let $0 < p < \infty$. Then the following statements are valid:*
(a) There exists a Schwartz function Φ with $\int_{\mathbf{R}^n} \Phi(x)\,dx \neq 0$ and a constant C_1 (which does not depend on any parameters) such that

$$\big\|M(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \le C_1 \big\|\vec{f}\,\big\|_{H^p(\mathbf{R}^n,\ell^2)} \tag{6.4.45}$$

for every sequence $\vec{f} = \{f_j\}_j$ of tempered distributions.
(b) For every $a > 0$ and Φ in $\mathscr{S}(\mathbf{R}^n)$ there exists a constant $C_2(n,p,a,\Phi)$ such that

$$\big\|M_a^*(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \le C_2(n,p,a,\Phi) \big\|M(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \tag{6.4.46}$$

for every sequence $\vec{f} = \{f_j\}_j$ of tempered distributions.
(c) For every $a > 0$, $b > n/p$, and Φ in $\mathscr{S}(\mathbf{R}^n)$ there exists a constant $C_3(n,p,a,b,\Phi)$ such that

$$\big\|M_b^{**}(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \le C_3(n,p,a,b,\Phi) \big\|M_a^*(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \tag{6.4.47}$$

for every sequence $\vec{f} = \{f_j\}_j$ of tempered distributions.
(d) For every $b > 0$ and Φ in $\mathscr{S}(\mathbf{R}^n)$ with $\int_{\mathbf{R}^n} \Phi(x)\,dx \neq 0$ there exists a constant $C_4(b,\Phi)$ such that if $N = [\frac{n}{p}] + 1$ we have

$$\big\|\mathscr{M}_N(\vec{f})\big\|_{L^p(\mathbf{R}^n,\ell^2)} \le C_4(b,\Phi) \big\|M_b^{**}(\vec{f};\Phi)\big\|_{L^p(\mathbf{R}^n,\ell^2)} \tag{6.4.48}$$

for every sequence $\vec{f} = \{f_j\}_j$ of tempered distributions.
(e) For every positive integer N there exists a constant $C_5(n,N)$ such that every

sequence $\vec{f} = \{f_j\}_j$ of tempered distributions that satisfies $\big\|\mathscr{M}_N(\vec{f})\big\|_{L^p(\mathbf{R}^n,\ell^2)} < \infty$ consists of bounded distributions and satisfies

$$\|\vec{f}\,\|_{H^p(\mathbf{R}^n,\ell^2)} \le C_5(n,N)\big\|\mathscr{M}_N(\vec{f})\big\|_{L^p(\mathbf{R}^n,\ell^2)}, \tag{6.4.49}$$

that is, it lies in the Hardy space $H^p(\mathbf{R}^n,\ell^2)$.

Proof. The proof of this theorem is obtained via a step-by-step repetition of the proof of Theorem 6.4.4 in which the scalar absolute values are replaced by ℓ^2 norms. This is small notational change in our point of view but yields a significant improvement of the scalar version of the theorem. Moreover, this perspective provides an example of the power of Hilbert space techniques. The verification of the details of this step-by-step repetition of the proof of Theorem 6.4.4 are left to the reader. □

We end this subsection by observing the validity of the following vector-valued analogue of (6.4.41):

$$\Big(\sum_j \big|\langle f_j,\varphi\rangle\big|^2\Big)^{\frac{1}{2}} \le \mathfrak{N}_N(\varphi;x_0,R) \inf_{|z-x_0|\le R} \mathscr{M}_N(\vec{f})(z). \tag{6.4.50}$$

The proof of (6.4.50) is identical to the corresponding estimate for scalar-valued functions. Set $\psi(x) = \varphi(-Rx+x_0)$. It follows directly from Definition 6.4.11 that for any fixed z with $|z-x_0| \le R$ we have

$$\begin{aligned}\Big(\sum_j \big|\langle f_j,\varphi\rangle\big|^2\Big)^{\frac{1}{2}} &= R^n\big\|\{(f_j*\psi_R)(x_0)\}_j\big\|_{\ell^2}\\ &\le \sup_{y:\,|y-z|\le R} R^n\big\|\{(f_j*\psi_R)(y)\}_j\big\|_{\ell^2}\\ &\le R^n\,\mathfrak{N}_N(\psi)\,\mathscr{M}_N(\vec{f})(z),\end{aligned}$$

which, combined with the observation

$$R^n\,\mathfrak{N}_N(\psi) = \mathfrak{N}_N(\varphi;x_0,R),$$

yields the required conclusion by taking the infimum over all z with $|z-x_0| \le R$.

6.4.5 Singular Integrals on Hardy Spaces

To obtain the Littlewood–Paley characterization of Hardy spaces, we need a multiplier theorem for vector-valued Hardy spaces.

Suppose that $K_j(x)$ is a family of functions defined on $\mathbf{R}^n\setminus\{0\}$ that satisfies the following: There exist constants $A,B<\infty$ and an integer N such that for all multi-indices α with $|\alpha| \le N$ we have

$$\Big|\sum_{j\in\mathbf{Z}}\partial^\alpha K_j(x)\Big| \le A\,|x|^{-n-|\alpha|} < \infty \tag{6.4.51}$$

and also

$$\sup_{\xi\in\mathbf{R}^n}\Big|\sum_{j\in\mathbf{Z}}\widehat{K_j}(\xi)\Big| \le B < \infty\,. \tag{6.4.52}$$

Theorem 6.4.14. *Suppose that a sequence of kernels $\{K_j\}_j$ satisfies (6.4.51) and (6.4.52) with $N=[\frac{n}{p}]+1$, for some $0<p\le 1$. Then there exists a constant $C_{n,p}$ that depends only on the dimension n and on p such that for all sequences of tempered distributions $\{f_j\}_j$ we have the estimate*

$$\Big\|\sum_j K_j * f_j\Big\|_{H^p(\mathbf{R}^n)} \le C_{n,p}(A+B)\big\|\{f_j\}_j\big\|_{H^p(\mathbf{R}^n,\ell^2)}\,.$$

Proof. We fix a smooth positive function Φ supported in the unit ball $B(0,1)$ with $\int_{\mathbf{R}^n}\Phi(x)\,dx=1$ and we consider the sequence of smooth maximal functions

$$M\Big(\sum_j K_j*f_j;\Phi\Big)=\sup_{\varepsilon>0}\Big|\Phi_\varepsilon*\sum_j K_j*f_j\Big|\,,$$

which will be shown to be an element of $L^p(\mathbf{R}^n,\ell^2)$. We work with a fixed sequence of integrable functions $\vec{f}=\{f_j\}_j$, since such functions are dense in $L^p(\mathbf{R}^n,\ell^2)$ in view of Proposition 6.4.7.

We now fix a $\lambda>0$ and we set $N=[\frac{n}{p}]+1$. We also fix $\gamma>0$ to be chosen later and we define the set

$$\Omega_\lambda=\{x\in\mathbf{R}^n:\ \mathscr{M}_N(\vec{f})(x)>\gamma\lambda\}\,.$$

The set Ω_λ is open, and we may use the Whitney decomposition (Appendix J) to write it is a union of cubes Q_k such that

(a) $\bigcup_k Q_k=\Omega_\lambda$ and the Q_k's have disjoint interiors;

(b) $\sqrt{n}\,\ell(Q_k)\le \operatorname{dist}(Q_k,(\Omega_\lambda)^c)\le 4\sqrt{n}\,\ell(Q_k)$.

We denote by $c(Q_k)$ the center of the cube Q_k. For each k we set

$$d_k=\operatorname{dist}(Q_k,(\Omega_\lambda)^c)+2\sqrt{n}\,\ell(Q_k)\approx \ell(Q_k)\,,$$

so that

$$B(c(Q_k),d_k)\cap(\Omega_\lambda)^c\neq\emptyset\,.$$

We now introduce a partition of unity $\{\varphi_k\}_k$ adapted to the sequence of cubes $\{Q_k\}_k$ such that

(c) $\chi_{\Omega_\lambda}=\sum_k\varphi_k$ and each φ_k satisfies $0\le\varphi_k\le 1$;

(d) each φ_k is supported in $\frac{6}{5}\,Q_k$ and satisfies $\int_{\mathbf{R}^n}\varphi_k\,dx\approx d_k^n$;

(e) $\big\|\partial^\alpha \varphi_k\big\|_{L^\infty} \le C_\alpha d_k^{-|\alpha|}$ for all multi-indices α and some constants C_α.

We decompose each f_j as

$$f_j = g_j + \sum_k b_{j,k}\,,$$

where g_j is the *good function* of the decomposition given by

$$g_j = f_j \chi_{\mathbf{R}^n\setminus\Omega_\lambda} + \sum_k \frac{\int_{\mathbf{R}^n} f_j \varphi_k\,dx}{\int_{\mathbf{R}^n} \varphi_k\,dx}\varphi_k$$

and $b_j = \sum_k b_{j,k}$ is the *bad function* of the decomposition given by

$$b_{j,k} = \left(f_j - \frac{\int_{\mathbf{R}^n} f_j \varphi_k\,dx}{\int_{\mathbf{R}^n} \varphi_k\,dx}\right)\varphi_k\,.$$

We note that each $b_{j,k}$ has integral zero. We define $\vec{g} = \{g_j\}_j$ and $\vec{b} = \{b_j\}_j$. At this point we appeal to (6.4.50) and to properties (d) and (e) to obtain

$$\Big(\sum_j \Big|\frac{\int_{\mathbf{R}^n} f_j \varphi_k\,dx}{\int_{\mathbf{R}^n} \varphi_k\,dx}\Big|^2\Big)^{\frac{1}{2}} \le \frac{\mathfrak{N}_N\big(\varphi_k; c(Q_k), d_k\big)}{\int_{\mathbf{R}^n} \varphi_k\,dx} \inf_{|z-c(Q_k)|\le d_k} \mathscr{M}_N(\vec{f})(z)\,. \tag{6.4.53}$$

But since

$$\frac{\mathfrak{N}_N\big(\varphi_k; c(Q_k), d_k\big)}{\int_{\mathbf{R}^n} \varphi_k\,dx} \le \left[\int_{Q_k}\Big(1+\frac{|x-c(Q_k)|}{d_k}\Big)^N \sum_{|\alpha|\le N+1} \frac{d_k^{|\alpha|} C_\alpha d_k^{-|\alpha|}}{\int_{\mathbf{R}^n} \varphi_k\,dx}\,dx\right] \le C_{N,n}\,,$$

it follows that (6.4.53) is at most a constant multiple of λ, since the ball $B(c(Q_k), d_k)$ meets the complement of Ω_λ. We conclude that

$$\big\|\vec{g}\big\|_{L^\infty(\Omega_\lambda,\ell^2)} \le C_{N,n}\,\gamma\lambda\,. \tag{6.4.54}$$

We now turn to estimating $M(\sum_j K_j * b_{j,k};\Phi)$. For fixed k and $\varepsilon > 0$ we have

$$\begin{aligned}
&\big(\Phi_\varepsilon * \sum_j K_j * b_{j,k}\big)(x)\\
&= \int_{\mathbf{R}^n} \Phi_\varepsilon * \sum_j K_j(x-y)\left[f_j(y)\varphi_k(y) - \frac{\int_{\mathbf{R}^n} f_j\varphi_k\,dx}{\int_{\mathbf{R}^n}\varphi_k\,dx}\,\varphi_k(y)\right]dy\\
&= \int_{\mathbf{R}^n} \sum_j \left\{\big(\Phi_\varepsilon * K_j\big)(x-z) - \int_{\mathbf{R}^n}\big(\Phi_\varepsilon * K_j\big)(x-y)\frac{\varphi_k(y)}{\int_{\mathbf{R}^n}\varphi_k\,dx}\,dy\right\}\varphi_k(z) f_j(z)\,dz\\
&= \int_{\mathbf{R}^n} \sum_j R_{j,k}(x,z)\varphi_k(z) f_j(z)\,dz,
\end{aligned}$$

where we set $R_{j,k}(x,z)$ for the expression inside the curly brackets. Using (6.4.41), we obtain

$$
\begin{aligned}
&\left|\int_{\mathbf{R}^n}\sum_j R_{j,k}(x,z)\varphi_k(z)f_j(z)\,dz\right| \\
&\qquad\le \sum_j \mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big)\inf_{|z-c(Q_k)|\le d_k}\mathscr{M}_N(f_j)(z) \\
&\qquad\le \sum_j \mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big)\inf_{|z-c(Q_k)|\le d_k}\mathscr{M}_N(\vec{f}\,)(z)\,.
\end{aligned}
\tag{6.4.55}
$$

Since $\varphi_k(z)$ is supported in $\frac{6}{5}Q_k$, the term $(1+\frac{|z-c(Q_k)|}{d_k})^N$ contributes only a constant factor in the integral defining $\mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big)$, and we obtain

$$
\begin{aligned}
&\mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big) \\
&\qquad\le C_{N,n}\int_{\frac{6}{5}Q_k}\sum_{|\alpha|\le N+1} d_k^{|\alpha|+n}\left|\frac{\partial^\alpha}{\partial z^\alpha}\big(R_{j,k}(x,z)\varphi_k(z)\big)\right|dz\,.
\end{aligned}
\tag{6.4.56}
$$

For notational convenience we set $K_j^\varepsilon=\Phi_\varepsilon * K_j$. We observe that the family $\{K_j^\varepsilon\}_j$ satisfies (6.4.51) and (6.4.52) with constants A' and B' that are only multiples of A and B, respectively, uniformly in ε. We now obtain a pointwise estimate for $\mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big)$ when $x\in\mathbf{R}^n\setminus\Omega_\lambda$. We have

$$
R_{j,k}(x,z)\varphi_k(z)=\int_{\mathbf{R}^n}\varphi_k(z)\Big\{K_j^\varepsilon(x-z)-K_j^\varepsilon(x-y)\Big\}\frac{\varphi_k(y)\,dy}{\int_{\mathbf{R}^n}\varphi_k\,dx}\,,
$$

from which it follows that

$$
\left|\frac{\partial^\alpha}{\partial z^\alpha}R_{j,k}(x,z)\varphi_k(z)\right|\le\int_{\mathbf{R}^n}\left|\frac{\partial^\alpha}{\partial z^\alpha}\bigg\{\varphi_k(z)\Big[K_j^\varepsilon(x-z)-K_j^\varepsilon(x-y)\Big]\bigg\}\right|\frac{\varphi_k(y)\,dy}{\int_{\mathbf{R}^n}\varphi_k\,dx}\,.
$$

Using hypothesis (6.4.51), we can now easily obtain the estimate

$$
\sum_j\left|\frac{\partial^\alpha}{\partial z^\alpha}\bigg\{\varphi_k(z)\Big\{K_{i,j}^\varepsilon(x-z)-K_{i,j}^\varepsilon(x-y)\Big\}\bigg\}\right|\le C_{N,n}A\frac{d_k d_k^{-|\alpha|}}{|x-c(Q_k)|^{n+1}}
$$

for all $|\alpha|\le N$ and for $x\in\mathbf{R}^n\setminus\Omega_\lambda$, since for such x we have $|x-c(Q_k)|\ge c_n\,d_k$. It follows that

$$
d_k^{|\alpha|+n}\sum_j\left|\frac{\partial^\alpha}{\partial z^\alpha}\big\{R_{j,k}(x,z)\varphi_k(z)\big\}\right|\le C_{N,n}A\,d_k^n\left(\frac{d_k}{|x-c(Q_k)|^{n+1}}\right).
$$

Inserting this estimate in the summation of (6.4.56) over all j yields

$$
\sum_j\mathfrak{N}_N\big(R_{j,k}(x,\cdot\,)\varphi_k;c(Q_k),d_k\big)\le C_{N,n}A\left(\frac{d_k^{n+1}}{|x-c(Q_k)|^{n+1}}\right).
\tag{6.4.57}
$$

Combining (6.4.57) with (6.4.55) gives for $x\in\mathbf{R}^n\setminus\Omega_\lambda$,

$$\sum_j \left| \int_{\mathbf{R}^n} R_{i,j,k}(x,z)\varphi_k(z)f_j(z)\,dz \right| \le \frac{C_{N,n}A\,d_k^{n+1}}{|x-c(Q_k)|^{n+1}} \inf_{|z-c(Q_k)|\le d_k} \mathscr{M}_N(\vec{f})(z)\,.$$

This provides the estimate

$$\sup_{\varepsilon>0} \Big|\sum_j (K_j^\varepsilon * b_{j,k})(x)\Big| \le \frac{C_{N,n}A\,d_k^{n+1}}{|x-c(Q_k)|^{n+1}}\,\gamma\lambda$$

for all $x \in \mathbf{R}^n \setminus \Omega_\lambda$, since the ball $B(c(Q_k),d_k)$ intersects $(\Omega_\lambda)^c$. Summing over k results in

$$M\Big(\sum_j K_j * b_j;\boldsymbol{\Phi}\Big)(x) \le \sum_k \frac{C_{N,n}A\,\gamma\lambda\,d_k^{n+1}}{|x-c(Q_k)|^{n+1}} \le \sum_k \frac{C_{N,n}A\,\gamma\lambda\,d_k^{n+1}}{(d_k+|x-c(Q_k)|)^{n+1}}$$

for all $x \in (\Omega_\lambda)^c$. The last sum is known as the *Marcinkiewicz function*. It is a simple fact that

$$\int_{\mathbf{R}^n} \sum_k \frac{d_k^{n+1}}{(d_k+|x-c(Q_k)|)^{n+1}}\,dx \le C_n \sum_k |Q_k| = C_n\,|\Omega_\lambda|;$$

see Exercise 4.6.6. We have therefore shown that

$$\int_{\mathbf{R}^n} M(\vec{K}*\vec{b};\boldsymbol{\Phi})(x)\,dx \le C_{N,n}A\gamma\lambda\,|\Omega_\lambda|\,, \tag{6.4.58}$$

where we used the notation $\vec{K}*\vec{b} = \sum_j K_j * b_j$.

We now combine the information we have acquired so far. First we have

$$\big|\{M(\vec{K}*\vec{f};\boldsymbol{\Phi})>\lambda\}\big| \le \big|\{M(\vec{K}*\vec{g};\boldsymbol{\Phi})>\tfrac{\lambda}{2}\}\big| + \big|\{M(\vec{K}*\vec{b};\boldsymbol{\Phi})>\tfrac{\lambda}{2}\}\big|\,.$$

For the good function $\vec{g}$ we have the estimate

$$\begin{aligned}
\big|\{M(\vec{K}*\vec{g};\boldsymbol{\Phi})>\tfrac{\lambda}{2}\}\big| &\le \frac{4}{\lambda^2}\int_{\mathbf{R}^n} M(\vec{K}*\vec{g};\boldsymbol{\Phi})(x)^2\,dx \\
&\le \frac{4}{\lambda^2}\sum_j \int_{\mathbf{R}^n} M(K_j*g_j)(x)^2\,dx \\
&\le \frac{C_nB^2}{\lambda^2}\int_{\mathbf{R}^n}\sum_j |g_j(x)|^2\,dx \\
&\le \frac{C_nB^2}{\lambda^2}\int_{\Omega_\lambda}\sum_j |g_j(x)|^2\,dx + \frac{C_nB^2}{\lambda^2}\int_{(\Omega_\lambda)^c}\sum_j |f_j(x)|^2\,dx \\
&\le B^2C_{N,n}\gamma^2\,|\Omega_\lambda| + \frac{C_nB^2}{\lambda^2}\int_{(\Omega_\lambda)^c} \mathscr{M}_N(\vec{f})(x)^2\,dx\,,
\end{aligned}$$

where we used Corollary 2.1.12, the L^2 boundedness of the Hardy–Littlewood maximal operator, hypothesis (6.4.52), the fact that $f_j = g_j$ on $(\Omega_\lambda)^c$, estimate (6.4.54), and the fact that $\|\vec{f}\|_{\ell^2} \le \mathscr{M}_N(\vec{f})$ in the sequence of estimates.

On the other hand, estimate (6.4.58) and Chebyshev's inequality gives

$$\big|\{M(\vec{K}*\vec{b};\Phi) > \tfrac{\lambda}{2}\}\big| \le C_{N,n} A\,\gamma|\Omega_\lambda|\,,$$

which, combined with the previously obtained estimate for $\vec{g}$, gives

$$\big|\big\{M(\vec{K}*\vec{f};\Phi) > \lambda\big\}\big| \le C_{N,n}(A\,\gamma + B^2\gamma^2)\,|\Omega_\lambda| + \frac{C_n B^2}{\lambda^2}\int_{(\Omega_\lambda)^c} \mathscr{M}_N(\vec{f})(x)^2\,dx\,.$$

Multiplying this estimate by $p\lambda^{p-1}$, recalling that $\Omega_\lambda = \{\mathscr{M}_N(\vec{f}) > \gamma\lambda\}$, and integrating in λ from 0 to ∞, we can easily obtain

$$\big\|M(\vec{K}*\vec{f};\Phi)\big\|^p_{L^p(\mathbf{R}^n,\ell^2)} \le C_{N,n}(A\gamma^{1-p} + B^2\gamma^{2-p})\big\|\mathscr{M}_N(\vec{f})\big\|^p_{L^p(\mathbf{R}^n,\ell^2)}\,. \tag{6.4.59}$$

Choosing $\gamma = (A+B)^{-1}$ and recalling that $N = [\frac{n}{p}]+1$ gives the required conclusion for some constant $C_{n,p}$ that depends only on n and p.

Finally, use density to extend this estimate to all $\vec{f}$ in $H^p(\mathbf{R}^n,\ell^2)$. □

6.4.6 The Littlewood–Paley Characterization of Hardy Spaces

We discuss an important characterization of Hardy spaces in terms of Littlewood–Paley square functions. The vector-valued Hardy spaces and the action of singular integrals on them are crucial tools in obtaining this characterization.

We first set up the notation. We fix a radial Schwartz function Ψ on $\mathbf{R}^n$ whose Fourier transform is nonnegative, supported in the annulus $\frac{1}{2}+\frac{1}{10} \le |\xi| \le 2-\frac{1}{10}$, and satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1 \tag{6.4.60}$$

for all $\xi \neq 0$. Associated with this bump, we define the Littlewood–Paley operators Δ_j given by multiplication on the Fourier transform side by the function $\widehat{\Psi}(2^{-j}\xi)$, that is,

$$\Delta_j(f) = \Delta_j^{\Psi}(f) = \Psi_{2^{-j}} * f\,. \tag{6.4.61}$$

We have the following.

Theorem 6.4.15. *Let Ψ be a radial Schwartz function on $\mathbf{R}^n$ whose Fourier transform is nonnegative, supported in $\frac{1}{2}+\frac{1}{10} \le |\xi| \le 2-\frac{1}{10}$, and satisfies* (6.4.60)*. Let Δ_j be the Littlewood–Paley operators associated with Ψ and let $0 < p \le 1$. Then there exists a constant $C = C_{n,p,\Psi}$ such that for all $f \in H^p(\mathbf{R}^n)$ we have*

$$\Big\| \Big(\sum_{j\in\mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \le C \big\| f \big\|_{H^p} . \tag{6.4.62}$$

Conversely, suppose that a tempered distribution f satisfies

$$\Big\| \Big(\sum_{j\in\mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} < \infty . \tag{6.4.63}$$

Then there exists a unique polynomial $Q(x)$ such that $f - Q$ lies in the Hardy space H^p and satisfies the estimate

$$\frac{1}{C} \big\| f - Q \big\|_{H^p} \le \Big\| \Big(\sum_{j\in\mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} . \tag{6.4.64}$$

Proof. We fix $\Phi \in \mathscr{S}(\mathbf{R}^n)$ with integral equal to 1 and we take $f \in H^p \cap L^1$ and M in $\mathbf{Z}^+$. Let r_j be the Rademacher functions, introduced in Appendix C.1, reindexed so that their index set is the set of all integers (not the set of nonnegative integers). We begin with the estimate

$$\Big| \sum_{j=-M}^{M} r_j(\omega) \Delta_j(f) \Big| \le \sup_{\varepsilon>0} \Big| \Phi_\varepsilon * \sum_{j=-M}^{M} r_j(\omega) \Delta_j(f) \Big| ,$$

which holds since $\{\Phi_\varepsilon\}_{\varepsilon>0}$ is an approximate identity. We raise this inequality to the power p, we integrate over $x \in \mathbf{R}^n$ and $\omega \in [0,1]$, and we use the maximal function characterization of H^p [Theorem 6.4.4 (a)] to obtain

$$\int_0^1 \int_{\mathbf{R}^n} \Big| \sum_{j=-M}^{M} r_j(\omega) \Delta_j(f)(x) \Big|^p dx\, d\omega \le C_{p,n}^p \int_0^1 \Big\| \sum_{j=-M}^{M} r_j(\omega) \Delta_j(f) \Big\|_{H^p}^p d\omega .$$

The lower inequality for the Rademacher functions in Appendix C.2 gives

$$\int_{\mathbf{R}^n} \Big(\sum_{j=-M}^{M} |\Delta_j(f)(x)|^2 \Big)^{\frac{p}{2}} dx \le C_p^p C_{p,n}^p \int_0^1 \Big\| \sum_{j=-M}^{M} r_j(\omega) \Delta_j(f) \Big\|_{H^p}^p d\omega ,$$

where the second estimate is a consequence of Theorem 6.4.14 (we need only the scalar version here), since the kernel

$$\sum_{k=-M}^{M} r_k(\omega) \Psi_{2^{-k}}(x)$$

satisfies (6.4.51) and (6.4.52) with constants A and B depending only on n and Ψ (and, in particular, independent of M). We have now proved that

$$\Big\| \Big(\sum_{j=-M}^{M} |\Delta_j(f)|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \le C_{n,p,\Psi} \big\| f \big\|_{H^p} ,$$

from which (6.4.62) follows directly by letting $M \to \infty$. We have now established (6.4.62) for $f \in H^p \cap L^1$. Using density, we can extend this estimate to all $f \in H^p$.

To obtain the converse estimate, for $r \in \{0,1,2\}$ we consider the sets

$$3\mathbf{Z}+r = \{3k+r:\ k \in \mathbf{Z}\},$$

and we observe that for $j,k \in 3\mathbf{Z}+r$ the Fourier transforms of $\Delta_j(f)$ and $\Delta_k(f)$ are disjoint if $j \neq k$. We fix a Schwartz function η whose Fourier transform is compactly supported away from the origin so that for all $j,k \in 3\mathbf{Z}$ we have

$$\Delta_j^\eta \Delta_k = \begin{cases} \Delta_j & \text{when } j = k, \\ 0 & \text{when } j \neq k, \end{cases} \tag{6.4.65}$$

where Δ_j^η is the Littlewood–Paley operator associated with the bump η, that is, $\Delta_j^\eta(f) = f * \eta_{2^{-j}}$. It follows from Theorem 6.4.14 that the map

$$\{f_j\}_{j\in\mathbf{Z}} \to \sum_{j\in 3\mathbf{Z}} \Delta_j^\eta(f_j)$$

maps $H^p(\mathbf{R}^n, \ell^2)$ to $H^p(\mathbf{R}^n)$. Indeed, we can see easily that

$$\Big| \sum_{j\in 3\mathbf{Z}} \widehat{\eta}(2^{-j}\xi) \Big| \le B$$

and

$$\sum_{j\in 3\mathbf{Z}} \big|\partial^\alpha\big(2^{jn}\eta(2^j x)\big)\big| \le A_\alpha |x|^{-n-|\alpha|}$$

for all multi-indices α and for constants depending only on B and A_α. Applying this estimate with $f_j = \Delta_j(f)$ and using (6.4.65) yields the estimate

$$\Big\| \sum_{j\in 3\mathbf{Z}} \Delta_j(f) \Big\|_{H^p} \le C_{n,p,\Psi} \Big\| \Big(\sum_{j\in 3\mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac12} \Big\|_{L^p}$$

for all distributions f that satisfy (6.4.63). Applying the same idea with $3\mathbf{Z}+1$ and $3\mathbf{Z}+2$ replacing $3\mathbf{Z}$ and summing the corresponding estimates gives

$$\Big\| \sum_{j\in \mathbf{Z}} \Delta_j(f) \Big\|_{H^p} \le 3^{\frac1p} C_{n,p,\Psi} \Big\| \Big(\sum_{j\in \mathbf{Z}} |\Delta_j(f)|^2 \Big)^{\frac12} \Big\|_{L^p}.$$

But note that $f - \sum_j \Delta_j(f)$ is equal to a polynomial $Q(x)$, since its Fourier transform is supported at the origin. It follows that $f - Q$ lies in H^p and satisfies (6.4.64). □

We show in the next section that the square function characterization of H^p is independent of the choice of the underlying function Ψ.

Exercises

6.4.1. Prove that if v is a bounded tempered distribution and h_1, h_2 are in $\mathscr{S}(\mathbf{R}^n)$, then

$$(h_1 * h_2) * v = h_1 * (h_2 * v).$$

6.4.2. (a) Show that the H^1 norm remains invariant under the L^1 dilation $f_t(x) = t^{-n} f(t^{-1}x)$.
(b) Show that the H^p norm remains invariant under the L^p dilation $t^{n-n/p} f_t(x)$ interpreted in the sense of distributions.

6.4.3. (a) Let $1 < q \le \infty$ and let g in $L^q(\mathbf{R}^n)$ be a compactly supported function with integral zero. Show that g lies in the Hardy space $H^1(\mathbf{R}^n)$.
(b) Prove the same conclusion when L^q is replaced by $L\log^+ L$.
$\big[$*Hint:* Part (a): Pick a $\mathscr{C}_0^\infty$ function Φ supported in the unit ball with nonvanishing integral and suppose that the support of g is contained in the ball $B(0,R)$. For $|x| \le 2R$ we have that $M(f;\Phi)(x) \le C_\Phi M(g)(x)$, and since $M(g)$ lies in L^q, it also lies in $L^1(B(0,2R))$. For $|x| > 2R$, write $(\Phi_t * g)(x) = \int_{\mathbf{R}^n} \big(\Phi_t(x-y) - \Phi_t(x)\big) g(y)\, dy$ and use the mean value theorem to estimate this expression by $t^{-n-1}\big\|\nabla\Phi\big\|_{L^\infty}\big\|g\big\|_{L^1} \le |x|^{-n-1} C_\Phi \big\|g\big\|_{L^q}$, since $t \ge |x-y| \ge |x| - |y| \ge |x|/2$ whenever $|x| \ge 2R$ and $|y| \le R$. Thus $M(f;\Phi)$ lies in $L^1(\mathbf{R}^n)$. Part (b): Use Exercise 2.1.4(a) to deduce that $M(g)$ is integrable over $B(0,2R)$.$\big]$

6.4.4. Show that the function $\psi(s)$ defined in (6.4.19) is continuous and integrable over $[1,\infty)$, decays faster than the reciprocal of any polynomial, and satisfies (6.4.18), that is,

$$\int_1^\infty s^k\, \psi(s)\, ds = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k = 1,2,3,\dots. \end{cases}$$

$\big[$*Hint:* Apply Cauchy's theorem over a suitable contour.$\big]$

6.4.5. Let $0 < a < \infty$ be fixed. Show that a bounded tempered distribution f lies in H^p if and only if the nontangential Poisson maximal function

$$M_a^*(f;P)(x) = \sup_{t>0} \sup_{\substack{y\in\mathbf{R}^n \\ |y-x|\le at}} |(P_t * f)(y)|$$

lies in L^p, and in this case we have $\big\|f\big\|_{H^p} \approx \big\|M_a^*(f;P)\big\|_{L^p}$.
$\big[$*Hint:* Observe that $M(f;P)$ can be replaced with $M_a^*(f;P)$ in the proof of parts (a) and (e) of Theorem 6.4.4).$\big]$

6.4.6. Show that for every integrable function g with mean value zero and support inside a ball B, we have $M(g;\Phi) \in L^p((3B)^c)$ for $p > n/(n+1)$. Here Φ is in $\mathscr{S}$.

6.4.7. Show that the space of all Schwartz functions whose Fourier transform is supported away from a neighborhood of the origin is dense in H^p.
$\big[$*Hint:* Use the square function characterization of H^p.$\big]$

6.4.8. (a) Suppose that $f \in H^p(\mathbf{R}^n)$ for some $0 < p \le 1$ and Φ in $\mathscr{S}(\mathbf{R}^n)$. Then show that for all $t > 0$ the function $\Phi_t * f$ belongs to $L^r(\mathbf{R}^n)$ for all $p \le r \le \infty$. Find an estimate for the L^r norm of $\Phi_t * f$ in terms of $\|f\|_{H^p}$ and $t > 0$.
(b) Let $0 < p \le 1$. Show that there exists a constant $C_{n,p}$ such that for all f in $H^p(\mathbf{R}^n) \cap L^1(\mathbf{R}^n)$ we have

$$|\widehat{f}(\xi)| \le C_{n,p} |\xi|^{\frac{n}{p}-n} \|f\|_{H^p}.$$

[*Hint:* Obtain that

$$\|\Phi_t * f\|_{L^1} \le C t^{-n/p+n} \|f\|_{H^p},$$

using an idea from the proof of Proposition 6.4.7.]

6.4.9. Show that $H^p(\mathbf{R}^n, \ell^2) = L^p(\mathbf{R}^n, \ell^2)$ whenever $1 < p < \infty$ and that $H^1(\mathbf{R}^n, \ell^2)$ is contained in $L^1(\mathbf{R}^n, \ell^2)$.

6.4.10. For a sequence of tempered distributions $\vec{f} = \{f_j\}_j$, define the following variant of the grand maximal function:

$$\widetilde{\mathscr{M}}_N(\vec{f})(x) = \sup_{\{\varphi_j\}_j \in \mathscr{F}_N} \sup_{\varepsilon > 0} \sup_{\substack{y \in \mathbf{R}^n \\ |y-x| < \varepsilon}} \Big(\sum_j \big| ((\varphi_j)_\varepsilon * f_j)(y) \big|^2 \Big)^{\frac{1}{2}},$$

where $N \ge [\frac{n}{p}] + 1$ and

$$\mathscr{F}_N = \Big\{ \{\varphi_j\}_j \in \mathscr{S}(\mathbf{R}^n) : \sum_j \mathfrak{N}_N(\varphi_j) \le 1 \Big\}.$$

Show that for all sequences of tempered distributions $\vec{f} = \{f_j\}_j$ we have

$$\big\|\widetilde{\mathscr{M}}_N(\vec{f})\big\|_{L^p(\mathbf{R}^n,\ell^2)} \approx \big\|\mathscr{M}_N(\vec{f})\big\|_{L^p(\mathbf{R}^n,\ell^2)}$$

with constants depending only on n and p.
[*Hint:* Fix Φ in $\mathscr{S}(\mathbf{R}^n)$ with integral 1. Using Lemma 6.4.5, write

$$(\varphi_j)_t(y) = \int_0^1 ((\Theta_j^{(s)})_t * \Phi_{ts})(y)\, ds$$

and apply a vector-valued extension of the proof of part (d) of Theorem 6.4.4 to obtain the pointwise estimate

$$\widetilde{\mathscr{M}}_N(\vec{f}) \le C_{n,p} M_m^{**}(\vec{f}; \Phi),$$

where $m > n/p$.]

6.5 Besov–Lipschitz and Triebel–Lizorkin Spaces

The main achievement of the previous sections was the remarkable characterization of Sobolev, Lipschitz, and Hardy spaces using the Littlewood–Paley operators Δ_j. These characterizations motivate the introduction of classes of spaces defined in terms of expressions involving the operators Δ_j. These scales furnish a general framework within which one can launch a study of function spaces from a unified perspective.

We have encountered two expressions involving the operators Δ_j in the characterizations of the function spaces obtained in the previous sections. Some spaces were characterized by an L^p norm of the Littlewood–Paley square function

$$\Big(\sum_j |2^{j\alpha}\Delta_j(f)|^2\Big)^{\frac{1}{2}},$$

and other spaces were characterized by an ℓ^q norm of the sequence of quantities $\big\|2^{j\alpha}\Delta_j(f)\big\|_{L^p}$. Examples of spaces in the first case are the homogeneous Sobolev spaces, Hardy spaces, and, naturally, L^p spaces. We have studied only one example of spaces in the second category, the Lipschitz spaces, in which case $p = q = \infty$. These examples motivate the introduction of two fundamental scales of function spaces, called the Triebel–Lizorkin and Besov–Lipschitz spaces, respectively.

6.5.1 Introduction of Function Spaces

Before we give the pertinent definitions, we recall the setup that we developed in Section 6.2 and used in Section 6.3. Throughout this section we fix a radial Schwartz function Ψ on $\mathbf{R}^n$ whose Fourier transform is nonnegative, is supported in the annulus $1-\frac{1}{7} \le |\xi| \le 2$, is equal to one on the smaller annulus $1 \le |\xi| \le 2-\frac{2}{7}$, and satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1\,, \qquad \xi \neq 0\,. \tag{6.5.1}$$

Associated with this bump, we define the Littlewood–Paley operators $\Delta_j = \Delta_j^{\Psi}$ given by multiplication on the Fourier transform side by the function $\widehat{\Psi}(2^{-j}\xi)$. We also define a Schwartz function Φ such that

$$\widehat{\Phi}(\xi) = \begin{cases} \sum_{j\le 0} \widehat{\Psi}(2^{-j}\xi) & \text{when } \xi \neq 0, \\ 1 & \text{when } \xi = 0. \end{cases} \tag{6.5.2}$$

Note that $\widehat{\Phi}(\xi)$ is equal to 1 for $|\xi| \le 2-\frac{2}{7}$ and vanishes when $|\xi| \ge 2$. It follows from these definitions that

$$S_0 + \sum_{j=1}^{\infty} \Delta_j = I\,, \tag{6.5.3}$$

where $S_0 = S_0^{\Psi}$ is the operator given by convolution with the bump Φ and the convergence of the series in (6.5.3) is in $\mathscr{S}'(\mathbf{R}^n)$. Moreover, we also have the identity

$$\sum_{j\in\mathbf{Z}} \Delta_j = I, \tag{6.5.4}$$

where the convergence of the series in (6.5.4) is in the sense of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$.

Definition 6.5.1. Let $\alpha \in \mathbf{R}$ and $0 < p,q \le \infty$. For $f \in \mathscr{S}'(\mathbf{R}^n)$ we set

$$\big\|f\big\|_{B_p^{\alpha,q}} = \big\|S_0(f)\big\|_{L^p} + \Big(\sum_{j=1}^{\infty} \big(2^{j\alpha}\big\|\Delta_j(f)\big\|_{L^p}\big)^q\Big)^{\frac{1}{q}}$$

with the obvious modification when $p,q=\infty$. When $p,q<\infty$ we also set

$$\big\|f\big\|_{F_p^{\alpha,q}} = \big\|S_0(f)\big\|_{L^p} + \Big\|\Big(\sum_{j=1}^{\infty} \big(2^{j\alpha}|\Delta_j(f)|\big)^q\Big)^{\frac{1}{q}}\Big\|_{L^p}.$$

The space of all tempered distributions f for which the quantity $\big\|f\big\|_{B_p^{\alpha,q}}$ is finite is called the (inhomogeneous) *Besov–Lipschitz* space with indices α, p, q and is denoted by $B_p^{\alpha,q}$. The space of all tempered distributions f for which the quantity $\big\|f\big\|_{F_p^{\alpha,q}}$ is finite is called the (inhomogeneous) *Triebel–Lizorkin* space with indices α, p, q and is denoted by $F_p^{\alpha,q}$.

We now define the corresponding homogeneous versions of these spaces. For an element f of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ we let

$$\big\|f\big\|_{\dot B_p^{\alpha,q}} = \Big(\sum_{j\in\mathbf{Z}} \big(2^{j\alpha}\big\|\Delta_j(f)\big\|_{L^p}\big)^q\Big)^{\frac{1}{q}}$$

and

$$\big\|f\big\|_{\dot F_p^{\alpha,q}} = \Big\|\Big(\sum_{j\in\mathbf{Z}} \big(2^{j\alpha}|\Delta_j(f)|\big)^q\Big)^{\frac{1}{q}}\Big\|_{L^p}.$$

The space of all f in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ for which the quantity $\big\|f\big\|_{\dot B_p^{\alpha,q}}$ is finite is called the (homogeneous) *Besov–Lipschitz* space with indices α, p, q and is denoted by $\dot B_p^{\alpha,q}$. The space of f in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$ such that $\big\|f\big\|_{\dot F_p^{\alpha,q}} < \infty$ is called the (homogeneous) *Triebel–Lizorkin* space with indices α, p, q and is denoted by $\dot F_p^{\alpha,q}$.

We now make several observations related to these definitions. First we note that the expressions $\big\|\cdot\big\|_{\dot F_p^{\alpha,q}}$, $\big\|\cdot\big\|_{F_p^{\alpha,q}}$, $\big\|\cdot\big\|_{\dot B_p^{\alpha,q}}$, and $\big\|\cdot\big\|_{B_p^{\alpha,q}}$ are built in terms of L^p quasinorms of ℓ^q quasinorms of $2^{j\alpha}\Delta_j$ or ℓ^q quasinorms of L^p quasinorms of the same expressions. As a result, we can see that these quantities satisfy the triangle inequality with a constant (which may be taken to be 1 when $1 \le p,q < \infty$). To determine whether these quantities are indeed quasinorms, we need to check whether the following property holds:

$$\|f\|_X = 0 \implies f = 0\,, \tag{6.5.5}$$

where X is one of the $\dot{F}_p^{\alpha,q}$, $F_p^{\alpha,q}$, $\dot{B}_p^{\alpha,q}$, and $B_p^{\alpha,q}$. Since these are spaces of distributions, the identity $f=0$ in (6.5.5) should be interpreted in the sense of distributions. If $\|f\|_X = 0$ for some inhomogeneous space X, then $S_0(f)=0$ and $\Delta_j(f)=0$ for all $j \geq 1$. Using (6.5.3), we conclude that $f=0$; thus the quantities $\|\cdot\|_{F_p^{\alpha,q}}$ and $\|\cdot\|_{B_p^{\alpha,q}}$ are indeed quasinorms. Let us investigate what happens when $\|f\|_X = 0$ for some homogeneous space X. In this case we must have $\Delta_j(f)=0$, and using (6.5.4) we conclude that $\widehat{f}$ must be supported at the origin. Proposition 2.4.1 yields that f must be a polynomial and thus f must be zero (since distributions whose difference is a polynomial are identified in homogeneous spaces).

Remark 6.5.2. We interpret the previous definition in certain cases. According to what we have seen so far, we have

$$\begin{aligned}
\dot{F}_p^{0,2} &\approx F_p^{0,2} \approx L^p\,, && 1<p<\infty,\\
\dot{F}_p^{0,2} &\approx H^p\,, && 0<p\leq 1,\\
F_p^{s,2} &\approx L_s^p\,, && 1<p<\infty,\\
\dot{F}_p^{s,2} &\approx \dot{L}_s^p\,, && 1<p<\infty,\\
B_\infty^{\gamma,\infty} &\approx \Lambda_\gamma\,, && \gamma>0,\\
\dot{B}_\infty^{\gamma,\infty} &\approx \dot{\Lambda}_\gamma\,, && \gamma>0\,,
\end{aligned}$$

where $\approx$ indicates that the corresponding norms are equivalent.

Although in this text we restrict attention to the case $p<\infty$, it is noteworthy mentioning that when $p=\infty$, $\dot{F}_\infty^{0,q}$ can be defined as the space of all $f \in \mathscr{S}'/\mathscr{P}$ that satisfy

$$\|f\|_{\dot{F}_\infty^{\alpha,q}} = \sup_{Q \text{ dyadic cube}} \int_Q \frac{1}{|Q|} \bigg(\sum_{j=-\log_2 \ell(Q)}^{\infty} (2^{j\alpha}|\Delta_j(f)|)^q \bigg)^{\frac{1}{q}} < \infty\,.$$

In the particular case $q=2$ and $\alpha=0$, the space obtained in this way is called *BMO* and coincides with the space introduced and studied in Chapter 7; this space serves as a substitute for L^∞ and plays a fundamental role in analysis. It should now be clear that several important spaces in analysis can be thought of as elements of the scale of Triebel–Lizorkin spaces.

It would have been more natural to denote Besov–Lipschitz and Triebel–Lizorkin spaces by $B_{\alpha,q}^p$ and $F_{\alpha,q}^p$ to maintain the upper and lower placements of the corresponding indices analogous to those in the previously defined Lebesgue, Sobolev, Lipschitz, and Hardy spaces. However, the notation in Definition 6.5.1 is more or less prevalent in the field of function spaces, and we adhere to it.

6.5.2 Equivalence of Definitions

It is not clear from the definitions whether the finiteness of the quasinorms defining the spaces $B_p^{\alpha,q}$, $F_p^{\alpha,q}$,$\dot{B}_p^{\alpha,q}$, and $\dot{F}_p^{\alpha,q}$ depends on the choice of the function Ψ (recall that Φ is determined by Ψ). We show that if Ω is another function that satisfies (6.5.1) and Θ is defined in terms of Ω in the same way that Φ is defined in terms of Ψ, [i.e., via (6.5.2)], then the norms defined in Definition 6.5.1 with respect to the pairs (Φ,Ψ) and (Θ,Ω) are comparable. To prove this we need the following lemma.

Lemma 6.5.3. *Let $0<c_0<\infty$ and $0<r<\infty$. Then there exist constants C_1 and C_2 (which depend only on n, c_0, and r) such that for all $t>0$ and for all $\mathscr{C}^1$ functions u on $\mathbf{R}^n$ whose Fourier transform is supported in the ball $|\xi|\le c_0 t$ and that satisfy $|u(z)|\le B(1+|z|)^{\frac{n}{r}}$ for some $B>0$ we have the estimate*

$$\sup_{z\in\mathbf{R}^n}\frac{1}{t}\frac{|\nabla u(x-z)|}{(1+t|z|)^{\frac{n}{r}}}\le C_1\sup_{z\in\mathbf{R}^n}\frac{|u(x-z)|}{(1+t|z|)^{\frac{n}{r}}}\le C_2 M(|u|^r)(x)^{\frac{1}{r}}\,, \tag{6.5.6}$$

where M denotes the Hardy–Littlewood maximal operator. (The constants C_1 and C_2 are independent of B.)

Proof. Select a Schwartz function ψ whose Fourier transform is supported in the ball $|\xi|\le 2c_0$ and is equal to 1 on the smaller ball $|\xi|\le c_0$. Then $\widehat{\psi}(\frac{\xi}{t})$ is equal to 1 on the support of $\widehat{u}$ and we can write

$$u(x-z)=\int_{\mathbf{R}^n}t^n\psi(t(x-z-y))u(y)\,dy\,.$$

Taking partial derivatives and using that ψ is a Schwartz function, we obtain

$$|\nabla u(x-z)|\le C_N\int_{\mathbf{R}^n}t^{n+1}(1+t|x-z-y|)^{-N}|u(y)|\,dy\,,$$

where N is arbitrarily large. Using that for all $x,y,z\in\mathbf{R}^n$ we have

$$1\le(1+t|x-z-y|)^{\frac{n}{r}}\frac{(1+t|z|)^{\frac{n}{r}}}{(1+t|x-y|)^{\frac{n}{r}}}\,,$$

we obtain

$$\frac{1}{t}\frac{|\nabla u(x-z)|}{(1+t|z|)^{\frac{n}{r}}}\le C_N\int_{\mathbf{R}^n}t^n(1+t|x-z-y|)^{\frac{n}{r}-N}\frac{|u(y)|}{(1+t|x-y|)^{\frac{n}{r}}}\,dy\,,$$

from which the first estimate in (6.5.6) follows easily.

Let $|y|\le\delta$ for some $\delta>0$ to be chosen later. We now use the mean value theorem to write

$$u(x-z)=(\nabla u)(x-z-\xi_y)\cdot y+u(x-z-y)$$

for some ξ_y satisfying $|\xi_y|\le|y|\le\delta$. This implies that

$$|u(x-z)| \le \sup_{|w|\le|z|+\delta} |(\nabla u)(x-w)|\,\delta + |u(x-z-y)|.$$

Raising to the power r, averaging over the ball $|y| \le \delta$, and then raising to the power $\frac{1}{r}$ yields

$$|u(x-z)| \le c_r \left[\sup_{|w|\le|z|+\delta} |(\nabla u)(x-w)|\,\delta + \left(\frac{1}{v_n \delta^n} \int_{|y|\le\delta} |u(x-z-y)|^r \, dy \right)^{\frac{1}{r}} \right]$$

with $c_r = \max(2^{1/r}, 2^r)$. Here v_n is the volume of the unit ball in $\mathbf{R}^n$. Then

$$\frac{|u(x-z)|}{(1+t|z|)^{\frac{n}{r}}} \le c_r \left[\sup_{|w|\le|z|+\delta} \frac{|(\nabla u)(x-w)|}{(1+t|z|)^{\frac{n}{r}}}\,\delta\, \frac{\left(\frac{1}{v_n \delta^n} \int_{|y|\le\delta+|z|} |u(x-y)|^r \, dy \right)^{\frac{1}{r}}}{(1+t|z|)^{\frac{n}{r}}} \right].$$

We now set $\delta = \varepsilon/t$ for some $\varepsilon \le 1$. Then we have

$$|w| \le |z| + \frac{\varepsilon}{t} \implies \frac{1}{1+t|z|} \le \frac{2}{1+t|w|},$$

and we can use this to obtain the estimate

$$\frac{|u(x-z)|}{(1+t|z|)^{\frac{n}{r}}} \le c_{r,n} \left[\sup_{w\in\mathbf{R}^n} \frac{1}{t} \frac{|(\nabla u)(x-w)|}{(1+t|w|)^{\frac{n}{r}}}\,\varepsilon\, \frac{\left(\frac{t^n}{v_n \varepsilon^n} \int_{|y|\le\frac{1}{t}+|z|} |u(x-y)|^r \, dy \right)^{\frac{1}{r}}}{(1+t|z|)^{\frac{n}{r}}} \right]$$

with $c_{r,n} = \max(2^{1/r}, 2^r) 2^{n/r}$. It follows that

$$\sup_{z\in\mathbf{R}^n} \frac{|u(x-z)|}{(1+t|z|)^{\frac{n}{r}}} \le c_{r,n} \left[\sup_{w\in\mathbf{R}^n} \frac{1}{t} \frac{|(\nabla u)(x-w)|}{(1+t|w|)^{\frac{n}{r}}}\,\varepsilon + \varepsilon^{-\frac{n}{r}} M(|u|^r)(x)^{\frac{1}{r}} \right].$$

Taking $\varepsilon = \frac{1}{2}(c_{r,n} C_1)^{-1}$, where C_1 is the constant in (6.5.6), we obtain the second estimate in (6.5.6) with $C_2 = 2\varepsilon^{-n/r}$. At this step we used the hypothesis that

$$\sup_{z\in\mathbf{R}^n} \frac{|u(x-z)|}{(1+t|z|)^{\frac{n}{r}}} \le \sup_{z\in\mathbf{R}^n} \frac{B(1+|x|+|z|)^{\frac{n}{r}}}{(1+t|z|)^{\frac{n}{r}}} < \infty.$$

This concludes the proof of the lemma. □

Remark 6.5.4. The reader is reminded that $\widehat{u}$ in Lemma 6.5.3 may not be a function; for example, this is the case when u is a polynomial (say of degree $[n/r]$). If $\widehat{u}$ were an integrable function, then u would be a bounded function, and condition $|u(x)| \le B(1+|x|)^{\frac{n}{r}}$ would not be needed.

We now return to a point alluded to earlier, that changing Ψ by another bump Ω that satisfies similar properties yields equivalent norms for the function spaces given in Definition 6.5.1. Suppose that Ω is another bump whose Fourier transform is supported in the annulus $1-\frac{1}{7} \le |\xi| \le 2$ and that satisfies (6.5.1). The support properties of Ψ and Ω imply the identity

$$\Delta_j^{\Omega} = \Delta_j^{\Omega}(\Delta_{j-1}^{\Psi} + \Delta_j^{\Psi} + \Delta_{j+1}^{\Psi}). \tag{6.5.7}$$

Let $0 < p < \infty$ and pick $r < p$ and $N > \frac{n}{r} + n$. Then we have

$$\begin{aligned} \big|\Delta_j^{\Omega}\Delta_j^{\Psi}(f)(x)\big| &\le C_{N,\Omega}\int_{\mathbf{R}^n} \frac{\big|\Delta_j^{\Psi}(f)(x-z)\big|}{(1+2^j|z|)^{\frac{n}{r}}}\frac{2^{jn}dz}{(1+2^j|z|)^{N-\frac{n}{r}}} \\ &\le C_{N,\Omega}\sup_{z\in\mathbf{R}^n}\frac{\big|\Delta_j^{\Psi}(f)(x-z)\big|}{(1+2^j|z|)^{\frac{n}{r}}}\int_{\mathbf{R}^n}\frac{2^{jn}dz}{(1+2^j|z|)^{N-\frac{n}{r}}} \\ &\le C_{N,r,\Omega}(M(|\Delta_j^{\Psi}(f)|^r)(x)^{\frac{1}{r}} \end{aligned} \tag{6.5.8}$$

where we applied Lemma 6.5.3. The same estimate is also valid for $\Delta_j^{\Omega}\Delta_{j\pm1}^{\Psi}(f)$ and thus for $\Delta_j^{\Omega}(f)$, in view of identity (6.5.7). Armed with this observation and recalling that $r < p$, the boundedness of the Hardy–Littlewood maximal operator on $L^{p/r}$ yields that the homogeneous Besov–Lipschitz norm defined in terms of the bump Ω is controlled by a constant multiple of the corresponding Besov–Lipschitz norm defined in terms of Ψ. A similar argument applies for the inhomogeneous Besov–Lipschitz norms. The equivalence constants depend on Ψ, Ω, n, p, q, and α.

The corresponding equivalence of norms for Triebel–Lizorkin spaces is more difficult to obtain, and it is a consequence of the characterization of these spaces proved later.

Definition 6.5.5. For $b > 0$ and $j \in \mathbf{R}$ we introduce the notation

$$M_{b,j}^{**}(f;\Psi)(x) = \sup_{y\in\mathbf{R}^n}\frac{|(\Psi_{2^{-j}} * f)(x-y)|}{(1+2^j|y|)^b},$$

so that we have

$$M_b^{**}(f;\Psi) = \sup_{t>0} M_{b,t}^{**}(f;\Psi),$$

in accordance with the notation in the previous section. The function $M_b^{**}(f;\Psi)$ is called the *Peetre maximal function of f (with respect to Ψ).*

We clearly have

$$|\Delta_j^{\Psi}(f)| \le M_{b,j}^{**}(f;\Psi),$$

but the next result shows that a certain converse is also valid.

Theorem 6.5.6. *Let $b > n(\min(p,q))^{-1}$ and $0 < p,q < \infty$. Let Ψ and Ω be Schwartz functions whose Fourier transforms are supported in the annulus $\frac{1}{2} \le |\xi| \le 2$ and satisfy (6.5.1). Then we have*

$$\Big\| \Big(\sum_{j\in\mathbf{Z}} \big|2^{j\alpha} M^{**}_{b,j}(f;\Omega)\big|^q \Big)^{\frac{1}{q}} \Big\|_{L^p} \le C \Big\| \Big(\sum_{j\in\mathbf{Z}} \big|2^{j\alpha} \Delta^{\Psi}_j(f)\big|^q \Big)^{\frac{1}{q}} \Big\|_{L^p} \tag{6.5.9}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$*, where* $C = C_{\alpha,p,q,n,b,\Psi,\Omega}$*.*

Proof. We start with a Schwartz function Θ whose Fourier transform is nonnegative, supported in the annulus $1-\frac{2}{7} \le |\xi| \le 2$, and satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\Theta}(2^{-j}\xi)^2 = 1, \qquad \xi \in \mathbf{R}^n \setminus \{0\}. \tag{6.5.10}$$

Using (6.5.10), we have

$$\Omega_{2^{-k}} * f = \sum_{j\in\mathbf{Z}} (\Omega_{2^{-k}} * \Theta_{2^{-j}}) * (\Theta_{2^{-j}} * f).$$

It follows that

$$\begin{aligned}
&2^{k\alpha} \frac{|(\Omega_{2^{-k}} * f)(x-z)|}{(1+2^k|z|)^b} \\
&\le \sum_{j\in\mathbf{Z}} 2^{k\alpha} \int_{\mathbf{R}^n} |(\Omega_{2^{-k}} * \Theta_{2^{-j}})(y)| \, \frac{|(\Theta_{2^{-j}} * f)(x-z-y)|}{(1+2^k|z|)^b} \, dy \\
&= \sum_{j\in\mathbf{Z}} 2^{k\alpha} \int_{\mathbf{R}^n} 2^{kn} |(\Omega * \Theta_{2^{-(j-k)}})(2^k y)| \, \frac{(1+2^j|y+z|)^b}{(1+2^k|z|)^b} \, \frac{|(\Theta_{2^{-j}} * f)(x-z-y)|}{(1+2^j|y+z|)^b} \, dy \\
&\le \sum_{j\in\mathbf{Z}} 2^{k\alpha} \int_{\mathbf{R}^n} |(\Omega * \Theta_{2^{-(j-k)}})(y)| \, \frac{(1+2^j|2^{-k}y+z|)^b}{(1+2^k|z|)^b} \, \frac{|(\Theta_{2^{-j}} * f)(x-z-y)|}{(1+2^j|y+z|)^b} \, dy \\
&\le \sum_{j\in\mathbf{Z}} 2^{(k-j)\alpha} \int_{\mathbf{R}^n} |(\Omega * \Theta_{2^{-(j-k)}})(y)| \, \frac{(1+2^{j-k}|y|+2^j|z|)^b}{(1+2^k|z|)^b} \, dy \, 2^{j\alpha} M^{**}_{b,j}(f;\Theta)(x) \\
&\le \sum_{j\in\mathbf{Z}} 2^{(k-j)\alpha} \int_{\mathbf{R}^n} |(\Omega * \Theta_{2^{-(j-k)}})(y)| \, (1+2^{j-k})^b (1+2^{j-k}|y|)^b dy 2^{j\alpha} M^{**}_{b,j}(f;\Theta)(x).
\end{aligned}$$

We conclude that

$$2^{k\alpha} M^{**}_{b,k}(f;\Omega)(x) \le \sum_{j\in\mathbf{Z}} V_{k-j} \, 2^{j\alpha} M^{**}_{b,j}(f;\Theta)(x), \tag{6.5.11}$$

where

$$V_j = 2^{-j\alpha}(1+2^j)^b \int_{\mathbf{R}^n} |(\Omega * \Theta_{2^{-j}})(y)| \, (1+2^j|y|)^b \, dy.$$

We now use the facts that both Ω and Θ have vanishing moments of all orders and the result in Appendix K.2 to obtain

$$|(\Omega * \Theta_{2^{-j}})(y)| \le C_{L,N,n,\Theta,\Omega} \frac{2^{-|j|L}}{(1+2^{\min(0,j)}|y|)^N}$$

for all $L,N>0$. We deduce the estimate

$$|V_j| \le C_{L,M,n,\Theta,\Omega} 2^{-|j|M}$$

for all M sufficiently large, which, in turn, yields the estimate

$$\sum_{j\in\mathbf{Z}} |V_j|^{\min(1,q)} < \infty .$$

We deduce from (6.5.11) that for all $x \in \mathbf{R}^n$ we have

$$\big\|\{2^{k\alpha} M^{**}_{b,k}(f;\Omega)(x)\}_k\big\|_{\ell^q} \le C_{\alpha,p,q,n,\Psi,\Omega} \big\|\{2^{k\alpha} M^{**}_{b,k}(f;\Theta)(x)\}_k\big\|_{\ell^q} .$$

We now appeal to Lemma 6.5.3, which gives

$$2^{k\alpha} M^{**}_{b,k}(f;\Theta) \le C 2^{k\alpha} M(|\Delta^\Theta_k(f)|^r)^{\frac{1}{r}} = CM(|2^{k\alpha}\Delta^\Theta_k(f)|^r)^{\frac{1}{r}}$$

with $b=n/r$. We choose $r<\min(p,q)$. We use the $L^{p/r}(\mathbf{R}^n,\ell^{q/r})$ to $L^{p/r}(\mathbf{R}^n,\ell^{q/r})$ boundedness of the Hardy–Littlewood maximal operator, Theorem 4.6.6, to complete the proof of (6.5.9) with the exception that the function Ψ on the right-hand side of (6.5.9) is replaced by Θ. The passage to Ψ is a simple matter (at least when $p \ge 1$), since

$$\Delta^\Psi_j = \Delta^\Psi_j \big(\Delta^\Theta_{j-1} + \Delta^\Theta_j + \Delta^\Theta_{j+1}\big) .$$

For general $0<p<\infty$ the conclusion follows with the use of (6.5.8). □

We obtain as a corollary that a different choice of bumps gives equivalent Triebel–Lizorkin norms.

Corollary 6.5.7. *Let Ψ, Ω be Schwartz functions whose Fourier transforms are supported in the annulus $1-\frac{1}{7} \le |\xi| \le 2$ and satisfy (6.5.1). Let Φ be as in (6.5.2) and let*

$$\widehat{\Theta}(\xi) = \begin{cases} \sum_{j\le 0} \widehat{\Omega}(2^{-j}\xi) & \text{when } \xi \neq 0, \\ 1 & \text{when } \xi = 0. \end{cases}$$

Then the Triebel–Lizorkin quasinorms defined with respect to the pairs (Ψ,Φ) and (Ω,Θ) are equivalent.

Proof. We note that the quantity on the left in (6.5.9) is greater than or equal to

$$\Big\|\Big(\sum_{j\in\mathbf{Z}} |2^{j\alpha}\Delta^\Omega_j(f)|^q\Big)^{\frac{1}{q}}\Big\|_{L^p}$$

for all $f \in \mathscr{S}'(\mathbf{R}^n)$. This shows that the homogeneous Triebel–Lizorkin norm defined using Ω is bounded by a constant multiple of that defined using Ψ. This proves the equivalence of norms in the homogeneous case.

In the case of the inhomogeneous spaces, we let S^Ψ_0 and S^Ω_0 be the operators given by convolution with the bumps Φ and Θ, respectively (recall that these are defined in terms of Ψ and Ω). Then for $f \in \mathscr{S}'(\mathbf{R}^n)$ we have

$$\Theta * f = \Theta * (\Phi * f) + \Theta * (\Psi_{2^{-1}} * f), \tag{6.5.12}$$

since the Fourier transform of the function $\Phi + \Psi_{2^{-1}}$ is equal to 1 on the support of $\widehat{\Theta}$. Applying Lemma 6.5.3 (with $t = 1$), we obtain that

$$|\Theta * (\Phi * f)| \le C_r M(|\Phi * f|^r)^{\frac{1}{r}}$$

and also

$$|\Theta * (\Psi_{2^{-1}} * f)| \le C_r M(|\Psi_{2^{-1}} * f|^r)^{\frac{1}{r}}$$

for any $0 < r < \infty$. Picking $r < p$, we obtain that

$$\big\|\Theta * (\Phi * f)\big\|_{L^p} \le C \big\|S_0^{\Psi}(f)\big\|_{L^p}$$

and also

$$\big\|\Theta * (\Psi_{2^{-1}} * f)\big\|_{L^p} \le C \big\|\Delta_1^{\Psi}(f)\big\|_{L^p}.$$

Combining the last two estimates with (6.5.12), we obtain that $\big\|S_0^{\Omega}(f)\big\|_{L^p}$ is controlled by a multiple of the Triebel–Lizorkin norm of f defined using Ψ. This gives the equivalence of norms in the inhomogeneous case. □

Several other properties of these spaces are discussed in the exercises that follow.

Exercises

6.5.1. Let $0 < q_0 \le q_1 < \infty$, $0 < p < \infty$, $\varepsilon > 0$, and $\alpha \in \mathbf{R}$. Prove the embeddings

$$\begin{aligned} B_p^{\alpha,q_0} &\subseteq B_p^{\alpha,q_1}, \\ F_p^{\alpha,q_0} &\subseteq F_p^{\alpha,q_1}, \\ B_p^{\alpha+\varepsilon,q_0} &\subseteq B_p^{\alpha,q_1}, \\ F_p^{\alpha+\varepsilon,q_0} &\subseteq F_p^{\alpha,q_1}, \end{aligned}$$

where p and q_1 are allowed to be infinite in the case of Besov spaces.

6.5.2. Let $0 < q < \infty$, $0 < p < \infty$, and $\alpha \in \mathbf{R}$. Show that

$$B_p^{\alpha,\min(p,q)} \subseteq F_p^{\alpha,q} \subseteq B_p^{\alpha,\max(p,q)}.$$

[*Hint:* Consider the cases $p \ge q$ and $p < q$ and use the triangle inequality in the spaces $L^{p/q}$ and $\ell^{q/p}$, respectively.]

6.5.3. (a) Let $0 < p, q \le \infty$ and $\alpha \in \mathbf{R}$. Show that $\mathscr{S}(\mathbf{R}^n)$ is continuously embedded in $B_p^{\alpha,q}(\mathbf{R}^n)$ and that the latter is continuously embedded in $\mathscr{S}'(\mathbf{R}^n)$.
(b) Obtain the same conclusion for $F_p^{\alpha,q}(\mathbf{R}^n)$ when $p, q < \infty$.

6.5.4. $0 < p,q < \infty$ and $\alpha \in \mathbf{R}$. Show that the Schwartz functions are dense in all the spaces $B_p^{\alpha,q}(\mathbf{R}^n)$ and $F_p^{\alpha,q}(\mathbf{R}^n)$.
$\big[$*Hint:* Every Cauchy sequence $\{f_k\}_k$ in $B_p^{\alpha,q}$ is also Cauchy in $\mathscr{S}'(\mathbf{R}^n)$ and hence converges to some f in $\mathscr{S}'(\mathbf{R}^n)$. Then $\Delta_j(f_k) \to \Delta_j(f)$ in $\mathscr{S}'(\mathbf{R}^n)$. But $\Delta_j(f_k)$ is also Cauchy in L^p and therefore converges to $\Delta_j(f)$ in L^p. Argue similarly for $F_p^{\alpha,q}(\mathbf{R}^n)$.$\big]$

6.5.5. Let $\alpha \in \mathbf{R}$, let $0 < p,q < \infty$, and let $N = [\frac{n}{2} + \frac{n}{\min(p,q)}] + 1$. Assume that m is a $\mathscr{C}^N$ function on $\mathbf{R}^n \setminus \{0\}$ that satisfies

$$|\partial^\gamma m(\xi)| \le C_\gamma |\xi|^{-|\gamma|}$$

for all $|\gamma| \le N$. Show that there exists a constant C such that for all $f \in \mathscr{S}'(\mathbf{R}^n)$ we have

$$\big\|(m\widehat{f})^\vee\big\|_{\dot{B}_p^{\alpha,q}} \le C\big\|f\big\|_{\dot{B}_p^{\alpha,q}}.$$

$\big[$*Hint:* Pick $r < \min(p,q)$ such that $N > \frac{n}{2} + \frac{n}{r}$. Write $m = \sum_j m_j$, where $\widehat{m_j}(\xi) = \widehat{\Theta}(2^{-j}\xi)m(\xi)$ and $\widehat{\Theta}(2^{-j}\xi)$ is supported in an annulus $2^j \le |\xi| \le 2^{j+1}$. Obtain the estimate

$$\begin{aligned}\sup_{z\in\mathbf{R}^n} \frac{\big|\big(m_j\widehat{\Delta_j(f)}\big)^\vee(x-z)\big|}{(1+2^j|z|)^{\frac{n}{r}}} &\le C \sup_{z\in\mathbf{R}^n} \frac{\big|\Delta_j(f)(x-z)\big|}{(1+2^j|z|)^{\frac{n}{r}}} \int_{\mathbf{R}^n} |m_j^\vee(y)|(1+2^j|y|)^{\frac{n}{r}}\,dy \\ &\le C'\bigg(\int_{\mathbf{R}^n} |m_j(2^j(\cdot))^\vee(y)|^2(1+|y|)^{2N}\,dy\bigg)^{\frac{1}{2}}.\end{aligned}$$

Then use the hypothesis on m and apply Lemma 6.5.3.$\big]$

6.5.6. (*Peetre [258]*) Let m be as in Exercise 6.5.5. Show that there exists a constant C such that for all $f \in \mathscr{S}'(\mathbf{R}^n)$ we have

$$\big\|(m\widehat{f})^\vee\big\|_{\dot{F}_p^{\alpha,q}} \le C\big\|f\big\|_{\dot{F}_p^{\alpha,q}}.$$

$\big[$*Hint:* Use the hint of Exercise 6.5.5 and Theorem 4.6.6.$\big]$

6.5.7. (a) Suppose that $B_{p_0}^{\alpha_0,q_0} = B_{p_1}^{\alpha_1,q_1}$ with equivalent norms. Prove that $\alpha_0 = \alpha_1$ and $p_0 = p_1$. Prove the same result for the scale of F spaces.
(b) Suppose that $B_{p_0}^{\alpha_0,q_0} = B_{p_1}^{\alpha_1,q_1}$ with equivalent norms. Prove that $q_0 = q_1$. Argue similarly with the scale of F spaces.
$\big[$*Hint:* Part (a): Test the corresponding norms on the function $\Psi(2^j x)$, where Ψ is chosen so that its Fourier transform is supported in $\frac{1}{2} \le |\xi| \le 2$. Part (b): Try a function f of the form $\widehat{f}(\xi) = \sum_{j=1}^N a_j \widehat{\varphi}(\xi_1 - 2^j, \xi_2, \dots, \xi_n)$, where φ is a Schwartz function whose Fourier transform is supported in a small neighborhood of the origin.$\big]$

6.6 Atomic Decomposition

In this section we focus attention on the homogeneous Triebel–Lizorkin spaces $\dot{F}_p^{\alpha,q}$, which include the Hardy spaces discussed in Section 6.4. Most results discussed in this section are also valid for the inhomogeneous Triebel–Lizorkin spaces and for the Besov–Lipschitz via a similar or simpler analysis. We refer the interested reader to the relevant literature on the subject at the end of this chapter.

6.6.1 *The Space of Sequences* $\dot{f}_p^{\alpha,q}$

To provide further intuition in the understanding of the homogeneous Triebel–Lizorkin spaces we introduce a related space consisting of sequences of scalars. This space is denoted by $\dot{f}_p^{\alpha,q}$ and is related to $\dot{F}_p^{\alpha,q}$ in a way similar to that in which $\ell^2(\mathbf{Z})$ is related to $L^2([0,1])$.

Definition 6.6.1. Let $0 < q \le \infty$ and $\alpha \in \mathbf{R}$. Let $\mathscr{D}$ be the set of all dyadic cubes in $\mathbf{R}^n$. We consider the set of all sequences $\{s_Q\}_{Q\in\mathscr{D}}$ such that the function

$$g^{\alpha,q}(\{s_Q\}_Q) = \Big(\sum_{Q\in\mathscr{D}} (|Q|^{-\frac{\alpha}{n}-\frac{1}{2}}\,|s_Q|\,\chi_Q)^q\Big)^{\frac{1}{q}} \tag{6.6.1}$$

is in $L^p(\mathbf{R}^n)$. For such sequences $s = \{s_Q\}_Q$ we set

$$\big\|s\big\|_{\dot{f}_p^{\alpha,q}} = \big\|g^{\alpha,q}(s)\big\|_{L^p(\mathbf{R}^n)}\,.$$

6.6.2 *The Smooth Atomic Decomposition of* $\dot{F}_p^{\alpha,q}$

Next, we discuss the smooth atomic decomposition of these spaces. We begin with the definition of smooth atoms on $\mathbf{R}^n$.

Definition 6.6.2. Let Q be a dyadic cube and let L be a nonnegative integer. A $\mathscr{C}^\infty$ function a_Q on $\mathbf{R}^n$ is called a *smooth L-atom for Q* if it satisfies

(a) a_Q is supported in $3Q$ (the cube concentric with Q having three times its side length);

(b) $\displaystyle\int_{\mathbf{R}^n} x^\gamma a_Q(x)\,dx = 0$ for all multi-indices $|\gamma| \le L$;

(c) $|\partial^\gamma a_Q| \le |Q|^{-\frac{|\gamma|}{n}-\frac{1}{2}}$ for all multi-indices γ satisfying $|\gamma| \le L+n+1$.

The value of the constant $L+n+1$ in (c) may vary in the literature. Any sufficiently large constant depending on L will serve the purposes of the definition.

We now prove a theorem stating that elements of $\dot{F}_p^{\alpha,q}$ can be decomposed as sums of smooth atoms.

Theorem 6.6.3. *Let $0 < p,q < \infty$, $\alpha \in \mathbf{R}$, and let L be a nonnegative integer satisfying $L \geq [n\max(1, \frac{1}{p}, \frac{1}{q}) - n - \alpha]$. Then there is a constant $C_{n,p,q,\alpha}$ such that for every sequence of smooth L-atoms $\{a_Q\}_{Q\in\mathscr{D}}$ and every sequence of complex scalars $\{s_Q\}_{Q\in\mathscr{D}}$ we have*

$$\Big\| \sum_{Q\in\mathscr{D}} s_Q a_Q \Big\|_{\dot{F}_p^{\alpha,q}} \leq C_{n,p,q,\alpha} \big\| \{s_Q\}_Q \big\|_{\dot{f}_p^{\alpha,q}} . \tag{6.6.2}$$

Conversely, there is a constant $C'_{n,p,q,\alpha}$ such that given any distribution f in $\dot{F}_p^{\alpha,q}$ and any $L \geq 0$, there exist a sequence of smooth L-atoms $\{a_Q\}_{Q\in\mathscr{D}}$ and a sequence of complex scalars $\{s_Q\}_{Q\in\mathscr{D}}$ such that

$$f = \sum_{Q\in\mathscr{D}} s_Q a_Q ,$$

where the sum converges in $\mathscr{S}'/\mathscr{P}$ and moreover,

$$\big\| \{s_Q\}_Q \big\|_{\dot{f}_p^{\alpha,q}} \leq C'_{n,p,q,\alpha} \big\| f \big\|_{\dot{F}_p^{\alpha,q}} . \tag{6.6.3}$$

Proof. We begin with the first claim of the theorem. We let Δ_j^Ψ be the Littlewood–Paley operator associated with a Schwartz function Ψ whose Fourier transform is compactly supported away from the origin in $\mathbf{R}^n$. Let a_Q be a smooth L-atom supported in a cube $3Q$ with center C_Q and let the side length be $\ell(Q) = 2^{-\mu}$. It follows trivially from Definition 6.6.2 that a_Q satisfies

$$|\partial_y^\gamma a_Q(y)| \leq C_{N,n} 2^{-\frac{\mu n}{2}} \frac{2^{\mu|\gamma|+\mu n}}{(1+2^\mu|y-c_Q|)^N} \tag{6.6.4}$$

for all $N > 0$ and for all multi-indices γ satisfying $|\gamma| \leq L+n+1$. Moreover, the function $y \mapsto \Psi_{2^{-j}}(y-x)$ satisfies

$$|\partial_y^\delta \Psi_{2^{-j}}(y-x)| \leq C_{N,n,\delta} \frac{2^{j|\delta|+jn}}{(1+2^j|y-x|)^N} \tag{6.6.5}$$

for all $N > 0$ and for all multi-indices δ. Using first the facts that a_Q has vanishing moments of all orders up to and including $L = (L+1)-1$ and that the function $y \mapsto \Psi_{2^{-j}}(y-x)$ satisfies (6.6.5) for all multi-indices δ with $|\delta| = L$, secondly the facts that the function $y \mapsto \Psi_{2^{-j}}(y-x)$ has vanishing moments of all orders up to and including $L+n = (L+n+1)-1$ and that a_Q satisfies (6.6.4) for all multi-indices γ satisfying $|\gamma| = L+n+1$, and the result in Appendix K.2, we deduce the following estimate for all $N > 0$:

$$\big|\Delta_j^\Psi(a_Q)(x)\big| \leq C_{N,n,L'}\, 2^{-\frac{\mu n}{2}} \frac{2^{\min(j,\mu)n-|\mu-j|L'}}{(1+2^{\min(j,\mu)}|x-c_Q|)^N} , \tag{6.6.6}$$

where

$$L' = \begin{cases} L+1 & \text{when } j < \mu, \\ L+n & \text{when } \mu \le j. \end{cases}$$

Now fix $0 < b < \min(1,p,q)$ so that

$$L+1 > \tfrac{n}{b} - n - \alpha. \tag{6.6.7}$$

This can be achieved by taking b close enough to $\min(1,p,q)$, since our assumption $L \ge \big[n\max\big(1,\frac{1}{p},\frac{1}{q}\big) - n - \alpha\big]$ implies $L+1 > n\max\big(1,\frac{1}{p},\frac{1}{q}\big) - n - \alpha$.

Using Exercise 6.6.6, we obtain

$$\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \frac{|s_Q|}{(1+2^{\min(j,\mu)}|x-c_Q|)^N} \le c\,2^{\max(\mu-j,0)\frac{n}{b}} \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} |s_Q|^b \chi_Q \Big)(x) \Big\}^{\frac{1}{b}}$$

whenever $N > n/b$, where M is the Hardy–Littlewood maximal operator. It follows from the preceding estimate and (6.6.6) that

$$\begin{aligned} 2^{j\alpha} \sum_{\mu\in\mathbf{Z}} \sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} |s_Q|\,\big|\Delta_j^\Psi(a_Q)(x)\big| \le C \sum_{\mu\in\mathbf{Z}} 2^{\min(j,\mu)n} 2^{-|j-\mu|L'} 2^{-\mu n} 2^{(j-\mu)\alpha} \\ \times 2^{\max(\mu-j,0)\frac{n}{b}} \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big)(x) \Big\}^{\frac{1}{b}}. \end{aligned}$$

Raise the preceding inequality to the power q and sum over $j \in \mathbf{Z}$; then raise to the power $1/q$ and take $\big\| \cdot \big\|_{L^p}$ norms in x. We obtain

$$\big\|f\big\|_{\dot F_p^{\alpha,q}} \le \Bigg\| \Bigg\{ \sum_{j\in\mathbf{Z}} \Bigg[\sum_{\mu\in\mathbf{Z}} d(j-\mu) \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big) \Big\}^{\frac{1}{b}} \Bigg]^q \Bigg\}^{\frac{1}{q}} \Bigg\|_{L^p},$$

where $f = \sum_{Q\in\mathscr{D}} s_Q a_Q$ and

$$d(j-\mu) = C\,2^{\min(j-\mu,0)(n-\frac{n}{b}) + (j-\mu)\alpha - |j-\mu|L'}.$$

We now estimate the expression inside the last L^p norm by

$$\Big\{ \sum_{j\in\mathbf{Z}} d(j)^{\min(1,q)} \Big\}^{\frac{1}{\min(1,q)}} \Big\{ \sum_{\mu\in\mathbf{Z}} \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big) \Big\}^{\frac{q}{b}} \Big\}^{\frac{1}{q}},$$

and we note that the first term is a constant in view of (6.6.7). We conclude that

$$\begin{aligned}
\Big\| \sum_{Q\in\mathscr{D}} s_Q a_Q \Big\|_{\dot F_p^{\alpha,q}} &\le C \bigg\| \bigg\{ \sum_{\mu\in\mathbf{Z}} \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big) \Big\}^{\frac{q}{b}} \bigg\}^{\frac{1}{q}} \bigg\|_{L^p} \\
&= C \bigg\| \bigg\{ \sum_{\mu\in\mathbf{Z}} \Big\{ M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big) \Big\}^{\frac{q}{b}} \bigg\}^{\frac{b}{q}} \bigg\|_{L^{\frac{p}{b}}}^{\frac{1}{b}} \\
&\le C' \bigg\| \bigg\{ \sum_{\mu\in\mathbf{Z}} \Big\{ \sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^b \chi_Q \Big\}^{\frac{q}{b}} \bigg\}^{\frac{b}{q}} \bigg\|_{L^{\frac{p}{b}}}^{\frac{1}{b}} \\
&= C' \bigg\| \bigg\{ \sum_{\mu\in\mathbf{Z}} \sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \big(|s_Q|\,|Q|^{-\frac{1}{2}-\frac{\alpha}{n}}\big)^q \chi_Q \bigg\}^{\frac{1}{q}} \bigg\|_{L^p} \\
&= C' \big\| \{s_Q\}_Q \big\|_{\dot f_p^{\alpha,q}} ,
\end{aligned}$$

where in the last inequality we used Theorem 4.6.6, which is valid under the assumption $1 < \frac{p}{b}, \frac{q}{b} < \infty$. This proves (6.6.2).

We now turn to the converse statement of the theorem. It is not difficult to see that there exist Schwartz functions Ψ (unrelated to the previous one) and Θ such that $\widehat{\Psi}$ is supported in the annulus $\frac{1}{2} \le |\xi| \le 2$, $\widehat{\Psi}$ is at least $c > 0$ in the smaller annulus $\frac{3}{5} \le |\xi| \le \frac{5}{3}$, and Θ is supported in the ball $|x| \le 1$ and satisfies $\int_{\mathbf{R}^n} x^\gamma \Theta(x)\,dx = 0$ for all $|\gamma| \le L$, such that the identity

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi)\widehat{\Theta}(2^{-j}\xi) = 1 \tag{6.6.8}$$

holds for all $\xi \in \mathbf{R}^n \setminus \{0\}$. (See Exercise 6.6.1.)

Using identity (6.6.8), we can write

$$f = \sum_{j\in\mathbf{Z}} \Psi_{2^{-j}} * \Theta_{2^{-j}} * f .$$

Setting $\mathscr{D}_j = \{Q \in \mathscr{D} :\ \ell(Q) = 2^{-j}\}$, we now have

$$f = \sum_{j\in\mathbf{Z}} \sum_{Q\in\mathscr{D}_j} \int_Q \Theta_{2^{-j}}(x-y)(\Psi_{2^{-j}} * f)(y)\,dy = \sum_{j\in\mathbf{Z}} \sum_{Q\in\mathscr{D}_j} s_Q a_Q ,$$

where we also set

$$s_Q = |Q|^{\frac{1}{2}} \sup_{y\in Q} |(\Psi_{2^{-j}} * f)(y)| \sup_{|\gamma|\le L} \big\| \partial^\gamma \Theta \big\|_{L^1}$$

for Q in $\mathscr{D}_j$ and

$$a_Q(x) = \frac{1}{s_Q} \int_Q \Theta_{2^{-j}}(x-y)(\Psi_{2^{-j}} * f)(y)\, dy\,.$$

It is straightforward to verify that a_Q is supported in $3Q$ and that it has vanishing moments up to and including order L. Moreover, we have

$$|\partial^\gamma a_Q| \le \frac{1}{s_Q} \big\|\partial^\gamma \Theta\big\|_{L^1} 2^{j(n+|\gamma|)} \sup_Q |\Psi_{2^{-j}} * f| \le |Q|^{-\frac{1}{2}-\frac{|\gamma|}{n}}\,,$$

which makes the function a_Q a smooth L-atom. Now note that

$$\begin{aligned}
\sum_{\ell(Q)=2^{-j}} & \big(|Q|^{-\frac{\alpha}{n}-\frac{1}{2}} s_Q \chi_Q(x)\big)^q \\
&= C \sum_{\ell(Q)=2^{-j}} \big(2^{j\alpha} \sup_{y\in Q} |(\Psi_{2^{-j}} * f)(y)| \chi_Q(x)\big)^q \\
&\le C \sup_{|z|\le\sqrt{n}2^{-j}} \big(2^{j\alpha}(1+2^j|z|)^{-b} |(\Psi_{2^{-j}} * f)(x-z)|\big)^q (1+2^j|z|)^{bq} \\
&\le C\big(2^{j\alpha} M^{**}_{b,j}(f,\Psi)(x)\big)^q\,,
\end{aligned}$$

where we used the fact that in the first inequality there is only one nonzero term in the sum because of the appearance of the characteristic function. Summing over all $j \in \mathbf{Z}^n$, raising to the power $1/q$, and taking L^p norms yields the estimate

$$\big\|\{s_Q\}_Q\big\|_{\dot f_p^{\alpha,q}} \le C \Big\| \Big(\sum_{j\in\mathbf{Z}} |2^{j\alpha} M^{**}_{b,j}(f;\Psi)|^q \Big)^{\frac{1}{q}} \Big\|_{L^p} \le C \big\|f\big\|_{\dot F_p^{\alpha,q}}\,,$$

where the last inequality follows from Theorem 6.5.6. This proves (6.6.3). □

6.6.3 The Nonsmooth Atomic Decomposition of $\dot F_p^{\alpha,q}$

We now discuss the main theorem of this section, the nonsmooth atomic decomposition of the homogeneous Triebel–Lizorkin spaces $\dot F_p^{\alpha,q}$, which in particular includes that of the Hardy spaces H^p. We begin this task with a definition.

Definition 6.6.4. Let $0 < p \le 1$ and $1 \le q \le \infty$. A sequence of complex numbers $r = \{r_Q\}_{Q\in\mathscr{D}}$ is called an ∞-atom for $\dot f_p^{\alpha,q}$ if there exists a dyadic cube Q_0 such that

(a) $r_Q = 0$ if $Q \not\subseteq Q_0$;
(b) $\big\|g^{\alpha,q}(r)\big\|_{L^\infty} \le |Q_0|^{-\frac{1}{p}}$.

We observe that every ∞-atom $r = \{r_Q\}$ for $\dot f_p^{\alpha,q}$ satisfies $\big\|r\big\|_{\dot f_p^{\alpha,q}} \le 1$. Indeed,

$$\big\|r\big\|^p_{\dot f_p^{\alpha,q}} = \int_{Q_0} |g^{\alpha,q}(r)|^p\, dx \le |Q_0|^{-1}|Q_0| = 1\,.$$

The following theorem concerns the atomic decomposition of the spaces $\dot{f}_p^{\alpha,q}$.

Theorem 6.6.5. *Suppose $\alpha \in \mathbf{R}$, $0 < q < \infty$, $0 < p < \infty$, and $s = \{s_Q\}_Q$ is in $\dot{f}_p^{\alpha,q}$. Then there exist $C_{n,p,q} > 0$, a sequence of scalars λ_j, and a sequence of ∞-atoms $r_j = \{r_{j,Q}\}_Q$ for $\dot{f}_p^{\alpha,q}$ such that*

$$s = \{s_Q\}_Q = \sum_{j=1}^{\infty} \lambda_j \{r_{j,Q}\}_Q = \sum_{j=1}^{\infty} \lambda_j r_j$$

and such that

$$\Big(\sum_{j=1}^{\infty} |\lambda_j|^p \Big)^{\frac{1}{p}} \le C_{n,p,q} \big\| s \big\|_{\dot{f}_p^{\alpha,q}} . \tag{6.6.9}$$

Proof. We fix α, p, q, and a sequence $s = \{s_Q\}_Q$ as in the statement of the theorem. For a dyadic cube R in $\mathscr{D}$ we define the function

$$g_R^{\alpha,q}(s)(x) = \Big(\sum_{\substack{Q \in \mathscr{D} \\ R \subseteq Q}} \big(|Q|^{\frac{\alpha}{n} - \frac{1}{2}} |s_Q| \chi_Q(x) \big)^q \Big)^{\frac{1}{q}}$$

and we observe that this function is constant on R. We also note that for dyadic cubes R_1 and R_2 with $R_1 \subseteq R_2$ we have

$$g_{R_2}^{\alpha,q}(s) \le g_{R_1}^{\alpha,q}(s) .$$

Finally, we observe that

$$\begin{aligned} \lim_{\substack{\ell(R) \to \infty \\ x \in R}} g_R^{\alpha,q}(s)(x) &= 0 \\ \lim_{\substack{\ell(R) \to 0 \\ x \in R}} g_R^{\alpha,q}(s)(x) &= g^{\alpha,q}(s)(x) , \end{aligned}$$

where $g^{\alpha,q}(s)$ is the function defined in (6.6.1).

For $k \in \mathbf{Z}$ we set

$$\mathscr{A}_k = \big\{ R \in \mathscr{D} : \ g_R^{\alpha,q}(s)(x) > 2^k \quad \text{for all } x \in R \big\} .$$

We note that $\mathscr{A}_{k+1} \subseteq \mathscr{A}_k$ for all k in $\mathbf{Z}$ and that

$$\{ x \in \mathbf{R}^n : \ g^{\alpha,q}(s)(x) > 2^k \} = \bigcup_{R \in \mathscr{A}_k} R . \tag{6.6.10}$$

Moreover, we have for all $k \in \mathbf{Z}$,

$$\Big(\sum_{Q \in \mathscr{D} \setminus \mathscr{A}_k} \big(|Q|^{-\frac{\alpha}{n} - \frac{1}{2}} |s_Q| \chi_Q(x) \big)^q \Big)^{\frac{1}{q}} \le 2^k , \qquad \text{for all } x \in \mathbf{R}^n. \tag{6.6.11}$$

To prove (6.6.11) we assume that $g^{\alpha,q}(s)(x) > 2^k$; otherwise, the conclusion is trivial. Then there exists a maximal dyadic cube $R_{\max}$ in $\mathscr{A}_k$ such that $x \in R_{\max}$. Letting R_0 be the unique dyadic cube that contains $R_{\max}$ and has twice its side length, we have that the left-hand side of (6.6.11) is equal to $g^{\alpha,q}_{R_0}(s)(x)$, which is at most 2^k, since R_0 is not contained in $\mathscr{A}_k$.

Since $g^{\alpha,q}(s) \in L^p(\mathbf{R}^n)$, by our assumption, and $g^{\alpha,q}(s) > 2^k$ for all $x \in Q$ if $Q \in \mathscr{A}_k$, the cubes in $\mathscr{A}_k$ must have size bounded above by some constant. We set

$$\mathscr{B}_k = \big\{Q \in \mathscr{D}: \quad Q \text{ is a maximal dyadic cube in } \mathscr{A}_k \setminus \mathscr{A}_{k+1}\big\}.$$

For J in $\mathscr{B}_k$ we define a sequence $t(k,J) = \{t(k,J)_Q\}_{Q\in\mathscr{D}}$ by setting

$$t(k,J)_Q = \begin{cases} s_Q & \text{if } Q \subseteq J \text{ and } Q \in \mathscr{A}_k \setminus \mathscr{A}_{k+1}, \\ 0 & \text{otherwise.} \end{cases}$$

We can see that if

$$Q \notin \bigcup_{k\in\mathbf{Z}} \mathscr{A}_k, \qquad \text{then} \qquad s_Q = 0,$$

and the identity

$$s = \sum_{k\in\mathbf{Z}} \sum_{J\in\mathscr{B}_k} t(k,J) \tag{6.6.12}$$

is valid. For all $x \in \mathbf{R}^n$ we have

$$\begin{aligned} \big|g^{\alpha,q}(t(k,J))(x)\big| &= \Bigg(\sum_{\substack{Q\subseteq J \\ Q\in\mathscr{A}_k\setminus\mathscr{A}_{k+1}}} \big(|Q|^{-\frac{\alpha}{n}-\frac{1}{2}}|s_Q|\,\chi_Q(x)\big)^q \Bigg)^{\frac{1}{q}} \\ &\le \Bigg(\sum_{\substack{Q\subseteq J \\ Q\in\mathscr{D}\setminus\mathscr{A}_{k+1}}} \big(|Q|^{-\frac{\alpha}{n}-\frac{1}{2}}|s_Q|\,\chi_Q(x)\big)^q \Bigg)^{\frac{1}{q}} \\ &\le 2^{k+1}, \end{aligned} \tag{6.6.13}$$

where we used (6.6.11) in the last estimate. We define atoms $r(k,J) = \{r(k,J)_Q\}_{Q\in\mathscr{D}}$ by setting

$$r(k,J)_Q = 2^{-k-1}|J|^{-\frac{1}{p}}t(k,J)_Q, \tag{6.6.14}$$

and we also define scalars

$$\lambda_{k,J} = 2^{k+1}|J|^{\frac{1}{p}}.$$

To see that each $r(k,J)$ is an ∞-atom for $\dot{f}^{\alpha,q}_p$, we observe that $r(k,J)_Q = 0$ if $Q \not\subseteq J$ and that

$$\big|g^{\alpha,q}(t(k,J))(x)\big| \le |J|^{-\frac{1}{p}}, \qquad \text{for all } x \in \mathbf{R}^n,$$

in view of (6.6.13). Also using (6.6.12) and (6.6.14), we obtain that

$$s = \sum_{k\in\mathbf{Z}} \sum_{J\in\mathscr{B}_k} \lambda_{k,J}\, r(k,J),$$

which says that s can be written as a linear combination of atoms. Finally, we estimate the sum of the pth power of the coefficients $\lambda_{k,J}$. We have

$$\begin{aligned}
\sum_{k\in\mathbf{Z}}\sum_{J\in\mathscr{B}_k}|\lambda_{k,J}|^p &= \sum_{k\in\mathbf{Z}}2^{(k+1)p}\sum_{J\in\mathscr{B}_k}|J| \\
&\le 2^p\sum_{k\in\mathbf{Z}}2^{kp}\Big|\bigcup_{Q\in\mathscr{A}_k}Q\Big| \\
&= 2^p\sum_{k\in\mathbf{Z}}2^{k(p-1)}2^k|\{x\in\mathbf{R}^n:\ g^{\alpha,q}(s)(x)>2^k\}| \\
&\le 2^p\sum_{k\in\mathbf{Z}}\int_{2^k}^{2^{k+1}}2^{k(p-1)}|\{x\in\mathbf{R}^n:\ g^{\alpha,q}(s)(x)>\tfrac{\lambda}{2}\}|\,d\lambda \\
&\le 2^p\sum_{k\in\mathbf{Z}}\int_{2^k}^{2^{k+1}}\lambda^{p-1}|\{x\in\mathbf{R}^n:\ g^{\alpha,q}(s)(x)>\tfrac{\lambda}{2}\}|\,d\lambda \\
&= \frac{2^{2p}}{p}\big\|g^{\alpha,q}(s)\big\|_{L^p}^p \\
&= \frac{2^{2p}}{p}\big\|s\big\|_{\dot{f}_p^{\alpha,q}}^p\,.
\end{aligned}$$

Taking the pth root yields (6.6.9). The proof of the theorem is now complete. □

We now deduce a corollary regarding a new characterization of the space $\dot{f}_p^{\alpha,q}$.

Corollary 6.6.6. *Suppose $\alpha\in\mathbf{R}$, $0<p\le 1$, and $p\le q\le\infty$. Then we have*

$$\big\|s\big\|_{\dot{f}_p^{\alpha,q}}\approx\inf\left\{\Big(\sum_{j=1}^{\infty}|\lambda_j|^p\Big)^{\frac{1}{p}}:\ s=\sum_{j=1}^{\infty}\lambda_j r_j\,,\quad r_j \textit{ is an } \infty\textit{-atom for } \dot{f}_p^{\alpha,q}\right\}.$$

Proof. One direction in the previous estimate is a direct consequence of (6.6.9). The other direction uses the observation made after Definition 6.6.4 that every ∞-atom r for $\dot{f}_p^{\alpha,q}$ satisfies $\big\|r\big\|_{\dot{f}_p^{\alpha,q}}\le 1$ and that for $p\le 1$ and $p\le q$ the quantity $s\to\big\|s\big\|_{\dot{f}_p^{\alpha,q}}^p$ is subadditive; see Exercise 6.6.2. Then each $s=\sum_{j=1}^{\infty}\lambda_j r_j$ (with r_j ∞-atoms for $\dot{f}_p^{\alpha,q}$ and $\sum_{j=1}^{\infty}|\lambda_j|^p<\infty$) must be an element of $\dot{f}_p^{\alpha,q}$, since

$$\Big\|\sum_{j=1}^{\infty}\lambda_j r_j\Big\|_{\dot{f}_p^{\alpha,q}}^p\le\sum_{j=1}^{\infty}|\lambda_j|^p\,\big\|r_j\big\|_{\dot{f}_p^{\alpha,q}}^p\le\sum_{j=1}^{\infty}|\lambda_j|^p<\infty.$$

This concludes the proof of the corollary. □

The theorem we just proved allows us to obtain an atomic decomposition for the space $\dot{F}_p^{\alpha,q}$ as well. Indeed, we have the following result:

Corollary 6.6.7. *Let $\alpha\in\mathbf{R}$, $0<p\le 1$, $L\ge[\frac{n}{p}-n-\alpha]$ and let q satisfy $p\le q<\infty$. Then we have the following representation:*

$$\|f\|_{\dot F_p^{\alpha,q}} \approx \inf\Big\{\Big(\sum_{j=1}^{\infty}|\lambda_j|^p\Big)^{\frac{1}{p}} : f=\sum_{j=1}^{\infty}\lambda_j A_j, \quad A_j=\sum_{Q\in\mathscr{D}} r_Q a_Q, \quad a_Q \textit{ are}$$
$$\textit{smooth L-atoms for } \dot F_p^{\alpha,q} \textit{ and } \{r_Q\}_Q \textit{ is an } \infty\textit{-atom for } \dot f_p^{\alpha,q}\Big\}.$$

Proof. Let $f=\sum_{j=1}^{\infty}\lambda_j A_j$ as described previously. Using Exercise 6.6.2, we have

$$\|f\|_{\dot F_p^{\alpha,q}}^p \le \sum_{j=1}^{\infty}|\lambda_j|^p\|A_j\|_{\dot F_p^{\alpha,q}}^p \le c_{n,p}\sum_{j=1}^{\infty}|\lambda_j|^p\|r\|_{\dot f_p^{\alpha,q}}^p,$$

where in the last estimate we used Theorem 6.6.3. Using the fact that every ∞-atom $r=\{r_Q\}$ for $\dot f_p^{\alpha,q}$ satisfies $\|r\|_{\dot f_p^{\alpha,q}}\le 1$, we conclude that every element f in $\mathscr{S}'(\mathbf{R}^n)$ that has the form $\sum_{j=1}^{\infty}\lambda_j A_j$ lies in the homogeneous Triebel–Lizorkin space $\dot F_p^{\alpha,q}$ [and has norm controlled by a constant multiple of $\big(\sum_{j=1}^{\infty}|\lambda_j|^p\big)^{\frac{1}{p}}$].

Conversely, Theorem 6.6.3 gives that every element of $\dot F_p^{\alpha,q}$ has a smooth atomic decomposition. Then we can write

$$f=\sum_{Q\in\mathscr{D}} s_Q a_Q,$$

where each a_Q is a smooth L-atom for the cube Q. Using Theorem 6.6.5 we can now write $s=\{s_Q\}_Q$ as a sum of ∞-atoms for $\dot f_p^{\alpha,q}$, that is,

$$s=\sum_{j=1}^{\infty}\lambda_j r_j,$$

where

$$\Big(\sum_{j=1}^{\infty}|\lambda_j|^p\Big)^{\frac{1}{p}} \le c\|s\|_{\dot f_p^{\alpha,q}} \le c\|f\|_{\dot F_p^{\alpha,q}},$$

where the last step uses Theorem 6.6.3 again. It is simple to see that

$$f=\sum_{Q\in\mathscr{D}}\sum_{j=1}^{\infty}\lambda_j r_{j,Q}a_Q=\sum_{j=1}^{\infty}\lambda_j\Big(\sum_{Q\in\mathscr{D}} r_{j,Q}a_Q\Big),$$

and we set the expression inside the parentheses equal to A_j. □

6.6.4 Atomic Decomposition of Hardy Spaces

We now pass to one of the main theorems of this chapter, the atomic decomposition of $H^p(\mathbf{R}^n)$ for $0<p\le 1$. We begin by defining atoms for H^p.

Definition 6.6.8. Let $1<q\le\infty$. A function A is called *an L^q-atom for $H^p(\mathbf{R}^n)$* if there exists a cube Q such that

(a) A is supported in Q;

(b) $\|A\|_{L^q} \le |Q|^{\frac{1}{q}-\frac{1}{p}}$;

(c) $\int x^\gamma A(x)\,dx = 0$ for all multi-indices γ with $|\gamma| \le [\frac{n}{p} - n]$.

Notice that any L^r-atom for H^p is also an L^q-atom for H^p whenever $0 < p \le 1$ and $1 < q < r \le \infty$. It is also simple to verify that an L^q-atom A for H^p is in fact in H^p. We prove this result in the next theorem for $q = 2$, and we refer the reader to Exercise 6.6.4 for the case of a general q.

Theorem 6.6.9. *Let* $0 < p \le 1$. *There is a constant* $C_{n,p} < \infty$ *such that every* L^2*-atom* A *for* $H^p(\mathbf{R}^n)$ *satisfies*

$$\|A\|_{H^p} \le C_{n,p}\,.$$

Proof. We could prove this theorem either by showing that the smooth maximal function $M(A;\Phi)$ is in L^p or by showing that the square function $\big(\sum_j |\Delta_j(A)|^2\big)^{1/2}$ is in L^p. The operators Δ_j here are as in Theorem 5.1.2. Both proofs are similar; we present the second, and we refer to Exercise 6.6.3 for the first.

Let $A(x)$ be an atom that we assume is supported in a cube Q centered at the origin [otherwise apply the argument to the atom $A(x - c_Q)$, where c_Q is the center of Q]. We control the L^p quasinorm of $\big(\sum_j |\Delta_j(A)|^2\big)^{1/2}$ by estimating it over the cube Q^* and over $(Q^*)^c$, where $Q^* = 2\sqrt{n}\,Q$. We have

$$\left(\int_{Q^*} \Big(\sum_j |\Delta_j(A)|^2\Big)^{\frac{p}{2}} dx\right)^{\frac{1}{p}} \le \left(\int_{Q^*} \sum_j |\Delta_j(A)|^2\,dx\right)^{\frac{1}{2}} |Q^*|^{\frac{1}{p(2/p)'}}\,.$$

Using that the square function $f \mapsto \big(\sum_j |\Delta_j(f)|^2\big)^{\frac{1}{2}}$ is L^2 bounded, we obtain

$$\begin{aligned}\left(\int_{Q^*} \Big(\sum_j |\Delta_j(A)|^2\Big)^{\frac{p}{2}} dx\right)^{\frac{1}{p}} &\le C_n \|A\|_{L^2} |Q^*|^{\frac{1}{p(2/p)'}} \\ &\le C_n (2\sqrt{n})^{\frac{n}{p}-\frac{n}{2}} |Q|^{\frac{1}{2}-\frac{1}{p}} |Q|^{\frac{1}{p}-\frac{1}{2}} \\ &= C_n'\,.\end{aligned} \tag{6.6.15}$$

To estimate the contribution of the square function outside Q^*, we use the cancellation of the atoms. Let $k = [\frac{n}{p} - n] + 1$. We have

$$\begin{aligned}\Delta_j(A)(x) &= \int_Q A(y)\Psi_{2^{-j}}(x-y)\,dy \\ &= 2^{jn}\int_Q A(y)\left[\Psi(2^j x - 2^j y) - \sum_{|\beta|\le k-1} (\partial^\beta \Psi)(2^j x)\frac{(-2^j y)^\beta}{\beta!}\right] dy \\ &= 2^{jn}\int_Q A(y)\left[\sum_{|\beta|=k} (\partial^\beta \Psi)(2^j x - 2^j \theta y)\frac{(-2^j y)^\beta}{\beta!}\right] dy,\end{aligned}$$

where $0 \le \theta \le 1$. Taking absolute values, using the fact that $\partial^\beta \Psi$ are Schwartz functions, and that $|x-\theta y| \ge |x|-|y| \ge \frac{1}{2}|x|$ whenever $y \in Q$ and $x \notin Q^*$, we obtain the estimate

$$\begin{aligned}
|\Delta_j(A)(x)| &\le 2^{jn} \int_Q |A(y)| \sum_{|\beta|=k} \frac{C_N}{(1+2^j\frac{1}{2}|x|)^N} \frac{|2^j y|^k}{\beta!} dy \\
&\le \frac{C_{N,p,n} 2^{j(k+n)}}{(1+2^j|x|)^N} \left(\int_Q |A(y)|^2 dy \right)^{\frac{1}{2}} \left(\int_Q |y|^{2k} dy \right)^{\frac{1}{2}} \\
&\le \frac{C'_{N,p,n} 2^{j(k+n)}}{(1+2^j|x|)^N} |Q|^{\frac{1}{2}-\frac{1}{p}} |Q|^{\frac{k}{n}+\frac{1}{2}} \\
&= \frac{C_{N,p,n} 2^{j(k+n)}}{(1+2^j|x|)^N} |Q|^{1+\frac{k}{n}-\frac{1}{p}}
\end{aligned}$$

for $x \in (Q^*)^c$. For such x we now have

$$\left(\sum_{j\in\mathbf{Z}} |\Delta_j(A)(x)|^2 \right)^{\frac{1}{2}} \le C_{N,p,n} |Q|^{1+\frac{k}{n}-\frac{1}{p}} \left(\sum_{j\in\mathbf{Z}} \frac{2^{2j(k+n)}}{(1+2^j|x|)^{2N}} \right)^{\frac{1}{2}}. \tag{6.6.16}$$

It is a simple fact that the series in (6.6.16) converges. Indeed, considering the cases $2^j \le 1/|x|$ and $2^j > 1/|x|$ we see that both terms in the second series in (6.6.16) contribute at most a fixed multiple of $|x|^{-2k-2n}$. It remains to estimate the L^p quasinorm of the square root of the second series in (6.6.16) raised over $(Q^*)^c$. This is bounded by a constant multiple of

$$\left(\int_{(Q^*)^c} \frac{1}{|x|^{p(k+n)}} dx \right)^{\frac{1}{p}} \le C_{n,p} \left(\int_{c|Q|^{\frac{1}{n}}}^\infty r^{-p(k+n)+n-1} dr \right)^{\frac{1}{p}},$$

for some constant c, and the latter is easily seen to be bounded above by a constant multiple of $|Q|^{-1-\frac{k}{n}+\frac{1}{p}}$. Here we use the fact that $p(k+n) > n$ or, equivalently, $k > \frac{n}{p} - n$, which is certainly true, since k was chosen to be $[\frac{n}{p} - n] + 1$. Combining this estimate with that in (6.6.15), we conclude the proof of the theorem. □

We now know that L^q-atoms for H^p are indeed elements of H^p. The main result of this section is to obtain the converse (i.e., every element of H^p can be decomposed as a sum of L^2-atoms for H^p).

Applying the same idea as in Corollary 6.6.7 to H^p, we obtain the following result.

Theorem 6.6.10. *Let $0 < p \le 1$. Given a distribution $f \in H^p(\mathbf{R}^n)$, there exists a sequence of L^2-atoms for H^p, $\{A_j\}_{j=1}^\infty$, and a sequence of scalars $\{\lambda_j\}_{j=1}^\infty$ such that*

$$\sum_{j=1}^N \lambda_j A_j \to f \qquad \text{in } H^p.$$

Moreover, we have

$$\|f\|_{H^p} \approx \inf\Big\{ \big(\sum_{j=1}^{\infty} |\lambda_j|^p\big)^{\frac{1}{p}} : \ f = \lim_{N\to\infty} \sum_{j=1}^{N} \lambda_j A_j, \quad A_j \textit{ are } L^2\textit{-atoms for } H^p \textit{ and the limit is taken in } H^p \Big\}. \tag{6.6.17}$$

Proof. Let A_j be L^2-atoms for H^p and $\sum_{j=1}^{\infty} |\lambda_j|^p < \infty$. It follows from Theorem 6.6.9 that

$$\Big\| \sum_{j=1}^{N} \lambda_j A_j \Big\|_{H^p}^p \le C_{n,p}^p \sum_{j=1}^{N} |\lambda_j|^p .$$

Thus if the sequence $\sum_{j=1}^{N} \lambda_j A_j$ converges to f in H^p, then

$$\|f\|_{H^p} \le C_{n,p} \Big(\sum_{j=1}^{\infty} |\lambda_j|^p \Big)^{\frac{1}{p}},$$

which proves the direction $\le$ in (6.6.17). The gist of the theorem is contained in the converse statement.

Using Theorem 6.6.3 (with $L = [\frac{n}{p} - n]$), we can write every element f in $\dot F_p^{0,2} = H^p$ as a sum of the form $f = \sum_{Q\in\mathscr{D}} s_Q a_Q$, where each a_Q is a smooth L-atom for the cube Q and $s = \{s_Q\}_{Q\in\mathscr{D}}$ is a sequence in $\dot f_p^{0,2}$. We now use Theorem 6.6.5 to write the sequence $s = \{s_Q\}_Q$ as

$$s = \sum_{j=1}^{\infty} \lambda_j r_j ,$$

i.e., as a sum of ∞-atoms r_j for $\dot f_p^{0,2}$, such that

$$\big(\sum_{j=1}^{\infty} |\lambda_j|^p \big)^{\frac{1}{p}} \le C \|s\|_{\dot f_p^{0,2}} \le C \|f\|_{H^p} . \tag{6.6.18}$$

Then we have

$$f = \sum_{Q\in\mathscr{D}} s_Q a_Q = \sum_{Q\in\mathscr{D}} \sum_{j=1}^{\infty} \lambda_j r_{j,Q} a_Q = \sum_{j=1}^{\infty} \lambda_j A_j , \tag{6.6.19}$$

where we set

$$A_j = \sum_{Q\in\mathscr{D}} r_{j,Q} a_Q \tag{6.6.20}$$

and the series in (6.6.19) converges in $\mathscr{S}'(\mathbf{R}^n)$. Next we show that each A_j is a fixed multiple of an L^2-atom for H^p. Let us fix an index j. By the definition of the ∞-atom for $\dot f_p^{0,2}$, there exists a dyadic cube Q_0^j such that $r_{j,Q} = 0$ for all dyadic cubes Q not contained in Q_0^j. Then the support of each a_Q that appears in (6.6.20) is contained in $3Q$, hence in $3Q_0^j$. This implies that the function A_j is supported in $3Q_0^j$. The same is true for the function $g^{0,2}(r_j)$ defined in (6.6.1). Using this fact, we have

$$\begin{aligned}\big\|A_j\big\|_{L^2} &\approx \big\|A_j\big\|_{\dot F_2^{0,2}}\\ &\le c\,\big\|r_j\big\|_{\dot f_2^{0,2}}\\ &= c\,\big\|g^{0,2}(r_j)\big\|_{L^2}\\ &\le c\,\big\|g^{0,2}(r_j)\big\|_{L^\infty}|3Q_0^j|^{\frac12}\\ &\le c\,|3Q_0^j|^{-\frac1p+\frac12}\,.\end{aligned}$$

Since the series (6.6.20) defining A_j converges in L^2 and A_j is supported in some cube, this series also converges in L^1. It follows that the vanishing moment conditions of A_j are inherited from those of each a_Q. We conclude that each A_j is a fixed multiple of an L^2-atom for H^p.

Finally, we need to show that the series in (6.6.19) converges in $H^p(\mathbf{R}^n)$. But

$$\Big\|\sum_{j=N}^{M}\lambda_jA_j\Big\|_{H^p}\le C_{n,p}\Big(\sum_{j=N}^{M}|\lambda_j|^p\Big)^{\frac1p}\to 0$$

as $M,N\to\infty$ in view of the convergence of the series in (6.6.18). This implies that the series $\sum_{j=1}^{\infty}\lambda_jA_j$ is Cauchy in H^p, and since it converges to f in $\mathscr{S}'(\mathbf{R}^n)$, it must converge to f in H^p. Combining this fact with (6.6.18) yields the direction $\ge$ in (6.6.17). $\square$

Remark 6.6.11. Property (c) in Definition 6.6.8 can be replaced by

$$\int x^\gamma A(x)\,dx=0 \qquad \text{for all multi-indices } \gamma \text{ with } |\gamma|\le L,$$

for any $L\ge[\frac np-n]$, and the atomic decomposition of H^p holds unchanged. In fact, in the proof of Theorem 6.6.10 we may take $L\ge[\frac np-n]$ instead of $L=[\frac np-n]$ and then apply Theorem 6.6.3 for this L. Observe that Theorem 6.6.3 was valid for all $L\ge[\frac np-n]$.

This observation can be very useful in certain applications.

Exercises

6.6.1. (a) Prove that there exists a Schwartz function Θ supported in the unit ball $|x|\le 1$ such that $\int_{\mathbf{R}^n}x^\gamma\Theta(x)\,dx=0$ for all multi-indices γ with $|\gamma|\le N$ and such that $|\widehat{\Theta}|\ge\frac12$ on the annulus $\frac12\le|\xi|\le 2$.
(b) Prove there exists a Schwartz function Ψ whose Fourier transform is supported in the annulus $\frac12\le|\xi|\le 2$ and is at least $c>0$ in the smaller annulus $\frac35\le|\xi|\le\frac53$ such that we have

$$\sum_{j\in\mathbf{Z}}\widehat{\Psi}(2^{-j}\xi)\widehat{\Theta}(2^{-j}\xi)=1$$

for all $\xi \in \mathbf{R}^n \setminus \{0\}$.
$\big[$*Hint:* Part (a): Let θ be a real-valued Schwartz function supported in the ball $|x| \le 1$ and such that $\widehat{\theta}(0) = 1$. Then for some $\varepsilon > 0$ we have $\widehat{\theta}(\xi) \ge \frac{1}{2}$ for all ξ satisfying $|\xi| < 2\varepsilon < 1$. Set $\Theta = (-\Delta)^N(\theta_\varepsilon)$. Part (b): Define the function $\widehat{\Psi(\xi)} = \widehat{\eta}(\xi)\big(\sum_{j\in\mathbf{Z}}\widehat{\eta}(2^{-j}\xi)\widehat{\Theta}(2^{-j}\xi)\big)^{-1}$ for a suitable η.$\big]$

6.6.2. Let $\alpha \in \mathbf{R}$, $0 < p \le 1$, $p \le q \le +\infty$.
(a) For all f, g in $\mathscr{S}'(\mathbf{R}^n)$ show that

$$\big\|f+g\big\|^p_{\dot F_p^{\alpha,q}} \le \big\|f\big\|^p_{\dot F_p^{\alpha,q}} + \big\|g\big\|^p_{\dot F_p^{\alpha,q}}\,.$$

(b) For all sequences $\{s_Q\}_{Q\in\mathscr{D}}$ and $\{t_Q\}_{Q\in\mathscr{D}}$ show that

$$\big\|\{s_Q\}_Q+\{t_Q\}_Q\big\|^p_{\dot f_p^{\alpha,q}} \le \big\|\{s_Q\}_Q\big\|^p_{\dot f_p^{\alpha,q}} + \big\|\{t_Q\}_Q\big\|^p_{\dot f_p^{\alpha,q}}\,.$$

$\big[$*Hint:* Use $|a+b|^p \le |a|^p + |b|^p$ and apply Minkowski's inequality on $L^{q/p}$ (or on $\ell^{q/p}$).$\big]$

6.6.3. Let Φ be a smooth function supported in the unit ball of $\mathbf{R}^n$. Use the same idea as in Theorem 6.6.9 to show directly (without appealing to any other theorem) that the smooth maximal function $M(\cdot,\Phi)$ of an L^2-atom for H^p lies in L^p when $p < 1$. Recall that $M(f,\Phi) = \sup_{t>0}|\Phi_t * f|$.

6.6.4. Extend Theorem 6.6.9 to the case $1 < q \le \infty$. Precisely, prove that there is a constant $C_{n,p,q}$ such that every L^q-atom A for H^p satisfies

$$\big\|A\big\|_{H^p} \le C_{n,p,q}\,.$$

$\big[$*Hint:* If $1 < q < 2$, use the boundedness of the square function on L^q, and for $2 \le q \le \infty$, its boundedness on L^2.$\big]$

6.6.5. Show that the space H^p_F of all finite linear combinations of L^2-atoms for H^p is dense in H^p.
$\big[$*Hint:* Use Theorem 6.6.10.$\big]$

6.6.6. Show that for all $\mu, j \in \mathbf{Z}$, all $N, b > 0$ satisfying $N > n/b$ and $b < 1$, all scalars s_Q (indexed by dyadic cubes Q with centers c_Q), and all $x \in \mathbf{R}^n$ we have

$$\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}} \frac{|s_Q|}{(1+2^{\min(j,\mu)}|x-c_Q|)^N}$$
$$\le c(n,N,b)\,2^{\max(\mu-j,0)\frac{n}{b}}\bigg\{M\Big(\sum_{\substack{Q\in\mathscr{D}\\ \ell(Q)=2^{-\mu}}}|s_Q|^b\chi_Q\Big)(x)\bigg\}^{\frac{1}{b}},$$

where M is the Hardy–Littlewood maximal operator and $c(n,N,b)$ is a constant. $\big[$*Hint:* Define $\mathscr{F}_0 = \{Q \in \mathscr{D} : \ell(Q) = 2^{-\mu},\ |c_Q - x|\,2^{\min(j,\mu)} \le 1\}$ and for $k \ge 1$ define $\mathscr{F}_k = \{Q \in \mathscr{D} :\ \ell(Q) = 2^{-\mu},\ 2^{k-1} < |c_Q - x|\,2^{\min(j,\mu)} \le 2^k\}$. Break up the sum on the left as a sum over the families $\mathscr{F}_k$ and use that $\sum_{Q\in\mathscr{F}_k} |s_Q| \le \big(\sum_{Q\in\mathscr{F}_k} |s_Q|^b\big)^{1/b}$ and the fact that $\big|\bigcup_{Q\in\mathscr{F}_k} Q\big| \le c_n 2^{-\min(j,\mu)n+kn}$.$\big]$

6.6.7. Let A be an L^2-atom for $H^p(\mathbf{R}^n)$ for some $0 < p < 1$. Show that there is a constant C such that for all multi-indices α with $|\alpha| \le k = [\frac{n}{p} - n]$ we have

$$\sup_{\xi\in\mathbf{R}^n} |\xi|^{|\alpha|-k-1}\big|(\partial^\alpha \widehat{A})(\xi)\big| \le C\,\big\|A\big\|_{L^2(\mathbf{R}^n)}^{-\frac{2p}{2-p}(\frac{k+1}{n}+\frac{1}{2})-1}.$$

[*Hint:* Subtract the Taylor polynomial of degree $k - |\alpha|$ at 0 of the function $x \mapsto e^{-2\pi i x\cdot\xi}$.]

6.6.8. Let A be an L^2-atom for $H^p(\mathbf{R}^n)$ for some $0 < p < 1$. Show that for all multi-indices α and all $1 \le r \le \infty$ there is a constant C such that

$$\big\|\,|\partial^\alpha \widehat{A}|^2\big\|_{L^{r'}(\mathbf{R}^n)} \le C\,\big\|A\big\|_{L^2(\mathbf{R}^n)}^{-\frac{2p}{2-p}(\frac{2|\alpha|}{n}+\frac{1}{r})+2}.$$

$\big[$*Hint:* In the case $r = 1$ use the $L^1 \to L^\infty$ boundedness of the Fourier transform and in the case $r = \infty$ use Plancherel's theorem. For general r use interpolation.$\big]$

6.6.9. Let f be in $H^p(\mathbf{R}^n)$ for some $0 < p \le 1$. Then the Fourier transform of f, originally defined as a tempered distribution, is a continuous function that satisfies

$$|\widehat{f}(\xi)| \le C_{n,p}\big\|f\big\|_{H^p(\mathbf{R}^n)}|\xi|^{\frac{n}{p}-n}$$

for some constant $C_{n,p}$ independent of f.
$\big[$*Hint:* If f is an L^2-atom for H^p, combine the estimates of Exercises 6.6.7 and 6.6.8 with $\alpha = 0$ (and $r = 1$). In general, apply Theorem 6.6.10.$\big]$

6.6.10. Let A be an L^∞-atom for $H^p(\mathbf{R}^n)$ for some $0 < p < 1$ and let $\alpha = \frac{n}{p} - n$. Show that there is a constant $C_{n,p}$ such that for all g in $\dot{\Lambda}_\alpha(\mathbf{R}^n)$ we have

$$\left|\int_{\mathbf{R}^n} A(x)g(x)\,dx\right| \le C_{n,p}\,\big\|g\big\|_{\dot{\Lambda}_\alpha(\mathbf{R}^n)}.$$

$\big[$*Hint:* Suppose that A is supported in a cube Q of side length $2^{-\nu}$ and center c_Q. Write the previous integrand as $\sum_j \Delta_j(A)\Delta_j(g)$ for a suitable Littlewood–Paley operator Δ_j and apply the result of Appendix K.2 to obtain the estimate

$$\big|\Delta_j(A)(x)\big| \le C_N |Q|^{-\frac{1}{p}+1}\frac{2^{\min(j,\nu)n}2^{-|j-\nu|D}}{\big(1+2^{\min(j,\nu)}|x-c_Q|\big)^N},$$

where $D = [\alpha] + 1$ when $\nu \ge j$ and $D = 0$ when $\nu < j$. Use Theorem 6.3.6.$\big]$

6.7 Singular Integrals on Function Spaces

Our final task in this chapter is to investigate the action of singular integrals on function spaces. The emphasis of our study focuses on Hardy spaces, although with no additional effort the action of singular integrals on other function spaces can also be obtained.

6.7.1 Singular Integrals on the Hardy Space H^1

Before we discuss the main results in this topic, we review some background on singular integrals from Chapter 4.

Let $K(x)$ be a function defined away from the origin on $\mathbf{R}^n$ that satisfies the size estimate

$$\sup_{0<R<\infty} \frac{1}{R} \int_{|x|\le R} |K(x)|\, |x|\, dx \le A_1\,, \tag{6.7.1}$$

the smoothness estimate, expressed in terms of Hörmander's condition,

$$\sup_{y\in\mathbf{R}^n\setminus\{0\}} \int_{|x|\ge 2|y|} |K(x-y)-K(x)|\, dx \le A_2\,, \tag{6.7.2}$$

and the cancellation condition

$$\sup_{0<R_1<R_2<\infty} \left| \int_{R_1<|x|<R_2} K(x)\, dx \right| \le A_3\,, \tag{6.7.3}$$

for some $A_1, A_2, A_3 < \infty$. Condition (6.7.3) implies that there exists a sequence $\varepsilon_j \downarrow 0$ as $j\to\infty$ such that the following limit exists:

$$\lim_{j\to\infty} \int_{\varepsilon_j\le|x|\le 1} K(x)\, dx = L_0.$$

This gives that for a smooth and compactly supported function f on $\mathbf{R}^n$, the limit

$$\lim_{j\to\infty} \int_{|x-y|>\varepsilon_j} K(x-y) f(y)\, dy = T(f)(x) \tag{6.7.4}$$

exists and defines a linear operator T. This operator T is given by convolution with a tempered distribution W that coincides with the function K on $\mathbf{R}^n\setminus\{0\}$.

By the results of Chapter 4 we know that such a T, initially defined on $\mathscr{C}_0^\infty(\mathbf{R}^n)$, admits an extension that is L^p bounded for all $1<p<\infty$ and is also of weak type $(1,1)$. All these norms are bounded above by dimensional constant multiples of the quantity $A_1+A_2+A_3$ (cf. Theorem 4.4.1). Therefore, such a T is well defined on

$L^1(\mathbf{R}^n)$ and in particular on $H^1(\mathbf{R}^n)$, which is contained in $L^1(\mathbf{R}^n)$. We begin with the following result.

Theorem 6.7.1. *Let K satisfy (6.7.1), (6.7.2), and (6.7.3), and let T be defined as in (6.7.4). Then there is a constant C_n such that for all f in $H^1(\mathbf{R}^n)$ we have*

$$\big\|T(f)\big\|_{L^1} \le C_n(A_1+A_2+A_3)\big\|f\big\|_{H^1}. \tag{6.7.5}$$

Proof. To prove this theorem we have a powerful tool at our disposal, the atomic decomposition of $H^1(\mathbf{R}^n)$. It is therefore natural to start by checking the validity of (6.7.5) whenever f is an L^2-atom for H^1.

Since T is a convolution operator (i.e., it commutes with translations), it suffices to take the atom f supported in a cube Q centered at the origin. Let $f = a$ be such an atom, supported in Q, and let $Q^* = 2\sqrt{n}\,Q$. We write

$$\int_{\mathbf{R}^n} |T(a)(x)|\,dx = \int_{Q^*} |T(a)(x)|\,dx + \int_{(Q^*)^c} |T(a)(x)|\,dx \tag{6.7.6}$$

and we estimate each term separately. We have

$$\begin{aligned}
\int_{Q^*} |T(a)(x)|\,dx &\le |Q^*|^{\frac12}\bigg(\int_{Q^*} |T(a)(x)|^2\,dx\bigg)^{\frac12} \\
&\le C_n(A_1+A_2+A_3)|Q^*|^{\frac12}\bigg(\int_{Q} |a(x)|^2\,dx\bigg)^{\frac12} \\
&\le C_n(A_1+A_2+A_3)|Q^*|^{\frac12}|Q|^{\frac12-\frac11} = C_n'(A_1+A_2+A_3),
\end{aligned}$$

where we used property (b) of atoms in Definition 6.6.8. Now note that if $x \notin Q^*$ and $y \in Q$, then $|x| \ge 2|y|$ and $x-y$ stays away from zero; thus $K(x-y)$ is well defined. Moreover, in this case $T(a)(x)$ can be expressed as a convergent integral of $a(y)$ against $K(x-y)$. We have

$$\begin{aligned}
\int_{(Q^*)^c} |T(a)(x)|\,dx &= \int_{(Q^*)^c} \bigg|\int_Q K(x-y)a(y)\,dy\bigg|\,dx \\
&= \int_{(Q^*)^c} \bigg|\int_Q \big(K(x-y)-K(x)\big)a(y)\,dy\bigg|\,dx \\
&\le \int_Q \int_{(Q^*)^c} \big|K(x-y)-K(x)\big|\,dx\,|a(y)|\,dy \\
&\le \int_Q \int_{|x|\ge 2|y|} \big|K(x-y)-K(x)\big|\,dx\,|a(y)|\,dy \\
&\le A_2 \int_Q |a(x)|\,dx \\
&\le A_2|Q|^{\frac12}\bigg(\int_Q |a(x)|^2\,dx\bigg)^{\frac12} \\
&\le A_2|Q|^{\frac12}|Q|^{\frac12-\frac11} = A_2.
\end{aligned}$$

Combining this calculation with the previous one and inserting the final conclusions in (6.7.6) we deduce that L^2-atoms a for H^1 satisfy

$$\big\|T(a)\big\|_{L^1} \le (C_n'+1)(A_1+A_2+A_3). \tag{6.7.7}$$

We now pass to general functions in H^1. In view of Theorem 6.6.10 we can write an $f \in H^1$ as

$$f = \sum_{j=1}^{\infty} \lambda_j a_j ,$$

where the series converges in H^1, the a_j are L^2-atoms for H^1, and

$$\big\|f\big\|_{H^1} \approx \sum_{j=1}^{\infty} |\lambda_j| . \tag{6.7.8}$$

Since T maps L^1 to weak L^1 (Theorem 4.3.3), $T(f)$ is already a well defined $L^{1,\infty}$ function. We plan to prove that

$$T(f) = \sum_{j=1}^{\infty} \lambda_j T(a_j) \qquad \text{a.e.} \tag{6.7.9}$$

We observe that the series in (6.7.9) converges in L^1. Once (6.7.9) is established, the required conclusion (6.7.5) follows easily by taking L^1 norms in (6.7.9) and using (6.7.7) and (6.7.8).

To prove (6.7.9), we show that T is of weak type $(1,1)$. For a given $\delta > 0$ we have

$$\begin{aligned}
&\Big|\Big\{\Big|T(f) - \sum_{j=1}^{\infty} \lambda_j T(a_j)\Big| > \delta\Big\}\Big| \\
&\le \Big|\Big\{\Big|T(f) - \sum_{j=1}^{N} \lambda_j T(a_j)\Big| > \delta/2\Big\}\Big| + \Big|\Big\{\Big|\sum_{j=N+1}^{\infty} \lambda_j T(a_j)\Big| > \delta/2\Big\}\Big| \\
&\le \frac{2}{\delta}\big\|T\big\|_{L^1\to L^{1,\infty}}\Big\|f - \sum_{j=1}^{N}\lambda_j a_j\Big\|_{L^1} + \frac{2}{\delta}\Big\|\sum_{j=N+1}^{\infty}\lambda_j T(a_j)\Big\|_{L^1} \\
&\le \frac{2}{\delta}\big\|T\big\|_{L^1\to L^{1,\infty}}\Big\|f - \sum_{j=1}^{N}\lambda_j a_j\Big\|_{H^1} + \frac{2}{\delta}(C_n'+1)(A_1+A_2+A_3)\sum_{j=N+1}^{\infty}|\lambda_j| .
\end{aligned}$$

Since $\sum_{j=1}^{N}\lambda_j a_j$ converges to f in H^1 and $\sum_{j=1}^{\infty}|\lambda_j| < \infty$, both terms in the sum converge to zero as $N \to \infty$. We conclude that

$$\Big|\Big\{\Big|T(f) - \sum_{j=1}^{\infty} \lambda_j T(a_j)\Big| > \delta\Big\}\Big| = 0$$

for all $\delta > 0$, which implies (6.7.9). □

6.7.2 Singular Integrals on Besov–Lipschitz Spaces

We continue with a corollary concerning Besov–Lipschitz spaces.

Corollary 6.7.2. *Let K satisfy (6.7.1), (6.7.2), and (6.7.3), and let T be defined as in (6.7.4). Let $1 \le p \le \infty$, $0 < q \le \infty$, and $\alpha \in \mathbf{R}$. Then there is a constant $C_{n,p,q,\alpha}$ such that for all f in $\mathscr{S}(\mathbf{R}^n)$ we have*

$$\big\|T(f)\big\|_{\dot B_p^{\alpha,q}} \le C_n(A_1+A_2+A_3)\big\|f\big\|_{\dot B_p^{\alpha,q}}. \tag{6.7.10}$$

Therefore, T admits a bounded extension on all homogeneous Besov–Lipschitz spaces $\dot B_p^{\alpha,q}$ with $p \ge 1$, in particular, on all homogeneous Lipschitz spaces.

Proof. Let Ψ be a Schwartz function whose Fourier transform is supported in the annulus $1-\frac{1}{7} \le |\xi| \le 2$ and that satisfies

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1\,, \qquad \xi \ne 0\,.$$

Pick a Schwartz function ζ whose Fourier transform $\widehat{\zeta}$ is supported in the annulus $\frac{1}{4} < |\xi| < 8$ and that is equal to one on the support of $\widehat{\Psi}$. Let W be the tempered distribution that coincides with K on $\mathbf{R}^n \setminus \{0\}$ so that $T(f) = f * W$. Then we have $\zeta_{2^{-j}} * \Psi_{2^{-j}} = \Psi_{2^{-j}}$ for all j and hence

$$\begin{aligned}\big\|\Delta_j(T(f))\big\|_{L^p} &= \big\|\zeta_{2^{-j}} * \Psi_{2^{-j}} * W * f\big\|_{L^p}\\ &\le \big\|\zeta_{2^{-j}} * W\big\|_{L^1}\big\|\Delta_j(f)\big\|_{L^p},\end{aligned} \tag{6.7.11}$$

since $1 \le p \le \infty$. It is not hard to check that the function $\zeta_{2^{-j}}$ is in H^1 with norm independent of j. Therefore, $\zeta_{2^{-j}}$ is in H^1. Using Theorem 6.7.1, we conclude that

$$\big\|T(\zeta_{2^{-j}})\big\|_{L^1} = \big\|\zeta_{2^{-j}} * W\big\|_{L^1} \le C\big\|\zeta_{2^{-j}}\big\|_{H^1} = C'\,.$$

Inserting this in (6.7.11), multiplying by $2^{j\alpha}$, and taking ℓ^q norms, we obtain the required conclusion. □

6.7.3 Singular Integrals on $H^p(\mathbf{R}^n)$

We are now interested in extending Theorem 6.7.1 to other H^p spaces for $p < 1$. It turns out that this is possible, provided some additional smoothness assumptions on K are imposed.

For the purposes of this subsection, we fix a function $K(x)$ on $\mathbf{R}^n \setminus \{0\}$ that satisfies $|K(x)| \le A|x|^{-n}$ for $x \ne 0$ and we assume that there is a distribution W in $\mathscr{S}'(\mathbf{R}^n)$ that coincides with K on $\mathbf{R}^n \setminus \{0\}$. We make two assumptions about the distribution W: first, that its Fourier transform $\widehat{W}$ is a bounded function, i.e., it satisfies

$$|\widehat{W}(\xi)| \le B, \qquad \xi \in \mathbf{R}^n, \tag{6.7.12}$$

for some $B < \infty$; secondly, that W is obtained from the function K as a limit of its smooth truncations. This allows us to properly define the convolution of this distribution with elements of H^p. So we fix a nonnegative smooth function η that vanishes in the unit ball of $\mathbf{R}^n$ and is equal to one outside the ball $B(0,2)$. We assume that for some sequence $\varepsilon_j \in (0,1)$ with $\varepsilon_j \downarrow 0$ the distribution W has the form

$$\langle W, \varphi \rangle = \lim_{j\to\infty} \int_{\mathbf{R}^n} K(y)\eta(y/\varepsilon_j)\varphi(y)\,dy \tag{6.7.13}$$

for all $\varphi \in \mathscr{S}(\mathbf{R}^n)$. Then we define the smoothly truncated singular integral associated with K and η by

$$T_\eta^{(\varepsilon)}(f)(x) = \int_{\mathbf{R}^n} \eta(y/\varepsilon)K(y)f(x-y)\,dy$$

for Schwartz functions f [actually the integral is absolutely convergent for every $f \in L^p$ and any $p \in [1,\infty)$]. We also define an operator T given by convolution with W by

$$T(f) = \lim_{j\to\infty} T^{(\varepsilon_j)}(f) = f * W. \tag{6.7.14}$$

This provides a representation of the operator T. If the function K satisfies condition (4.4.3), this representation is also valid pointwise almost everywhere for functions $f \in L^2$, i.e., $\lim_{j\to\infty} T^{(\varepsilon_j)}(f)(x) = T(f)(x)$ for almost all $x \in \mathbf{R}^n$. This follows from Theorem 4.4.5, Exercise 4.3.10, and Theorem 2.1.14 (since the convergence holds for Schwartz functions).

Next we define $T(f)$ for $f \in H^p$. One can write $W = W_0 + K_\infty$, where $W_0 = \Phi W$ and $K_\infty = (1-\Phi)K$, where Φ is a smooth function equal to one on the ball $B(0,1)$ and vanishing off the ball $B(0,2)$. Then for f in $H^p(\mathbf{R}^n)$, $0 < p \le 1$, we may define a tempered distribution $T(f) = W * f$ by setting

$$\langle T(f), \phi \rangle = \langle f, \phi * \widetilde{W_0} \rangle + \langle \widetilde{\phi} * f, \widetilde{K_\infty} \rangle \tag{6.7.15}$$

for ϕ in $\mathscr{S}(\mathbf{R}^n)$. The function $\phi * \widetilde{W_0}$ is in $\mathscr{S}$, so the action of f on it is well defined. Also $\widetilde{\phi} * f$ is in L^1 (see Proposition 6.4.9), while $\widetilde{K_\infty}$ is in L^∞; hence the second term on the right above represents an absolutely convergent integral. Moreover, in view of Theorem 2.3.20 and Corollary 6.4.9, both terms on the right in (6.7.15) are controlled by a finite sum of seminorms $\rho_{\alpha,\beta}(\phi)$ (cf. Definition 2.2.1). This defines $T(f)$ as a tempered distribution.

The following is an extension of Theorem 6.7.1 for $p < 1$.

Theorem 6.7.3. *Let* $0 < p \le 1$ *and* $N = [\frac{n}{p} - n] + 1$. *Let* K *be a* $\mathscr{C}^N$ *function on* $\mathbf{R}^n \setminus \{0\}$ *that satisfies*

$$|\partial^\beta K(x)| \le A\,|x|^{-n-|\beta|} \tag{6.7.16}$$

for all multi-indices $|\beta| \le N$ *and all* $x \neq 0$. *Let* W *be a tempered distribution that coincides with* K *on* $\mathbf{R}^n \setminus \{0\}$ *and satisfies* (6.7.12) *and* (6.7.13). *Then there is a constant* $C_{n,p}$ *such that the operator* T *defined in* (6.7.15) *satisfies, for all* $f \in H^p$,

$$\big\|T(f)\big\|_{L^p} \le C_{n,p}(A+B)\big\|f\big\|_{H^p}.$$

Proof. The proof of this theorem is based on the atomic decomposition of H^p.

We first take $f = a$ to be an L^2-atom for H^p, and without loss of generality we may assume that a is supported in a cube Q centered at the origin. We let Q^* be the cube with side length $2\sqrt{n}\,\ell(Q)$, where $\ell(Q)$ is the side length of Q. We have

$$\begin{aligned}
\left(\int_{Q^*} |T(a)(x)|^p\,dx\right)^{\frac{1}{p}} &\le C|Q^*|^{\frac{1}{p}-\frac{1}{2}}\left(\int_{Q^*} |T(a)(x)|^2\,dx\right)^{\frac{1}{2}} \\
&\le C''B|Q|^{\frac{1}{p}-\frac{1}{2}}\left(\int_{Q} |a(x)|^2\,dx\right)^{\frac{1}{2}} \\
&\le C_nB|Q|^{\frac{1}{p}-\frac{1}{2}}|Q|^{\frac{1}{2}-\frac{1}{p}} \\
&= C_nB\,.
\end{aligned}$$

For $x \notin Q^*$ and $y \in Q$, we have $|x| \ge 2|y|$, and thus $x-y$ stays away from zero and $K(x-y)$ is well defined. We have

$$T(a)(x) = \int_Q K^{(t)}(x-y)\,a(y)\,dy\,.$$

Recall that $N = [\frac{n}{p} - n] + 1$. Using the cancellation of atoms for H^p, we deduce

$$\begin{aligned}
T(a)(x) &= \int_Q a(y)K(x-y)\,dy \\
&= \int_Q a(y)\left[K(x-y) - \sum_{|\beta|\le N-1} (\partial^\beta K(x)\frac{(y)^\beta}{\beta!}\right]dy \\
&= \int_Q a(y)\left[\sum_{|\beta|=N} (\partial^\beta K(x-\theta_y y)\frac{(y)^\beta}{\beta!}\right]dy
\end{aligned}$$

for some $0 \le \theta_y \le 1$. Using that $|x| \ge 2|y|$ and (6.7.23), we obtain the estimate

$$|T(a)(x)| \le c_{n,N}\frac{A}{|x|^{N+n}}\int_Q |a(y)|\,|y|^{|\beta|}\,dy\,,$$

from which it follows that for $x \notin Q^*$ we have

$$|T(a)(x)| \le c_{n,p}\frac{A}{|x|^{N+n}}|Q|^{1+\frac{N}{n}-\frac{1}{p}}$$

via a calculation using properties of atoms (see the proof of Theorem 6.6.9). Integrating over $(Q^*)^c$, we obtain that

$$\left(\int_{(Q^*)^c} |T(a)(x)|^p dx\right)^{\frac{1}{p}} \le c_{n,p} A\,|Q|^{1+\frac{N}{n}-\frac{1}{p}} \left(\int_{(Q^*)^c} \frac{1}{|x|^{p(N+n)}}\,dx\right)^{\frac{1}{p}} \le c'_{n,p} A\,.$$

We have now shown that there exists a constant $C_{n,p}$ such that

$$\big\|T(a)\big\|_{L^p} \le C_{n,p}\,(A+B) \tag{6.7.17}$$

whenever a is an L^2-atom for H^p. We need to extend this estimate to infinite sums of atoms. To achieve this, it convenient to use operators with more regular kernels and then approximate T by such operators.

Recall the smooth function η that vanishes when $|x| \le 1$ and is equal to 1 when $|x| \ge 2$. We fix a smooth function θ with support in the unit ball having integral equal to 1. We define $\theta_\delta(x) = \delta^{-n}\theta(x/\delta)$,

$$K_{\varepsilon,\mu}(x) = K(x)\big(\eta(x/\varepsilon) - \eta(\mu x)\big)$$

and

$$K_{\delta,\varepsilon,\mu} = \theta_\delta * K_{\varepsilon,\mu}$$

for $0 < 10\delta < \varepsilon < (10\mu)^{-1}$. We make the following observations: first $K_{\delta,\varepsilon,\mu}$ is $\mathscr{C}^\infty$; second, it has rapid decay at infinity, and hence it is a Schwartz function; third, it satisfies (6.7.16) for all $|\beta| \le N$ with constant a multiple of A, that is, independent of δ, ε, μ. Let $T_{\delta,\varepsilon,\mu}$ be the operator given by convolution with $K_{\delta,\varepsilon,\mu}$ and let $T_\eta^{(*)}$ be the maximal smoothly truncated singular integral associated with the bump η. Then for $h \in L^2$ we have

$$\big\|T_{\delta,\varepsilon,\mu}(h)\big\|_{L^2} \le 2\big\|T_\eta^{(*)}(\theta_\delta * h)\big\|_{L^2} \le C_n\,(A+B)\big\|\theta_\delta * h\big\|_{L^2} \le C_n\,(A+B)\big\|h\big\|_{L^2}\,;$$

hence $T_{\delta,\varepsilon,\mu}$ maps L^2 to L^2 with norm a fixed multiple of $A+B$. The proof of (6.7.17) thus yields for any L^2-atom a for H^p the estimate

$$\big\|T_{\delta,\varepsilon,\mu}(a)\big\|_{L^p} \le C'_{n,p}\,(A+B) \tag{6.7.18}$$

with a constant $C'_{n,p}$ that is independent of δ,ε,μ.

Let f be in $L^2 \cap H^p$, which is a dense subspace of H^p, and suppose that $f = \sum_j \lambda_j a_j$, where a_j are L^2-atoms for H^p, the series converges in H^p, and we have

$$\sum_j |\lambda_j|^p \le C_p^p \big\|f\big\|_{H^p(\mathbf{R}^n)}^p\,. \tag{6.7.19}$$

We set $f_M = \sum_{j=1}^M \lambda_j a_j$. Then f_M, f are in L^2 but $f_M \to f$ in H^p; hence by Proposition 6.4.10, $f_M \to f$ in $\mathscr{S}'$. Acting on the Schwartz functions $K_{\delta,\varepsilon,\mu}(x-\cdot)$, we obtain that

$$T_{\delta,\varepsilon,\mu}(f_M)(x) \to T_{\delta,\varepsilon,\mu}(f)(x) \qquad \text{as } M \to \infty \text{ for all } x \in \mathbf{R}^n\,. \tag{6.7.20}$$

Recall the discussion in the introduction of this section defining $T = \lim_{j\to\infty} T^{(\varepsilon_j)}$ in an appropriate sense. Let $h \in L^2(\mathbf{R}^n)$. Since $h * K_{\delta,\varepsilon,\mu}$ is a continuous function, Theorem 1.2.19 (b) gives that

$$T_{\delta,\varepsilon_j,\mu}(h) \to T_\eta^{(\varepsilon_j)}(h) - T_\eta^{(1/\mu)}(h) \tag{6.7.21}$$

pointwise as $\delta \to 0$, where $T_\eta^{(\varepsilon)}$ is the smoothly truncated singular integral associated with the bump η (cf. Exercise 4.3.10). The expressions on the right in (6.7.21) are obviously pointwise bounded by $2T_\eta^{(*)}(h)$. Since $T_\eta^{(*)}$ is an L^2 bounded operator, and $T_\eta^{(\varepsilon_j)}(\psi) - T_\eta^{(1/\mu)}(\psi) \to T(\psi)$ for every $\psi \in \mathscr{S}(\mathbf{R}^n)$, it follows from Theorem 2.1.14 that $T_\eta^{(\varepsilon_j)}(h) - T_\eta^{(1/\mu)}(h) \to T(h)$ pointwise a.e. as $\varepsilon_j, \mu \to 0$. Thus $T_{\delta,\varepsilon_j,\mu}(h) \to T(h)$ pointwise a.e. as $\delta \to 0$, $\mu \to 0$, and $\varepsilon_j \to 0$ in this order. Using this fact, (6.7.20), and Fatou's lemma, we deduce for the given $f, f_M \in L^2 \cap H^p$ that

$$\big\|T(f)\big\|_{L^p}^p \le \liminf_{\delta,\mu,\varepsilon_j\to 0} \big\|T_{\delta,\varepsilon_j,\mu}(f)\big\|_{L^p}^p \le \liminf_{\delta,\mu,\varepsilon_j\to 0} \liminf_{M\to\infty} \big\|T_{\delta,\varepsilon_j,\mu}(f_M)\big\|_{L^p}^p .$$

The last displayed expression is at most $(C_pC'_{n,p})^p(A+B)^p\big\|f\big\|_{H^p}^p$ using the sublinearity of the pth power of the L^p norm, (6.7.18), and (6.7.19).

This proves the required assertion for $f \in H^p \cap L^2$. The case of general $f \in H^p$ follows by density and the fact that $T(f)$ is well defined for all $f \in H^p$, as observed at the beginning of this subsection. □

We discuss another version of the previous theorem in which the target space is H^p.

Theorem 6.7.4. *Under the hypotheses of Theorem 6.7.3, we have the following conclusion: there is a constant $C_{n,p}$ such that the operator T satisfies, for all $f \in H^p$,*

$$\big\|T(f)\big\|_{H^p} \le C_{n,p}(A+B)\big\|f\big\|_{H^p} .$$

Proof. The proof of this theorem provides another classical application of the atomic decomposition of H^p. However, we use the atomic decomposition only for the domain Hardy space, while it is more convenient to use the maximal (or square function) characterization of H^p for the target H^p space.

We fix a smooth function Φ supported in the unit ball $B(0,1)$ in $\mathbf{R}^n$ whose mean value is not equal to zero. For $t > 0$ we define the smooth functions

$$W^{(t)} = \Phi_t * W$$

and we observe that they satisfy

$$\sup_{t>0} \big|\widehat{W^{(t)}}(\xi)\big| \le \big\|\widehat{\Phi}\big\|_{L^\infty} B \tag{6.7.22}$$

and that

$$\sup_{t>0} |\partial^\beta W^{(t)}(x)| \le C_\Phi A\, |x|^{-n-|\beta|} \tag{6.7.23}$$

for all $|\beta| \le N$, where

$$C_\Phi = \sup_{|\gamma|\le N} \int_{\mathbf{R}^n} |\xi|^{|\gamma|} |\widehat{\Phi}(\xi)|\, d\xi\, .$$

Indeed, assertion (6.7.22) is easily verified, while assertion (6.7.23) follows from the identity

$$W^{(t)}(x) = \big((\Phi_t * W)\widehat{\ }\big)^\vee(x) = \int_{\mathbf{R}^n} e^{2\pi i x\cdot\xi}\, \widehat{W}(\xi)\, \widehat{\Phi}(t\xi)\, d\xi$$

whenever $|x| \le 2t$ and from (6.7.16) and the fact that for $|x| \ge 2t$ we have the integral representation

$$\partial^\beta W^{(t)}(x) = \int_{|y|\le t} \partial^\beta K(x-y)\, \Phi_t(y)\, dy\, .$$

We now take $f = a$ to be an L^2-atom for H^p, and without loss of generality we may assume that a is supported in a cube Q centered at the origin. We let Q^* be the cube with side length $2\sqrt{n}\,\ell(Q)$, where $\ell(Q)$ is the side length of Q. Recall the smooth maximal function $M(f;\Phi)$ from Section 6.4. Then $M(T(a);\Phi)$ is pointwise controlled by the Hardy–Littlewood maximal function of $T(a)$. Using an argument similar to that in Theorem 6.7.1, we have

$$\begin{aligned}
\left(\int_{Q^*} |M(T(a);\Phi)(x)|^p\, dx\right)^{\frac{1}{p}} &\le \big\|\Phi\big\|_{L^1} \left(\int_{Q^*} |M(T(a))(x)|^p\, dx\right)^{\frac{1}{p}} \\
&\le C|Q^*|^{\frac{1}{p}-\frac{1}{2}} \left(\int_{Q^*} |M(T(a))(x)|^2\, dx\right)^{\frac{1}{2}} \\
&\le C'|Q|^{\frac{1}{p}-\frac{1}{2}} \left(\int_{\mathbf{R}^n} |T(a)(x)|^2\, dx\right)^{\frac{1}{2}} \\
&\le C''B|Q|^{\frac{1}{p}-\frac{1}{2}} \left(\int_{Q} |a(x)|^2\, dx\right)^{\frac{1}{2}} \\
&\le C_n B|Q|^{\frac{1}{p}-\frac{1}{2}} |Q|^{\frac{1}{2}-\frac{1}{p}} \\
&= C_n B\, .
\end{aligned}$$

It therefore remains to estimate the contribution of $M(T(a);\Phi)$ on the complement of Q^*.

If $x \notin Q^*$ and $y \in Q$, then $|x| \ge 2|y|$ and hence $x - y \ne 0$. Thus $K(x-y)$ is well defined as an integral. We have

$$(T(a) * \Phi_t)(x) = (a * W^{(t)})(x) = \int_Q K^{(t)}(x-y)\, a(y)\, dy\, .$$

Recall that $N = [\frac{n}{p} - n] + 1$. Using the cancellation of atoms for H^p we deduce

$$\begin{aligned}(T(a) * \Phi_t)(x) &= \int_Q a(y) \left[K^{(t)}(x-y) - \sum_{|\beta| \le N-1} (\partial^\beta K^{(t)})(x) \frac{(y)^\beta}{\beta!} \right] dy \\ &= \int_Q a(y) \left[\sum_{|\beta|=N} (\partial^\beta K^{(t)})(x - \theta_y y) \frac{(y)^\beta}{\beta!} \right] dy\end{aligned}$$

for some $0 \le \theta_y \le 1$. Using that $|x| \ge 2|y|$ and (6.7.23), we obtain the estimate

$$\big|(T(a) * \Phi_t)(x)\big| \le c_{n,N} \frac{A}{|x|^{N+n}} \int_Q |a(y)|\, |y|^{|\beta|}\, dy\,,$$

from which it follows that for $x \notin Q^*$ we have

$$\big|(T(a) * \Phi_t)(x)\big| \le c_{n,p} \frac{A}{|x|^{N+n}} |Q|^{1+\frac{N}{n}-\frac{1}{p}}$$

via a calculation using properties of atoms (see the proof of Theorem 6.6.9). Taking the supremum over all $t > 0$ and integrating over $(Q^*)^c$, we obtain that

$$\left(\int_{(Q^*)^c} \sup_{t>0} |(T(a) * \Phi_t)(x)|^p dx \right)^{\frac{1}{p}} \le c_{n,p} A\, |Q|^{1+\frac{N}{n}-\frac{1}{p}} \left(\int_{(Q^*)^c} \frac{1}{|x|^{p(N+n)}}\, dx \right)^{\frac{1}{p}},$$

and the latter is easily seen to be finite and controlled by a constant multiple of A. Combining this estimate with the previously obtained estimate for the integral of $M(T(a); \Phi) = \sup_{t>0} |(T(a) * \Phi_t|$ over Q^* yields the conclusion of the theorem when $f = a$ is an atom.

We have now shown that there exists a constant $C_{n,p}$ such that

$$\big\|T(a)\big\|_{H^p} \le C_{n,p}(A+B) \tag{6.7.24}$$

whenever a is an L^2-atom for H^p. We need to extend this estimate to infinite sums of atoms.

Let f be $L^2 \cap H^p$ which is a dense subspace of H^p, and suppose that $f = \sum_j \lambda_j a_j$ for some L^2-atoms a_j for H^p, where the series converges in H^p and we have

$$\sum_j |\lambda_j|^p \le C_p^p \big\|f\big\|_{H^p(\mathbf{R}^n)}^p. \tag{6.7.25}$$

We let $f_M = \sum_{j=1}^M \lambda_j a_j$ and we recall the smooth truncations $T_{\delta,\varepsilon_j,\mu}$ of T. As $f_M \to f$ in H^p, Proposition 6.4.10 gives that $f_M \to f$ in $\mathscr{S}'$, and since the functions $K_{\delta,\varepsilon_j,\mu}$ are smooth with compact support, it follows that for all $\delta, \varepsilon_j, \mu$,

$$T_{\delta,\varepsilon_j,\mu}(f_M) \to T_{\delta,\varepsilon_j,\mu}(f) \qquad \text{in } \mathscr{S}' \text{ as } M \to \infty. \tag{6.7.26}$$

We show that this convergence is also valid for T. Given $\varepsilon > 0$ and a Schwartz function φ, we find $\delta_0, \varepsilon_{j_0}, \mu_0$ such that

$$\big|\big\langle T(f_M),\varphi\big\rangle - \big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f_M),\varphi\big\rangle\big| < \varepsilon\, C_p \big\|f\big\|_{H^p} \quad \text{for all } M = 1,2,\dots. \tag{6.7.27}$$

To find such $\delta_0, \varepsilon_{j_0}, \mu_0$, we write

$$\begin{aligned}\big|\big\langle T(f_M),\varphi\big\rangle - \big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f_M),\varphi\big\rangle\big| &\le \Big|\sum_{j=1}^{M}\lambda_j\big\langle (K_{\delta_0,\varepsilon_{j_0},\mu_0} - W)*a_j,\varphi\big\rangle\Big| \\ &\le \bigg(\sum_{j=1}^{M}|\lambda_j|^p\big|\big\langle a_j,(\widetilde{K}_{\delta_0,\varepsilon_{j_0},\mu_0} - \widetilde{W})*\varphi\big\rangle\big|^p\bigg)^{\frac{1}{p}} \\ &\le \bigg(\sum_{j=1}^{M}|\lambda_j|^p\big\|a_j\big\|_{L^2}^p\big\|(\widetilde{K}_{\delta_0,\varepsilon_{j_0},\mu_0} - \widetilde{W})*\varphi\big\|_{L^2}^p\bigg)^{\frac{1}{p}} \\ &\le C_p\big\|f\big\|_{H^p}\big\|(K_{\delta_0,\varepsilon_{j_0},\mu_0} - W)*\widetilde{\varphi}\,\big\|_{L^2}.\end{aligned}$$

Now pick $\delta_0, \varepsilon_{j_0}, \mu_0$ such that

$$\big\|(K_{\delta_0,\varepsilon_{j_0},\mu_0} - W)*\widetilde{\varphi}\,\big\|_{L^2} = \big\|((K_{\delta_0,\varepsilon_{j_0},\mu_0})\widehat{\ } - \widehat{W})\widehat{\widetilde{\varphi}}\,\big\|_{L^2} < \varepsilon.$$

This is possible, since this expression tends to zero when $\delta_0, \varepsilon_{j_0}, \mu_0 \to 0$ by the Lebesgue dominated convergence theorem; indeed, the functions $(K_{\delta_0,\varepsilon_{j_0},\mu_0})\widehat{\ } - \widehat{W}$ are uniformly bounded and converge pointwise to zero as $\delta_0, \varepsilon_{j_0}, \mu_0 \to 0$, while $\widehat{\widetilde{\varphi}}$ is square integrable. This proves (6.7.27).

Next we show that for this choice of $\delta_0, \varepsilon_{j_0}, \mu_0$ we also have

$$\big|\big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f),\varphi\big\rangle - \big\langle T(f),\varphi\big\rangle\big| < \varepsilon\,\big\|f\big\|_{L^2}. \tag{6.7.28}$$

This is a consequence of the Cauchy–Schwarz inequality, since

$$\big|\big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f),\varphi\big\rangle - \big\langle T(f),\varphi\big\rangle\big| \le \big\|((K_{\delta_0,\varepsilon_{j_0},\mu_0})\widehat{\ } - \widehat{W})\widehat{\widetilde{\varphi}}\,\big\|_{L^2}\big\|f\big\|_{L^2}.$$

Using (6.7.26) we can find an M_0 such that for $M \ge M_0$ we have

$$\big|\big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f_M),\varphi\big\rangle - \big\langle T_{\delta_0,\varepsilon_{j_0},\mu_0}(f),\varphi\big\rangle\big| < \varepsilon. \tag{6.7.29}$$

Combining (6.7.27), (6.7.28), and (6.7.29) for $M \ge M_0$, we obtain

$$\big|\big\langle T(f_M),\varphi\big\rangle - \big\langle T(f),\varphi\big\rangle\big| < \varepsilon\big(1 + C_p\big\|f\big\|_{H^p} + \big\|f\big\|_{L^2}\big),$$

and this implies that $T(f_M)$ converges to $T(f)$ in $\mathscr{S}'(\mathbf{R}^n)$.

Using the inequality,

$$\big\|T(f_M)-T(f_{M'})\big\|_{H^p}^p \le C_{n,p}^p (A+B)^p \sum_{M<j\le M'} |\lambda_j|^p\,,$$

one easily shows that the sequence $\{T(f_M)\}_M$ is Cauchy in H^p. Thus $T(f_M)$ converges in H^p to some element $G\in H^p$ as $M\to\infty$. By Proposition 6.4.10, $T(f_M)$ converges to G in $\mathscr{S}'$. But as we saw, $T(f_M)$ converges to $T(f)$ in $\mathscr{S}'$ as $M\to\infty$. Hence $T(f)=G$ and we conclude that $T(f_M)$ converges to $T(f)$ in H^p, i.e., the series $\sum_j \lambda_j T(a_j)$ converges to $T(f)$ in H^p. This allows us to estimate the H^p quasi-norm of $T(f)$ as follows:

$$\begin{aligned}\big\|T(f)\big\|_{H^p(\mathbf{R}^n)}^p &= \Big\|\sum_j \lambda_j T(a_j)\Big\|_{H^p(\mathbf{R}^n)}^p\\ &\le \sum_j |\lambda_j|^p \big\|T(a_j)\big\|_{H^p(\mathbf{R}^n)}^p\\ &\le (C'_{n,p})^p (A+B)^p \sum_j |\lambda_j|^p\\ &\le (C'_{n,p}C_p)^p (A+B)^p \big\|f\big\|_{H^p(\mathbf{R}^n)}^p\,.\end{aligned}$$

This concludes the proof for $f\in H^p\cap L^2$. The extension to general $f\in H^p$ follows by density and the fact that $T(f)$ is well defined for all $f\in H^p$, as observed at the beginning of this subsection. □

6.7.4 A Singular Integral Characterization of $H^1(\mathbf{R}^n)$

We showed in Section 6.7.1 that singular integrals map H^1 to L^1. In particular, the Riesz transforms have this property. In this subsection we obtain a converse to this statement. We show that if $R_j(f)$ are integrable functions for some $f\in L^1$ and all $j=1,\dots,n$, then f must be an element of the Hardy space H^1. This provides a characterization of $H^1(\mathbf{R}^n)$ in terms of the Riesz transforms.

Theorem 6.7.5. *For $n\ge 2$, there exists a constant C_n such that for f in $L^1(\mathbf{R}^n)$ we have*

$$C_n\big\|f\big\|_{H^1} \le \big\|f\big\|_{L^1} + \sum_{k=1}^n \big\|R_k(f)\big\|_{L^1}\,. \tag{6.7.30}$$

When $n=1$ the corresponding statement is

$$C_1\big\|f\big\|_{H^1} \le \big\|f\big\|_{L^1} + \big\|H(f)\big\|_{L^1} \tag{6.7.31}$$

for all $f\in L^1(\mathbf{R})$. Naturally, these statements are interesting when the expressions on the right in (6.7.30) *and* (6.7.31) *are finite.*

Before we prove this theorem we discuss two corollaries.

Corollary 6.7.6. *An integrable function on the line lies in the Hardy space $H^1(\mathbf{R})$ if and only if its Hilbert transform is integrable. For $n \geq 2$, an integrable function on $\mathbf{R}^n$ lies in the Hardy space $H^1(\mathbf{R}^n)$ if and only its Riesz transforms are also in $L^1(\mathbf{R}^n)$.*

Proof. The corollary follows by combining Theorems 6.7.1 and 6.7.5. $\square$

Corollary 6.7.7. *Functions in $H^1(\mathbf{R}^n)$, $n \geq 1$, have integral zero.*

Proof. Indeed, if $f \in H^1(\mathbf{R}^n)$, we must have $R_1(f) \in L^1(\mathbf{R}^n)$; thus $\widehat{R_1(f)}$ is uniformly continuous. But since

$$\widehat{R_1(f)}(\xi) = -i\widehat{f}(\xi)\frac{\xi_1}{|\xi|},$$

it follows that $\widehat{R_1(f)}$ is continuous at zero if and only if $\widehat{f}(\xi) = 0$. But this happens exactly when f has integral zero. $\square$

We now discuss the proof of Theorem 6.7.5.

Proof. We consider the case $n \geq 2$, although the argument below also works in the case $n = 1$ with a suitable change of notation. Let P_t be the Poisson kernel. In the proof we may assume that f is real-valued, since it can be written as $f = f_1 + if_2$, where f_k are real-valued and $R_j(f_k)$ are also integrable. Given a real-valued function $f \in L^1(\mathbf{R}^n)$ such that $R_j(f)$ are integrable over $\mathbf{R}^n$ for every $j = 1, \dots, n$, we associate with it the $n+1$ functions

$$\begin{aligned} u_1(x,t) &= (P_t * R_1(f))(x), \\ \dots &= \dots, \\ u_n(x,t) &= (P_t * R_n(f))(x), \\ u_{n+1}(x,t) &= (P_t * f)(x), \end{aligned}$$

which are harmonic on the space $\mathbf{R}^{n+1}_+$ (see Example 2.1.13). It is convenient to denote the last variable t by x_{n+1}. One may check using the Fourier transform that these harmonic functions satisfy the following system:

$$\begin{aligned} \sum_{j=1}^{n+1} \frac{\partial u_j}{\partial x_j} &= 0, \\ \frac{\partial u_j}{\partial x_k} - \frac{\partial u_k}{\partial x_j} &= 0, \qquad k, j \in \{1, \dots, n+1\}, \quad k \neq j. \end{aligned} \tag{6.7.32}$$

This system of equations may also be expressed as div $F = 0$ and curl $F = \vec{0}$, where $F = (u_1, \dots, u_{n+1})$ is a vector field in $\mathbf{R}^{n+1}_+$. Note that when $n = 1$, the equations in (6.7.32) are the usual Cauchy–Riemann equations, which assert that the function $F = (u_1, u_2) = u_1 + iu_2$ is holomorphic in the upper half-space. For this reason, when

$n \geq 2$ the equations in (6.7.32) are often referred to as the *system of generalized Cauchy–Riemann equations*.

The function $|F|$ enjoys a crucial property in the study of this problem.

Lemma 6.7.8. *Let u_j be real-valued harmonic functions on $\mathbf{R}^{n+1}$ satisfying the system of equations* (6.7.32) *and let $F = (u_1, \dots, u_{n+1})$. Then the function*

$$|F|^q = \Big(\sum_{j=1}^{n+1} |u_j|^2\Big)^{q/2}$$

is subharmonic when $q \geq (n-1)/n$, i.e., it satisfies $\Delta(|F|^q) \geq 0$, on $\mathbf{R}^{n+1}_+$.

Lemma 6.7.9. *Let $0 < q < p < \infty$. Suppose that the function $|F(x,t)|^q$ defined on $\mathbf{R}^{n+1}_+$ is subharmonic and satisfies*

$$\sup_{t>0} \left(\int_{\mathbf{R}^n} |F(x,t)|^p \, dx \right)^{1/p} \leq A < \infty. \tag{6.7.33}$$

Then there is a constant $C_{n,p,q} < \infty$ such that the nontangential maximal function $|F|^*(x) = \sup_{t>0} \sup_{|y-x|<t} |F(y,t)|$, $x \in \mathbf{R}^n$, (cf. Definition 7.3.1) satisfies

$$\big\| |F|^* \big\|_{L^p(\mathbf{R}^n)} \leq C_{n,p,q} A \, .$$

Assuming these lemmas, whose proofs are postponed until the end of this section, we return to the proof of the theorem.

Since the Poisson kernel is an approximate identity, the function $x \mapsto u_{n+1}(x,t)$ converges to $f(x)$ in L^1 as $t \to 0$. To show that $f \in H^1(\mathbf{R}^n)$, it suffices to show that the Poisson maximal function

$$M(f;P)(x) = \sup_{t>0} |(P_t * f)(x)| = \sup_{t>0} |u_{n+1}(x,t)|$$

is integrable. But this maximal function is pointwise controlled by

$$\sup_{t>0} |F(x,t)| \leq \sup_{t>0} \Big[|(P_t * f)(x)| + \sum_{j=1}^{n} |(P_t * R_j(f))(x)| \Big] ,$$

and certainly it satisfies

$$\sup_{t>0} \int_{\mathbf{R}^n} |F(x,t)| \, dx \leq A_f \, , \tag{6.7.34}$$

where

$$A_f = \big\| f \big\|_{L^1} + \sum_{k=1}^{n} \big\| R_k(f) \big\|_{L^1} \, .$$

We now have

$$M(f;P)(x) \leq \sup_{t>0} |u_{n+1}(x,t)| \leq \sup_{t>0} |F(x,t)| \leq |F|^*(x) \, , \tag{6.7.35}$$

and using Lemma 6.7.8 with $q = \frac{n-1}{n}$ and Lemma 6.7.9 with $p = 1$ we obtain that

$$\big\||F|^*\big\|_{L^1(\mathbf{R}^n)} \le C_n A_f . \tag{6.7.36}$$

Combining (6.7.34), (6.7.35), and (6.7.36), one deduces that

$$\big\|M(f;P)(x)\big\|_{L^1(\mathbf{R}^n)} \le C_n\bigg(\big\|f\big\|_{L^1} + \sum_{k=1}^{n}\big\|R_k(f)\big\|_{L^1}\bigg),$$

from which (6.7.30) follows. This proof is also valid when $n = 1$, provided one replaces the Riesz transforms with the Hilbert transform; hence the proof of (6.7.31) is subsumed in that of (6.7.30). □

See Exercise 6.7.1 for an extension of this result to H^p for $\frac{n-1}{n} < p < 1$. We now give a proof of Lemma 6.7.8

Proof. Denoting the variable t by x_{n+1}, we have

$$\frac{\partial}{\partial x_j}|F|^q = q|F|^{q-2}\Big(F\cdot\frac{\partial F}{\partial x_j}\Big)$$

and also

$$\frac{\partial^2}{\partial x_j^2}|F|^q = q|F|^{q-2}\bigg[F\cdot\frac{\partial^2 F}{\partial x_j^2} + \frac{\partial F}{\partial x_j}\cdot\frac{\partial F}{\partial x_j}\bigg] + q(q-2)|F|^{q-4}\Big(F\cdot\frac{\partial F}{\partial x_j}\Big)^2$$

for all $j = 1,2,\dots,n+1$. Summing over all these j's, we obtain

$$\Delta(|F|^q) = q|F|^{q-4}\bigg[|F|^2\sum_{j=1}^{n+1}\Big|\frac{\partial F}{\partial x_j}\Big|^2 + (q-2)\sum_{j=1}^{n+1}\Big|F\cdot\frac{\partial F}{\partial x_j}\Big|^2\bigg], \tag{6.7.37}$$

since the term containing $F\cdot\Delta(F) = \sum_{j=1}^{n+1}u_j\Delta(u_j)$ vanishes because each u_j is harmonic. The only term that could be negative in (6.7.37) is that containing the factor $q-2$ and naturally, if $q \ge 2$, the conclusion is obvious. Let us assume that $\frac{n-1}{n} \le q < 2$. Since $q \ge \frac{n-1}{n}$, we must have that $2-q \le \frac{n+1}{n}$. Thus (6.7.37) is non-negative if

$$\sum_{j=1}^{n+1}\Big|F\cdot\frac{\partial F}{\partial x_j}\Big|^2 \le \frac{n}{n+1}|F|^2\sum_{j=1}^{n+1}\Big|\frac{\partial F}{\partial x_j}\Big|^2. \tag{6.7.38}$$

This is certainly valid for points (x,t) such that $F(x,t) = 0$. To prove (6.7.38) for points (x,t) with $F(x,t) \ne 0$, it suffices to show that for every vector $v \in \mathbf{R}^{n+1}$ with Euclidean norm $|v| = 1$, we have

$$\sum_{j=1}^{n+1}\Big|v\cdot\frac{\partial F}{\partial x_j}\Big|^2 \le \frac{n}{n+1}\sum_{j=1}^{n+1}\Big|\frac{\partial F}{\partial x_j}\Big|^2. \tag{6.7.39}$$

Denoting by A the $(n+1)\times(n+1)$ matrix whose entries are $a_{j,k} = \partial u_k/\partial x_j$, we rewrite (6.7.39) as

$$\big|Av\big|^2 \le \frac{n}{n+1}\big\|A\big\|^2, \tag{6.7.40}$$

where

$$\big\|A\big\|^2 = \sum_{j=1}^{n+1}\sum_{k=1}^{n+1}|a_{j,k}|^2.$$

By assumption, the functions u_j are real-valued and thus the numbers $a_{j,k}$ are real. In view of identities (6.7.32), the matrix A is real symmetric and has zero trace (i.e., $\sum_{j=1}^{n+1} a_{j,j} = 0$). A real symmetric matrix A can be written as $A = PDP^t$, where P is an orthogonal matrix and D is a real diagonal matrix. Since orthogonal matrices preserve the Euclidean distance, estimate (6.7.40) follows from the corresponding one for a diagonal matrix D. If $A = PDP^t$, then the traces of A and D are equal; hence $\sum_{j=1}^{n+1}\lambda_j = 0$, where λ_j are entries on the diagonal of D. Notice that estimate (6.7.40) with the matrix D in the place of A is equivalent to

$$\sum_{j=1}^{n+1}|\lambda_j|^2|v_j|^2 \le \frac{n}{n+1}\Big(\sum_{j=1}^{n+1}|\lambda_j|^2\Big), \tag{6.7.41}$$

where we set $v = (v_1,\dots,v_{n+1})$ and we are assuming that $|v|^2 = \sum_{j=1}^{n+1}|v_j|^2 = 1$. Estimate (6.7.41) is certainly a consequence of

$$\sup_{1\le j\le n+1}|\lambda_j|^2 \le \frac{n}{n+1}\Big(\sum_{j=1}^{n+1}|\lambda_j|^2\Big). \tag{6.7.42}$$

But this is easy to prove. Let $|\lambda_{j_0}| = \max_{1\le j\le n+1}|\lambda_j|$. Then

$$|\lambda_{j_0}|^2 = \Big|-\sum_{j\ne j_0}\lambda_j\Big|^2 \le \Big(\sum_{j\ne j_0}|\lambda_j|\Big)^2 \le n\sum_{j\ne j_0}|\lambda_j|^2. \tag{6.7.43}$$

Adding $n|\lambda_{j_0}|^2$ to both sides of (6.7.43), we deduce (6.7.42) and thus (6.7.38). □

We now give the proof of Lemma 6.7.9.

Proof. A consequence of the subharmonicity of $|F|^q$ is that

$$|F(x,t+\varepsilon)|^q \le (|F(\cdot,\varepsilon)|^q * P_t)(x) \tag{6.7.44}$$

for all $x\in\mathbf{R}^n$ and $t,\varepsilon>0$. To prove (6.7.44), fix $\varepsilon>0$ and consider the functions

$$U(x,t) = |F(x,t+\varepsilon)|^q, \qquad V(x,t) = (|F(\cdot,\varepsilon)|^q * P_t)(x).$$

Given $\eta>0$, we find a half-ball

$$B_{R_0} = \{(x,t)\in\mathbf{R}^{n+1}_+ : |x|^2+t^2 < R_0^2\}$$

such that for $(x,t) \in \mathbf{R}^{n+1}_+ \setminus B_{R_0}$ we have

$$U(x,t) - V(x,t) \leq \eta\,. \tag{6.7.45}$$

Suppose that this is possible. Since $U(x,0) = V(x,0)$, then (6.7.45) actually holds on the entire boundary of B_{R_0}. The function V is harmonic and U is subharmonic; thus $U - V$ is subharmonic. The maximum principle for subharmonic functions implies that (6.7.45) holds in the interior of B_{R_0}, and since it also holds on the exterior, it must be valid for all (x,t) with $x \in \mathbf{R}^n$ and $t \geq 0$. Since η was arbitrary, letting $\eta \to 0+$ implies (6.7.44).

We now prove that R_0 exists such that (6.7.45) is possible for $(x,t) \in \mathbf{R}^{n+1}_+ \setminus B_{R_0}$. Let $B((x,t),t/2)$ be the $(n+1)$-dimensional ball of radius $t/2$ centered at (x,t). The subharmonicity of $|F|^q$ is reflected in the inequality

$$|F(x,t)|^q \leq \frac{1}{|B((x,t),\, t/2)|} \int_{B((x,t),t/2)} |F(y,s)|^q\, dyds\,,$$

which by Hölder's inequality and the fact $p > q$ gives

$$|F(x,t)|^q \leq \left(\frac{1}{|B((x,t),\, t/2)|} \int_{B((x,t),t/2)} |F(y,s)|^p\, dyds \right)^{\frac{q}{p}}.$$

From this we deduce that

$$|F(x,t+\varepsilon)|^q \leq \left[\frac{2^{n+1}/v_{n+1}}{(t+\varepsilon)^{n+1}} \int_{\frac{1}{2}(t+\varepsilon)}^{\frac{3}{2}(t+\varepsilon)} \int_{|y| \geq |x| - \frac{1}{2}(t+\varepsilon)} |F(y,s)|^p\, dy\, ds \right]^{\frac{q}{p}}. \tag{6.7.46}$$

If $t + \varepsilon \geq |x|$, using (6.7.33), we see that the expression on the right in (6.7.46) is bounded by $c'\varepsilon^{-n} A^q t^{-nq/p}$, and thus it can be made smaller than $\eta/2$ by taking $t \geq R_1$, for some $R_1 > \varepsilon$ large enough. Since $R_1 > \varepsilon$, we must have $2t \geq t + \varepsilon \geq |x|$, which implies that $t \geq |x|/2$, and thus with $R_0' = \sqrt{5}R_1$, if $|(x,t)| > R_0'$ then $t \geq R_1$. Hence, the expression in (6.7.46) can be made smaller than $\eta/2$ for $|(x,t)| > R_0'$.

If $t + \varepsilon < |x|$ we estimate the expression on the right in (6.7.46) by

$$\left(\frac{2^{n+1}}{v_{n+1}} \frac{1}{(t+\varepsilon)^{n+1}} \int_{\frac{1}{2}(t+\varepsilon)}^{\frac{3}{2}(t+\varepsilon)} \left[\int_{|y| \geq \frac{1}{2}|x|} |F(y,s)|^p\, dy \right] ds \right)^{\frac{q}{p}},$$

and we notice that the preceding expression is bounded by

$$\left(\frac{3^{n+1}}{v_{n+1}} \int_{\frac{1}{2}\varepsilon}^{\infty} \left[\int_{|y| \geq \frac{1}{2}|x|} |F(y,s)|^p\, dy \right] \frac{ds}{s^{n+1}} \right)^{\frac{q}{p}}. \tag{6.7.47}$$

Let $G_{|x|}(s)$ be the function inside the square brackets in (6.7.47). Then $G_{|x|}(s) \to 0$ as $|x| \to \infty$ for all s. The hypothesis (6.7.33) implies that $G_{|x|}$ is bounded by a constant and it is therefore integrable over the interval $\left[\frac{1}{2}\varepsilon, \infty\right)$ with respect to the measure $s^{-n-1}ds$. By the Lebesgue dominated convergence theorem we deduce that the expression in (6.7.47) converges to zero as $|x| \to \infty$ and thus it can be made smaller that $\eta/2$ for $|x| \ge R_2$, for some constant R_2. Then with $R_0'' = \sqrt{2}R_2$ we have that if $|(x,t)| \ge R_0''$ then (6.7.47) is at most $\eta/2$. Since $U - V \le U$, we deduce the validity of (6.7.45) for $|(x,t)| > R_0 = \max(R_0', R_0'')$.

Let $r = p/q > 1$. Assumption (6.7.33) implies that the functions $x \mapsto |F(x,\varepsilon)|^q$ are in L^r uniformly in t. Since any closed ball of L^r is weak* compact, there is a sequence $\varepsilon_k \to 0$ such that $|F(x,\varepsilon_k)|^q \to h$ weakly in L^r as $k \to \infty$ to some function $h \in L^r$. Since $P_t \in L^{r'}$, this implies that

$$(|F(\cdot,\varepsilon_k)|^q * P_t)(x) \to (h * P_t)(x)$$

for all $x \in \mathbf{R}^n$. Using (6.7.44) we obtain

$$|F(x,t)|^q = \limsup_{k\to\infty} |F(x,t+\varepsilon_k)|^p \le \limsup_{k\to\infty} \big(|F(x,\varepsilon_k)|^q * P_t\big)(x) = (h * P_t)(x)\,,$$

which gives for all $x \in \mathbf{R}^n$,

$$|F|^*(x) \le \Big[\sup_{t>0}\sup_{|y-x|<t} (|h| * P_t)(x)\Big]^{1/q} \le C_n' M(h)(x)^{1/q}\,. \tag{6.7.48}$$

Let $g \in L^{r'}(\mathbf{R}^n)$ with $L^{r'}$ norm at most one. The weak convergence yields

$$\int_{\mathbf{R}^n} |F(x,\varepsilon_k)|^q g(x)\,dx \to \int_{\mathbf{R}^n} h(x)\,g(x)\,dx$$

as $k \to \infty$, and consequently we have

$$\left|\int_{\mathbf{R}^n} h(x)\,g(x)\,dx\right| \le \sup_k \int_{\mathbf{R}^n} |F(x,\varepsilon_k)|^q |g(x)|\,dx \le \big\|g\big\|_{L^{r'}} \sup_{t>0}\left(\int_{\mathbf{R}^n} |F(x,t)|^p\,dx\right)^{\frac{1}{r}}.$$

Since g is arbitrary with $L^{r'}$ norm at most one, this implies that

$$\big\|h\big\|_{L^r} \le \sup_{t>0}\left(\int_{\mathbf{R}^n} |F(x,t)|^p\,dx\right)^{\frac{1}{r}}. \tag{6.7.49}$$

Putting things together, we have

$$\begin{aligned}
\big\||F|^*\big\|_{L^p} &\le C_n'\big\|M(h)^{1/q}\big\|_{L^p} \\
&= C_n'\big\|M(h)\big\|_{L^r}^{1/q} \\
&= C_{n,p,q}\big\|h\big\|_{L^r}^{1/q} \\
&= C_{n,p,q}\sup_{t>0}\left(\int_{\mathbf{R}^n} |F(x,t)|^p\,dx\right)^{1/qr} \\
&\le C_{n,p,q} A\,,
\end{aligned}$$

where we have used (6.7.48) and (6.7.49) in the last two displayed inequalities. □

Exercises

6.7.1. Prove the following generalization of Theorem 6.7.4. Let φ be a nonnegative Schwartz function with integral one on $\mathbf{R}^n$ and let $\frac{n-1}{n} < p < 1$. Prove that there are constants c_1, c_n, C_1, C_n such that for bounded tempered distributions f on $\mathbf{R}^n$ (cf. Section 6.4.1) we have

$$c_n \big\|f\big\|_{H^p} \le \sup_{\delta>0} \Big[\big\|\varphi_\delta * f\big\|_{L^p} + \sum_{k=1}^{n} \big\|\varphi_\delta * R_k(f)\big\|_{L^p}\Big] \le C_n \big\|f\big\|_{H^p}$$

when $n \ge 2$ and

$$c_1 \big\|f\big\|_{H^p} \le \sup_{\delta>0} \Big[\big\|\varphi_\delta * f\big\|_{L^p} + \big\|\varphi_\delta * H(f)\big\|_{L^p}\Big] \le C_1 \big\|f\big\|_{H^p}$$

when $n = 1$.
[*Hint:* One direction is a consequence of Theorem 6.7.4. For the other direction, define $F_\delta = (u_1 * \varphi_\delta, \dots, u_{n+1} * \varphi_\delta)$, where $u_j(x,t) = (P_t * R_j(f))(x)$, $j = 1, \dots, n$, and $u_{n+1}(x,t) = (P_t * f)(x)$. Each $u_j * \varphi_\delta$ is a harmonic function on $\mathbf{R}^{n+1}_+$ and continuous up to the boundary. The subharmonicity of $|F_\delta(x,t)|^p$ has as a consequence that $|F_\delta(x,t+\varepsilon)|^p \le |(F_\delta(\cdot,\varepsilon)|^p * P_t)(x)$ in view of (6.7.44). Letting $\varepsilon \to 0$ implies that $|F_\delta(x,t)|^p \le |(F_\delta(\cdot,0)|^p * P_t)(x)$, by the continuity of F_δ up to the boundary. Since $F_\delta(x,0) = (R_1(f) * \varphi_\delta, \dots, R_n(f) * \varphi_\delta, f * \varphi_\delta)$, the hypothesis that $f * \varphi_\delta, R_j(f) * \varphi_\delta$ are in L^p uniformly in $\delta > 0$ gives that $\sup_{t,\delta>0} \int_{\mathbf{R}^n} |F_\delta(x,t)|^p\, dx < \infty$. Fatou's lemma yields (6.7.33) for $F(x,t) = (u_1, \dots, u_{n+1})$. Then Lemma 6.7.9 implies the required conclusion.]

6.7.2. (a) Let h be a function on $\mathbf{R}$ such that $h(x)$ and $xh(x)$ are in $L^2(\mathbf{R})$. Show that h is integrable over $\mathbf{R}$ and satisfies

$$\big\|h\big\|_{L^1}^2 \le 8 \big\|h\big\|_{L^2} \big\|xh(x)\big\|_{L^2}.$$

(b) Suppose that g is an integrable function on $\mathbf{R}$ with vanishing integral and $g(x)$ and $xg(x)$ are in $L^2(\mathbf{R})$. Show that g lies in $H^1(\mathbf{R})$ and that for some constant C we have

$$\big\|g\big\|_{H^1}^2 \le C \big\|g\big\|_{L^2} \big\|xg(x)\big\|_{L^2}.$$

[*Hint:* Part (a): split the integral of $|h(x)|$ over the regions $|x| \le R$ and $|x| > R$ and pick a suitable R. Part (b): Show that both $H(g)$ and $H(yg(y))$ lie in L^2. But since g has vanishing integral, we have $xH(g)(x) = H(yg(y))(x)$.]

6.7.3. (a) Let H be the Hilbert transform. Prove the identity

$$H(fg - H(f)H(g)) = fH(g) + gH(f)$$

for all f,g in $\bigcup_{1\le p<\infty} L^p(\mathbf{R})$.
(b) Show that the bilinear operators

$$(f,g) \mapsto f H(g) + H(f)\, g\,,$$
$$(f,g) \mapsto f\, g - H(f) H(g)\,,$$

map $L^p(\mathbf{R}) \times L^{p'}(\mathbf{R}) \to H^1(\mathbf{R})$ whenever $1 < p < \infty$.
$\big[$*Hint:* Part (a): Consider the boundary values of the product of the analytic extensions of $f + iH(f)$ and $g + iH(g)$ on the upper half-space. Part (b): Use part (a) and Theorem 6.7.5.$\big]$

6.7.4. Follow the steps given to prove the following interpolation result. Let $1 < p_1 \le \infty$ and let T be a subadditive operator that maps $H^1(\mathbf{R}^n) + L^{p_1}(\mathbf{R}^n)$ into measurable functions on $\mathbf{R}^n$. Suppose that there is $A_0 < \infty$ such that for all $f \in H^1(\mathbf{R}^n)$ we have

$$\sup_{\lambda>0} \lambda \left|\left\{x \in \mathbf{R}^n :\ |T(f)(x)| > \lambda\right\}\right| \le A_0 \big\|f\big\|_{H^1}$$

and that it also maps $L^{p_1}(\mathbf{R}^n)$ to $L^{p_1,\infty}(\mathbf{R}^n)$ with norm at most A_1. Show that for any $1 < p < p_1$, T maps $L^p(\mathbf{R}^n)$ to itself with norm at most

$$C A_0^{\frac{\frac{1}{p}-\frac{1}{p_1}}{1-\frac{1}{p_1}}} A_1^{\frac{1-\frac{1}{p}}{1-\frac{1}{p_1}}}\,,$$

where $C = C(n,p,p_1)$.
(a) Fix $1 < q < p < p_1 < \infty$ and f and let Q_j be the family of all maximal dyadic cubes such that $\lambda^q < |Q_j|^{-1}\int_{Q_j} |f|^q\,dx$. Write $E_\lambda = \bigcup Q_j$ and note that $E_\lambda \subseteq \left\{M(|f|^q)^{\frac{1}{q}} > \lambda\right\}$ and that $|f| \le \lambda$ a.e. on $(E_\lambda)^c$. Write f as the sum of the *good function*

$$g_\lambda = f\chi_{(E_\lambda)^c} + \sum_j (\operatorname*{Avg}_{Q_j} f)\,\chi_{Q_j}$$

and the *bad function*

$$b_\lambda = \sum_j b_\lambda^j\,, \qquad \text{where} \qquad b_\lambda^j = \big(f - \operatorname*{Avg}_{Q_j} f\big)\,\chi_{Q_j}\,.$$

(b) Show that g_λ lies in $L^{p_1}(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$, $\big\|g_\lambda\big\|_{L^\infty} \le 2^{\frac{n}{q}}\lambda$, and that

$$\big\|g_\lambda\big\|_{L^{p_1}}^{p_1} \le \int_{|f|\le\lambda} |f(x)|^{p_1}\,dx + 2^{\frac{np_1}{q}}\lambda^{p_1}|E_\lambda| < \infty\,.$$

(c) Show that for $c = 2^{\frac{n}{q}+1}$, each $c^{-1}\lambda^{-1}|Q_j|^{-1}b_\lambda^j$ is an L^q-atom for H^1. Conclude that b_λ lies in $H^1(\mathbf{R}^n)$ and satisfies

$$\big\|b_\lambda\big\|_{H^1} \le c\lambda \sum_j |Q_j| \le c\lambda\,|E_\lambda| < \infty\,.$$

(d) Start with

$$\|T(f)\|_{L^p}^p \le p\gamma^p \int_0^\infty \lambda^{p-1} \big|\{T(g_\lambda)| > \tfrac{1}{2}\gamma\lambda\}\big|\, d\lambda$$
$$+ p\gamma^p \int_0^\infty \lambda^{p-1} \big|\{T(b_\lambda)| > \tfrac{1}{2}\gamma\lambda\}\big|\, d\lambda$$

and use the results in parts (b) and (c) to obtain that the preceding expression is at most $C(n,p,q,p_1)\max(A_1\gamma^{p-p_1}, \gamma^{p-1}A_0)$. Select $\gamma = A_1^{\frac{p_1}{p_1-1}} A_0^{-\frac{1}{p_1-1}}$ to obtain the required conclusion.
(e) In the case $p_1 = \infty$ we have $|T(g_\lambda)| \le A_1 2^{\frac{n}{q}}\lambda$ and pick $\gamma > 2A_1 2^{\frac{n}{q}}$ to make the integral involving g_λ vanishing.

6.7.5. Let f be an integrable function on the line whose Fourier transform is also integrable and vanishes on the negative half-line. Show that f lies in $H^1(\mathbf{R})$.

HISTORICAL NOTES

The strong type $L^p \to L^q$ estimates in Theorem 6.1.3 were obtained by Hardy and Littlewood [157] (see also [158]) when $n = 1$ and by Sobolev [285] for general n. The weak type estimate $L^1 \to L^{\frac{n}{n-s},\infty}$ first appeared in Zygmund [339]. The proof of Theorem 6.1.3 using estimate (6.1.11) is taken from Hedberg [161]. The best constants in this theorem when $p = \frac{2n}{n+s}$, $q = \frac{2n}{n-s}$, and $0 < s < n$ were precisely evaluated by Lieb [213]. A generalization of Theorem 6.1.3 for nonconvolution operators was obtained by Folland and Stein [132].

The Riesz potentials were systematically studied by Riesz [270] on $\mathbf{R}^n$ although their one-dimensional version appeared in earlier work of Weyl [330]. The Bessel potentials were introduced by Aronszajn and Smith [7] and also by Calderón [41], who was the first to observe that the potential space $\mathscr{L}_s^p$ (i.e., the Sobolev space L_s^p) coincides with the space L_k^p given in the classical Definition 6.2.1 when $s = k$ is an integer. Theorem 6.2.4 is due to Sobolev [285] when s is a positive integer. The case $p = 1$ of Sobolev's theorem (Exercise 6.2.9) was later obtained independently by Gagliardo [139] and Nirenberg [249]. We refer to the books of Adams [2], Lieb and Loss [214], and Maz'ya [229] for a more systematic study of Sobolev spaces and their use in analysis.

An early characterization of Lipschitz spaces using Littlewood–Paley type operators (built from the Poisson kernel) appears in the work of Hardy and Littlewood [160]. These and other characterizations were obtained and extensively studied in higher dimensions by Taibleson [300], [301], [302] in his extensive study. Lipschitz spaces can also be characterized via mean oscillation over cubes. This idea originated in the simultaneous but independent work of Campanato [39], [40] and Meyers [234] and led to duality theorems for these spaces. Incidentally, the predual of the space $\dot{\Lambda}_\alpha$ is the Hardy space H^p with $p = \frac{n}{n+\alpha}$, as shown by Duren, Romberg, and Shields [118] for the unit circle and by Walsh [327] for higher-dimensional spaces; see also Fefferman and Stein [130]. We refer to the book of García-Cuerva and Rubio de Francia [141] for a nice exposition of these results. An excellent expository reference on Lipschitz spaces is the article of Krantz [199].

Taibleson in his aforementioned work also studied the generalized Lipschitz spaces $\Lambda_\alpha^{p,q}$ called today Besov spaces. These spaces were named after Besov, who obtained a trace theorem and embeddings for them [24], [25]. The spaces $B_p^{\alpha,q}$, as defined in Section 6.5, were introduced by Peetre [255], although the case $p = q = 2$ was earlier considered by Hörmander [166]. The connection of Besov spaces with modern Littlewood–Paley theory was brought to the surface by Peetre [255]. The extension of the definition of Besov spaces to the case $p < 1$ is also due to Peetre [256], but there was a forerunner by Flett [131]. The spaces $F_p^{\alpha,q}$ with $1 < p,q < \infty$ were introduced by Triebel [316] and independently by Lizorkin [218]. The extension of the spaces $F_p^{\alpha,q}$ to the case $0 < p < \infty$ and $0 < q \le \infty$ first appeared in Peetre [258], who also obtained a maximal characterization for all of these spaces. Lemma 6.5.3 originated in Peetre [258]; the version given in the text is based on a refinement of Triebel [317]. The article of Lions, Lizorkin, and Nikol'skij [216] presents an account of the treatment of the spaces $F_p^{\alpha,q}$ introduced by Triebel and Lizorkin as well as the equivalent characterizations obtained by Lions, using interpolation between Banach spaces, and by Nikol'skij, using best approximation.

The theory of Hardy spaces is vast and complicated. In classical complex analysis, the Hardy spaces H^p were spaces of analytic functions and were introduced to characterize boundary values of analytic functions on the unit disk. Precisely, the space $H^p(\mathbb{D})$ was introduced by Hardy [156] to consist of all analytic functions F on the unit disk $\mathbb{D}$ with the property that $\sup_{0<r<1}\int_0^1 |F(re^{2\pi i\theta})|^p\,d\theta < \infty$, $0 < p < \infty$. When $1 < p < \infty$, this space coincides with the space of analytic functions whose real parts are Poisson integrals of functions in $L^p(\mathbf{T}^1)$. But for $0 < p \le 1$ this characterization fails and for several years a satisfactory characterization was missing. For a systematic treatment of these spaces we refer to the books of Duren [117] and Koosis [195].

With the illuminating work of Stein and Weiss [293] on systems of conjugate harmonic functions the road opened to higher-dimensional extensions of Hardy spaces. Burkholder, Gundy, and Silverstein [38] proved the fundamental theorem that an analytic function F lies in $H^p(\mathbf{R}^2_+)$ [i.e., $\sup_{y>0}\int_{\mathbf{R}} |F(x+iy)|^p\,dx < \infty$] if and only if the nontangential maximal function of its real part lies in $L^p(\mathbf{R})$. This result was proved using Brownian motion, but later Koosis [194] obtained another proof using complex analysis. This theorem spurred the development of the modern theory of Hardy spaces by providing the first characterization without the notion of conjugacy and indicating that Hardy spaces are intrinsically defined. The pioneering article of Fefferman and Stein [130] furnished three new characterizations of Hardy spaces: using a maximal function associated with a general approximate identity, using the grand maximal function, and using the area function of Luzin. From this point on, the role of the Poisson kernel faded into the background, when it turned out that it was not essential in the study of Hardy spaces. A previous characterization of Hardy spaces using the g-function, a radial analogue of the Luzin area function, was obtained by Calderón [42]. Two alternative characterizations of Hardy spaces were obtained by Uchiyama in terms of the generalized Littlewood–Paley g-function [319] and in terms of Fourier multipliers [320]. Necessary and sufficient conditions for systems of singular integral operators to characterize $H^1(\mathbf{R}^n)$ were also obtained by Uchiyama [318]. The characterization of H^p using Littlewood–Paley theory was observed by Peetre [257]. The case $p = 1$ was later independently obtained by Rubio de Francia, Ruiz, and Torrea [276].

The one-dimensional atomic decomposition of Hardy spaces is due to Coifman [72] and its higher-dimensional extension to Latter [206]. A simplification of some of the technical details in Latter's proof was subsequently obtained by Latter and Uchiyama [207]. Using the atomic decomposition Coifman and Weiss [86] extended the definition of Hardy spaces to more general structures. The idea of obtaining the atomic decomposition from the reproducing formula (6.6.8) goes back to Calderón [44]. Another simple proof of the L^2-atomic decomposition for H^p (starting from the nontangential Poisson maximal function) was obtained by Wilson [332]. With only a little work, one can show that L^q-atoms for H^p can be written as sums of L^∞-atoms for H^p. We refer to the book of García-Cuerva and Rubio de Francia [141] for a proof of this fact. Although finite sums of atoms are dense in H^1, an example due to Y. Meyer (contained in [233]) shows that the H^1 norm of a function may not be comparable to $\inf\sum_{j=1}^N |\lambda_j|$, where the infimum is taken over all representations of the function as finite linear combinations $\sum_{j=1}^N \lambda_j a_j$ with the a_j being L^∞-atoms for H^1. Based on this idea, Bownik [34] constructed an example of a linear functional on a dense subspace of H^1 that is uniformly bounded on L^∞-atoms for H^1 but does not extend to a bounded linear functional on the whole H^1. However, if a Banach-valued linear operator is bounded uniformly on all L^q-atoms for H^p with $1 < q < \infty$ and $0 < p \le 1$, then it is bounded on the entire H^p as shown by Meda, Sjögren, and Vallarino [230]. This fact is also valid for quasi-Banach-valued linear operators, and when $q = 2$ it was obtained independently by Yang and Zhou [338]. A related general result says that a sublinear operator maps the Triebel–Lizorkin space $\dot{F}^{s}_{p,q}(\mathbf{R}^n)$ to a quasi-Banach space if and only if it is uniformly bounded on certain infinitely differentiable atoms of the space; see Liu and Yang [217]. Atomic decompositions of general function spaces were obtained in the fundamental work of Frazier and Jawerth [135], [136]. The exposition in Section 6.6 is based on the article of Frazier and Jawerth [137]. The work of these authors provides a solid manifestation that atomic decompositions are intrinsically related to Littlewood–Paley theory and not wedded to a particular space. Littlewood–Paley theory therefore provides a comprehensive and unifying perspective on function spaces.

Main references on H^p spaces and their properties are the books of Baernstein and Sawyer [12], Folland and Stein [133] in the context of homogeneous groups, Lu [219] (on which the proofs of Lemma 6.4.5 and Theorem 6.4.4 are based), Strömberg and Torchinsky [298] (on weighted Hardy spaces), and Uchiyama [321]. The articles of Calderón and Torchinsky [45], [46] develop and extend the theory of Hardy spaces to the nonisotropic setting. Hardy spaces can also be defined in terms of nonstandard convolutions, such as the "twisted convolution" on $\mathbf{R}^{2n}$. Characterizations of the space H^1 in this context have been obtained by Mauceri, Picardello, and Ricci [226]

The localized Hardy spaces h_p, $0 < p \le 1$, were introduced by Goldberg [146] as spaces of distributions for which the maximal operator $\sup_{0<t<1} |\Phi_t * f|$ lies in $L^p(\mathbf{R}^n)$ (here Φ is a Schwartz function with nonvanishing integral). These spaces can be characterized in ways analogous to those of the homogeneous Hardy spaces H^p; in particular, they admit an atomic decomposition. It was shown by Bui [37] that the space h^p coincides with the Triebel–Lizorkin space $F^{0,2}_p(\mathbf{R}^n)$; see also Meyer [232]. For the local theory of Hardy spaces one may consult the articles of Dafni [100] and Chang, Krantz, and Stein [59].

Interpolation of operators between Hardy spaces was originally based on complex function theory; see the articles of Calderón and Zygmund [48] and Weiss [328]. The real-interpolation approach discussed in Exercise 6.7.4 can be traced in the article of Igari [174]. Interpolation between Hardy spaces was further studied and extended by Riviere and Sagher [271] and Fefferman, Riviere, and Sagher [128].

The action of singular integrals on periodic spaces was studied by Calderón and Zygmund [47]. The preservation of Lipschitz spaces under singular integral operators is due to Taibleson [299]. The case $0 < \alpha < 1$ was earlier considered by Privalov [268] for the conjugate function on the circle. Fefferman and Stein [130] were the first to show that singular integrals map Hardy spaces to themselves. The boundedness of fractional integrals on H^p was obtained by Krantz [198]. The case $p = 1$ was earlier considered by Stein and Weiss [293]. The action of multilinear singular integrals on Hardy spaces was studied by Coifman and Grafakos [75] and Grafakos and Kalton [149]. An exposition on the subject of function spaces and the action of singular integrals on them was written by Frazier, Jawerth, and Weiss [138]. For a careful study of the action of singular integrals on function spaces, we refer to the book of Torres [315]. The study of anisotropic function spaces and the action of singular integrals on them has been studied by Bownik [33]. Weighted anisotropic Hardy spaces have been studied by Bownik, Li, Yang, and Zhou [35].