$$
\begin{aligned}
&\le \int_{\mathbf{R}^n} S(f)(x)\,S(g)(x)\,dx && \text{(Cauchy–Schwarz inequality)}\\
&\le \big\|S(f)\big\|_{L^p}\,\big\|S(g)\big\|_{L^{p'}} && \text{(Hölder's inequality)}\\
&\le \big\|S(f)\big\|_{L^p}\,c_{p',n}\big\|g\big\|_{L^{p'}}\,.
\end{aligned}
$$

Taking the supremum over all functions g on $\mathbf{R}^n$ with $L^{p'}$ norm at most 1, we obtain that f gives rise to a bounded linear functional on $L^{p'}$. It follows by the Riesz representation theorem that f must be an L^p function that satisfies the lower estimate in (6.4.12). □

6.4.4 *Almost Orthogonality Between the Littlewood–Paley Operators and the Dyadic Martingale Difference Operators*

Next, we discuss connections between the Littlewood–Paley operators Δ_j and the dyadic martingale difference operators D_k. It turns out that these operators are almost orthogonal in the sense that the L^2 operator norm of the composition $D_k\Delta_j$ decays exponentially as the indices j and k get farther away from each other.

For the purposes of the next theorem we define the Littlewood–Paley operators Δ_j as convolution operators with the function $\Psi_{2^{-j}}$, where

$$\widehat{\Psi}(\xi) = \widehat{\Phi}(\xi) - \widehat{\Phi}(2\xi)$$

and Φ is a fixed radial Schwartz function whose Fourier transform $\widehat{\Phi}$ is real-valued, supported in the ball $|\xi| < 2$, and equal to 1 on the ball $|\xi| < 1$. In this case we clearly have the identity

$$\sum_{j\in\mathbf{Z}} \widehat{\Psi}(2^{-j}\xi) = 1, \qquad \xi \neq 0\,.$$

Then we have the following theorem.

Theorem 6.4.8. *There exists a constant C such that for every k, j in $\mathbf{Z}$ the following estimate on the operator norm of $D_k\Delta_j : L^2(\mathbf{R}^n) \to L^2(\mathbf{R}^n)$ is valid:*

$$\big\|D_k\Delta_j\big\|_{L^2(\mathbf{R}^n)\to L^2(\mathbf{R}^n)} = \big\|\Delta_j D_k\big\|_{L^2(\mathbf{R}^n)\to L^2(\mathbf{R}^n)} \le C\,2^{-\frac{9}{20}|j-k|}. \tag{6.4.16}$$

Proof. Since Ψ is a radial function, it follows that Δ_j is equal to its transpose operator on L^2. Moreover, the operator D_k is also equal to its transpose. Thus

$$(D_k\Delta_j)^t = \Delta_j D_k$$

and it therefore suffices to prove only that

$$\big\|D_k\Delta_j\big\|_{L^2\to L^2} \le C2^{-\frac{1}{2}|j-k|}\,. \tag{6.4.17}$$

By a simple dilation argument it suffices to prove (6.4.17) when $k=0$. In this case we have the estimate

$$\begin{aligned}\big\|D_0\Delta_j\big\|_{L^2\to L^2} &= \big\|E_0\Delta_j - E_{-1}\Delta_j\big\|_{L^2\to L^2}\\ &\le \big\|E_0\Delta_j - \Delta_j\big\|_{L^2\to L^2} + \big\|E_{-1}\Delta_j - \Delta_j\big\|_{L^2\to L^2},\end{aligned}$$

and since the D_k's and Δ_j's are self-transposes, we have

$$\begin{aligned}\big\|D_0\Delta_j\big\|_{L^2\to L^2} = \big\|\Delta_j D_0\big\|_{L^2\to L^2} &= \big\|\Delta_j E_0 - \Delta_j E_{-1}\big\|_{L^2\to L^2}\\ &\le \big\|\Delta_j E_0\big\|_{L^2\to L^2} + \big\|\Delta_j E_{-1}\big\|_{L^2\to L^2}.\end{aligned}$$

Estimate (6.4.17) when $k=0$ will be a consequence of the pair of inequalities

$$\big\|E_0\Delta_j - \Delta_j\big\|_{L^2\to L^2} + \big\|E_{-1}\Delta_j - \Delta_j\big\|_{L^2\to L^2} \le C2^{\frac{j}{2}} \quad \text{for } j\le 0, \tag{6.4.18}$$

$$\big\|\Delta_j E_0\big\|_{L^2\to L^2} + \big\|\Delta_j E_{-1}\big\|_{L^2\to L^2} \le C2^{-\frac{1}{2}j} \quad \text{for } j\ge 0. \tag{6.4.19}$$

We start by proving (6.4.18). We consider only the term $E_0\Delta_j - \Delta_j$, since the term $E_{-1}\Delta_j - \Delta_j$ is similar. Let $f\in L^2(\mathbf{R}^n)$. Then

$$\begin{aligned}&\big\|E_0\Delta_j(f) - \Delta_j(f)\big\|_{L^2}^2\\ &\quad= \sum_{Q\in\mathscr{D}_0}\big\|f*\Psi_{2^{-j}} - \operatorname*{Avg}_Q(f*\Psi_{2^{-j}})\big\|_{L^2(Q)}^2\\ &\quad\le \sum_{Q\in\mathscr{D}_0}\int_Q\int_Q |(f*\Psi_{2^{-j}})(x) - (f*\Psi_{2^{-j}})(t)|^2\,dt\,dx\\ &\quad\le 3\sum_{Q\in\mathscr{D}_0}\int_Q\int_Q\bigg(\int_{5\sqrt{n}Q}|f(y)||\Psi_{2^{-j}}(x-y)|\,dy\bigg)^2 dt\,dx\\ &\qquad\quad +3\sum_{Q\in\mathscr{D}_0}\int_Q\int_Q\bigg(\int_{5\sqrt{n}Q}|f(y)||\Psi_{2^{-j}}(t-y)|\,dy\bigg)^2 dt\,dx\\ &\qquad\quad +3\sum_{Q\in\mathscr{D}_0}\int_Q\int_Q\bigg(\int_{(5\sqrt{n}Q)^c}|f(y)|2^{jn+j}|\nabla\Psi(2^j(\xi_{x,t}-y))|\,dy\bigg)^2 dt\,dx,\end{aligned}$$

where $\xi_{x,t}$ lies on the line segment joining x and t. Applying the Cauchy-Schwarz inequality to the first two terms, we see that the last expression is bounded by

$$C2^{jn}\sum_{Q\in\mathscr{D}_0}\int_{5\sqrt{n}Q}|f(y)|^2\,dy + C_M 2^{2j}\sum_{Q\in\mathscr{D}_0}\int_Q\bigg(\int_{\mathbf{R}^n}\frac{2^{jn}|f(y)|\,dy}{(1+2^j|x-y|)^M}\bigg)^2 dx,$$

which is clearly controlled by $C(2^{jn}+2^{2j})\|f\|_{L^2}^2 \le 2C2^j\|f\|_{L^2}^2$. This proves (6.4.18).

We now turn to the proof of (6.4.19). We work only with the term $\Delta_j E_0$, since the other term can be treated similarly. We have

$$\begin{aligned}\big\|\Delta_j E_0(f)\big\|_{L^2}^2 &= \Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)\,(\Psi_{2^{-j}}*\chi_Q)\Big\|_{L^2}^2\\ &\le 2\Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)\,(\Psi_{2^{-j}}*\chi_Q)\chi_{5\sqrt{n}Q}\Big\|_{L^2}^2\\ &\qquad +2\Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)\,(\Psi_{2^{-j}}*\chi_Q)\chi_{(5\sqrt{n}Q)^c}\Big\|_{L^2}^2.\end{aligned}$$

Since the functions appearing inside the sum in the first term have supports with bounded overlap, we obtain

$$\Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)(\Psi_{2^{-j}}*\chi_Q)\chi_{5\sqrt{n}Q}\Big\|_{L^2}^2 \le C\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q|f|)^2\big\|\Psi_{2^{-j}}*\chi_Q\big\|_{L^2}^2,$$

and the crucial observation is that for any $Q\in\mathscr{D}_0$ we have

$$\big\|\Psi_{2^{-j}}*\chi_Q\big\|_{L^2}^2 \le C'\sum_{r=1}^{n}\int_{|\xi_r|\approx 2^j}\frac{|e^{2\pi i\xi_r}-1|^2}{|2\pi i\xi_r|^2}d\xi_r\Big[\prod_{l\neq r}\int\frac{|e^{2\pi i\xi_l}-1|^2}{|2\pi i\xi_l|^2}d\xi_l\Big]\le C2^{-\frac{9}{10}j},$$

a consequence of Plancherel's identity and of the fact that in the region where ξ_r is the largest variable of $\xi=(\xi_1,\dots,\xi_n)$ we have $|\xi_r|\approx|\xi|\approx 2^j$ on the support of $\widehat{\Psi_{2^{-j}}}(\xi)$. Putting these observations together, we deduce

$$\Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)\,(\Psi_{2^{-j}}*\chi_Q)\chi_{3Q}\Big\|_{L^2}^2 \le C\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q|f|)^2 2^{-\frac{9}{10}j}\le C2^{-\frac{9}{10}j}\|f\|_{L^2}^2,$$

and the required conclusion will be proved if we can show that

$$\Big\|\sum_{Q\in\mathscr{D}_0}(\operatorname*{Avg}_Q f)\,(\Psi_{2^{-j}}*\chi_Q)\chi_{(3Q)^c}\Big\|_{L^2}^2\le C2^{-j}\|f\|_{L^2}^2. \tag{6.4.20}$$

We prove (6.4.20) by using an estimate based purely on size. Let c_Q be the center of the dyadic cube Q. For $x\notin 3Q$ we have the estimate

$$|(\Psi_{2^{-j}}*\chi_Q)(x)|\le\frac{C_M 2^{jn}}{(1+2^j|x-c_Q|)^M}\le\frac{C_M 2^{jn}}{(1+2^j)^{M/2}}\frac{1}{(1+|x-c_Q|)^{M/2}},$$

since both $2^j\ge 1$, and $|x-c_Q|\ge 1$. We now control the left-hand side of (6.4.20) by

$$\begin{aligned}&2^{j(2n-M)}\sum_{Q\in\mathscr{D}_0}\sum_{Q'\in\mathscr{D}_0}(\operatorname*{Avg}_Q|f|)(\operatorname*{Avg}_{Q'}|f|)\int_{\mathbf{R}^n}\frac{C_M\,dx}{(1+|x-c_Q|)^{\frac{M}{2}}(1+|x-c_{Q'}|)^{\frac{M}{2}}}\\ &\le 2^{j(2n-M)}\sum_{Q\in\mathscr{D}_0}\sum_{Q'\in\mathscr{D}_0}\frac{(\operatorname*{Avg}_Q|f|)(\operatorname*{Avg}_{Q'}|f|)}{(1+|c_Q-c_{Q'}|)^{\frac{M}{4}}}\int_{\mathbf{R}^n}\frac{C_M\,dx}{(1+|x-c_Q|)^{\frac{M}{4}}(1+|x-c_{Q'}|)^{\frac{M}{4}}}\\ &\le 2^{j(2n-M)}\sum_{Q\in\mathscr{D}_0}\sum_{Q'\in\mathscr{D}_0}\frac{C_M}{(1+|c_Q-c_{Q'}|)^{\frac{M}{4}}}\Big(\int_Q|f(y)|^2\,dy+\int_{Q'}|f(y)|^2\,dy\Big)\end{aligned}$$

$$
\begin{aligned}
&\le C_M 2^{j(2n-M)} \sum_{Q\in\mathscr{D}_0} \int_Q |f(y)|^2\,dy \\
&= C_M 2^{j(2n-M)} \|f\|_{L^2}^2 .
\end{aligned}
$$

By taking M large enough, we obtain (6.4.20) and thus (6.4.19). □

Exercises

6.4.1. (a) Prove that no dyadic cube in $\mathbf{R}^n$ contains the point 0 in its interior.
(b) Prove that every interval $[a,b]$ is contained in the union of three dyadic intervals of length less than $b-a$.
(c) Prove that every cube of length l in $\mathbf{R}^n$ is contained in the union of 3^n dyadic cubes, each having length less than l.

6.4.2. Let $k\in\mathbf{Z}$. Show that the set $[m2^{-k},(m+s)2^{-k})$ is a dyadic interval if and only if $s=2^p$ for some $p\in\mathbf{Z}$ and m is an integer multiple of s.

6.4.3. Given a cube Q in $\mathbf{R}^n$ of side length $\ell(Q)\le 2^{k-1}$ for some integer k, prove that there is a dyadic cube D_Q of side length 2^k such that $Q\subseteqq \sigma+D_Q$ for some $\sigma=(\sigma_1,\dots,\sigma_n)$, where $\sigma_j\in\{0,2^k/3,-2^k/3\}$.

6.4.4. Show that the martingale maximal function $f\mapsto \sup_{k\in\mathbf{Z}}|E_k(f)|$ is weak type $(1,1)$ with constant at most 1.
$\big[$*Hint:* Use Exercise 2.1.12.$\big]$

6.4.5. (a) Show that $E_N(f)\to f$ a.e. as $N\to\infty$ for all $f\in L^1_{\mathrm{loc}}(\mathbf{R}^n)$.
(b) Prove that $E_N(f)\to f$ in L^p as $N\to\infty$ for all $f\in L^p(\mathbf{R}^n)$ whenever $1<p<\infty$.

6.4.6. (a) Let $k,k'\in\mathbf{Z}$ be such that $k\neq k'$. Show that for functions f and g in $L^2(\mathbf{R}^n)$ we have

$$\big\langle D_k(f), D_{k'}(g)\big\rangle = 0 .$$

(b) Conclude that for functions f_j in $L^2(\mathbf{R}^n)$ we have

$$\Big\|\sum_{j\in\mathbf{Z}} D_j(f_j)\Big\|_{L^2(\mathbf{R}^n)} = \Big(\sum_{j\in\mathbf{Z}} \big\|D_j(f_j)\big\|_{L^2(\mathbf{R}^n)}^2\Big)^{\frac12} .$$

(c) Let Δ_j and C be as in the statement of Theorem 6.4.8. Show that for any $r\in\mathbf{Z}$ we have

$$\Big\|\sum_{j\in\mathbf{Z}} D_j\Delta_{j+r}D_j\Big\|_{L^2(\mathbf{R}^n)\to L^2(\mathbf{R}^n)} \le C\,2^{-\frac{9}{20}|r|} .$$

6.4.7. ([133]) Let D_j, Δ_j be as in Theorem 6.4.8.
(a) Prove that the operator

$$V_r=\sum_{j\in\mathbf{Z}} D_j\Delta_{j+r}$$

is bounded from $L^2(\mathbf{R}^n)$ to itself with norm at most a multiple of $2^{-\frac{9}{20}|r|}$.

(b) Show that V_r is $L^p(\mathbf{R}^n)$ bounded for all $1 < p < \infty$ with a constant depending only on p and n.
(c) Conclude that for each $1 < p < \infty$ there is a constant $c_p > 0$ such that V_r is bounded on $L^p(\mathbf{R}^n)$ with norm at most a multiple of $2^{-c_p|r|}$.
$\big[$*Hint:* Part (a): Write $\Delta_j = \Delta_j \widetilde{\Delta}_j$, where $\widetilde{\Delta}_j$ is another family of Littlewood–Paley operators and use Exercise 6.4.6 (b). Part (b): Use duality and (6.1.21).$\big]$

6.5 The Spherical Maximal Function

In this section we discuss yet another consequence of the Littlewood–Paley theory, the boundedness of the spherical maximal operator.

6.5.1 *Introduction of the Spherical Maximal Function*

We denote throughout this section by $d\sigma$ the ~~normalized~~ Lebesgue measure on the sphere $\mathbf{S}^{n-1}$. For f in $\mathscr{S}(\mathbf{R}^n)$, $1 \le p \le \infty$, we define the maximal operator

$$\mathscr{M}(f)(x) = \sup_{t>0} \left| \int_{\mathbf{S}^{n-1}} f(x - t\theta)\, d\sigma(\theta) \right|. \tag{6.5.1}$$

The operator $\mathscr{M}$ is called the *spherical maximal function*. It is unclear at this point for which classes of L^p functions f the definition of $\mathscr{M}$ extends and for which values of $p < \infty$ the maximal inequality

$$\big\|\mathscr{M}(f)\big\|_{L^p(\mathbf{R}^n)} \le C_p \big\|f\big\|_{L^p(\mathbf{R}^n)} \tag{6.5.2}$$

holds for all functions $f \in L^p(\mathbf{R}^n)$.

Spherical averages often make their appearance as solutions of partial differential equations. For instance, the spherical average

$$u(x,t) = \frac{1}{4\pi} \int_{\mathbf{S}^2} t\, f(x - ty)\, d\sigma(y) \tag{6.5.3}$$

is a solution of the *wave equation*

$$\begin{aligned} \Delta_x(u)(x,t) &= \frac{\partial^2 u}{\partial t^2}(x,t)\,, \\ u(x,0) &= 0\,, \\ \frac{\partial u}{\partial t}(x,0) &= f(x)\,, \end{aligned}$$

in $\mathbf{R}^3$. The introduction of the spherical maximal function is motivated by the fact that the related spherical average

$$u(x,t) = \frac{1}{4\pi}\int_{\mathbf{S}^2} f(x-ty)\,d\sigma(y) \tag{6.5.4}$$

solves *Darboux's equation*

$$\begin{aligned} \Delta_x(u)(x,t) &= \frac{\partial^2 u}{\partial t^2}(x,t) + \frac{2}{t}\frac{\partial u}{\partial t}(x,t)\,,\\ u(x,0) &= f(x)\,,\\ \frac{\partial u}{\partial t}(x,0) &= 0\,, \end{aligned}$$

in $\mathbf{R}^3$. It is rather remarkable that the Fourier transform can be used to study almost everywhere convergence for several kinds of maximal averaging operators such as the spherical averages in (6.5.4). This is achieved via the boundedness of the corresponding maximal operator; the maximal operator controlling the averages over $\mathbf{S}^{n-1}$ is given in (6.5.1).

Before we begin the analysis of the spherical maximal function, we recall that

$$\widehat{d\sigma}(\xi) = \frac{2\pi}{|\xi|^{\frac{n-2}{2}}} J_{\frac{n-2}{2}}(2\pi|\xi|)\,,$$

as shown in Appendix B.4. Using the estimates in Appendices B.6 and B.7 and the identity

$$\frac{d}{dt}J_\nu(t) = \frac{1}{2}\big(J_{\nu-1}(t) - J_{\nu+1}(t)\big)$$

derived in Appendix B.2, we deduce the crucial estimate

$$|\widehat{d\sigma}(\xi)| + |\nabla\widehat{d\sigma}(\xi)| \le \frac{C_n}{(1+|\xi|)^{\frac{n-1}{2}}}\,. \tag{6.5.5}$$

Theorem 6.5.1. *Let $n \ge 3$. For each $\frac{n}{n-1} < p \le \infty$, there is a constant C_p such that*

$$\big\|\mathscr{M}(f)\big\|_{L^p(\mathbf{R}^n)} \le C_p\big\|f\big\|_{L^p(\mathbf{R}^n)} \tag{6.5.6}$$

holds for all f in $\mathscr{S}(\mathbf{R}^n)$. Consequently, for all $\frac{n}{n-1} < p < \infty$, $\mathscr{M}$ admits a bounded extension on L^p, and for $f \in L^p(\mathbf{R}^n)$ we have

$$\lim_{t\to 0}\frac{1}{\omega_{n-1}}\int_{\mathbf{S}^{n-1}} f(x-t\theta)\,d\sigma(\theta) = f(x) \tag{6.5.7}$$

for almost all $x \in \mathbf{R}^n$. Here we set $\omega_{n-1} = |\mathbf{S}^{n-1}|$.