Chapter 8
Singular Integrals of Nonconvolution Type

Up to this point we have studied singular integrals given by convolution with certain tempered distributions. These operators commute with translations. We are now ready to broaden our perspective and study a class of more general singular integrals that are not necessarily translation invariant. Such operators appear in many places in harmonic analysis and partial differential equations. For instance, a large class of pseudodifferential operators falls under the scope of this theory.

This broader point of view does not necessarily bring additional complications in the development of the subject except at one point, the study of L^2 boundedness, where Fourier transform techniques are lacking. The L^2 boundedness of convolution operators is easily understood via a careful examination of the Fourier transform of the kernel, but for nonconvolution operators different tools are required in this study. The main result of this chapter is the derivation of a set of necessary and sufficient conditions for nonconvolution singular integrals to be L^2 bounded. This result is referred to as the $T(1)$ theorem and owes its name to a condition expressed in terms of the action of the operator T on the function 1.

An extension of the $T(1)$ theorem, called the $T(b)$ theorem, is obtained in Section 8.6 and is used to deduce the L^2 boundedness of the Cauchy integral along Lipschitz curves. A variant of the $T(b)$ theorem is also used in the boundedness of the square root of a divergence form elliptic operator discussed in Section 8.7.

8.1 General Background and the Role of *BMO*

We begin by recalling the notion of the adjoint and transpose operator. One may choose to work with either a real or a complex inner product on pairs of functions. For f,g complex-valued functions with integrable product, we denote the real inner product by

$$\langle f,g\rangle = \int_{\mathbf{R}^n} f(x)g(x)\,dx\,.$$

L. Grafakos, *Modern Fourier Analysis*, DOI: 10.1007/978-0-387-09434-2_8,

This notation is suitable when we think of f as a distribution acting on a test function g. We also have the complex inner product

$$\langle f \,|\, g \rangle = \int_{\mathbf{R}^n} f(x)\overline{g(x)}\,dx,$$

which is an appropriate notation when we think of f and g as elements of a Hilbert space over the complex numbers. Now suppose that T is a linear operator bounded on L^p. Then the *adjoint* operator T^* of T is uniquely defined via the identity

$$\langle T(f) \,|\, g \rangle = \langle f \,|\, T^*(g) \rangle$$

for all f in L^p and g in $L^{p'}$. The *transpose* operator T^t of T is uniquely defined via the identity

$$\langle T(f), g \rangle = \langle f, T^t(g) \rangle = \langle T^t(g), f \rangle$$

for all functions f in L^p and g in $L^{p'}$. The name *transpose* comes from matrix theory, where if A^t denotes the transpose of a complex $n \times n$ matrix A, then we have the identity

$$\langle Ax, y \rangle = \sum_{j=1}^{n} (Ax)_j y_j = Ax \cdot y = x \cdot A^t y = \sum_{j=1}^{n} x_j (A^t y)_j = \langle x, A^t y \rangle$$

for all column vectors $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ in $\mathbf{C}^n$. We may easily check the following intimate relationship between the transpose and the adjoint of a linear operator T:

$$T^*(f) = \overline{T^t(\overline{f}\,)},$$

indicating that they have almost interchangeable use. However, in many cases, it is convenient to avoid complex conjugates and work with the transpose operator for simplicity. Observe that if a linear operator T has kernel $K(x,y)$, that is,

$$T(f)(x) = \int K(x,y) f(y)\,dy,$$

then the kernel of T^t is $K^t(x,y) = K(y,x)$ and that of T^* is $K^*(x,y) = \overline{K(y,x)}$.

An operator is called *self-adjoint* if $T = T^*$ and *self-transpose* if $T = T^t$. For example, the operator iH, where H is the Hilbert transform, is self-adjoint but not self-transpose, and the operator with kernel $i(x+y)^{-1}$ is self-transpose but not self-adjoint.

8.1.1 Standard Kernels

The singular integrals we study in this chapter have kernels that satisfy size and regularity properties similar to those encountered in Chapter 4 for convolution-type

Calderón–Zygmund operators. Let us be specific and introduce the relevant background. We consider functions $K(x,y)$ defined on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) : x \in \mathbf{R}^n\}$ that satisfy for some $A > 0$ the size condition

$$|K(x,y)| \le \frac{A}{|x-y|^n} \tag{8.1.1}$$

and for some $\delta > 0$ the regularity conditions

$$|K(x,y) - K(x',y)| \le \frac{A\,|x-x'|^\delta}{(|x-y|+|x'-y|)^{n+\delta}}, \tag{8.1.2}$$

whenever $|x-x'| \le \frac{1}{2}\max\big(|x-y|,|x'-y|\big)$ and

$$|K(x,y) - K(x,y')| \le \frac{A\,|y-y'|^\delta}{(|x-y|+|x-y'|)^{n+\delta}}, \tag{8.1.3}$$

whenever $|y-y'| \le \frac{1}{2}\max\big(|x-y|,|x-y'|\big)$.

Remark 8.1.1. Observe that if

$$|x-x'| \le \frac{1}{2}\max\big(|x-y|,|x'-y|\big),$$

then

$$\max\big(|x-y|,|x'-y|\big) \le 2\min\big(|x-y|,|x'-y|\big),$$

implying that the numbers $|x-y|$ and $|x'-y|$ are comparable. This fact is useful in specific calculations.

Another important observation is that if (8.1.1) holds and we have

$$|\nabla_x K(x,y)| + |\nabla_y K(x,y)| \le \frac{A}{|x-y|^{n+1}}$$

for all $x \ne y$, then K is in $SK(1, 4^{n+1}A)$.

Definition 8.1.2. Functions on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) : x \in \mathbf{R}^n\}$ that satisfy (8.1.1), (8.1.2), and (8.1.3) are called *standard kernels* with constants δ, A. The class of all standard kernels with constants δ, A is denoted by $SK(\delta,A)$. Given a kernel $K(x,y)$ in $SK(\delta,A)$, we observe that the functions $K(y,x)$ and $\overline{K(y,x)}$ are also in $SK(\delta,A)$. These functions have special names. The function

$$K^t(x,y) = K(y,x)$$

is called the *transpose kernel* of K, and the function

$$K^*(x,y) = \overline{K(y,x)}$$

is called the *adjoint kernel* of K.

Example 8.1.3. The function $K(x,y) = |x-y|^{-n}$ defined away from the diagonal of $\mathbf{R}^n \times \mathbf{R}^n$ is in $SK(1, n4^{n+1})$. Indeed, for

$$|x-x'| \le \frac{1}{2}\max\big(|x-y|, |x'-y|\big)$$

the mean value theorem gives

$$\big|\,|x-y|^{-n} - |x'-y|^{-n}\big| \le \frac{n|x-x'|}{|\theta-y|^{n+1}}$$

for some θ that lies on the line segment joining x and x'. But then we have $|\theta - y| \ge \frac{1}{2}\max\big(|x-y|, |x'-y|\big)$, which gives (8.1.2) with $A = n4^{n+1}$.

Remark 8.1.4. The previous example can be modified to give that if $K(x,y)$ satisfies

$$|\nabla_x K(x,y)| \le A'|x-y|^{-n-1}$$

for all $x \neq y$ in $\mathbf{R}^n$, then $K(x,y)$ also satisfies (8.1.2) with $\delta = 1$ and A controlled by a constant multiple of A'. Likewise, if

$$|\nabla_y K(x,y)| \le A'|x-y|^{-n-1}$$

for all $x \neq y$ in $\mathbf{R}^n$, then $K(x,y)$ satisfies (8.1.3) with with $\delta = 1$ and A bounded by a multiple of A'.

We are interested in standard kernels K that can be extended to tempered distributions on $\mathbf{R}^n \times \mathbf{R}^n$. We begin by observing that given a standard kernel $K(x,y)$, there may not exist a tempered distribution W on $\mathbf{R}^n \times \mathbf{R}^n$ that coincides with the given $K(x,y)$ on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) : x \in \mathbf{R}^n\}$. For example, the function $K(x,y) = |x-y|^{-n}$ does not admit such an extension; see Exercise 8.1.2.

We are concerned with kernels $K(x,y)$ in $SK(\delta, A)$ for which there are tempered distributions W on $\mathbf{R}^n \times \mathbf{R}^n$ that coincide with K on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) : x \in \mathbf{R}^n\}$. This means that the convergent integral representation

$$\big\langle W, F\big\rangle = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)F(x,y)\,dx\,dy \tag{8.1.4}$$

is valid whenever the Schwartz function F on $\mathbf{R}^n \times \mathbf{R}^n$ is supported away from the diagonal $\{(x,x) : x \in \mathbf{R}^n\}$. Note that the integral in (8.1.4) is well defined and absolutely convergent whenever F is a Schwartz function that vanishes in a neighborhood of the set $\{(x,x) : x \in \mathbf{R}^n\}$. Also observe that there may be several distributions W coinciding with a fixed function $K(x,y)$. In fact, if W is such a distribution, then so is $W + \delta_{x=y}$, where $\delta_{x=y}$ denotes Lebesgue measure on the diagonal of $\mathbf{R}^{2n}$. (This is some sort of a Dirac distribution.)

We now consider continuous linear operators

$$T : \mathscr{S}(\mathbf{R}^n) \to \mathscr{S}'(\mathbf{R}^n)$$

from the space of Schwartz functions $\mathscr{S}(\mathbf{R}^n)$ to the space of all tempered distributions $\mathscr{S}'(\mathbf{R}^n)$. By the *Schwartz kernel theorem* (see Hörmander [168, p. 129]), for such an operator T there is a distribution W in $\mathscr{S}'(\mathbf{R}^{2n})$ that satisfies

$$\langle T(f),\varphi\rangle = \langle W, f\otimes\varphi\rangle \qquad \text{when} \quad f,\varphi\in\mathscr{S}(\mathbf{R}^n), \tag{8.1.5}$$

where $(f\otimes\varphi)(x,y) = f(x)\varphi(y)$. Furthermore, as a consequence of the same theorem, there exist constants C,N,M such that for all $f,g\in\mathscr{S}(\mathbf{R}^n)$ we have

$$|\langle T(f),g\rangle| = |\langle W, f\otimes g\rangle| \le C\Big[\sum_{|\alpha|,|\beta|\le N}\rho_{\alpha,\beta}(f)\Big]\Big[\sum_{|\alpha|,|\beta|\le M}\rho_{\alpha,\beta}(g)\Big], \tag{8.1.6}$$

where $\rho_{\alpha,\beta}(\varphi) = \sup_{x\in\mathbf{R}^n}|\partial_x^\alpha(x^\beta\varphi)(x)|$ is the set of seminorms for the topology in $\mathscr{S}$. A distribution W that satisfies (8.1.5) and (8.1.6) is called a *Schwartz kernel*.

We study continuous linear operators $T:\mathscr{S}(\mathbf{R}^n)\to\mathscr{S}'(\mathbf{R}^n)$ whose Schwartz kernels coincide with standard kernels $K(x,y)$ on $\mathbf{R}^n\times\mathbf{R}^n\setminus\{(x,x):\ x\in\mathbf{R}^n\}$. This means that (8.1.5) admits the absolutely convergent integral representation

$$\langle T(f),\varphi\rangle = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)f(y)\varphi(x)\,dx\,dy \tag{8.1.7}$$

whenever f and φ are Schwartz functions whose supports do not intersect.

We make some remarks concerning duality in this context. Given a continuous linear operator $T:\ \mathscr{S}(\mathbf{R}^n)\to\mathscr{S}'(\mathbf{R}^n)$ with a Schwartz kernel W, we can define another distribution W^t as follows:

$$\langle W^t, F\rangle = \langle W, F^t\rangle,$$

where $F^t(x,y) = F(y,x)$. This means that for all $f,\varphi\in\mathscr{S}(\mathbf{R}^n)$ we have

$$\langle W, f\otimes\varphi\rangle = \langle W^t,\varphi\otimes f\rangle.$$

It is a simple fact that the transpose operator T^t of T, which satisfies

$$\langle T(\varphi),f\rangle = \langle T^t(f),\varphi\rangle \tag{8.1.8}$$

for all f,φ in $\mathscr{S}(\mathbf{R}^n)$, is the unique continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ whose Schwartz kernel is the distribution W^t, that is, we have

$$\langle T^t(f),\varphi\rangle = \langle T(\varphi),f\rangle = \langle W,\varphi\otimes f\rangle = \langle W^t, f\otimes\varphi\rangle. \tag{8.1.9}$$

We now observe that a large class of standard kernels admits extensions to tempered distributions W on $\mathbf{R}^{2n}$.

Example 8.1.5. Suppose that $K(x,y)$ satisfies (8.1.1) and (8.1.2) and is *antisymmetric*, in the sense that

$$K(x,y) = -K(y,x)$$

for all $x \neq y$ in $\mathbf{R}^n$. Then K also satisfies (8.1.3), and moreover, there is a distribution W on $\mathbf{R}^{2n}$ that extends K on $\mathbf{R}^n \times \mathbf{R}^n$.

Indeed, define

$$\langle W, F \rangle = \lim_{\varepsilon \to 0} \iint_{|x-y|>\varepsilon} K(x,y) F(x,y)\, dy\, dx \tag{8.1.10}$$

for all F in the Schwartz class of $\mathbf{R}^{2n}$. In view of antisymmetry, we may write

$$\iint_{|x-y|>\varepsilon} K(x,y) F(x,y)\, dy\, dx = \frac{1}{2} \iint_{|x-y|>\varepsilon} K(x,y) \big(F(x,y) - F(y,x)\big)\, dy\, dx.$$

Using (8.1.1), the observation that

$$|F(x,y) - F(y,x)| \le \frac{2\,|x-y|}{(1+|x|^2+|y|^2)^{n+1}} \sup_{(x,y)\in \mathbf{R}^{2n}} \Big| \nabla_{x,y} \Big((1+|x|^2+|y|^2)^{n+1} F(x,y) \Big) \Big|,$$

and the fact that the preceding supremum is controlled by a finite sum of Schwartz seminorms of F, it follows that the limit in (8.1.10) exists and gives a tempered distribution on $\mathbf{R}^{2n}$. We can therefore define an operator $T : \mathscr{S}(\mathbf{R}^n) \to \mathscr{S}'(\mathbf{R}^n)$ with kernel W as follows:

$$\langle T(f), \varphi \rangle = \lim_{\varepsilon \to 0} \iint_{|x-y|>\varepsilon} K(x,y) f(x) \varphi(y)\, dy\, dx.$$

Example 8.1.6. Let A be a Lipschitz function on $\mathbf{R}$. This means that it satisfies the estimate $|A(x) - A(y)| \le L|x-y|$ for some $L < \infty$ and all $x, y \in \mathbf{R}$. For $x, y \in \mathbf{R}$, $x \neq y$, we let

$$K(x,y) = \frac{1}{x - y + i(A(x) - A(y))} \tag{8.1.11}$$

and we observe that $K(x,y)$ is a standard kernel in $SK(1, 4+4L)$. The details are left to the reader. Note that the kernel K defined in (8.1.11) is antisymmetric.

Example 8.1.7. Let the function A be as in the previous example. For each integer $m \ge 1$ we set

$$K_m(x,y) = \left(\frac{A(x) - A(y)}{x-y} \right)^m \frac{1}{x-y}, \qquad x, y \in \mathbf{R}. \tag{8.1.12}$$

Clearly, K_m is an antisymmetric function. To see that each K_m is a standard kernel, we use the simple fact that

$$\max\big(|\nabla_x K_m(x,y)|, |\nabla_y K_m(x,y)| \big) \le \frac{(2m+1)L^m}{|x-y|^2}$$

and the observation made in Remark 8.1.1. It follows that K_m lies in $SK(\delta,C)$ with $\delta = 1$ and $C = 16(2m+1)L^m$. The linear operator with kernel $(\pi i)^{-1}K_m$ is called the mth *Calderón commutator*.

8.1.2 Operators Associated with Standard Kernels

Having introduced standard kernels, we are in a position to define linear operators associated with them.

Definition 8.1.8. Let $0 < \delta, A < \infty$ and K in $SK(\delta,A)$. A continuous linear operator T from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ is said to be *associated with* K if it satisfies

$$T(f)(x) = \int_{\mathbf{R}^n} K(x,y)f(y)\,dy \tag{8.1.13}$$

for all $f \in \mathscr{C}_0^\infty$ and x not in the support of f. If T is associated with K, then the Schwartz kernel W of T coincides with K on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) :\ x \in \mathbf{R}^n\}$.

If T is associated with K and admits a bounded extension on $L^2(\mathbf{R}^n)$, that is, it satisfies

$$\big\|T(f)\big\|_{L^2} \le B\big\|f\big\|_{L^2} \tag{8.1.14}$$

for all $f \in \mathscr{S}(\mathbf{R}^n)$, then T is called a *Calderón–Zygmund operator* associated with the standard kernel K. In this case we use the same notation for the L^2 extension.

In the sequel we denote by $CZO(\delta,A,B)$ the class of all Calderón–Zygmund operators associated with standard kernels in $SK(\delta,A)$ that admit L^2 bounded extensions with norm at most B.

We make the point that there may be several Calderón–Zygmund operators associated with a given standard kernel K. For instance, we may check that the zero operator and the identity operator have the same kernel $K(x,y) = 0$. We investigate connections between any two such operators in Proposition 8.1.11. Next we discuss the important fact that once an operator T admits an extension that is L^2 bounded, then (8.1.13) holds for all f that are bounded and compactly supported whenever x does not lie in its support.

Proposition 8.1.9. *Let T be an element of $CZO(\delta,A,B)$ associated with a standard kernel K. Then for all f in L^∞ with compact support and every $x \notin \operatorname{supp} f$ we have the absolutely convergent integral representation*

$$T(f)(x) = \int_{\mathbf{R}^n} K(x,y)f(y)\,dy. \tag{8.1.15}$$

Proof. Identity (8.1.15) can be deduced from the fact that whenever f and φ are bounded and compactly supported functions that satisfy

$$\operatorname{dist}(\operatorname{supp}\varphi, \operatorname{supp} f) > 0, \tag{8.1.16}$$

then we have the integral representation

$$\int_{\mathbf{R}^n} T(f)(x)\,\varphi(x)\,dx = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)f(y)\varphi(x)\,dy\,dx. \tag{8.1.17}$$

To see this, given f and φ as previously, select $f_j, \varphi_j \in \mathscr{C}_0^\infty$ such that φ_j are uniformly bounded and supported in a small neighborhood of the support of φ, $\varphi_j \to \varphi$ in L^2 and almost everywhere, $f_j \to f$ in L^2 and almost everywhere, and

$$\operatorname{dist}(\operatorname{supp}\varphi_j, \operatorname{supp} f_j) \geq \frac{1}{2}\operatorname{dist}(\operatorname{supp}\varphi, \operatorname{supp} f) > 0$$

for all j. Because of (8.1.7), identity (8.1.17) is valid for the functions f_j and φ_j in place of f and φ. By the boundedness of T, it follows that $T(f_j)$ converges to $T(f)$ in L^2 and thus

$$\int_{\mathbf{R}^n} T(f_j)(x)\varphi_j(x)\,dx \to \int_{\mathbf{R}^n} T(f)(x)\varphi(x)\,dx.$$

Now write $f_j\varphi_j - f\varphi = (f_j - f)\varphi_j + f(\varphi_j - \varphi)$ and observe that

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)f(y)(\varphi_j(x) - \varphi(x))\,dy\,dx \to 0\,,$$

since it is controlled by a multiple of $\big\|T(f)\big\|_{L^2}\big\|\varphi_j - \varphi\big\|_{L^2}$, while

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)(f_j(y) - f(y))\varphi_j(x)\,dy\,dx \to 0\,,$$

since it is controlled by a multiple of $\sup_j \big\|T^t(\varphi_j)\big\|_{L^2}\big\|f_j - f\big\|_{L^2}$. This gives that

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)f_j(y)\varphi_j(x)\,dy\,dx \to \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)f(y)\varphi(x)\,dy\,dx$$

as $j \to \infty$, which proves the validity of (8.1.17). □

We now define truncated kernels and operators.

Definition 8.1.10. Given a kernel K in $SK(\delta,A)$ and $\varepsilon > 0$, we define the *truncated kernel*

$$K^{(\varepsilon)}(x,y) = K(x,y)\chi_{|x-y|>\varepsilon}\,.$$

Given a continuous linear operator T from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ and $\varepsilon > 0$, we define the *truncated operator* $T^{(\varepsilon)}$ by

$$T^{(\varepsilon)}(f)(x) = \int_{\mathbf{R}^n} K^{(\varepsilon)}(x,y)\,f(y)\,dy$$

and the *maximal singular operator* associated with T as follows:

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

Note that both $T^{(\varepsilon)}$ and $T^{(*)}$ are well defined for f in $\bigcup_{1\le p<\infty} L^p(\mathbf{R}^n)$.

We investigate a certain connection between the boundedness of T and the boundedness of the family $\{T^{(\varepsilon)}\}_{\varepsilon>0}$ uniformly in $\varepsilon > 0$.

Proposition 8.1.11. *Let K be a kernel in $SK(\delta,A)$ and let T in $CZO(\delta,A,B)$ be associated with K. For $\varepsilon > 0$, let $T^{(\varepsilon)}$ be the truncated operators obtained from T. Assume that there exists a constant $B' < \infty$ such that*

$$\sup_{\varepsilon>0} \big\|T^{(\varepsilon)}\big\|_{L^2\to L^2} \le B'. \tag{8.1.18}$$

Then there exists a linear operator T_0 defined on $L^2(\mathbf{R}^n)$ such that

(1) The Schwartz kernel of T_0 coincides with K on

$$\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x):\ x\in \mathbf{R}^n\}.$$

(2) For some subsequence $\varepsilon_j \downarrow 0$, we have

$$\int_{\mathbf{R}^n} T^{(\varepsilon_j)}(f)(x)g(x)\,dx \to \int_{\mathbf{R}^n} (T_0 f)(x)g(x)\,dx$$

as $j\to\infty$ for all f,g in $L^2(\mathbf{R}^n)$.

(3) T_0 is bounded on $L^2(\mathbf{R}^n)$ with norm

$$\big\|T_0\big\|_{L^2\to L^2} \le B'.$$

(4) There exists a measurable function b on $\mathbf{R}^n$ with $\big\|b\big\|_{L^\infty} \le B+B'$ such that

$$T(f) - T_0(f) = bf\,,$$

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

Proof. Consider the Banach space $X = \mathscr{B}(L^2,L^2)$ of all bounded linear operators from $L^2(\mathbf{R}^n)$ to itself. Then X is isomorphic to $\mathscr{B}((L^2)^*,(L^2)^*)^*$, which is a dual space. Since the unit ball of a dual space is weak* compact, and the operators $T^{(\varepsilon)}$ lie in a multiple of this unit ball, the Banach–Alaoglu theorem gives the existence of a sequence $\varepsilon_j \downarrow 0$ such that $T^{(\varepsilon_j)}$ converges to some T_0 in the weak* topology of $\mathscr{B}(L^2,L^2)$ as $j\to\infty$. This means that

$$\int_{\mathbf{R}^n} T^{(\varepsilon_j)}(f)(x)g(x)\,dx \to \int_{\mathbf{R}^n} T_0(f)(x)g(x)\,dx \tag{8.1.19}$$

for all f,g in $L^2(\mathbf{R}^n)$ as $j\to\infty$. This proves (2). The L^2 boundedness of T_0 is a consequence of (8.1.19), hypothesis (8.1.18), and duality, since

$$\big\|T_0(f)\big\|_{L^2} \le \sup_{\|g\|_{L^2}\le 1}\ \limsup_{j\to\infty} \left| \int_{\mathbf{R}^n} T^{(\varepsilon_j)}(f)(x)g(x)\,dx\right| \le B'\big\|f\big\|_{L^2}.$$

This proves (3). Finally, (1) is a consequence of the integral representation

$$\int_{\mathbf{R}^n} T^{(\varepsilon_j)}(f)(x)g(x)\,dx = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K^{(\varepsilon_j)}(x,y)f(y)\,dy\,g(x)\,dx,$$

whenever f, g are Schwartz functions with disjoint supports, by letting $j \to \infty$.

We finally prove (4). We first observe that if g is a bounded function with compact support and Q is an open cube in $\mathbf{R}^n$, we have

$$(T^{(\varepsilon)} - T)(g\chi_Q)(x) = \chi_Q(x)\,(T^{(\varepsilon)} - T)(g)(x)\,, \tag{8.1.20}$$

whenever $x \notin \partial Q$ and ε is small enough. Indeed, take first $x \notin \overline{Q}$; then x is not in the support of $g\chi_Q$. Note that since $g\chi_Q$ is bounded and has compact support, we can use the integral representation formula (8.1.15) obtained in Proposition 8.1.9. Then we have that for $\varepsilon < \operatorname{dist}(x, \operatorname{supp} g\chi_Q)$, the left-hand side in (8.1.20) is zero. Moreover, for $x \in Q$, we have that x does not lie in the support of $g\chi_{Q^c}$, and again because of (8.1.15) we obtain $(T^{(\varepsilon)} - T)(g\chi_{Q^c})(x) = 0$ whenever $\varepsilon < \operatorname{dist}(x, \operatorname{supp} g\chi_{Q^c})$. This proves (8.1.20) for all x not in the boundary ∂Q of Q. Taking weak limits in (8.1.20) as $\varepsilon \to 0$, we obtain that

$$(T_0 - T)(g\chi_Q) = \chi_Q\,(T_0 - T)(g) \qquad \text{a.e.} \tag{8.1.21}$$

for all open cubes Q in $\mathbf{R}^n$. By linearity we extend (8.1.21) to simple functions. Using the fact that $T_0 - T$ is L^2 bounded and a simple density argument, we obtain

$$(T_0 - T)(gf) = f\,(T_0 - T)(g) \qquad \text{a.e.} \tag{8.1.22}$$

whenever f is in L^2 and g is bounded and has compact support. If $B(0,j)$ is the open ball with center 0 and radius j on $\mathbf{R}^n$, when $j \le j'$ we have

$$(T_0 - T)(\chi_{B(0,j)}) = (T_0 - T)(\chi_{B(0,j)}\chi_{B(0,j')}) = \chi_{B(0,j)}\,(T_0 - T)(\chi_{B(0,j')})\,.$$

Therefore, the sequence of functions $(T_0 - T)(\chi_{B(0,j)})$ satisfies the "consistency" property

$$(T_0 - T)(\chi_{B(0,j)}) = (T_0 - T)(\chi_{B(0,j')}) \quad \text{in } B(0,j)$$

when $j \le j'$. It follows that there exists a well defined function b such that

$$b = (T_0 - T)(\chi_{B(0,j)}) \quad \text{a.e. in } B(0,j)\,.$$

Applying (8.1.22) with f supported in $B(0,j)$ and $g = \chi_{B(0,j)}$, we obtain

$$(T_0 - T)(f) = (T_0 - T)(f\chi_{B(0,j)}) = f\,(T_0 - T)(\chi_{B(0,j)}) = f\,b \qquad \text{a.e.},$$

from which it follows that $(T_0 - T)(f) = b\,f$ for all $f \in L^2$. Since the norm of $T - T_0$ on L^2 is at most $B + B'$, it follows that the norm of the linear map $f \mapsto b\,f$ from L^2 to itself is at most $B + B'$. From this we obtain that $\big\|b\big\|_{L^\infty} \le B + B'$. □

Remark 8.1.12. We show in the next section (cf. Corollary 8.2.4) that if a Calderón–Zygmund operator maps L^2 to L^2, then so do all of its truncations $T^{(\varepsilon)}$ uniformly in $\varepsilon > 0$. By Proposition 8.1.11, there exists a linear operator T_0 that has the form

$$T_0(f)(x) = \lim_{j\to\infty} \int_{|x-y|>\varepsilon_j} K(x,y) f(y)\, dy,$$

where the limit is taken in the weak topology of L^2, so that T is equal to T_0 plus a bounded function times the identity operator.

We give a special name to operators of this form.

Definition 8.1.13. Suppose that for a given T in $CZO(\delta,A,B)$ there is a sequence ε_j of positive numbers that tends to zero as $j \to \infty$ such that for all $f \in L^2(\mathbf{R}^n)$,

$$T^{(\varepsilon_j)}(f) \to T(f)$$

weakly in L^2. Then T is called a *Calderón–Zygmund singular integral operator*. Thus Calderón–Zygmund singular integral operators are special kinds of Calderón–Zygmund operators. The subclass of $CZO(\delta,A,B)$ consisting of all Calderón–Zygmund singular integral operators is denoted by $CZSIO(\delta,A,B)$.

In view of Proposition 8.1.11 and Remark 8.1.12, a Calderón–Zygmund operator is equal to a Calderón–Zygmund singular integral operator plus a bounded function times the identity operator. For this reason, the study of Calderón–Zygmund operators is equivalent to the study of Calderón–Zygmund singular integral operators, and in almost all situations it suffices to restrict attention to the latter.

8.1.3 Calderón–Zygmund Operators Acting on Bounded Functions

We are now interested in defining the action of a Calderón–Zygmund operator T on bounded and smooth functions. To achieve this we first need to define the space of special test functions $\mathscr{D}_0$.

Definition 8.1.14. Recall the space $\mathscr{D}(\mathbf{R}^n) = \mathscr{C}_0^\infty(\mathbf{R}^n)$ of all smooth functions with compact support on $\mathbf{R}^n$. We define $\mathscr{D}_0(\mathbf{R}^n)$ to be the space of all smooth functions with compact support and integral zero. We equip $\mathscr{D}_0(\mathbf{R}^n)$ with the same topology as the space $\mathscr{D}(\mathbf{R}^n)$ (cf. Definition 2.3.1). The dual space of $\mathscr{D}_0(\mathbf{R}^n)$ under this topology is denoted by $\mathscr{D}_0'(\mathbf{R}^n)$. This is a space of distributions larger than $\mathscr{D}'(\mathbf{R}^n)$.

Example 8.1.15. *BMO* functions are examples of elements of $\mathscr{D}_0'(\mathbf{R}^n)$. Indeed, given $b \in BMO(\mathbf{R}^n)$, for any compact set K there is a constant $C_K = \big\|b\big\|_{L^1(K)}$ such that

$$\left| \int_{\mathbf{R}^n} b(x)\varphi(x)\, dx \right| \le C_K \big\|\varphi\big\|_{L^\infty}$$

for any $\varphi \in \mathscr{D}_0(\mathbf{R}^n)$. Moreover, observe that the preceding integral remains unchanged if the BMO function b is replaced by $b+c$, where c is a constant.

Definition 8.1.16. Let T be a continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ that satisfies (8.1.5) for some distribution W that coincides with a standard kernel $K(x,y)$ satisfying (8.1.1), (8.1.2), and (8.1.3). Given f bounded and smooth, we define an element $T(f)$ of $\mathscr{D}_0'(\mathbf{R}^n)$ as follows: For a given φ in $\mathscr{D}_0(\mathbf{R}^n)$, select η in $\mathscr{C}_0^\infty$ with $0 \le \eta \le 1$ and equal to 1 in a neighborhood of the support of φ. Since T maps $\mathscr{S}$ to $\mathscr{S}'$, the expression $T(f\eta)$ is a tempered distribution, and its action on φ is well defined. We define the action of $T(f)$ on φ via

$$\big\langle T(f),\varphi\big\rangle = \big\langle T(f\eta),\varphi\big\rangle + \int_{\mathbf{R}^n}\left[\int_{\mathbf{R}^n} K(x,y)\varphi(x)\,dx\right] f(y)(1-\eta(y))\,dy, \quad (8.1.23)$$

provided we make sense of the double integral as an absolutely convergent integral. To do this, we pick x_0 in the support of φ and we split the y-integral in (8.1.23) into the sum of integrals over the regions $I_0 = \{y \in \mathbf{R}^n :\ |x-x_0| > \frac{1}{2}|x_0-y|\}$ and $I_\infty = \{y \in \mathbf{R}^n :\ |x-x_0| \le \frac{1}{2}|x_0-y|\}$. By the choice of η we must necessarily have $\operatorname{dist}(\operatorname{supp}\eta, \operatorname{supp}\varphi) > 0$, and hence the part of the double integral in (8.1.23) when y is restricted to I_0 is absolutely convergent in view of (8.1.1). For $y \in I_\infty$ we use the mean value property of φ to write the expression inside the square brackets in (8.1.23) as

$$\int_{\mathbf{R}^n} \big(K(x,y) - K(x_0,y)\big)\,\varphi(x)\,dx.$$

With the aid of (8.1.2) we deduce the absolute convergence of the double integral in (8.1.23) as follows:

$$\begin{aligned}
\iint_{|y-x_0|\ge 2|x-x_0|} &|K(x,y)-K(x_0,y)|\,|\varphi(x)|\,(1-\eta(y))\,|f(y)|\,dx\,dy \\
&\le \int_{\mathbf{R}^n} A|x-x_0|^\delta \int_{|y-x_0|\ge 2|x-x_0|} |x_0-y|^{-n-\delta}\,|f(y)|\,dy\,|\varphi(x)|\,dx \\
&\le A\,\frac{\omega_{n-1}}{\delta\, 2^\delta}\,\big\|\varphi\big\|_{L^1}\,\big\|f\big\|_{L^\infty} < \infty.
\end{aligned}$$

This completes the definition of $T(f)$ as an element of $\mathscr{D}_0'$ when $f \in \mathscr{C}^\infty \cap L^\infty$ but leaves two points open. We need to show that this definition is independent of η and secondly that whenever f is a Schwartz function, the distribution $T(f)$ defined in (8.1.23) coincides with the original element of $\mathscr{S}'(\mathbf{R}^n)$ given in Definition 8.1.8.

Remark 8.1.17. We show that the definition of $T(f)$ is independent of the choice of the function η. Indeed, if ζ is another function satisfying $0 \le \zeta \le 1$ that is also equal to 1 in a neighborhood of the support of φ, then $f(\eta-\zeta)$ and φ have disjoint supports, and by (8.1.7) we have the absolutely convergent integral realization

$$\big\langle T(f(\eta-\zeta)),\varphi\big\rangle = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y) f(y)(\eta-\zeta)(y)\,dy\,\varphi(x)\,dx.$$

It follows that the expression in (8.1.23) coincides with the corresponding expression obtained when η is replaced by ζ.

Next, if f is a Schwartz function, then both ηf and $(1-\eta)f$ are Schwartz functions; by the linearity of T one has $\langle T(f),\varphi\rangle = \langle T(\eta f),\varphi\rangle + \langle T((1-\eta)f),\varphi\rangle$, and by (8.1.7) the second expression can be written as the double absolutely convergent integral in (8.1.23), since φ and $(1-\eta)f$ have disjoint supports. Thus the distribution $T(f)$ defined in (8.1.23) coincides with the original element of $\mathscr{S}'(\mathbf{R}^n)$ given in Definition 8.1.8.

Remark 8.1.18. When T has a bounded extension that maps L^2 to itself, we may define $T(f)$ for all $f \in L^\infty(\mathbf{R}^n)$, not necessarily smooth. Simply observe that under this assumption, the expression $T(f\eta)$ is a well defined L^2 function and thus

$$\langle T(f\eta),\varphi\rangle = \int_{\mathbf{R}^n} T(f\eta)(x)\varphi(x)\,dx$$

is given by an absolutely convergent integral for all $\varphi \in \mathscr{D}_0$.

Finally, observe that although $\langle T(f),\varphi\rangle$ is defined for f in L^∞ and φ in $\mathscr{D}_0$, this definition is valid for all square integrable functions φ with compact support and integral zero; indeed, the smoothness of φ was never an issue in the definition of $\langle T(f),\varphi\rangle$.

In summary, if T is a Calderón–Zygmund operator and f lies in $L^\infty(\mathbf{R}^n)$, then $T(f)$ has a well defined *action* $\langle T(f),\varphi\rangle$ on square integrable functions φ with compact support and integral zero. This action satisfies

$$\big|\langle T(f),\varphi\rangle\big| \le \big\|T(f\eta)\big\|_{L^2}\big\|\varphi\big\|_{L^2} + C_{n,\delta}A\,\big\|\varphi\big\|_{L^1}\big\|f\big\|_{L^\infty} < \infty. \tag{8.1.24}$$

In the next section we show that in this case, $T(f)$ is in fact an element of *BMO*.

Exercises

8.1.1. Suppose that K is a function defined away from the diagonal on $\mathbf{R}^n \times \mathbf{R}^n$ that satisfies for some $\delta > 0$ the condition

$$|K(x,y) - K(x',y)| \le A' \frac{|x-x'|^\delta}{|x-y|^{n+\delta}}$$

whenever $|x-x'| \le \frac{1}{2}|x-y|$. Prove that K satisfies (8.1.2) with constant $A = (\frac{5}{2})^{n+\delta}A'$. Obtain an analogous statement for condition (8.1.3).

8.1.2. Prove that there does not exist a tempered distribution W on $\mathbf{R}^{2n}$ that extends the function $|x-y|^{-n}$ defined on $\mathbf{R}^{2n} \setminus \{(x,x) :\ x \in \mathbf{R}^n\}$.
$\big[$*Hint:* Apply such a distribution to a positive smooth bump that does not vanish at the origin.$\big]$

8.1.3. Let $\varphi(x)$ be a smooth radial function that is equal to 1 when $|x| \ge 1$ and vanishes when $|x| \le \frac{1}{2}$. Prove that if K lies in $SK(\delta,A)$, then all the smooth truncations $K_\varphi^{(\varepsilon)}(x,y) = K(x,y)\varphi(\frac{x-y}{\varepsilon})$ lie in $SK(\delta,cA)$ for some $c>0$ independent of $\varepsilon>0$.

8.1.4. Suppose that A is a Lipschitz map from $\mathbf{R}^n$ to $\mathbf{R}^m$. This means that there exists a constant L such that $|A(x)-A(y)| \le L|x-y|$ for all $x,y \in \mathbf{R}^n$. Suppose that F is a $\mathscr{C}^\infty$ odd function defined on $\mathbf{R}^m$. Show that the kernel

$$K(x,y) = \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right)$$

is in $SK(1,C)$ for some $C>0$.

8.1.5. Extend the result of Proposition 8.1.11 to the case that the space L^2 is replaced by L^q for some $1<q<\infty$.

8.1.6. Observe that for an operator T as in Definition 8.1.16, the condition $T(1)=0$ is equivalent to the statement that for all φ smooth with compact support and integral zero we have $\int_{\mathbf{R}^n} T^t(\varphi)(x)\,dx = 0$. A similar statement holds for T^t.

8.1.7. Suppose that $K(x,y)$ is continuous, bounded, and nonnegative on $\mathbf{R}^n \times \mathbf{R}^n$ and satisfies $\int_{\mathbf{R}^n} K(x,y)\,dy = 1$ for all $x \in \mathbf{R}^n$. Define a linear operator T by setting $T(f)(x) = \int_{\mathbf{R}^n} K(x,y) f(y)\,dy$ for $f \in L^1(\mathbf{R}^n)$.
(a) Suppose that h is a continuous and integrable function on $\mathbf{R}^n$ that has a global minimum [i.e., there exists $x_0 \in \mathbf{R}^n$ such that $h(x_0) \le h(x)$ for all $x \in \mathbf{R}^n$]. If we have

$$T(h)(x) = h(x)$$

for all $x \in \mathbf{R}^n$, prove that h is a constant function.
(b) Show that T preserves the set of integrable functions that are bounded below by a fixed constant.
(c) Suppose that $T(T(f)) = f$ for some everywhere positive and continuous function f on $\mathbf{R}^n$. Show that $T(f) = f$.
[*Hint:* Part (c): Let $L(x,y)$ be the kernel of $T \circ T$. Show that

$$\int_{\mathbf{R}^n} L(x,y) \frac{f(y)}{f(x)} \frac{T(f)(y)}{f(y)}\,dy = \frac{T(f)(x)}{f(x)}$$

and conclude by part (a) that $\frac{T(f)(y)}{f(y)}$ is a constant.]

8.2 Consequences of L^2 Boundedness

Calderón–Zygmund singular integral operators admit L^2 bounded extensions. As in the case of convolution operators, L^2 boundedness has several consequences. In this section we are concerned with consequences of the L^2 boundedness of Calderón–Zygmund singular integral operators. Throughout the entire discussion, we assume that $K(x,y)$ is a kernel defined away from the diagonal in $\mathbf{R}^{2n}$ that satisfies the standard size and regularity conditions (8.1.1), (8.1.2), and (8.1.3). These conditions may be relaxed; see the exercises at the end of this section.

8.2.1 Weak Type $(1,1)$ and L^p Boundedness of Singular Integrals

We begin by proving that operators in $CZO(\delta,A,B)$ are bounded from L^1 to weak L^1. This result is completely analogous to that in Theorem 4.3.3.

Theorem 8.2.1. *Assume that $K(x,y)$ is in $SK(\delta,A)$ and let T be an element of $CZO(\delta,A,B)$ associated with the kernel K. Then T has a bounded extension that maps $L^1(\mathbf{R}^n)$ to $L^{1,\infty}(\mathbf{R}^n)$ with norm*

$$\big\|T\big\|_{L^1\to L^{1,\infty}} \le C_n(A+B),$$

and also maps $L^p(\mathbf{R}^n)$ to itself for $1<p<\infty$ with norm

$$\big\|T\big\|_{L^p\to L^p} \le C_n \max(p,(p-1)^{-1})(A+B),$$

where C_n is a dimensional constant.

Proof. The proof of this theorem is a reprise of the argument of the proof of Theorem (4.3.3). Fix $\alpha>0$ and let f be in $L^1(\mathbf{R}^n)$. Since $T(f)$ may not be defined when f is a general integrable function, we take f to be a Schwartz class function. Once we obtain a weak type $(1,1)$ estimate for Schwartz functions, it is only a matter of density to extend it to all f in L^1.

We apply the Calderón–Zygmund decomposition to f at height $\gamma\alpha$, where γ is a positive constant to be chosen later. Write $f=g+b$, where $b=\sum_j b_j$ and conditions (1)–(6) of Theorem 4.3.1 are satisfied with the constant α replaced by $\gamma\alpha$. Since we are assuming that f is Schwartz function, it follows that each bad function b_j is bounded and compactly supported. Thus $T(b_j)$ is an L^2 function, and when x is not in the support of b_j we have the integral representation

$$T(b_j)(x) = \int_{Q_j} b_j(y)K(x,y)\,dy$$

in view of Proposition 8.1.9.

As usual, we denote by $\ell(Q)$ the side length of a cube Q. Let Q_j^* be the unique cube with sides parallel to the axes having the same center as Q_j and having side length

$$\ell(Q_j^*) = 2\sqrt{n}\,\ell(Q_j)\,.$$

We have

$$
\begin{aligned}
&|\{x\in\mathbf{R}^n : |T(f)(x)|>\alpha\}| \\
&\quad\le \Big|\Big\{x\in\mathbf{R}^n:\ |T(g)(x)|>\frac{\alpha}{2}\Big\}\Big| + \Big|\Big\{x\in\mathbf{R}^n:\ |T(b)(x)|>\frac{\alpha}{2}\Big\}\Big| \\
&\quad\le \frac{2^2}{\alpha^2}\big\|T(g)\big\|_{L^2}^2 + \Big|\bigcup_j Q_j^*\Big| + \Big|\Big\{x\notin \bigcup_j Q_j^*:\ |T(b)(x)|>\frac{\alpha}{2}\Big\}\Big| \\
&\quad\le \frac{2^2}{\alpha^2}B^2\big\|g\big\|_{L^2}^2 + \sum_j |Q_j^*| + \frac{2}{\alpha}\int_{(\bigcup_j Q_j^*)^c} |T(b)(x)|\,dx \\
&\quad\le \frac{2^2}{\alpha^2}2^nB^2(\gamma\alpha)\big\|f\big\|_{L^1} + (2\sqrt{n})^n\frac{\big\|f\big\|_{L^1}}{\gamma\alpha} + \frac{2}{\alpha}\sum_j\int_{(Q_j^*)^c}|T(b_j)(x)|\,dx \\
&\quad\le \left(\frac{(2^{n+1}B\gamma)^2}{2^n\gamma} + \frac{(2\sqrt{n})^n}{\gamma}\right)\frac{\big\|f\big\|_{L^1}}{\alpha} + \frac{2}{\alpha}\sum_j\int_{(Q_j^*)^c}|T(b_j)(x)|\,dx\,.
\end{aligned}
$$

It suffices to show that the last sum is bounded by some constant multiple of $\big\|f\big\|_{L^1}$. Let y_j be the center of the cube Q_j. For $x\in(Q_j^*)^c$, we have $|x-y_j|\ge\frac{1}{2}\ell(Q_j^*)=\sqrt{n}\,\ell(Q_j)$. But if $y\in Q_j$ we have $|y-y_j|\le\sqrt{n}\,\ell(Q_j)/2$; thus $|y-y_j|\le\frac{1}{2}|x-y_j|$, since the diameter of a cube is equal to $\sqrt{n}$ times its side length. We now estimate the last displayed sum as follows:

$$
\begin{aligned}
\sum_j\int_{(Q_j^*)^c}|T(b_j)(x)|\,dx &= \sum_j\int_{(Q_j^*)^c}\left|\int_{Q_j} b_j(y)K(x,y)\,dy\right|dx \\
&= \sum_j\int_{(Q_j^*)^c}\left|\int_{Q_j} b_j(y)\big(K(x,y)-K(x,y_j)\big)\,dy\right|dx \\
&\le \sum_j\int_{Q_j}|b_j(y)|\int_{(Q_j^*)^c}|K(x,y)-K(x,y_j)|\,dx\,dy \\
&\le \sum_j\int_{Q_j}|b_j(y)|\int_{|x-y_j|\ge 2|y-y_j|}|K(x,y)-K(x,y_j)|\,dx\,dy \\
&\le A_2\sum_j\int_{Q_j}|b_j(y)|\,dy \\
&= A_2\sum_j\big\|b_j\big\|_{L^1} \\
&\le A_2 2^{n+1}\big\|f\big\|_{L^1}\,.
\end{aligned}
$$

Combining the facts proved and choosing $\gamma=B^{-1}$, we deduce a weak type $(1,1)$ estimate for $T(f)$ when f is in the Schwartz class. We obtain that T has a bounded extension from L^1 to $L^{1,\infty}$ with bound at most $C_n(A+B)$. The L^p result for $1<p<2$ follows by interpolation and Exercise 1.3.2. The result for $2<p<\infty$ follows by duality; one uses here that the dual operator T^t has a kernel $K^t(x,y)=K(y,x)$ that satisfies the same estimates as K, and by the result just proved, it is also bounded on

L^p for $1 < p < 2$ with norm at most $C_n(A+B)$. Thus T must be bounded on L^p for $2 < p < \infty$ with norm at most a constant multiple of $A+B$. $\square$

Consequently, for operators T in $CZO(\delta,A,B)$ and L^p functions f, $1 \le p < \infty$, the expressions $T(f)$ make sense as L^p (or $L^{1,\infty}$ when $p=1$) functions. The following result addresses the question whether these functions can be expressed as integrals.

Proposition 8.2.2. *Let T be an operator in $CZO(\delta,A,B)$ associated with a kernel K. Then for $g \in L^p(\mathbf{R}^n)$, $1 \le p < \infty$, the following absolutely convergent integral representation is valid:*

$$T(g)(x) = \int_{\mathbf{R}^n} K(x,y)\, g(y)\, dy \tag{8.2.1}$$

for almost all $x \in \mathbf{R}^n \setminus \operatorname{supp} g$, provided that $\operatorname{supp} g \subsetneq \mathbf{R}^n$.

Proof. Set $g_k(x) = g(x)\chi_{|g(x)|\le k}\chi_{|x|\le k}$. These are L^p functions with compact support that is contained in the support of g. Also, the g_k converge to g in L^p as $k \to \infty$. In view of Proposition 8.1.9, for every k we have

$$T(g_k)(x) = \int_{\mathbf{R}^n} K(x,y)\, g_k(y)\, dy$$

for all $x \in \mathbf{R}^n \setminus \operatorname{supp} g$. Since T maps L^p to L^p (or to weak L^1 when $p=1$), it follows that $T(g_k)$ converges to $T(g)$ in weak L^p and hence in measure. By Proposition 1.1.9, a subsequence of $T(g_k)$ converges to $T(g)$ almost everywhere. On the other hand, for $x \in \mathbf{R}^n \setminus \operatorname{supp} g$ we have

$$\int_{\mathbf{R}^n} K(x,y)\, g_k(y)\, dy \to \int_{\mathbf{R}^n} K(x,y)\, g(y)\, dy$$

when $k \to \infty$, since the absolute value of the difference is bounded by $B\big\|g_k - g\big\|_{L^p}$, which tends to zero. The constant B is the $L^{p'}$ norm of the function $|x-y|^{-n-\delta}$ on the support of g; one has $|x-y| \ge c > 0$ for all y in the support of g and thus $B < \infty$. Therefore $T(g_k)(x)$ converges a.e. to both sides of the identity (8.2.1) for x not in the support of g. This concludes the proof of this identity. $\square$

8.2.2 Boundedness of Maximal Singular Integrals

We pose the question whether there is an analogous boundedness result to Theorem 8.2.1 concerning the maximal singular integral operator $T^{(*)}$. We note that given f in $L^p(\mathbf{R}^n)$ for some $1 \le p < \infty$, the expression $T^{(*)}(f)(x)$ is well defined for all $x \in \mathbf{R}^n$. This is a simple consequence of estimate (8.1.1) and Hölder's inequality.

Theorem 8.2.3. *Let K be in $SK(\delta,A)$ and T in $CZO(\delta,A,B)$ be associated with K. Let $r \in (0,1)$. Then there is a constant $C(n,r)$ such that*

$$|T^{(*)}(f)(x)| \le C(n,r)\Big[M(|T(f)|^r)(x)^{\frac{1}{r}} + (A+B)M(f)(x)\Big] \tag{8.2.2}$$

is valid for all functions in $\bigcup_{1\le p<\infty} L^p(\mathbf{R}^n)$. *Also, there exist dimensional constants* C_n, C_n' *such that*

$$\big\|T^{(*)}(f)\big\|_{L^{1,\infty}(\mathbf{R}^n)} \le C_n'(A+B)\big\|f\big\|_{L^1(\mathbf{R}^n)}, \tag{8.2.3}$$

$$\big\|T^{(*)}(f)\big\|_{L^p(\mathbf{R}^n)} \le C_n(A+B)\max(p,(p-1)^{-1})\big\|f\big\|_{L^p(\mathbf{R}^n)}, \tag{8.2.4}$$

for all $1 \le p < \infty$ *and all* f *in* $L^p(\mathbf{R}^n)$.

Estimate (8.2.2) is referred to as *Cotlar's inequality*.

Proof. We fix r so that $0 < r < 1$ and $f \in L^p(\mathbf{R}^n)$ for some p satisfying $1 \le p < \infty$. To prove (8.2.2), we also fix $\varepsilon > 0$ and we set $f_0^{\varepsilon,x} = f\chi_{B(x,\varepsilon)}$ and $f_\infty^{\varepsilon,x} = f\chi_{B(x,\varepsilon)^c}$. Since $x \notin \operatorname{supp} f_\infty^{\varepsilon,x}$ whenever $|x-y| \ge \varepsilon$, using Proposition 8.2.2 we can write

$$T(f_\infty^{\varepsilon,x})(x) = \int_{\mathbf{R}^n} K(x,y)\, f_\infty^{\varepsilon,x}(y)\,dy = \int_{|x-y|\ge\varepsilon} K(x,y)\,f(y)\,dy = T^{(\varepsilon)}(f)(x).$$

In view of (8.1.2), for $z \in B(x,\frac{\varepsilon}{2})$ we have $|z-x| \le \frac{1}{2}|x-y|$ whenever $|x-y| \ge \varepsilon$ and thus

$$\begin{aligned}
|T(f_\infty^{\varepsilon,x})(x) - T(f_\infty^{\varepsilon,x})(z)| &= \left|\int_{|x-y|\ge\varepsilon} \big(K(z,y)-K(x,y)\big) f(y)\,dy\right| \\
&\le |z-x|^\delta \int_{|x-y|\ge\varepsilon} \frac{A\,|f(y)|}{(|x-y|+|y-z|)^{n+\delta}}\,dy \\
&\le \Big(\frac{\varepsilon}{2}\Big)^\delta \int_{|x-y|\ge\varepsilon} \frac{A\,|f(y)|}{(|x-y|+\varepsilon/2)^{n+\delta}}\,dy \\
&\le C_{n,\delta}\,A\,M(f)(x),
\end{aligned}$$

where the last estimate is a consequence of Theorem 2.1.10. We conclude that for all $z \in B(x,\frac{\varepsilon}{2})$ we have

$$\begin{aligned}
|T^{(\varepsilon)}(f)(x)| &= |T(f_\infty^{\varepsilon,x})(x)| \\
&\le |T(f_\infty^{\varepsilon})(x) - T(f_\infty^{\varepsilon,x})(z)| + |T(f_\infty^{\varepsilon,x})(z)| \\
&\le C_{n,\delta}\,A\,M(f)(x) + |T(f_0^{\varepsilon,x})(z)| + |T(f)(z)|.
\end{aligned} \tag{8.2.5}$$

For $0 < r < 1$ it follows from (8.2.5) that for $z \in B(x,\frac{\varepsilon}{2})$ we have

$$|T^{(\varepsilon)}(f)(x)|^r \le C_{n,\delta}^r A^r M(f)(x)^r + |T(f_0^{\varepsilon,x})(z)|^r + |T(f)(z)|^r. \tag{8.2.6}$$

Integrating over $z \in B(x,\frac{\varepsilon}{2})$, dividing by $|B(x,\frac{\varepsilon}{2})|$, and raising to the power $\frac{1}{r}$, we obtain

$$|T^{(\varepsilon)}(f)(x)| \le 3^{\frac{1}{r}}\left[C_{n,\delta}AM(f)(x)+\left(\frac{1}{|B(x,\frac{\varepsilon}{2})|}\int_{B(x,\frac{\varepsilon}{2})}|T(f_0^{\varepsilon,x})(z)|^r dz\right)^{\frac{1}{r}} + M(|T(f)|^r)(x)^{\frac{1}{r}}\right].$$

Using Exercise 2.1.5, we estimate the middle term on the right-hand side of the preceding equation by

$$\left(\frac{1}{|B(x,\frac{\varepsilon}{2})|}\frac{\big\|T\big\|^r_{L^1\to L^{1,\infty}}}{1-r}|B(x,\tfrac{\varepsilon}{2})|^{1-r}\big\|f_0^{\varepsilon,x}\big\|^r_{L^1}\right)^{\frac{1}{r}} \le C_{n,r}(A+B)M(f)(x).$$

This proves (8.2.2).

We now use estimate (8.2.2) to show that T is L^p bounded and of weak type $(1,1)$. To obtain the weak type $(1,1)$ estimate for $T^{(*)}$ we need to use that the Hardy–Littlewood maximal operator maps $L^{p,\infty}$ to $L^{p,\infty}$ for all $1<p<\infty$. See Exercise 2.1.13. We also use the trivial fact that for all $0<p,q<\infty$ we have

$$\big\||f|^q\big\|_{L^{p,\infty}} = \big\|f\big\|^q_{L^{pq,\infty}}.$$

Take any $r<1$ in (8.2.2). Then we have

$$\begin{aligned}
\big\|M(|T(f)|^r)^{\frac{1}{r}}\big\|_{L^{1,\infty}} &= \big\|M(|T(f)|^r)\big\|^{\frac{1}{r}}_{L^{\frac{1}{r},\infty}} \\
&\le C_{n,r}\big\||T(f)|^r\big\|^{\frac{1}{r}}_{L^{\frac{1}{r},\infty}} \\
&= C_{n,r}\big\|T(f)\big\|_{L^{1,\infty}} \\
&\le \widetilde{C}_{n,r}(A+B)\big\|f\big\|_{L^1},
\end{aligned}$$

where we used the weak type $(1,1)$ bound for T in the last estimate.

To obtain the L^p boundedness of $T^{(*)}$ for $1<p<\infty$, we use the same argument as before. We fix $r=\frac{1}{2}$. Recall that the maximal function is bounded on L^{2p} with norm at most $3^{\frac{n}{2p}}\frac{2p}{2p-1}\le 2\cdot 3^{\frac{n}{2}}$ [see (2.1.5)]. We have

$$\begin{aligned}
\big\|M(|T(f)|^{\frac{1}{2}})^2\big\|_{L^p} &= \big\|M(|T(f)|^{\frac{1}{2}})\big\|^2_{L^{2p}} \\
&\le \big(3^{\frac{n}{2p}}\tfrac{2p}{2p-1}\big)^2\big\||T(f)|^{\frac{1}{2}}\big\|^2_{L^{2p}} \\
&\le 4\cdot 3^n\big\|T(f)\big\|_{L^p} \\
&\le C_n\max(\tfrac{1}{p-1},p)(A+B)\big\|f\big\|_{L^p},
\end{aligned}$$

where we used the L^p boundedness of T in the last estimate. □

We end this section with two corollaries, the first of which confirms a fact mentioned in Remark 8.1.12.

Corollary 8.2.4. *Let K be in $SK(\delta,A)$ and T in $CZO(\delta,A,B)$ be associated with K. Then there exists a dimensional constant C_n such that*

$$\sup_{\varepsilon>0} \big\|T^{(\varepsilon)}\big\|_{L^2\to L^2} \le C_n\big(A+\big\|T\big\|_{L^2\to L^2}\big)\,.$$

Corollary 8.2.5. *Let K be in $SK(\delta,A)$ and let $T=\lim_{\varepsilon_j\to 0} T^{(\varepsilon_j)}$ be an element of $CZSIO(\delta,A,B)$ associated with K. Then for $1\le p<\infty$ and all $f\in L^p(\mathbf{R}^n)$ we have that*

$$T^{(\varepsilon_j)}(f)\to T(f)$$

almost everywhere.

Proof. Using (8.1.1), (8.1.2), and (8.1.3), we see that the alleged convergence holds (everywhere) for smooth functions with compact support. The general case follows from Theorem 8.2.3 and Theorem 2.1.14. □

8.2.3 $H^1\to L^1$ and $L^\infty\to BMO$ Boundedness of Singular Integrals

Theorem 8.2.6. *Let T be an element of $CZO(\delta,A,B)$. Then T has an extension that maps $H^1(\mathbf{R}^n)$ to $L^1(\mathbf{R}^n)$. Precisely, there is a constant $C_{n,\delta}$ such that*

$$\big\|T\big\|_{H^1\to L^1} \le C_{n,\delta}\big(A+\big\|T\big\|_{L^2\to L^2}\big)\,.$$

Proof. The proof is analogous to that of Theorem 6.7.1. Let $B=\big\|T\big\|_{L^2\to L^2}$. We start by examining the action of T on L^2 atoms for H^1. Let $f=a$ be such an atom, supported in a cube Q. Let c_Q be the center of Q and let $Q^*=2\sqrt{n}\,Q$. We write

$$\int_{\mathbf{R}^n} |T(a)(x)|\,dx = \int_{Q^*} |T(a)(x)|\,dx + \int_{(Q^*)^c} |T(a)(x)|\,dx \tag{8.2.7}$$

and we estimate each term separately. We have

$$\begin{aligned}
\int_{Q^*} |T(a)(x)|\,dx &\le |Q^*|^{\frac12}\bigg(\int_{Q^*} |T(a)(x)|^2\,dx\bigg)^{\frac12}\\
&\le B|Q^*|^{\frac12}\bigg(\int_{Q} |a(x)|^2\,dx\bigg)^{\frac12}\\
&\le B|Q^*|^{\frac12}|Q|^{-\frac12}\\
&= C_nB\,,
\end{aligned}$$

where we used property (b) of atoms in Definition 6.6.8. Now observe that if $x\notin Q^*$ and $y\in Q$, then

$$|y-c_Q|\le \frac{1}{2}|x-c_Q|\,;$$

hence $x-y$ stays away from zero and $T(a)(x)$ can be expressed as a convergent integral by Proposition 8.2.2. We have

$$\begin{aligned}\int_{(Q^*)^c} |T(a)(x)|\,dx &= \int_{(Q^*)^c} \Big| \int_Q K(x,y)a(y)\,dy \Big|\,dx \\ &= \int_{(Q^*)^c} \Big| \int_Q \big(K(x,y)-K(x,c_Q)\big)a(y)\,dy \Big|\,dx \\ &\le \int_Q \int_{(Q^*)^c} \big|K(x,y)-K(x,c_Q)\big|\,dx\,|a(y)|\,dy \\ &\le \int_Q \int_{(Q^*)^c} \frac{A|y-c_Q|^\delta}{|x-c_Q|^{n+\delta}}\,dx\,|a(y)|\,dy \\ &\le C'_{n,\delta}A\int_Q |a(y)|\,dy \\ &\le C'_{n,\delta}A|Q|^{\frac12}\big\|a\big\|_{L^2} \\ &\le C'_{n,\delta}A|Q|^{\frac12}|Q|^{-\frac12} \\ &= C'_{n,\delta}A\,.\end{aligned}$$

Combining this calculation with the previous one and inserting the final conclusions in (8.2.7), we deduce that L^2 atoms for H^1 satisfy

$$\big\|T(a)\big\|_{L^1} \le C_{n,\delta}\,(A+B)\,. \tag{8.2.8}$$

To pass to general functions in H^1, we use Theorem 6.6.10 to 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

$$\big\|f\big\|_{H^1} \approx \sum_{j=1}^\infty |\lambda_j|\,. \tag{8.2.9}$$

Since T maps L^1 to weak L^1 by Theorem 8.2.1, $T(f)$ is already a well defined $L^{1,\infty}$ function. We plan to prove that

$$T(f) = \sum_{j=1}^\infty \lambda_j T(a_j) \qquad \text{a.e.} \tag{8.2.10}$$

Note that the series in (8.2.10) converges in L^1 and defines an integrable function almost everywhere. Once (8.2.10) is established, the required conclusion (6.7.5) follows easily by taking L^1 norms in (8.2.10) and using (8.2.8) and (8.2.9).

To prove (8.2.10), we use that T is of weak type $(1,1)$. For a given $\mu > 0$ we have

$$
\begin{aligned}
&\Big|\Big\{\Big|T(f)-\sum_{j=1}^{\infty}\lambda_j T(a_j)\Big|>\mu\Big\}\Big| \\
&\quad\le \Big|\Big\{\Big|T(f)-\sum_{j=1}^{N}\lambda_j T(a_j)\Big|>\mu/2\Big\}\Big|+\Big|\Big\{\Big|\sum_{j=N+1}^{\infty}\lambda_j T(a_j)\Big|>\mu/2\Big\}\Big| \\
&\quad\le \frac{2}{\mu}\big\|T\big\|_{L^1\to L^{1,\infty}}\Big\|f-\sum_{j=1}^{N}\lambda_j a_j\Big\|_{L^1}+\frac{2}{\mu}\Big\|\sum_{j=N+1}^{\infty}\lambda_j T(a_j)\Big\|_{L^1} \\
&\quad\le \frac{2}{\mu}\big\|T\big\|_{L^1\to L^{1,\infty}}\Big\|f-\sum_{j=1}^{N}\lambda_j a_j\Big\|_{H^1}+\frac{2}{\mu}C_{n,\delta}\,(A+B)\sum_{j=N+1}^{\infty}|\lambda_j|\,.
\end{aligned}
$$

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

$$
\Big|\Big\{\Big|T(f)-\sum_{j=1}^{\infty}\lambda_j T(a_j)\Big|>\mu\Big\}\Big|=0
$$

for all $\mu>0$, which implies (8.2.10). □

Theorem 8.2.7. *Let T be in $CZO(\delta,A,B)$. Then for any bounded function f, the distribution $T(f)$ can be identified with a BMO function that satisfies*

$$
\big\|T(f)\big\|_{BMO}\le C'_{n,\delta}(A+B)\big\|f\big\|_{L^\infty}\,, \tag{8.2.11}
$$

where $C_{n,\delta}$ is a constant.

Proof. Let $L^2_{0,c}$ be the space of all square integrable functions with compact support and integral zero on $\mathbf{R}^n$. This space is contained in $H^1(\mathbf{R}^n)$ (cf. Exercise 6.4.3) and contains the set of finite sums of L^2 atoms for H^1, which is dense in H^1 (cf. Exercise 6.6.5); thus $L^2_{0,c}$ is dense in H^1. Recall that for $f\in L^\infty$, $T(f)$ has a well defined action $\langle T(f),\varphi\rangle$ on functions φ in $L^2_{0,c}$ that satisfies (8.1.24).

Suppose we have proved the identity

$$
\big\langle T(f),\varphi\big\rangle=\int_{\mathbf{R}^n}T^t(\varphi)(x)f(x)\,dx\,, \tag{8.2.12}
$$

for all bounded functions f and all φ in $L^2_{0,c}$. Since such a φ is in H^1, Theorem 8.2.6 yields that $T^t(\varphi)$ is in L^1, and consequently, the integral in (8.2.12) converges absolutely. Assuming (8.2.12) and using Theorem 8.2.6 we obtain that

$$
\big|\big\langle T(f),\varphi\big\rangle\big|\le\big\|T^t(\varphi)\big\|_{L^1}\big\|f\big\|_{L^\infty}\le C_{n,\delta}(A+B)\,\big\|\varphi\big\|_{H^1}\,\big\|f\big\|_{L^\infty}\,.
$$

We conclude that $L(\varphi)=\langle T(f),\varphi\rangle$ is a bounded linear functional on $L^2_{0,c}$ with norm at most $C_{n,\delta}(A+B)\,\|f\|_{L^\infty}$. Obviously, L has a bounded extension on H^1 with the same norm. By Theorem 7.2.2 there exists a BMO function b_f that satisfies $\|b_f\|_{BMO}\le C'_n\|L\|_{H^1\to\mathbf{C}}$ such that the linear functional L has the form L_{b_f} (using the notation of Theorem 7.2.2). In other words, the distribution $T(f)$ can be identified with a BMO function that satisfies (8.2.11) with $C_{n,\delta} = C_n' C_{n,\delta}$, i.e.,

$$\big\|T(f)\big\|_{BMO} \le C_n' C_{n,\delta}(A+B)\,\big\|f\big\|_{L^\infty}.$$

We return to the proof of identity (8.2.12). Pick a smooth function with compact support η that satisfies $0 \le \eta \le 1$ and is equal to 1 in a neighborhood of the support of φ. We write the right-hand side of (8.2.12) as

$$\int_{\mathbf{R}^n} T^t(\varphi)\eta f\,dx + \int_{\mathbf{R}^n} T^t(\varphi)(1-\eta)f\,dx = \big\langle T(\eta f),\varphi\big\rangle + \int_{\mathbf{R}^n} T^t(\varphi)(1-\eta)f\,dx.$$

In view of Definition 8.1.16, to prove (8.2.12) it will suffice to show that

$$\int_{\mathbf{R}^n} T^t(\varphi)(1-\eta)f\,dx = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \big(K(x,y)-K(x_0,y)\big)\varphi(x)\,dx(1-\eta(y))f(y)\,dy,$$

where x_0 lies in the support of φ. But the inner integral above is absolutely convergent and equal to

$$\int_{\mathbf{R}^n} \big(K(x,y)-K(x_0,y)\big)\varphi(x)\,dx = \int_{\mathbf{R}^n} K^t(y,x)\varphi(x)\,dx = T^t(\varphi)(y),$$

since $y \notin \operatorname{supp}\varphi$, by Proposition 8.1.9. Thus (8.2.12) is valid. □

Exercises

8.2.1. Let $T:\mathscr{S}(\mathbf{R}^n)\to\mathscr{S}'(\mathbf{R}^n)$ be a continuous linear operator whose Schwartz kernel coincides with a function $K(x,y)$ on $\mathbf{R}^n\times\mathbf{R}^n$ minus its diagonal. Suppose that the function $K(x,y)$ satisfies

$$\sup_{R>0}\int_{R\le|x-y|\le 2R} |K(x,y)|\,dy \le A < \infty.$$

(a) Show that the previous condition is equivalent to

$$\sup_{R>0}\frac{1}{R}\int_{|x-y|\le R} |x-y|\,|K(x,y)|\,dy \le A' < \infty$$

by proving that $A' \le A \le 2A'$.
(b) For $\varepsilon>0$, let $T^{(\varepsilon)}$ be the truncated linear operators with kernels $K^{(\varepsilon)}(x,y) = K(x,y)\chi_{|x-y|>\varepsilon}$. Show that $T^{(\varepsilon)}(f)$ is well defined for Schwartz functions.
$\big[$*Hint:* Consider the annuli $\varepsilon 2^j \le |x| \le \varepsilon 2^{j+1}$ for $j\ge 0$.$\big]$

8.2.2. Let T be as in Exercise 8.2.1. Prove that the limit $T^{(\varepsilon)}(f)(x)$ exists for all f in the Schwartz class and for almost all $x\in\mathbf{R}^n$ as $\varepsilon\to 0$ if and only if the limit

$$\lim_{\varepsilon \to 0} \int_{\varepsilon < |x-y| < 1} K(x,y)\, dy$$

exists for almost all $x \in \mathbf{R}^n$.

8.2.3. Let $K(x,y)$ be a function defined away from the diagonal in $\mathbf{R}^{2n}$ that satisfies

$$\sup_{R>0} \int_{R \le |x-y| \le 2R} |K(x,y)|\, dy \le A < \infty$$

and also *Hörmander's condition*

$$\sup_{\substack{y,y' \in \mathbf{R}^n \\ y \neq y'}} \int_{|x-y| \ge 2|y-y'|} |K(x,y) - K(x,y')|\, dx \le A'' < \infty .$$

Show that all the truncations $K^{(\varepsilon)}(x,y)$ also satisfy Hörmander's condition uniformly in $\varepsilon > 0$ with a constant $A + A$".

8.2.4. Let T be as in Exercise 8.2.1 and assume that T maps $L^r(\mathbf{R}^n)$ to itself for some $1 < r \le \infty$.
(a) Assume that $K(x,y)$ satisfies Hörmander's condition, Then T has an extension that maps $L^1(\mathbf{R}^n)$ to $L^{1,\infty}(\mathbf{R}^n)$ with norm

$$\big\|T\big\|_{L^1 \to L^{1,\infty}} \le C_n(A+B),$$

and therefore T maps $L^p(\mathbf{R}^n)$ to itself for $1 < p < r$ with norm

$$\big\|T\big\|_{L^p \to L^p} \le C_n(p-1)^{-1}(A+B),$$

where C_n is a dimensional constant.
(b) Assuming that $K^t(x,y) = K(y,x)$ satisfies Hörmander's condition, prove that T maps $L^p(\mathbf{R}^n)$ to itself for $r < p < \infty$ with norm

$$\big\|T\big\|_{L^p \to L^p} \le C_n\, p\,(A+B),$$

where C_n is independent of p.

8.2.5. Show that estimate (8.2.2) also holds when $r = 1$.
$\big[$*Hint:* Estimate (8.2.6) holds when $r = 1$. For fixed $\varepsilon > 0$, take $0 < b < |T^{(\varepsilon)}(f)(x)|$ and define $B_1^\varepsilon(x) = B(x, \frac{\varepsilon}{2}) \cap \{|T(f)| > \frac{b}{3}\}$, $B_2^\varepsilon(x) = B(x, \frac{\varepsilon}{2}) \cap \{|T(f_0^{\varepsilon,x})| > \frac{b}{3}\}$, and $B_3^\varepsilon(x) = B(x, \frac{\varepsilon}{2})$ if $C_{n,\delta} M(f)(x) > \frac{b}{3}$ and empty otherwise. Then $|B(x, \frac{\varepsilon}{2})| \le |B_1^\varepsilon(x)| + |B_2^\varepsilon(x)| + |B_3^\varepsilon(x)|$. Use the weak type $(1,1)$ property of T to show that $b \le C(n)\big(M(|T(f)|)(x) + M(f)(x)\big)$, and take the supremum over all $b < |T^{(\varepsilon)}(f)(x)|$.$\big]$

8.2.6. Prove that if $|f| \log^+ |f|$ is integrable over a ball, then $T^{(*)}(f)$ is integrable over the same ball.
$\big[$*Hint:* Use the behavior of the norm of $T^{(*)}$ on L^p as $p \to 1$ and use Exercise 1.3.7.$\big]$

8.3 The $T(1)$ Theorem

We now turn to one of the main results of this chapter, the so-called $T(1)$ theorem. This theorem gives necessary and sufficient conditions for linear operators T with standard kernels to be bounded on $L^2(\mathbf{R}^n)$. In this section we obtain several such equivalent conditions. The name of theorem $T(1)$ is due to the fact that one of the conditions that we derive is expressed in terms of properties of the distribution $T(1)$, which was introduced in Definition 8.1.16.

8.3.1 Preliminaries and Statement of the Theorem

We begin with some preliminary facts and definitions.

Definition 8.3.1. A *normalized bump* is a smooth function φ supported in the ball $B(0,10)$ that satisfies

$$|(\partial_x^\alpha \varphi)(x)| \le 1$$

for all multi-indices $|\alpha| \le 2\left[\frac{n}{2}\right]+2$, where $[x]$ denotes here the integer part of x.

Observe that every smooth function supported inside the ball $B(0,10)$ is a constant multiple of a normalized bump. Also note that if a normalized bump is supported in a compact subset of $B(0,10)$, then small translations of it are also normalized bumps.

Given a function f on $\mathbf{R}^n$, $R>0$, and $x_0 \in \mathbf{R}^n$, we use the notation f_R to denote the function $f_R(x) = R^{-n} f(R^{-1}x)$ and $\tau^{x_0}(f)$ to denote the function $\tau^{x_0}(f)(x) = f(x-x_0)$. Thus

$$\tau^{x_0}(f_R)(y) = f_R(y-x_0) = R^{-n} f\big(R^{-1}(y-x_0)\big).$$

Set $N = \left[\frac{n}{2}\right]+1$. Using that all derivatives up to order $2N$ of normalized bumps are bounded by 1, we easily deduce that for all $x_0 \in \mathbf{R}^n$, all $R>0$, and all normalized bumps φ we have the estimate

$$\begin{aligned} R^n \int_{\mathbf{R}^n} & \big|\widehat{\tau^{x_0}(\varphi_R)}(\xi)\big|\, d\xi \\ &= \int_{\mathbf{R}^n} \big|\widehat{\varphi}(\xi)\big|\, d\xi \\ &= \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} \varphi(y) e^{-2\pi i y\cdot\xi}\, dy \right| d\xi \\ &= \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} (I-\Delta)^N(\varphi)(y) e^{-2\pi i y\cdot\xi}\, dy \right| \frac{d\xi}{(1+4\pi^2|\xi|^2)^N} \\ &\le C_n, \end{aligned} \tag{8.3.1}$$

since $|(\partial_x^\alpha \varphi)(x)| \le 1$ for all multi-indices α with $|\alpha| \le \left[\frac{n}{2}\right]+1$, and C_n is independent of the bump φ. Here $I-\Delta$ denotes the operator

$$(I-\Delta)(\varphi) = \varphi + \sum_{j=1}^{n} \frac{\partial^2 \varphi}{\partial x_j^2}.$$

Definition 8.3.2. We say that a continuous linear operator

$$T:\ \mathscr{S}(\mathbf{R}^n) \to \mathscr{S}'(\mathbf{R}^n)$$

satisfies the *weak boundedness property* (WBP) if there is a constant C such that for all f and g normalized bumps and for all $x_0 \in \mathbf{R}^n$ and $R > 0$ we have

$$|\langle T(\tau^{x_0}(f_R)), \tau^{x_0}(g_R)\rangle| \le CR^{-n}. \tag{8.3.2}$$

The smallest constant C in (8.3.2) is denoted by $\|T\|_{WB}$.

Note that $\|\tau^{x_0}(f_R)\|_{L^2} = \|f\|_{L^2} R^{-n/2}$ and thus if T has a bounded extension from $L^2(\mathbf{R}^n)$ to itself, then T satisfies the weak boundedness property with bound

$$\|T\|_{WB} \le 10^n v_n \|T\|_{L^2\to L^2},$$

where v_n is the volume of the unit ball in $\mathbf{R}^n$.

We now state one of the main theorems in this chapter.

Theorem 8.3.3. *Let T be a continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ whose Schwartz kernel coincides with a function K on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x):\ x \in \mathbf{R}^n\}$ that satisfies (8.1.1), (8.1.2), and (8.1.3) for some $0<\delta, A<\infty$. Let $K^{(\varepsilon)}$ and $T^{(\varepsilon)}$ be the usual truncated kernel and operator for $\varepsilon > 0$. Assume that there exists a sequence $\varepsilon_j \downarrow 0$ such that for all $f,g \in \mathscr{S}(\mathbf{R}^n)$ we have*

$$\langle T^{(\varepsilon_j)}(f), g\rangle \to \langle T(f), g\rangle. \tag{8.3.3}$$

Consider the assertions:

(i) The following statement is valid:

$$B_1 = \sup_B \sup_{\varepsilon>0} \left[\frac{\|T^{(\varepsilon)}(\chi_B)\|_{L^2}}{|B|^{\frac12}} + \frac{\|(T^{(\varepsilon)})^t(\chi_B)\|_{L^2}}{|B|^{\frac12}} \right] < \infty,$$

where the first supremum is taken over all balls B in $\mathbf{R}^n$.

(ii) The following statement is valid:

$$B_2 = \sup_{\varepsilon,N,x_0} \Bigg[\frac{1}{N^n} \int_{B(x_0,N)} \bigg| \int_{|x-y|<N} K^{(\varepsilon)}(x,y)\,dy \bigg|^2 dx + \frac{1}{N^n} \int_{B(x_0,N)} \bigg| \int_{|x-y|<N} K^{(\varepsilon)}(y,x)\,dy \bigg|^2 dx \Bigg]^{\frac12} < \infty,$$

where the supremum is taken over all $0 < \varepsilon < N < \infty$ *and all* $x_0 \in \mathbf{R}^n$.

(iii) The following statement is valid:

$$B_3 = \sup_{\varphi} \sup_{x_0 \in \mathbf{R}^n} \sup_{R>0} R^{\frac{n}{2}} \Big[\big\| T(\tau^{x_0}(\varphi_R)) \big\|_{L^2} + \big\| T^t(\tau^{x_0}(\varphi_R)) \big\|_{L^2} \Big] < \infty,$$

where the first supremum is taken over all normalized bumps φ.

(iv) The operator T *satisfies the weak boundedness property and the distributions* $T(1)$ *and* $T^t(1)$ *coincide with BMO functions, that is,*

$$B_4 = \big\| T(1) \big\|_{BMO} + \big\| T^t(1) \big\|_{BMO} + \big\| T \big\|_{WB} < \infty.$$

(v) For every $\xi \in \mathbf{R}^n$ *the distributions* $T(e^{2\pi i (\cdot)\cdot\xi})$ *and* $T^t(e^{2\pi i (\cdot)\cdot\xi})$ *coincide with BMO functions such that*

$$B_5 = \sup_{\xi \in \mathbf{R}^n} \big\| T(e^{2\pi i (\cdot)\cdot\xi}) \big\|_{BMO} + \sup_{\xi \in \mathbf{R}^n} \big\| T^t(e^{2\pi i (\cdot)\cdot\xi}) \big\|_{BMO} < \infty.$$

(vi) The following statement is valid:

$$B_6 = \sup_{\varphi} \sup_{x_0 \in \mathbf{R}^n} \sup_{R>0} R^n \Big[\big\| T(\tau^{x_0}(\varphi_R)) \big\|_{BMO} + \big\| T^t(\tau^{x_0}(\varphi_R)) \big\|_{BMO} \Big] < \infty,$$

where the first supremum is taken over all normalized bumps φ.

Then assertions (i)–(vi) *are all equivalent to each other and to the* L^2 *boundedness of* T*, and we have the following equivalence of the previous quantities:*

$$c_{n,\delta}(A + B_j) \le \big\| T \big\|_{L^2 \to L^2} \le C_{n,\delta}(A + B_j),$$

for all $j \in \{1,2,3,4,5,6\}$*, for some constants* $c_{n,\delta}, C_{n,\delta}$ *that depend only on the dimension* n *and on the parameter* $\delta > 0$.

Remark 8.3.4. Condition (8.3.3) says that the operator T is the weak limit of a sequence of its truncations. We already know that if T is bounded on L^2, then it must be equal to an operator that satisfies (8.3.3) plus a bounded function times the identity operator. (See Proposition 8.1.11.) Therefore, it is not a serious restriction to assume this. See Remark 8.3.6 for a version of Theorem 8.3.3 in which this assumption is not imposed. However, the reader should always keep in mind the following pathological situation: Let K be a function on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) : x \in \mathbf{R}^n\}$ that satisfies condition (ii) of the theorem. Then nothing prevents the Schwartz kernel W of T from having the form

$$W = K(x,y) + h(x)\delta_{x=y},$$

where $h(x)$ is an unbounded function and $\delta_{x=y}$ is Lebesgue measure on the subspace $x = y$. In this case, although the $T^{(\varepsilon)}$'s are uniformly bounded on L^2, T cannot be L^2 bounded, since h is not a bounded function.

Before we begin the lengthy proof of this theorem, we state a lemma that we need.

Lemma 8.3.5. *Under assumptions (8.1.1), (8.1.2), and (8.1.3), there is a constant* C_n *such that for all normalized bumps* φ *we have*

$$\sup_{x_0\in\mathbf{R}^n}\int_{|x-x_0|\geq 20R}\left|\int_{\mathbf{R}^n}K(x,y)\tau^{x_0}(\varphi_R)(y)\,dy\right|^2 dx \leq \frac{C_nA^2}{R^n}. \tag{8.3.4}$$

Proof. Note that the interior integral in (8.3.4) is absolutely convergent, since $\tau^{x_0}(\varphi_R)$ is supported in the ball $B(x_0,10R)$ and x lies in the complement of the double of this ball. To prove (8.3.4), simply observe that since $|K(x,y)|\leq A|x-y|^{-n}$, we have that

$$|T(\tau^{x_0}(\varphi_R))(x)|\leq\frac{C_nA}{|x-x_0|^n}$$

whenever $|x-x_0|\geq 20R$. The estimate follows easily. □

8.3.2 The Proof of Theorem 8.3.3

This subsection is dedicated to the proof of Theorem 8.3.3.

Proof. The proof is based on a series of steps. We begin by showing that condition (iii) implies condition (iv).

(iii) $\Longrightarrow$ (iv)

Fix a $\mathscr{C}_0^\infty$ function ϕ with $0\leq\phi\leq 1$, supported in the ball $B(0,4)$, and equal to 1 on the ball $B(0,2)$. We consider the functions $\phi(\cdot/R)$ that tend to 1 as $R\to\infty$ and we show that $T(1)$ is the weak limit of the functions $T(\phi(\cdot/R))$. This means that for all $g\in\mathscr{D}_0'$ (smooth functions with compact support and integral zero) one has

$$\big\langle T(\phi(\cdot/R)),g\big\rangle\to\big\langle T(1),g\big\rangle \tag{8.3.5}$$

as $R\to\infty$. To prove (8.3.5) we fix a $\mathscr{C}_0^\infty$ function η that is equal to one on the support of g. Then we write

$$\begin{aligned}\big\langle T(\phi(\cdot/R)),g\big\rangle &= \big\langle T(\eta\phi(\cdot/R)),g\big\rangle+\big\langle T((1-\eta)\phi(\cdot/R)),g\big\rangle\\ &= \big\langle T(\eta\phi(\cdot/R)),g\big\rangle\\ &\qquad+\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\phi(y/R)\,dy\,dx,\end{aligned}$$

where x_0 is a point in the support of g. There exists an $R_0>0$ such that for $R\geq R_0$, $\phi(\cdot/R)$ is equal to 1 on the support of η, and moreover the expressions

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\phi(y/R)\,dy\,dx$$

converge to

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\,dy\,dx$$

as $R\to\infty$ by the Lebesgue dominated convergence theorem. Using Definition 8.1.16, we obtain the validity of (8.3.5).

Next we observe that the functions $\phi(\cdot/R)$ are in L^2, since $\phi(x/R)=R^{-n}\phi_R(x)$, and by hypothesis (iii), ϕ_R are in L^2. We show that

$$\big\|T(\phi(\cdot/R))\big\|_{BMO}\le C_{n,\delta}(A+B_3) \tag{8.3.6}$$

uniformly in $R>0$. Once (8.3.6) is established, then the sequence $\{T(\phi(\cdot/j))\}_{j=1}^{\infty}$ lies in a multiple of the unit ball of $BMO=(H^1)^*$, and by the Banach–Alaoglou theorem, there is a subsequence of the positive integers R_j such that $T(\phi(\cdot/R_j))$ converges weakly to an element b in BMO. This means that

$$\big\langle T(\phi(\cdot/R_j)),g\big\rangle\to\big\langle b,g\big\rangle \tag{8.3.7}$$

as $j\to\infty$ for all $g\in\mathscr{D}_0$. Using (8.3.5), we conclude that $T(1)$ can be identified with the BMO function b, and as a consequence of (8.3.6) it satisfies

$$\big\|T(1)\big\|_{BMO}\le C_{n,\delta}(A+B_3)\,.$$

In a similar fashion, we identify $T^t(1)$ with a BMO function with norm satisfying

$$\big\|T^t(1)\big\|_{BMO}\le C_{n,\delta}(A+B_3)\,.$$

We return to the proof of (8.3.6). We fix a ball $B=B(x_0,r)$ with radius $r>0$ centered at $x_0\in\mathbf{R}^n$. If we had a constant c_B such that

$$\frac{1}{|B|}\int_B|T(\phi(\cdot/R))(x)-c_B|\,dx\le c_{n,\delta}B_3 \tag{8.3.8}$$

for all $R>0$, then property (3) in Proposition 7.1.2 (adapted to balls) would yield (8.3.6). Obviously, (8.3.8) is a consequence of the two estimates

$$\frac{1}{|B|}\int_B\big|T\big[\phi(\tfrac{\cdot-x_0}{r})\phi(\tfrac{\cdot}{R})\big](x)\big|\,dx\le c_nB_3\,, \tag{8.3.9}$$

$$\frac{1}{|B|}\int_B\Big|T\big[(1-\phi(\tfrac{\cdot-x_0}{r}))\phi(\tfrac{\cdot}{R})\big](x)-T\big[(1-\phi(\tfrac{\cdot-x_0}{r}))\phi(\tfrac{\cdot}{R})\big](x_0)\Big|\,dx\le\frac{c_n}{\delta}A\,. \tag{8.3.10}$$

We bound the double integral in (8.3.10) by

$$\frac{1}{|B|}\int_B\int_{|y-x_0|\ge 2r}|K(x,y)-K(x_0,y)|\,\phi(y/R)\,dy\,dx\,, \tag{8.3.11}$$

since $1-\phi((y-x_0)/r)=0$ when $|y-x_0|\le 2r$. Since $|x-x_0|\le r\le \frac{1}{2}|y-x_0|$, condition (8.1.2) gives that (8.3.11) holds with $c_n=\omega_{n-1}=|\mathbf{S}^{n-1}|$.

It remains to prove (8.3.9). It is easy to verify that there is a constant $C_0 = C_0(n,\phi)$ such that for $0<\varepsilon\le 1$ and for all $a\in\mathbf{R}^n$ the functions

$$C_0^{-1}\phi(\varepsilon(x+a))\phi(x), \qquad C_0^{-1}\phi(x)\phi(-a+\varepsilon x) \tag{8.3.12}$$

are normalized bumps. The important observation is that with $a=x_0/r$ we have

$$\phi(\tfrac{x}{R})\phi(\tfrac{x-x_0}{r}) = r^n\tau^{x_0}\Big[\Big(\phi\big(\tfrac{r}{R}(\cdot+a)\big)\phi(\cdot)\Big)_r\Big](x) \tag{8.3.13}$$

$$= R^n\Big(\phi(\cdot)\phi\big(-a+\tfrac{R}{r}(\cdot)\big)\Big)_R(x), \tag{8.3.14}$$

and thus in either case $r\le R$ or $R\le r$, one may express the product $\phi(\frac{x}{R})\phi(\frac{x-x_0}{r})$ as a multiple of a translation of an L^1-dilation of a normalized bump.

Let us suppose that $r\le R$. In view of (8.3.13) we write

$$T\big[\phi(\tfrac{\cdot-x_0}{r})\phi(\tfrac{\cdot}{R})\big](x) = C_0\, r^n T\big[\tau^{x_0}(\varphi_r)\big](x)$$

for some normalized bump φ. Using this fact and the Cauchy–Schwarz inequality, we estimate the expression on the left in (8.3.9) by

$$\frac{C_0\, r^{n/2}}{|B|^{\frac{1}{2}}} r^{n/2}\left(\int_B \big|T\big[\tau^{x_0}(\varphi_r)\big](x)\big|^2\,dx\right)^{\frac{1}{2}} \le \frac{C_0\, r^{n/2}}{|B|^{\frac{1}{2}}} B_3 = c_n B_3\,,$$

where the first inequality follows by applying hypothesis (iii).

We now consider the case $R\le r$. In view of (8.3.14) we write

$$T\big[\phi(\tfrac{\cdot-x_0}{r})\phi(\tfrac{\cdot}{R})\big](x) = C_0\, R^n T\big(\varphi_R\big)(x)$$

for some other normalized bump φ. Using this fact and the Cauchy–Schwarz inequality, we estimate the expression on the left in (8.3.9) by

$$\frac{C_0\, R^{n/2}}{|B|^{\frac{1}{2}}} R^{n/2}\left(\int_B |T(\zeta_R)(x)|^2\,dx\right)^{\frac{1}{2}} \le \frac{C_0\, R^{n/2}}{|B|^{\frac{1}{2}}} B_3 \le c_n B_3$$

by applying hypothesis (iii) and recalling that $R\le r$. This proves (8.3.9).

To finish the proof of (iv), we need to prove that T satisfies the weak boundedness property. But this is elementary, since for all normalized bumps φ and ψ and all $x\in\mathbf{R}^n$ and $R>0$ we have

$$\begin{aligned} \big|\big\langle T(\tau^x(\psi_R)),\tau^x(\varphi_R)\big\rangle\big| &\le \big\|T(\tau^x(\psi_R))\big\|_{L^2}\big\|\tau^x(\varphi_R)\big\|_{L^2} \\ &\le B_3 R^{-\frac{n}{2}}\big\|\tau^x(\varphi_R)\big\|_{L^2} \\ &\le C_n B_3 R^{-n}. \end{aligned}$$

This gives $\|T\|_{WB} \le C_n B_3$, which implies the estimate $B_4 \le C_{n,\delta}(A+B_3)$ and concludes the proof of the fact that condition (iii) implies (iv).

(iv) $\implies$ (L^2 boundedness of T)

We now assume condition (iv) and we present the most important step of the proof, establishing the fact that T has an extension that maps $L^2(\mathbf{R}^n)$ to itself. The assumption that the distributions $T(1)$ and $T^t(1)$ coincide with BMO functions leads to the construction of Carleson measures that provide the key tool in the boundedness of T.

We pick a smooth radial function Φ with compact support that is supported in the ball $B(0,\frac{1}{2})$ and that satisfies $\int_{\mathbf{R}^n} \Phi(x)\,dx = 1$. For $t>0$ we define $\Phi_t(x) = t^{-n}\Phi(\frac{x}{t})$. Since Φ is a radial function, the operator

$$P_t(f) = f * \Phi_t \tag{8.3.15}$$

is self-transpose. The operator P_t is a continuous analogue of $S_j = \sum_{k\le j}\Delta_k$, where the Δ_j's are the Littlewood–Paley operators.

We now fix a Schwartz function f whose Fourier transform is supported away from a neighborhood of the origin. We discuss an integral representation for $T(f)$. We begin with the facts, which can be found in Exercises 8.3.1 and 8.3.2, that

$$\begin{aligned} T(f) &= \lim_{s\to 0} P_s^2 T P_s^2(f)\,, \\ 0 &= \lim_{s\to\infty} P_s^2 T P_s^2(f)\,, \end{aligned}$$

where the limits are interpreted in the topology of $\mathscr{S}'(\mathbf{R}^n)$. Thus, with the use of the fundamental theorem of calculus and the product rule, we are able to write

$$\begin{aligned} T(f) &= \lim_{s\to 0} P_s^2 T P_s^2(f) - \lim_{s\to\infty} P_s^2 T P_s^2(f) \\ &= -\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\frac{1}{\varepsilon}} s\frac{d}{ds}\big(P_s^2 T P_s^2\big)(f)\,\frac{ds}{s} \\ &= -\lim_{\varepsilon\to 0}\int_{\varepsilon}^{\frac{1}{\varepsilon}} \left[s\left(\frac{d}{ds}P_s^2\right) T P_s^2(f) + P_s^2\left(Ts\frac{d}{ds}P_s^2\right)(f)\right]\frac{ds}{s}\,. \end{aligned} \tag{8.3.16}$$

For a Schwartz function g we have

$$\begin{aligned} \left(s\frac{d}{ds}P_s^2(g)\right)^{\widehat{}}(\xi) &= \widehat{g}(\xi)\, s\frac{d}{ds}\widehat{\Phi}(s\xi)^2 \\ &= \widehat{g}(\xi)\,\widehat{\Phi}(s\xi)\,\big(2s\xi\cdot\nabla\widehat{\Phi}(s\xi)\big) \\ &= \widehat{g}(\xi)\sum_{k=1}^{n}\widehat{\Psi_k}(s\xi)\widehat{\Theta_k}(s\xi) \\ &= \sum_{k=1}^{n}\Big(\widetilde{Q}_{k,s}Q_{k,s}(g)\Big)^{\widehat{}}(\xi) = \sum_{k=1}^{n}\Big(Q_{k,s}\widetilde{Q}_{k,s}(g)\Big)^{\widehat{}}(\xi)\,, \end{aligned}$$

where for $1 \le k \le n$, $\widehat{\Psi_k}(\xi) = 2\xi_k \widehat{\Phi}(\xi)$, $\widehat{\Theta_k}(\xi) = \partial_k \widehat{\Phi}(\xi)$ and $Q_{k,s}$, $\widetilde{Q}_{k,s}$ are operators defined by

$$Q_{k,s}(g) = g * (\Psi_k)_s\,, \qquad \widetilde{Q}_{k,s}(g) = g * (\Theta_k)_s\,;$$

here $(\Theta_k)_s(x) = s^{-n}\Theta_k(s^{-1}x)$ and $(\Psi_k)_s$ are defined similarly. Observe that Ψ_k and Θ_k are smooth odd bumps supported in $B(0,\frac{1}{2})$ and have integral zero. Since Ψ_k and Θ_k are odd, they are anti-self-transpose, meaning that $(Q_{k,s})^t = -Q_{k,s}$ and $(\widetilde{Q}_{k,s})^t = -\widetilde{Q}_{k,s}$. We now write the expression in (8.3.16) as

$$-\lim_{\varepsilon\to 0}\sum_{k=1}^{n}\left[\int_{\varepsilon}^{\frac{1}{\varepsilon}} \widetilde{Q}_{k,s}Q_{k,s}TP_sP_s(f)\frac{ds}{s} + \int_{\varepsilon}^{\frac{1}{\varepsilon}} P_sP_sTQ_{k,s}\widetilde{Q}_{k,s}(f)\frac{ds}{s}\right], \tag{8.3.17}$$

where the limit converges in $\mathscr{S}'(\mathbf{R}^n)$. We set

$$T_{k,s} = Q_{k,s}TP_s\,,$$

and we observe that the operator $P_sTQ_{k,s}$ is equal to $-((T^t)_{k,s})^t$.

Recall the notation $\tau^x(h)(z) = h(z-x)$. In view of identity (2.3.21) and the convergence of the Riemann sums to the integral defining $f * \Phi_s$ in the topology of $\mathscr{S}$ (see the proof of Theorem 2.3.20), we deduce that the operator $T_{k,s}$ has kernel

$$K_{k,s}(x,y) = -\big\langle T(\tau^y(\Phi_s)), \tau^x((\Psi_k)_s)\big\rangle = -\big\langle T^t(\tau^x((\Psi_k)_s)), \tau^y(\Phi_s)\big\rangle\,. \tag{8.3.18}$$

Likewise, the operator $-(T^t)^t_{k,s}$ has kernel

$$\big\langle T^t(\tau^x(\Phi_s)), \tau^y((\Psi_k)_s)\big\rangle = \big\langle T(\tau^y((\Psi_k)_s)), \tau^x(\Phi_s)\big\rangle\,.$$

For $1 \le k \le n$ we need the following facts regarding the kernels of these operators:

$$\big|\big\langle T(\tau^x((\Psi_k)_s)), \tau^y(\Phi_s)\big\rangle\big| \le C_{n,\delta}\big(\big\|T\big\|_{WB} + A\big)\, p_s(x-y)\,, \tag{8.3.19}$$

$$\big|\big\langle T^t(\tau^x((\Psi_k)_s)), \tau^y(\Phi_s)\big\rangle\big| \le C_{n,\delta}\big(\big\|T\big\|_{WB} + A\big)\, p_s(x-y)\,, \tag{8.3.20}$$

where

$$p_t(u) = \frac{1}{t^n}\frac{1}{(1+|\frac{u}{t}|)^{n+\delta}}$$

is the L^1 dilation of the function $p(u) = (1+|u|)^{-n-\delta}$.

To prove (8.3.20), we consider the following two cases: If $|x-y| \le 5s$, then the weak boundedness property gives

$$\big|\big\langle T(\tau^y(\Phi_s)), \tau^x((\Psi_k)_s)\big\rangle\big| = \big|\big\langle T(\tau^x((\tau^{\frac{y-x}{s}}(\Phi))_s)), \tau^x((\Psi_k)_s)\big\rangle\big| \le \frac{C_n\big\|T\big\|_{WB}}{s^n}\,,$$

since both Ψ_k and $\tau^{\frac{y-x}{s}}(\Phi)$ are multiples of normalized bumps. Notice here that both of these functions are supported in $B(0,10)$, since $\frac{1}{s}|x-y| \le 5$. This estimate proves (8.3.20) when $|x-y| \le 5s$.

We now turn to the case $|x-y| \ge 5s$. Then the functions $\tau^y(\Phi_s)$ and $\tau^x((\Psi_k)_s)$ have disjoint supports and so we have the integral representation

$$\big\langle T^t(\tau^x((\Psi_k)_s)), \tau^y(\Phi_s)\big\rangle = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_s(v-y)K(u,v)(\Psi_k)_s(u-x)\,du\,dv\,.$$

Using that Ψ_k has mean value zero, we can write the previous expression as

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_s(v-y)\big(K(u,v)-K(x,v)\big)(\Psi_k)_s(u-x)\,du\,dv\,.$$

We observe that $|u-x| \le s$ and $|v-y| \le s$ in the preceding double integral. Since $|x-y| \ge 5s$, this makes $|u-v| \ge |x-y|-2s \ge 3s$, which implies that $|u-x| \le \frac{1}{2}|u-v|$. Using (8.1.2), we obtain

$$|K(u,v)-K(x,v)| \le \frac{A|x-u|^\delta}{(|u-v|+|x-v|)^{n+\delta}} \le C_{n,\delta}A\frac{s^\delta}{|x-y|^{n+\delta}}\,,$$

where we used the fact that $|u-v| \approx |x-y|$. Inserting this estimate in the double integral, we obtain (8.3.20). Estimate (8.3.19) is proved similarly.

At this point we drop the dependence of $Q_{k,s}$ and $\widetilde{Q}_{k,s}$ on the index k, since we can concentrate on one term of the sum in (8.3.17). We have managed to express $T(f)$ as a finite sum of operators of the form

$$\int_0^\infty \widetilde{Q}_s T_s P_s(f)\,\frac{ds}{s} \tag{8.3.21}$$

and of the form

$$\int_0^\infty P_s T_s \widetilde{Q}_s(f)\,\frac{ds}{s}\,, \tag{8.3.22}$$

where the preceding integrals converge in $\mathscr{S}'(\mathbf{R}^n)$ and the T_s's have kernels $K_s(x,y)$, which are pointwise dominated by a constant multiple of

$$(A+B_4)p_s(x-y)\,.$$

It suffices to obtain L^2 bounds for an operator of the form (8.3.21) with constant at most a multiple of $A+B_4$. Then by duality the same estimate also holds for the operators of the form (8.3.22). We make one more observation. Using (8.3.18) (recall that we have dropped the indices k), we obtain

$$T_s(1)(x) = \int_{\mathbf{R}^n} K_s(x,y)\,dy = -\big\langle T^t(\tau^x(\Psi_s)),1\big\rangle = -(\Psi_s * T(1))(x)\,, \tag{8.3.23}$$

where all integrals converge absolutely.

We can therefore concentrate on the L^2 boundedness of the operator in (8.3.21). We pair this operator with a Schwartz function g and we use the convergence of the integral in $\mathscr{S}'$ and the property $(\widetilde{Q}_s)^t = -\widetilde{Q}_s$ to obtain

$$\left\langle \int_0^\infty \widetilde{Q}_s T_s P_s(f)\,\frac{ds}{s}, g\right\rangle = \int_0^\infty \left\langle \widetilde{Q}_s T_s P_s(f), g\right\rangle \frac{ds}{s} = -\int_0^\infty \left\langle T_s P_s(f), \widetilde{Q}_s(g)\right\rangle \frac{ds}{s}.$$

The intuition here is as follows: T_s is an averaging operator at scale s and $P_s(f)$ is essentially constant on that scale. Therefore, the expression $T_s P_s(f)$ must look like $T_s(1)P_s(f)$. To be precise, we introduce this term and try to estimate the error that occurs. We have

$$T_s P_s(f) = T_s(1)P_s(f) + \big[T_s P_s(f) - T_s(1)P_s(f)\big]. \tag{8.3.24}$$

We estimate the terms that arise from this splitting. Recalling (8.3.23), we write

$$\left|\int_0^\infty \Big\langle (\Psi_s * T(1))P_s(f), \widetilde{Q}_s(g)\Big\rangle \frac{ds}{s}\right| \tag{8.3.25}$$

$$\le \left(\int_0^\infty \big\|P_s(f)\,(\Psi_s * T(1))\big\|_{L^2}^2 \frac{ds}{s}\right)^{\frac{1}{2}} \left(\int_0^\infty \big\|\widetilde{Q}_s(g)\big\|_{L^2}^2 \frac{ds}{s}\right)^{\frac{1}{2}}$$

$$= \left\|\left(\int_0^\infty \big|P_s(f)\,(\Psi_s * T(1))\big|^2 \frac{ds}{s}\right)^{\frac{1}{2}}\right\|_{L^2} \left\|\left(\int_0^\infty \big|\widetilde{Q}_s(g)\big|^2 \frac{ds}{s}\right)^{\frac{1}{2}}\right\|_{L^2}. \tag{8.3.26}$$

Since $T(1)$ is a BMO function, $|(\Psi_s * T(1))(x)|^2 dx\frac{ds}{s}$ is a Carleson measure on $\mathbf{R}^{n+1}_+$. Using Theorem 7.3.8 and the Littlewood–Paley theorem (Exercise 5.1.4), we obtain that (8.3.26) is controlled by

$$C_n \big\|T(1)\big\|_{BMO} \big\|f\big\|_{L^2} \big\|g\big\|_{L^2} \le C_n B_4 \big\|f\big\|_{L^2} \big\|g\big\|_{L^2}.$$

This gives the sought estimate for the first term in (8.3.24). For the second term in (8.3.24) we have

$$\left|\int_0^\infty \int_{\mathbf{R}^n} \widetilde{Q}_s(g)(x)\big[T_s P_s(f) - T_s(1)\,P_s(f)\big](x)\,dx\frac{ds}{s}\right|$$

$$\le \left(\int_0^\infty \int_{\mathbf{R}^n} |\widetilde{Q}_s(g)(x)|^2\,dx\frac{ds}{s}\right)^{\frac{1}{2}} \left(\int_0^\infty \int_{\mathbf{R}^n} |(T_s P_s(f) - T_s(1)P_s(f))(x)|^2\,dx\frac{ds}{s}\right)^{\frac{1}{2}}$$

$$\le C_n \big\|g\big\|_{L^2} \left(\int_0^\infty \int_{\mathbf{R}^n} \left|\int_{\mathbf{R}^n} K_s(x,y)[P_s(f)(y) - P_s(f)(x)]\,dy\right|^2 dx\frac{ds}{s}\right)^{\frac{1}{2}}$$

$$\le C_n(A+B_4)\big\|g\big\|_{L^2} \left(\int_0^\infty \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} p_s(x-y)\big|P_s(f)(y) - P_s(f)(x)\big|^2\,dy\,dx\,\frac{ds}{s}\right)^{\frac{1}{2}},$$

where in the last estimate we used the fact that the measure $p_t(x-y)\,dy$ is a multiple of a probability measure. It suffices to estimate the last displayed square root. Changing variables $u = x-y$ and applying Plancherel's theorem, we express this square root as

$$\begin{aligned}
&\left(\int_0^\infty\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} p_s(u)\big|P_s(f)(y)-P_s(f)(y+u)\big|^2\,du\,dy\,\frac{ds}{s}\right)^{\frac12}\\
&\quad=\left(\int_0^\infty\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} p_s(u)\big|\widehat{\Phi}(s\xi)-\widehat{\Phi}(s\xi)e^{2\pi i u\cdot\xi}\big|^2\big|\widehat{f}(\xi)\big|^2\,du\,d\xi\,\frac{ds}{s}\right)^{\frac12}\\
&\quad\le\left(\int_0^\infty\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} p_s(u)\big|\widehat{\Phi}(s\xi)\big|^2 4\pi^{\frac{\delta}{2}}|u|^{\frac{\delta}{2}}|\xi|^{\frac{\delta}{2}}\big|\widehat{f}(\xi)\big|^2\,du\,d\xi\,\frac{ds}{s}\right)^{\frac12}\\
&\quad=2\pi^{\frac{\delta}{4}}\left(\int_{\mathbf{R}^n}\int_0^\infty\left(\int_{\mathbf{R}^n} p_s(u)\big|\tfrac{u}{s}\big|^{\frac{\delta}{2}}du\right)\big|\widehat{\Phi}(s\xi)\big|^2|s\xi|^{\frac{\delta}{2}}\frac{ds}{s}\,\big|\widehat{f}(\xi)\big|^2\,d\xi\right)^{\frac12},
\end{aligned}$$

and we claim that this last expression is bounded by $C_{n,\delta}\big\|f\big\|_{L^2}$. Indeed, we first bound the quantity $\int_{\mathbf{R}^n} p_s(u)\big|\frac{u}{s}\big|^{\delta/2}du$ by a constant, and then we use the estimate

$$\int_0^\infty\big|\widehat{\Phi}(s\xi)\big|^2|s\xi|^{\frac{\delta}{2}}\frac{ds}{s}=\int_0^\infty\big|\widehat{\Phi}(se_1)\big|^2 s^{\frac{\delta}{2}}\frac{ds}{s}\le C'_{n,\delta}<\infty$$

and Plancherel's theorem to obtain the claim. [Here $e_1=(1,0,\dots,0)$.] Taking g to be an arbitrary Schwartz function with L^2 norm at most 1 and using duality, we deduce the estimate $\big\|T(f)\big\|_{L^2}\le C_{n,\delta}(A+B_4)\big\|f\big\|_{L^2}$ for all Schwartz functions f whose Fourier transform does not contain a neighborhood of the origin. Such functions are dense in $L^2(\mathbf{R}^n)$ (cf. Exercise 5.2.9) and thus T admits an extension on L^2 that satisfies $\big\|T\big\|_{L^2\to L^2}\le C_{n,\delta}(A+B_4)$.

<u>(L^2 boundedness of T) $\implies$ (v)</u>

If T has an extension that maps L^2 to itself, then by Theorem 8.2.7 we have

$$B_5\le C_{n,\delta}\big(A+\big\|T\big\|_{L^2\to L^2}\big)<\infty.$$

Thus the boundedness of T on L^2 implies condition (v).

<u>(v) $\implies$ (vi)</u>

At a formal level the proof of this fact is clear, since we can write a normalized bump as the inverse Fourier transform of its Fourier transform and interchange the integrations with the action of T to obtain

$$T(\tau^{x_0}(\varphi_R))=\int_{\mathbf{R}^n}\widehat{\tau^{x_0}(\varphi_R)}(\xi)\,T(e^{2\pi i\xi\cdot(\cdot)})\,d\xi. \tag{8.3.27}$$

The conclusion follows by taking BMO norms. To make identity (8.3.27) precise we provide the following argument.

Let us fix a normalized bump φ and a smooth and compactly supported function g with mean value zero. We pick a smooth function η with compact support that is equal to 1 on the double of a ball containing the support of g and vanishes off the triple of that ball. Define $\eta_k(\xi)=\eta(\xi/k)$ and note that η_k tends pointwise to 1 as

$k \to \infty$. Observe that $\eta_k \tau^{x_0}(\varphi_R)$ converges to $\tau^{x_0}(\varphi_R)$ in $\mathscr{S}(\mathbf{R}^n)$ as $k \to \infty$, and by the continuity of T we obtain

$$\lim_{k\to\infty} \big\langle T(\eta_k \tau^{x_0}(\varphi_R)), g\big\rangle = \big\langle T(\tau^{x_0}(\varphi_R)), g\big\rangle.$$

The continuity and linearity of T also allow us to write

$$\big\langle T(\tau^{x_0}(\varphi_R)), g\big\rangle = \lim_{k\to\infty} \int_{\mathbf{R}^n} \widehat{\tau^{x_0}(\varphi_R)}(\xi)\, \big\langle T\big(\eta_k e^{2\pi i \xi\cdot(\,\cdot\,)}\big), g\big\rangle\, d\xi. \tag{8.3.28}$$

Let W be the Schwartz kernel of T. By (8.1.5) we have

$$\big\langle T(\eta_k e^{2\pi i \xi\cdot(\,\cdot\,)}), g\big\rangle = \big\langle W, g \otimes \eta_k e^{2\pi i \xi\cdot(\,\cdot\,)}\big\rangle. \tag{8.3.29}$$

Using (8.1.6), we obtain that the expression in (8.3.29) is controlled by a finite sum of L^∞ norms of derivatives of the function

$$g(x)\,\eta_k(y) e^{2\pi i \xi\cdot y}$$

on some compact set (that depends on g). Then for some $M > 0$ and some constant $C(g)$ depending on g, we have that this sum of L^∞ norms of derivatives is controlled by

$$C(g)\,(1+|\xi|)^M$$

uniformly in $k \ge 1$. Since $\widehat{\tau^{x_0}(\varphi_R)}$ is integrable, the Lebesgue dominated convergence theorem allows us to pass the limit inside the integrals in (8.3.28) to obtain

$$\big\langle T(\tau^{x_0}(\varphi_R)), g\big\rangle = \int_{\mathbf{R}^n} \widehat{\tau^{x_0}(\varphi_R)}(\xi)\, \big\langle T\big(e^{2\pi i \xi\cdot(\,\cdot\,)}\big), g\big\rangle\, d\xi.$$

We now use assumption (v). The distributions $T\big(e^{2\pi i \xi\cdot(\,\cdot\,)}\big)$ coincide with BMO functions whose norm is at most B_5. It follows that

$$\begin{aligned} \big|\big\langle T(\tau^{x_0}(\varphi_R)), g\big\rangle\big| &\le \big\|\widehat{\tau^{x_0}(\varphi_R)}\big\|_{L^1} \sup_{\xi\in\mathbf{R}^n} \big\|T\big(e^{2\pi i \xi\cdot(\,\cdot\,)}\big)\big\|_{BMO} \big\|g\big\|_{H^1} \\ &\le C_n B_5 R^{-n} \big\|g\big\|_{H^1}, \end{aligned} \tag{8.3.30}$$

where the constant C_n is independent of the normalized bump φ in view of (8.3.1). It follows from (8.3.30) that

$$g \mapsto \big\langle T(\tau^{x_0}(\varphi_R)), g\big\rangle$$

is a bounded linear functional on BMO with norm at most a multiple of $B_5 R^{-n}$. It follows from Theorem 7.2.2 that $T(\tau^{x_0}(\varphi_R))$ coincides with a BMO function that satisfies

$$R^n \big\|T(\tau^{x_0}(\varphi_R))\big\|_{BMO} \le C_n B_5.$$

The same argument is valid for T^t, and this shows that

$$B_6 \le C_{n,\delta}(A+B_5)$$

and concludes the proof that (v) implies (vi).

(vi) $\Longrightarrow$ (iii)

We fix $x_0 \in \mathbf{R}^n$ and $R>0$. Pick z_0 in $\mathbf{R}^n$ such that $|x_0-z_0| = 40R$. Then if $|y-x_0| \le 10R$ and $|x-z_0| \le 20R$ we have

$$\begin{aligned} 10R &\le |z_0-x_0| - |x-z_0| - |y-x_0| \\ &\le |x-y| \\ &\le |x-z_0| + |z_0-x_0| + |x_0-y| \le 70R. \end{aligned}$$

From this it follows that when $|x-z_0| \le 20R$ we have

$$\left| \int_{|y-x_0|\le 10R} K(x,y)\tau^{x_0}(\varphi_R)(y)\,dy \right| \le \int_{10R\le|x-y|\le 70R} |K(x,y)|\,\frac{dy}{R^n} \le \frac{C_{n,\delta}A}{R^n}$$

and thus

$$\Big| \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}(\varphi_R)) \Big| \le \frac{C_{n,\delta}A}{R^n}, \tag{8.3.31}$$

where $\mathrm{Avg}_B\, g$ denotes the average of g over B. Because of assumption (vi), the BMO norm of the function $T(\tau^{x_0}(\varphi_R))$ is bounded by a multiple of B_6R^{-n}, a fact used in the following sequence of implications. We have

$$\begin{aligned} &\big\|T(\tau^{x_0}(\varphi_R))\big\|_{L^2(B(x_0,20R))} \\ &\quad\le \Big\| T(\tau^{x_0}(\varphi_R)) - \operatorname*{Avg}_{B(x_0,20R)} T(\tau^{x_0}(\varphi_R)) \Big\|_{L^2(B(x_0,20R))} \\ &\qquad\quad + v_n^{\frac12}(20R)^{\frac n2} \Big| \operatorname*{Avg}_{B(x_0,20R)} T(\tau^{x_0}(\varphi_R)) - \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}(\varphi_R)) \Big| \\ &\qquad\quad + v_n^{\frac12}(20R)^{\frac n2} \Big| \operatorname*{Avg}_{B(z_0,20R)} T(\tau^{x_0}(\varphi_R)) \Big| \\ &\quad\le C_{n,\delta}\Big(R^{\frac n2} \big\|T(\tau^{x_0}(\varphi_R))\big\|_{BMO} + R^{\frac n2} \big\|T(\tau^{x_0}(\varphi_R))\big\|_{BMO} + R^{-\frac n2}A \Big) \\ &\quad\le C_{n,\delta}R^{-\frac n2}\big(B_6+A\big), \end{aligned}$$

where we used (8.3.31) and Exercise 7.1.6. Now we have that

$$\big\|T(\tau^{x_0}(\varphi_R))\big\|_{L^2(B(x_0,20R)^c)} \le C_{n,\delta}AR^{-\frac n2}$$

in view of Lemma 8.3.5. Since the same computations apply to T^t, it follows that

$$R^{\frac n2}\big(\big\|T(\tau^{x_0}(\varphi_R))\big\|_{L^2} + \big\|T^t(\tau^{x_0}(\varphi_R))\big\|_{L^2} \big) \le C_{n,\delta}(A+B_6), \tag{8.3.32}$$

which proves that $B_3 \le C_{n,\delta}(A+B_6)$ and hence (iii). This concludes the proof of the fact that (vi) implies (iii)

We have now completed the proof of the following equivalence of statements:

$$\big(L^2 \text{ boundedness of } T\big) \iff \text{(iii)} \iff \text{(iv)} \iff \text{(v)} \iff \text{(vi)}. \tag{8.3.33}$$

<u>(i) $\iff$ (ii)</u>

We show that the quantities $A+B_1$ and $A+B_2$ are controlled by constant multiples of each other. Let us set

$$I_{\varepsilon,N}(x) = \int_{\varepsilon<|x-y|<N} K(x,y)\,dy \qquad \text{and} \qquad I^t_{\varepsilon,N}(x) = \int_{\varepsilon<|x-y|<N} K^t(x,y)\,dy\,.$$

We work with a ball $B(x_0,N)$. Observe that

$$\begin{aligned} I_{\varepsilon,N}(x) - T^{(\varepsilon)}(\chi_{B(x_0,N)})(x) &= \int_{\varepsilon<|x-y|<N} K(x,y)\,dy - \int_{\substack{\varepsilon<|x-y| \\ |x_0-y|<N}} K(x,y)\,dy \\ &= -\int_{S_{\varepsilon,N}(x,x_0)} K(x,y)\,dy, \end{aligned} \tag{8.3.34}$$

where $S_{\varepsilon,N}(x,x_0)$ is the set of all $y \in \mathbf{R}^n$ that satisfy $\varepsilon < |x-y|$ and $|x_0-y| < N$ but do not satisfy $\varepsilon < |x-y| < N$. But observe that when $|x_0-x| < N$, then

$$S_{\varepsilon,N}(x,x_0) \subseteq \{y \in \mathbf{R}^n :\ N \le |x-y| < 2N\}. \tag{8.3.35}$$

Using (8.3.34), (8.3.35), and (8.1.1), we obtain

$$\big|I_{\varepsilon,N}(x) - T^{(\varepsilon)}(\chi_{B(x_0,N)})(x)\big| \le \int_{N\le|x-y|\le 2N} |K(x,y)|\,dy \le (\omega_{n-1}\log 2)A \tag{8.3.36}$$

whenever $|x_0-x| < N$. It follows that

$$\big\|I_{\varepsilon,N} - T^{(\varepsilon)}(\chi_{B(x_0,N)})\big\|_{L^2(B(x_0,N))} \le C_n A N^{\frac{n}{2}}\,,$$

and similarly, it follows that

$$\big\|I^t_{\varepsilon,N} - (T^{(\varepsilon)})^t(\chi_{B(x_0,N)})\big\|_{L^2(B(x_0,N))} \le C_n A N^{\frac{n}{2}}\,.$$

These two estimates easily imply the equivalence of conditions (i) and (ii).

We now consider the following condition analogous to (iii):

$$\text{(iii)}' \quad B_3' = \sup_{\varphi} \sup_{x_0 \in \mathbf{R}^n} \sup_{\substack{\varepsilon>0 \\ R>0}} R^{\frac{n}{2}} \Big[\big\| T^{(\varepsilon)}(\tau^{x_0}(\varphi_R)) \big\|_{L^2} + \big\| (T^{(\varepsilon)})^t (\tau^{x_0}(\varphi_R)) \big\|_{L^2} \Big] < \infty,$$

where the first supremum is taken over all normalized bumps φ. We continue by showing that this condition is a consequence of (ii).

<u>(ii) $\implies$ (iii)$'$</u>

More precisely, we prove that $B_3' \le C_{n,\delta}(A+B_2)$. To prove (iii)$'$, fix a normalized bump φ, a point $x_0 \in \mathbf{R}^n$, and $R > 0$. Also fix $x \in \mathbf{R}^n$ with $|x-x_0| \le 20R$. Then we have

$$T^{(\varepsilon)}(\tau^{x_0}(\varphi_R))(x) = \int_{\varepsilon<|x-y|\le 30R} K^{(\varepsilon)}(x,y)\tau^{x_0}(\varphi_R)(y)\,dx = U_1(x) + U_2(x),$$

where

$$\begin{aligned} U_1(x) &= \int_{\varepsilon<|x-y|\le 30R} K(x,y)\big(\tau^{x_0}(\varphi_R)(y) - \tau^{x_0}(\varphi_R)(x)\big)\,dy, \\ U_2(x) &= \tau^{x_0}(\varphi_R)(x) \int_{\varepsilon<|x-y|\le 30R} K(x,y)\,dy. \end{aligned}$$

But we have that $|\tau^{x_0}(\varphi_R)(y) - \tau^{x_0}(\varphi_R)(x)| \le C_n R^{-1-n}|x-y|$; thus we obtain

$$|U_1(x)| \le C_n A R^{-n}$$

on $B(x_0,20R)$; hence $\big\|U_1\big\|_{L^2(B(x_0,20R))} \le C_n A R^{-\frac{n}{2}}$. Condition (ii) gives that

$$\big\|U_2\big\|_{L^2(B(x_0,20R))} \le R^{-n} \big\|I_{\varepsilon,30R}\big\|_{L^2(B(x_0,30R))} \le B_2 (30R)^{\frac{n}{2}} R^{-n}.$$

Combining these two, we obtain

$$\big\|T^{(\varepsilon)}(\tau^{x_0}(\varphi_R))\big\|_{L^2(B(x_0,20R))} \le C_n(A+B_2)R^{-\frac{n}{2}} \tag{8.3.37}$$

and likewise for $(T^{(\varepsilon)})^t$. It follows from Lemma 8.3.5 that

$$\big\|T^{(\varepsilon)}(\tau^{x_0}(\varphi_R))\big\|_{L^2(B(x_0,20R)^c)} \le C_{n,\delta} A R^{-\frac{n}{2}},$$

which combined with (8.3.37) gives condition (iii)$'$ with constant

$$B_3' \le C_{n,\delta}(A+B_2).$$

This concludes the proof that condition (ii) implies (iii)$'$.

<u>(iii)$'$ $\implies$ [$T^{(\varepsilon)} : L^2 \to L^2$ uniformly in $\varepsilon > 0$]</u>

For $\varepsilon > 0$ we introduce the smooth truncations $T_\zeta^{(\varepsilon)}$ of T by setting

$$T_\zeta^{(\varepsilon)}(f)(x) = \int_{\mathbf{R}^n} K(x,y)\zeta(\tfrac{x-y}{\varepsilon})\, f(y)\, dy,$$

where $\zeta(x)$ is a smooth function that is equal to 1 for $|x| \geq 1$ and vanishes for $|x| \leq \frac{1}{2}$. We observe that

$$\big|T_\zeta^{(\varepsilon)}(f) - T^{(\varepsilon)}(f)\big| \leq C_n A M(f); \tag{8.3.38}$$

thus the uniform boundedness of $T^{(\varepsilon)}$ on L^2 is equivalent to the uniform boundedness of $T_\zeta^{(\varepsilon)}$. In view of Exercise 8.1.3, the kernels of the operators $T_\zeta^{(\varepsilon)}$ lie in $SK(\delta, cA)$ uniformly in $\varepsilon > 0$ (for some constant c). Moreover, because of (8.3.38), we see that the operators $T_\zeta^{(\varepsilon)}$ satisfy (iii)$'$ with constant $C_n A + B_3'$. The point to be noted here is that condition (iii) for T (with constant B_3) is identical to condition (iii)$'$ for the operators $T_\zeta^{(\varepsilon)}$ uniformly in $\varepsilon > 0$ (with constant $C_n A + B_3'$).

A careful examination of the proof of the implications

$$\text{(iii)} \implies \text{(iv)} \implies (L^2 \text{ boundedness of } T)$$

reveals that all the estimates obtained depend only on the constants B_3, B_4, and A, but not on the specific operator T. Therefore, these estimates are valid for the operators $T_\zeta^{(\varepsilon)}$ that satisfy condition (iii)$'$. This gives the uniform boundedness of the $T_\zeta^{(\varepsilon)}$ on $L^2(\mathbf{R}^n)$ with bounds at most a constant multiple of $A + B_3'$. The same conclusion also holds for the operators $T^{(\varepsilon)}$.

$[T^{(\varepsilon)} :\ L^2 \to L^2$ uniformly in $\varepsilon > 0] \implies$ (i)

This implication holds trivially.

We have now established the following equivalence of statements:

$$\text{(i)} \iff \text{(ii)} \iff \text{(iii)}' \iff [T^{(\varepsilon)} :\ L^2 \to L^2 \text{ uniformly in } \varepsilon > 0] \tag{8.3.39}$$

(iii) $\iff$ (iii)$'$

Finally, we link the sets of equivalent conditions (8.3.33) and (8.3.39). We first observe that (iii)$'$ implies (iii). Indeed, using duality and (8.3.3), we obtain

$$\begin{aligned}
\big\|T(\tau^{x_0}(\varphi_R))\big\|_{L^2} &= \sup_{\substack{h\in\mathscr{S}\\ \|h\|_{L^2}\leq 1}} \left| \int_{\mathbf{R}^n} T(\tau^{x_0}(\varphi_R))(x)\, h(x)\, dx \right| \\
&\leq \sup_{\substack{h\in\mathscr{S}\\ \|h\|_{L^2}\leq 1}} \limsup_{j\to\infty} \left| \int_{\mathbf{R}^n} T^{(\varepsilon_j)}(\tau^{x_0}(\varphi_R))(x)\, h(x)\, dx \right| \\
&\leq B_3' R^{-\frac{n}{2}},
\end{aligned}$$

which gives $B_3 \le B_3'$. Thus under assumption (8.3.3), (ii) implies (iii) and as we have shown, (iii) implies the boundedness of T on L^2. But in view of Corollary 8.2.4, the boundedness of T on L^2 implies the boundedness of $T^{(\varepsilon)}$ on L^2 uniformly in $\varepsilon > 0$, which implies (iii)$'$.

This completes the proof of the equivalence of the six statements (i)–(vi) in such a way that

$$\big\|T\big\|_{L^2\to L^2} \approx (A+B_j)$$

for all $j \in \{1,2,3,4,5,6\}$. The proof of the theorem is now complete. □

Remark 8.3.6. Suppose that condition (8.3.3) is removed from the hypothesis of Theorem 8.3.3. Then the given proof of Theorem 8.3.3 actually shows that (i) and (ii) are equivalent to each other and to the statement that the $T^{(\varepsilon)}$'s have bounded extensions on $L^2(\mathbf{R}^n)$ that satisfy

$$\sup_{\varepsilon>0}\big\|T^{(\varepsilon)}\big\|_{L^2\to L^2} < \infty .$$

Also, without hypothesis (8.3.3), conditions (iii), (iv), (v), and (vi) are equivalent to each other and to the statement that T has an extension that maps $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$.

8.3.3 An Application

We end this section with one application of the $T(1)$ theorem. We begin with the following observation.

Corollary 8.3.7. *Let K be a standard kernel that is antisymmetric, i.e., it satisfies $K(x,y) = -K(y,x)$ for all $x \neq y$. Then a linear continuous operator T associated with K is L^2 bounded if and only if $T(1)$ is in BMO.*

Proof. In view of Exercise 8.3.3, T automatically satisfies the weak boundedness property. Moreover, $T^t = -T$. Therefore, the three conditions of Theorem 8.3.3 (iv) reduce to the single condition $T(1) \in BMO$. □

Example 8.3.8. Let us recall the kernels K_m of Example 8.1.7. These arise in the expansion of the kernel in Example 8.1.6 in geometric series

$$\frac{1}{x-y+i(A(x)-A(y))} = \frac{1}{x-y}\sum_{m=0}^{\infty}\left(i\frac{A(x)-A(y)}{x-y}\right)^m \tag{8.3.40}$$

when $L = \sup_{x\neq y}\frac{|A(x)-A(y)|}{|x-y|} < 1$. The operator with kernel $(i\pi)^{-1}K_m(x,y)$, i.e.,

$$\mathscr{C}_m(f)(x) = \frac{1}{\pi i}\lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon}\left(\frac{A(x)-A(y)}{x-y}\right)^m\frac{1}{x-y}f(y)\,dy, \tag{8.3.41}$$

is called the mth *Calderón commutator*. We use the $T(1)$ theorem to show that the operators $\mathscr{C}_m$ are L^2 bounded.

We show that there exists a constant $R > 0$ such that for all $m \geq 0$ we have

$$\big\|\mathscr{C}_m\big\|_{L^2\to L^2} \leq R^m L^m\,. \tag{8.3.42}$$

We prove (8.3.42) by induction. We note that (8.3.42) is trivially true when $m = 0$, since $\mathscr{C}_0 = -iH$, where H is the Hilbert transform.

Assume that (8.3.42) holds for a certain m. We show its validity for $m+1$. Recall that K_m is a kernel in $SK(1,16(2m+1)L^m)$ by the discussion in Example 8.1.7. We need the following estimate proved in Theorem 8.2.7:

$$\big\|\mathscr{C}_m\big\|_{L^\infty\to BMO} \leq C_2\big[16(2m+1)L^m + \big\|\mathscr{C}_m\big\|_{L^2\to L^2}\big]\,, \tag{8.3.43}$$

which holds for some absolute constant C_2.

We start with the following consequence of Theorem 8.3.3:

$$\big\|\mathscr{C}_{m+1}\big\|_{L^2\to L^2} \leq C_1\big[\big\|\mathscr{C}_{m+1}(1)\big\|_{BMO} + \big\|(C_{m+1})^t(1)\big\|_{BMO} + \big\|\mathscr{C}_{m+1}\big\|_{WB}\big]\,, \tag{8.3.44}$$

valid for some absolute constant C_1. The key observation is that

$$\mathscr{C}_{m+1}(1) = \mathscr{C}_m(A')\,, \tag{8.3.45}$$

for which we refer to Exercise 8.3.4. Here A' denotes the derivative of A, which exists almost everywhere, since Lipschitz functions are differentiable almost everywhere. Note that the kernel of $\mathscr{C}_m$ is antisymmetric; consequently, $(\mathscr{C}_m)^t = -\mathscr{C}_m$ and Exercise 8.3.3 gives that $\big\|\mathscr{C}_m\big\|_{WB} \leq C_3\,16\,(2m+1)\,L^m$ for some other absolute constant C_3. Using all these facts we deduce from (8.3.44) that

$$\big\|\mathscr{C}_{m+1}\big\|_{L^2\to L^2} \leq C_1\big[2\big\|\mathscr{C}_m(A')\big\|_{BMO} + C_3\,16\,(2m+3)\,L^{m+1}\big]\,.$$

Using (8.3.43) and the fact that $\big\|A'\big\|_{L^\infty} \leq L$ we obtain that

$$\big\|\mathscr{C}_{m+1}\big\|_{L^2\to L^2} \leq C_1\Big[2C_2L\big\{16\,(2m+1)L^m + \big\|\mathscr{C}_m\big\|_{L^2\to L^2}\big\} + C_3\,16\,(2m+3)\,L^{m+1}\Big]\,.$$

Combining this estimate with the induction hypothesis (8.3.42), we obtain

$$\big\|\mathscr{C}_{m+1}(1)\big\|_{BMO} \leq R^{m+1}L^{m+1},$$

provided that R is chosen so that

$$\begin{aligned} R^{m+1} &> 96C_1C_2(2m+1)\,,\\ R &> 6C_1C_2\,,\\ R^{m+1} &> 48C_1C_3(2m+3) \end{aligned}$$

for all $m \geq 0$. Such an R exists independent of m. This completes the proof of (8.3.42) by induction.

Exercises

8.3.1. Let T be a continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ and let f be in $\mathscr{S}(\mathbf{R}^n)$. Let P_t be as in (8.3.15).
(a) Show that $P_t(f)$ converges to f in $\mathscr{S}(\mathbf{R}^n)$ as $t \to 0$.
(b) Conclude that $TP_t(f) \to T(f)$ in $\mathscr{S}'(\mathbf{R}^n)$ as $t \to 0$.
(c) Conclude that $P_tTP_t(f) \to T(f)$ in $\mathscr{S}'(\mathbf{R}^n)$ as $t \to 0$.
(d) Observe that (a)–(c) are also valid if P_t is replaced by P_t^2.
$\big[$*Hint:* Part (a): Use that $g_k \to g$ in $\mathscr{S}$ if and only if $\widehat{g_k} \to \widehat{g}$ in $\mathscr{S}$.$\big]$

8.3.2. Let T and P_t be as in Exercise 8.3.1 and let f be a Schwartz function whose Fourier transform vanishes in a neighborhood of the origin.
(a) Show that $P_t(f)$ converges to 0 in $\mathscr{S}(\mathbf{R}^n)$ as $t \to \infty$.
(b) Conclude that $TP_t(f) \to 0$ in $\mathscr{S}'(\mathbf{R}^n)$ as $t \to \infty$.
(c) Conclude that $P_tTP_t(f) \to 0$ in $\mathscr{S}'(\mathbf{R}^n)$ as $t \to \infty$.
(d) Observe that (a)–(c) are also valid if P_t is replaced by P_t^2.
$\big[$*Hint:* Part (a): Use the hint in Exercise 8.3.1 and the observation that $|\widehat{\Phi}(t\xi)\widehat{f}(\xi)| \le C(1+tc_0)^{-1}|\widehat{f}(\xi)|$ if $\widehat{f}$ is supported outside the ball $B(0,c_0)$. Part (c): Pair with a Schwartz function g and use part (a) and the fact that all Schwartz seminorms of $P_t(g)$ are bounded uniformly in $t > 0$. To prove the latter you may need that all Schwartz seminorms of $P_t(g)$ are bounded uniformly in $t > 0$ if and only if all Schwartz seminorms of $\widehat{P_t(g)}$ are bounded uniformly in $t > 0$.$\big]$

8.3.3. (a) Prove that every linear operator T from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ associated with an antisymmetric kernel in $SK(\delta,A)$ satisfies the weak boundedness property. Precisely, for some dimensional constant C_n we have

$$\big\|T\big\|_{WB} \le C_n A\,.$$

(b) Conclude that for some $c < \infty$, the Calderón commutators satisfy

$$\big\|\mathscr{C}_m\big\|_{WB} \le c\,16(2m+1)L^m\,.$$

$\big[$*Hint:* Write $\big\langle T(\tau^{x_0}(f_R)), \tau^{x_0}(g_R)\big\rangle$ as

$$\frac{1}{2}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} K(x,y)\big(\tau^{x_0}(f_R)(y)\tau^{x_0}(g_R)(x) - \tau^{x_0}(f_R)(x)\tau^{x_0}(g_R)(y)\big)\,dy\,dx$$

and use the mean value theorem.$\big]$

8.3.4. Prove identity (8.3.45). This identity is obvious by a formal integration by parts, but to prove it properly, one should interpret things in the sense of distributions.

8.3.5. Suppose that a standard kernel $K(x,y)$ has the form $k(x-y)$ for some function k on $\mathbf{R}^n \setminus \{0\}$. Suppose that k extends to a tempered distribution on $\mathbf{R}^n$ whose Fourier transform is a bounded function. Let T be a continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$

to $\mathscr{S}'(\mathbf{R}^n)$ associated with K.
(a) Identify the functions $T(e^{2\pi i\xi\cdot()})$ and $T^t(e^{2\pi i\xi\cdot()})$ and restrict to $\xi = 0$ to obtain $T(1)$ and $T^t(1)$.
(b) Use Theorem 8.3.3 to obtain the L^2 boundedness of T.
(c) What are $H(1)$ and $H^t(1)$ equal to when H is the Hilbert transform?

8.3.6. (*A. Calderón*) Let A be a Lipschitz function on $\mathbf{R}$. Use expansion (8.3.40) and estimate (8.3.42) to show that the operator

$$\mathscr{C}_A(f)(x) = \frac{1}{\pi i}\lim_{\varepsilon\to 0}\int_{|x-y|>\varepsilon}\frac{f(y)\,dy}{x-y+i(A(x)-A(y))}$$

is bounded on $L^2(\mathbf{R})$ when $\left\|A'\right\|_{L^\infty} < R^{-1}$, where R satisfies (8.3.43).

8.3.7. Prove that condition (i) of Theorem 8.3.3 is equivalent to the statement that

$$\sup_Q\sup_{\varepsilon>0}\left(\frac{\left\|T^{(\varepsilon)}(\chi_Q)\right\|_{L^2}}{|Q|^{\frac{1}{2}}}+\frac{\left\|(T^{(\varepsilon)})^t(\chi_Q)\right\|_{L^2}}{|Q|^{\frac{1}{2}}}\right) = B_1' < \infty,$$

where the first supremum is taken over all cubes Q in $\mathbf{R}^n$.
[*Hint:* You may repeat the argument in the equivalence (i) $\iff$ (ii) replacing the ball $B(x_0,N)$ by a cube centered at x_0 with side length N.]

8.4 Paraproducts

In this section we study a useful class of operators called paraproducts. Their name suggests they are related to products; in fact, they are "half products" in some sense that needs to be made precise. Paraproducts provide interesting examples of nonconvolution operators with standard kernels whose L^2 boundedness was discussed in the Section 8.3. They have found use in many situations, including a proof of the main implication in Theorem 8.3.3. This proof is discussed in the present section.

8.4.1 Introduction to Paraproducts

Throughout this section we fix a Schwartz radial function Ψ whose Fourier transform is supported in the annulus $\frac{1}{2} \le |\xi| \le 2$ and that satisfies

$$\sum_{j\in\mathbf{Z}}\widehat{\Psi}(2^{-j}\xi) = 1, \qquad \text{when} \quad \xi \in \mathbf{R}^n\setminus\{0\}. \tag{8.4.1}$$

Associated with this Ψ we define the Littlewood–Paley operator $\Delta_j(f) = f * \Psi_{2^{-j}}$, where $\Psi_t(x) = t^{-n}\Psi(t^{-1}x)$. Using (8.4.1), we easily obtain

$$\sum_{j\in\mathbf{Z}} \Delta_j = I, \tag{8.4.2}$$

where (8.4.2) is interpreted as an identity on Schwartz functions with mean value zero. See Exercise 8.4.1. Note that by construction, the function Ψ is radial and thus even. This makes the operator Δ_j equal to its transpose.

We now observe that in view of the properties of Ψ, the function

$$\xi \mapsto \sum_{j\le 0} \widehat{\Psi}(2^{-j}\xi) \tag{8.4.3}$$

is supported in $|\xi| \le 2$, and is equal to 1 when $0 < |\xi| \le \frac{1}{2}$. But $\widehat{\Psi}(0) = 0$, which implies that the function in (8.4.3) also vanishes at the origin. We can easily fix this discontinuity by introducing the Schwartz function whose Fourier transform is equal to

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

Definition 8.4.1. We define the *partial sum operator* S_j as

$$S_j = \sum_{k\le j} \Delta_k. \tag{8.4.4}$$

In view of the preceding discussion, S_j is given by convolution with $\Phi_{2^{-j}}$, that is,

$$S_j(f)(x) = (f * \Phi_{2^{-j}})(x), \tag{8.4.5}$$

and the expression in (8.4.5) is well defined for all f in $\bigcup_{1\le p\le\infty} L^p(\mathbf{R}^n)$. Since Φ is a radial function by construction, the operator S_j is self-transpose.

Similarly, $\Delta_j(g)$ is also well defined for all g in $\bigcup_{1\le p\le\infty} L^p(\mathbf{R}^n)$. Moreover, since Δ_j is given by convolution with a function with mean value zero, it also follows that $\Delta_j(b)$ is well defined when $b \in BMO(\mathbf{R}^n)$. See Exercise 8.4.2 for details.

Definition 8.4.2. Given a function g on $\mathbf{R}^n$, we define the *paraproduct operator* P_g as follows:

$$P_g(f) = \sum_{j\in\mathbf{Z}} \Delta_j(g) S_{j-3}(f) = \sum_{j\in\mathbf{Z}} \sum_{k\le j-3} \Delta_j(g) \Delta_k(f), \tag{8.4.6}$$

for f in $L^1_{\mathrm{loc}}(\mathbf{R}^n)$. It is not clear for which functions g and in what sense the series in (8.4.6) converges even when f is a Schwartz function. One may verify that the series in (8.4.6) converges absolutely almost everywhere when g is a Schwartz function with mean value zero; in this case, by Exercise 8.4.1 the series $\sum_j \Delta_j(g)$ converges absolutely (everywhere) and $S_j(f)$ is uniformly bounded by the Hardy–Littlewood maximal function $M(f)$, which is finite almost everywhere.

One of the main goals of this section is to show that the series in (8.4.6) converges in L^2 when f is in $L^2(\mathbf{R}^n)$ and g is a BMO function.

The name *paraproduct* is derived from the fact that $P_g(f)$ is essentially "half" the product of fg. Namely, in view of the identity in (8.4.2) the product fg can be written as

$$fg = \sum_j \sum_k \Delta_j(f)\,\Delta_k(g)\,.$$

Restricting the summation of the indices to $k < j$ defines an operator that corresponds to "half" the product of fg. It is only for minor technical reasons that we take $k \le j-3$ in (8.4.6).

The main feature of the paraproduct operator P_g is that it is essentially a sum of orthogonal L^2 functions. Indeed, the Fourier transform of the function $\widehat{\Delta_j(g)}$ is supported in the set

$$\{\xi \in \mathbf{R}^n :\ 2^{j-1} \le |\xi| \le 2^{j+1}\},$$

while the Fourier transform of the function $\widehat{S_{j-3}(f)}$ is supported in the set

$$\bigcup_{k\le j-3} \{\xi \in \mathbf{R}^n :\ 2^{k-1} \le |\xi| \le 2^{k+1}\}.$$

This implies that the Fourier transform of the function $\Delta_j(g)\,S_{j-3}(f)$ is supported in the algebraic sum

$$\{\xi \in \mathbf{R}^n :\ 2^{j-1} \le |\xi| \le 2^{j+1}\} + \{\xi \in \mathbf{R}^n :\ |\xi| \le 2^{j-2}\}.$$

But this sum is contained in the set

$$\{\xi \in \mathbf{R}^n :\ 2^{j-2} \le |\xi| \le 2^{j+2}\}, \tag{8.4.7}$$

and the family of sets in (8.4.7) is "almost disjoint" as j varies. This means that every point in $\mathbf{R}^n$ belongs to at most four annuli of the form (8.4.7). Therefore, the paraproduct $P_g(f)$ can be written as a sum of functions h_j such that the families $\{h_j :\ j \in 4\mathbf{Z}+r\}$ are mutually orthogonal in L^2, for all $r \in \{0,1,2,3\}$. This orthogonal decomposition of the paraproduct has as an immediate consequence its L^2 boundedness when g is an element of BMO.

8.4.2 L^2 Boundedness of Paraproducts

The following theorem is the main result of this subsection.

Theorem 8.4.3. *For fixed $b \in BMO(\mathbf{R}^n)$ and $f \in L^2(\mathbf{R}^n)$ the series*

$$\sum_{|j|\le M} \Delta_j(b)\,S_{j-3}(f)$$

converges in L^2 as $M \to \infty$ to a function that we denote by $P_b(f)$. The operator P_b thus defined is bounded on $L^2(\mathbf{R}^n)$, and there is a dimensional constant C_n such that *for all* $b \in BMO(\mathbf{R}^n)$ *we have*

$$\big\|P_b\big\|_{L^2\to L^2} \le C_n\big\|b\big\|_{BMO}.$$

Proof. The proof of this result follows by putting together some of the powerful ideas developed in Chapter 7. First we define a measure on $\mathbf{R}^{n+1}_+$ by setting

$$d\mu(x,t) = \sum_{j\in\mathbf{Z}} |\Delta_j(b)(x)|^2\,dx\,\delta_{2^{-(j-3)}}(t)\,.$$

By Theorem 7.3.8 we have that μ is a Carleson measure on $\mathbf{R}^{n+1}_+$ whose norm is controlled by a constant multiple of $\big\|b\big\|^2_{BMO}$. Now fix $f \in L^2(\mathbf{R}^n)$ and recall that $\Phi(x) = \sum_{r\le 0}\Psi_{2^{-r}}(x)$. We define a function $F(x,t)$ on $\mathbf{R}^{n+1}_+$ by setting

$$F(x,t) = (\Phi_t * f)(x)\,.$$

Observe that $F(x,2^{-k}) = S_k(f)(x)$ for all $k\in\mathbf{Z}$. We estimate the L^2 norm of a finite sum of terms of the form $\Delta_j(b)\,S_{j-3}(f)$. For $M,N\in\mathbf{Z}^+$ with $M\ge N$ we have

$$\begin{aligned}\int_{\mathbf{R}^n}\bigg|\sum_{N\le|j|\le M}\Delta_j(b)(x)\,S_{j-3}(f)(x)\bigg|^2dx\\ = \int_{\mathbf{R}^n}\bigg|\sum_{N\le|j|\le M}\big(\Delta_j(b)\,S_{j-3}(f)\big)\widehat{\,}(\xi)\bigg|^2d\xi\,.\end{aligned}\tag{8.4.8}$$

It is a simple fact that every $\xi\in\mathbf{R}^n$ belongs to at most four annuli of the form (8.4.7). It follows that at most four terms in the last sum in (8.4.8) are nonzero. Thus

$$\begin{aligned}\int_{\mathbf{R}^n}\bigg|&\sum_{N\le|j|\le M}\big(\Delta_j(b)S_{j-3}(f)\big)\widehat{\,}(\xi)\bigg|^2d\xi\\ &\le 4\sum_{N\le|j|\le M}\int_{\mathbf{R}^n}\big|\big(\Delta_j(b)\,S_{j-3}(f)\big)\widehat{\,}(\xi)\big|^2d\xi\\ &\le 4\sum_{j\in\mathbf{Z}}\int_{\mathbf{R}^n}\big|\Delta_j(b)(x)S_{j-3}(f)(x)\big|^2dx\\ &= 4\int_{\mathbf{R}^n}\big|F(x,t)\big|^2d\mu(x,t)\\ &\le C_n\big\|b\big\|^2_{BMO}\int_{\mathbf{R}^n}F^*(x)^2\,dx,\end{aligned}\tag{8.4.9}$$

where we used Corollary 7.3.6 in the last inequality.

Next we note that the nontangential maximal function F^* of F is controlled by the Hardy–Littlewood maximal function of f. Indeed, since Φ is a Schwartz function, we have

$$F^*(x) \le C_n \sup_{t>0} \sup_{|y-x|<t} \int_{\mathbf{R}^n} \frac{1}{t^n} \frac{|f(z)|}{(1+\frac{|z-y|}{t})^{n+1}}\, dz. \tag{8.4.10}$$

Now break the previous integral into parts such that $|z-y| \ge 3t$ and $|z-y| \le 3t$. In the first case we have $|z-y| \ge |z-x| - t \ge \frac{1}{2}|z-x|$, and the last inequality is valid, since $|z-x| \ge |z-y| - t \ge 2t$. Using this estimate together with Theorem 2.1.10 we obtain that this part of the integral is controlled by a constant multiple of $M(f)(x)$. The part of the integral in (8.4.10) where $|z-y| \le 3t$ is controlled by the integral over the larger set $|z-x| \le 4t$, and since the denominator in (8.4.10) is always bounded by 1, we also obtain that this part of the integral is controlled by a constant multiple of $M(f)(x)$. We conclude that

$$\int_{\mathbf{R}^n} F^*(x)^2\, dx \le C_n \int_{\mathbf{R}^n} M(f)(x)^2\, dx \le C_n \int_{\mathbf{R}^n} |f(x)|^2\, dx. \tag{8.4.11}$$

Combining (8.4.9) and (8.4.11), we obtain the estimate

$$4 \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} \left|\big(\Delta_j(b)\, S_{j-3}(f)\big)\widehat{\;}(\xi)\right|^2 d\xi \le C_n \big\|b\big\|_{BMO}^2 \big\|f\big\|_{L^2}^2 < \infty.$$

This implies that given $\varepsilon > 0$, we can find an $N_0 > 0$ such that

$$M \ge N \ge N_0 \implies \sum_{N\le|j|\le M} \int_{\mathbf{R}^n} \left|\big(\Delta_j(b)\, S_{j-3}(f)\big)\widehat{\;}(\xi)\right|^2 d\xi < \varepsilon.$$

But recall from (8.4.8) and (8.4.9) that

$$\int_{\mathbf{R}^n} \bigg| \sum_{N\le|j|\le M} \Delta_j(b)(x)\, S_{j-3}(f)(x) \bigg|^2 dx \le 4 \sum_{N\le|j|\le M} \int_{\mathbf{R}^n} \left|\big(\Delta_j(b) S_{j-3}(f)\big)\widehat{\;}(\xi)\right|^2 d\xi.$$

We conclude that the sequence

$$\bigg\{ \sum_{|j|\le M} \Delta_j(b) S_{j-3}(f) \bigg\}_M$$

is Cauchy in $L^2(\mathbf{R}^n)$, and therefore it converges in L^2 to a function $P_b(f)$. The boundedness of P_b on L^2 follows from the sequence of inequalities already proved. □

8.4.3 Fundamental Properties of Paraproducts

Having established the L^2 boundedness of paraproducts, we turn to some properties that they possess. First we study their kernels. Paraproducts are not operators of convolution type but are more general integral operators of the form discussed in Section 8.1. We show that the kernel of P_b is a tempered distribution L_b that coincides with a standard kernel on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) :\ x \in \mathbf{R}^n\}$.

First we study the kernel of the operator $f \mapsto \Delta_j(b)S_{j-3}(f)$ for any $j \in \mathbf{Z}$. We have that

$$\Delta_j(b)(x)S_{j-3}(f)(x) = \int_{\mathbf{R}^n} L_j(x,y)f(y)\,dy,$$

where L_j is the integrable function

$$L_j(x,y) = (b * \Psi_{2^{-j}})(x)2^{(j-3)n}\Phi(2^{j-3}(x-y)).$$

Next we can easily verify the following size and regularity estimates for L_j:

$$|L_j(x,y)| \le C_n \|b\|_{BMO} \frac{2^{nj}}{(1+2^j|x-y|)^{n+1}}, \tag{8.4.12}$$

$$|\partial_x^\alpha \partial_y^\beta L_j(x,y)| \le C_{n,\alpha,\beta,N} \|b\|_{BMO} \frac{2^{j(n+|\alpha|+|\beta|)}}{(1+2^j|x-y|)^{n+1+N}}, \tag{8.4.13}$$

for all multi-indices α and β and all $N \ge |\alpha|+|\beta|$.

It follows from (8.4.12) that when $x \neq y$ the series

$$\sum_{j\in\mathbf{Z}} L_j(x,y) \tag{8.4.14}$$

converges absolutely and is controlled in absolute value by

$$C_n \|b\|_{BMO} \sum_{j\in\mathbf{Z}} \frac{2^{nj}}{(1+2^j|x-y|)^{n+1}} \le \frac{C_n' \|b\|_{BMO}}{|x-y|^n}.$$

Similarly, by taking $N \ge |\alpha|+|\beta|$, it can be shown that the series

$$\sum_{j\in\mathbf{Z}} \partial_x^\alpha \partial_y^\beta L_j(x,y) \tag{8.4.15}$$

converges absolutely when $x \neq y$ and is controlled in absolute value by

$$C_{n,\alpha,\beta,N} \|b\|_{BMO} \sum_{j\in\mathbf{Z}} \frac{2^{j(n+|\alpha|+|\beta)}}{(1+2^j|x-y|)^{n+1+N}} \le \frac{C_{n,\alpha,\beta}' \|b\|_{BMO}}{|x-y|^{n+|\alpha|+|\beta|}}$$

for all multi-indices α and β.

The Schwartz kernel of P_b is a distribution W_b on $\mathbf{R}^{2n}$. It follows from the preceding discussion that the distribution W_b coincides with the function

$$L_b(x,y) = \sum_{j\in\mathbf{Z}} L_j(x,y)$$

on $\mathbf{R}^n \times \mathbf{R}^n \setminus \{(x,x) :\ x \in \mathbf{R}^n\}$, and also that the function L_b satisfies the estimates

$$|\partial_x^\alpha \partial_y^\beta L_b(x,y)| \le \frac{C'_{n,\alpha,\beta}\|b\|_{BMO}}{|x-y|^{n+|\alpha|+|\beta|}} \tag{8.4.16}$$

away from the diagonal $x = y$.

We note that the transpose of the operator P_b is formally given by the identity

$$P_b^t(f) = \sum_{j\in\mathbf{Z}} S_{j-3}(f\Delta_j(b)).$$

As remarked in the previous section, the kernel of the operator P_b^t is a distribution W_b^t that coincides with the function

$$L_b^t(x,y) = L_b(y,x)$$

away from the diagonal of $\mathbf{R}^{2n}$. It is trivial to observe that L_b^t satisfies the same size and regularity estimates (8.4.16) as L_b. Moreover, it follows from Theorem 8.4.3 that the operator P_b^t is bounded on $L^2(\mathbf{R}^n)$ with norm at most a multiple of the *BMO* norm of b.

We now turn to two important properties of paraproducts. In view of Definition 8.1.16, we have a meaning for $P_b(1)$ and $P_b^t(1)$, where P_b is the paraproduct operator. The first property we prove is that $P_b(1) = b$. Observe that this statement is trivially valid at a formal level, since $S_j(1) = 1$ for all j and $\sum_j \Delta_j(b) = b$. The second property is that $P_b^t(1) = 0$. This is also trivially checked at a formal level, since $S_{j-3}(\Delta_j(b)) = 0$ for all j, as a Fourier transform calculation shows. We make both of these statements precise in the following proposition.

Proposition 8.4.4. *Given $b \in BMO(\mathbf{R}^n)$, let P_b be the paraproduct operator defined as in (8.4.6). Then the distributions $P_b(1)$ and $P_b^t(1)$ coincide with elements of BMO. Precisely, we have*

$$P_b(1) = b \qquad \textit{and} \qquad P_b^t(1) = 0. \tag{8.4.17}$$

Proof. Let φ be an element of $\mathscr{D}_0(\mathbf{R}^n)$. Find a uniformly bounded sequence of smooth functions with compact support $\{\eta_N\}_{N=1}^\infty$ that converges to the function 1 as $N \to \infty$. Without loss of generality assume that all the functions η_N are equal to 1 on the ball $B(y_0, 3R)$, where $B(y_0, R)$ is a ball that contains the support of φ. As we observed in Remark 8.1.17, the definition of $P_b(1)$ is independent of the choice of sequence η_N, so we have the following identity for all $N \ge 1$:

$$\begin{aligned}\big\langle P_b(1), \varphi\big\rangle = &\int_{\mathbf{R}^n} \sum_{j\in\mathbf{Z}} \Delta_j(b)(x)\, S_{j-3}(\eta_N)(x)\, \varphi(x)\, dx \\ &+ \int_{\mathbf{R}^n} \left[\int_{\mathbf{R}^n} L_b(x,y)\, \varphi(x)\, dx\right] (1-\eta_N(y))\, dy.\end{aligned} \tag{8.4.18}$$

Since φ has mean value zero, we can subtract the constant $L_b(y_0, y)$ from $L_b(x,y)$ in the integral inside the square brackets in (8.4.18). Then we estimate the absolute value of the double integral in (8.4.18) by

$$\int_{|y-y_0|\geq 3R}\int_{|x-y_0|\leq R} A\frac{|y_0-x|}{|y_0-y|^{n+1}}|1-\eta_N(y)|\,|\varphi(x)|\,dx\,dy,$$

which tends to zero as $N\to\infty$ by the Lebesgue dominated convergence theorem.

It suffices to prove that the first integral in (8.4.18) tends to $\int_{\mathbf{R}^n} b(x)\varphi(x)\,dx$ as $N\to\infty$. Let us make some preliminary observations. Since the Fourier transform of the product $\Delta_j(b)S_{j-3}(\eta_N)$ is supported in the annulus

$$\{\xi\in\mathbf{R}^n:\ 2^{j-2}\leq|\xi|\leq 2^{j+2}\},\tag{8.4.19}$$

we may introduce a smooth and compactly supported function $\widehat{Z}(\xi)$ such that for all $j\in\mathbf{Z}$ the function $\widehat{Z}(2^{-j}\xi)$ is equal to 1 on the annulus (8.4.19) and vanishes outside the annulus $\{\xi\in\mathbf{R}^n:\ 2^{j-3}\leq|\xi|\leq 2^{j+3}\}$. Let us denote by Q_j the operator given by multiplication on the Fourier transform by the function $\widehat{Z}(2^{-j}\xi)$.

Note that $S_j(1)$ is well defined and equal to 1 for all j. This is because Φ has integral equal to 1. Also, the duality identity

$$\int f\,S_j(g)\,dx=\int g\,S_j(f)\,dx\tag{8.4.20}$$

holds for all $f\in L^1$ and $g\in L^\infty$. For φ in $\mathscr{D}_0(\mathbf{R}^n)$ we have

$$\begin{aligned}
&\int_{\mathbf{R}^n}\sum_{j\in\mathbf{Z}}\Delta_j(b)\,S_{j-3}(\eta_N)\,\varphi\,dx\\
&=\sum_{j\in\mathbf{Z}}\int_{\mathbf{R}^n}\Delta_j(b)\,S_{j-3}(\eta_N)\,\varphi\,dx && \text{(series converges in } L^2 \text{ and } \varphi\in L^2)\\
&=\sum_{j\in\mathbf{Z}}\int_{\mathbf{R}^n}\Delta_j(b)\,S_{j-3}(\eta_N)\,Q_j(\varphi)\,dx && \big[\widehat{Q_j(\varphi)}=\widehat{\varphi}\text{ on the}\\
& && \quad\text{support of }\big((\Delta_j(b)\,S_{j-3}(\eta_N))\widehat{\ }\big]\\
&=\sum_{j\in\mathbf{Z}}\int_{\mathbf{R}^n}\eta_N\,S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)\,dx && \text{(duality)}\\
&=\int_{\mathbf{R}^n}\eta_N\sum_{j\in\mathbf{Z}}S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)\,dx && \text{(series converges in } L^1 \text{ and } \eta_N\in L^\infty).
\end{aligned}$$

We now explain why the last series of the foregoing converges in L^1. Since φ is in $\mathscr{D}_0(\mathbf{R}^n)$, Exercise 8.4.1 gives that the series $\sum_{j\in\mathbf{Z}}Q_j(\varphi)$ converges in L^1. Since S_j preserves L^1 and

$$\sup_j\big\|\Delta_j(b)\big\|_{L^\infty}\leq C_n\big\|b\big\|_{BMO}$$

by Exercise 8.4.2, it follows that the series $\sum_{j\in\mathbf{Z}}S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)$ also converges in L^1.

We now use the Lebesgue dominated convergence theorem to obtain that the expression

$$\int_{\mathbf{R}^n} \eta_N \sum_{j\in\mathbf{Z}} S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)\,dx$$

converges as $N \to \infty$ to

$$\begin{aligned}
&\int_{\mathbf{R}^n} \sum_{j\in\mathbf{Z}} S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)\,dx \\
&\qquad = \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} S_{j-3}\big(\Delta_j(b)Q_j(\varphi)\big)\,dx && \text{(series converges in } L^1\text{)} \\
&\qquad = \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} S_{j-3}(1)\,\Delta_j(b)\,Q_j(\varphi)\,dx && \text{(in view of (8.4.20))} \\
&\qquad = \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} \Delta_j(b)\,Q_j(\varphi)\,dx && \text{(since } S_{j-3}(1)=1\text{)} \\
&\qquad = \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} \Delta_j(b)\,\varphi\,dx && \big(\widehat{Q_j(\varphi)} = \widehat{\varphi} \text{ on support } \widehat{\Delta_j(b)}\big) \\
&\qquad = \sum_{j\in\mathbf{Z}} \big\langle b, \Delta_j(\varphi)\big\rangle && \text{(duality)} \\
&\qquad = \Big\langle b, \sum_{j\in\mathbf{Z}} \Delta_j(\varphi)\Big\rangle && \text{(series converges in } H^1,\ b\in BMO\text{)} \\
&\qquad = \big\langle b, \varphi\big\rangle && \text{(Exercise 8.4.1(a)).}
\end{aligned}$$

Regarding the fact that the series $\sum_j \Delta_j(\varphi)$ converges in H^1, we refer to Exercise 8.4.1. We now obtain that the first integral in (8.4.18) tends to $\langle b, \varphi\rangle$ as $N \to \infty$. We have therefore proved that

$$\big\langle P_b(1), \varphi\big\rangle = \big\langle b, \varphi\big\rangle$$

for all φ in $\mathscr{D}_0(\mathbf{R}^n)$. In other words, we have now identified $P_b(1)$ as an element of $\mathscr{D}_0'$ with the BMO function b.

For the transpose operator P_b^t we observe that we have the identity

$$\begin{aligned}
\big\langle P_b^t(1), \varphi\big\rangle = &\int_{\mathbf{R}^n} \sum_{j\in\mathbf{Z}} S_{j-3}^t\big(\Delta_j(b)\,\eta_N\big)(x)\,\varphi(x)\,dx \\
&\qquad + \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} L_b^t(x,y)(1-\eta_N(y))\,\varphi(x)\,dy\,dx.
\end{aligned} \tag{8.4.21}$$

As before, we can use the Lebesgue dominated convergence theorem to show that the double integral in (8.4.21) tends to zero. As for the first integral in (8.4.21), we have the identity

$$\int_{\mathbf{R}^n} P_b^t(\eta_N)\,\varphi\,dx = \int_{\mathbf{R}^n} \eta_N\,P_b(\varphi)\,dx.$$

Since φ is a multiple of an L^2-atom for H^1, Theorem 8.2.6 gives that $P_b(\varphi)$ is an L^1 function. The Lebesgue dominated convergence theorem now implies that

$$\int_{\mathbf{R}^n} \eta_N P_b(\varphi)\,dx \to \int_{\mathbf{R}^n} P_b(\varphi)\,dx = \int_{\mathbf{R}^n} \sum_{j\in\mathbf{Z}} \Delta_j(b)\,S_{j-3}(\varphi)\,dx$$

as $N \to \infty$. The required conclusion would follow if we could prove that the function $P_b(\varphi)$ has integral zero. Since $\Delta_j(b)$ and $S_{j-3}(\varphi)$ have disjoint Fourier transforms, it follows that

$$\int_{\mathbf{R}^n} \Delta_j(b)\,S_{j-3}(\varphi)\,dx = 0$$

for all j in $\mathbf{Z}$. But the series

$$\sum_{j\in\mathbf{Z}} \Delta_j(b)\,S_{j-3}(\varphi) \tag{8.4.22}$$

defining $P_b(\varphi)$ converges in L^2 and not necessarily in L^1, and for this reason we need to justify the interchange of the following integrals:

$$\int_{\mathbf{R}^n} \sum_{j\in\mathbf{Z}} \Delta_j(b)\,S_{j-3}(\varphi)\,dx = \sum_{j\in\mathbf{Z}} \int_{\mathbf{R}^n} \Delta_j(b)\,S_{j-3}(\varphi)\,dx\,. \tag{8.4.23}$$

To complete the proof, it suffices to show that when φ is in $\mathscr{D}_0(\mathbf{R}^n)$, the series in (8.4.22) converges in L^1. To prove this, pick a ball $B(y_0,R)$ that contains the support of φ. The series in (8.4.22) converges in $L^2(3B)$ and hence converges in $L^1(3B)$. It remains to prove that it converges in $L^1((3B)^c)$. For a fixed $x \in (3B)^c$ and a finite subset F of $\mathbf{Z}$, we have

$$\sum_{j\in F} \int_{\mathbf{R}^n} L_j(x,y)\varphi(y)\,dy = \sum_{j\in F} \int_B \big(L_j(x,y) - L_j(x,y_0)\big)\,\varphi(y)\,dy\,. \tag{8.4.24}$$

Using estimates (8.4.13), we obtain that the expression in (8.4.24) is controlled by a constant multiple of

$$\int_B \sum_{j\in F} \frac{|y-y_0|2^{nj}2^j}{(1+2^j|x-y_0|)^{n+2}}\,|\varphi(y)|\,dy \le c\,\frac{1}{|x-y_0|^{n+1}} \int_{\mathbf{R}^n} |y-y_0|\,|\varphi(y)|\,dy.$$

Integrating this estimate with respect to $x \in (3B)^c$, we obtain that

$$\sum_{j\in F} \big\|\Delta_j(b)S_{j-3}(\varphi)\big\|_{L^1((3B)^c)} \le C_n \big\|\varphi\big\|_{L^1} < \infty$$

for all finite subsets F of $\mathbf{Z}$. This proves that the series in (8.4.22) converges in L^1.

We have now proved that $\langle P_b^t(1),\varphi\rangle = 0$ for all $\varphi \in \mathscr{D}_0(\mathbf{R}^n)$. This shows that the distribution $P_b^t(1)$ is a constant function, which is of course identified with zero if considered as an element of BMO. □

Remark 8.4.5. The boundedness of P_b on L^2 is a consequence of Theorem 8.3.3, since hypothesis (iv) is satisfied. Indeed, $P_b(1) = b$, $P_b^t(1) = 0$ are both BMO functions, and see Exercise 8.4.4 for a sketch of a proof of the estimate $\big\|P_b\big\|_{WB} \le C_n\big\|b\big\|_{BMO}$. This provides another proof of the fact that $\big\|P_b\big\|_{L^2\to L^2} \le C_n\big\|b\big\|_{BMO}$, bypassing Theorem 8.3.3. We use this result to obtain a different proof of the main direction in the $T(1)$ theorem in the next section.

Exercises

8.4.1. Let $f \in \mathscr{S}(\mathbf{R}^n)$ have mean value zero, and consider the series

$$\sum_{j\in\mathbf{Z}} \Delta_j(f)\,.$$

(a) Show that this series converges to f absolutely everywhere.
(b) Show that this series converges in L^1.
(b) Show that this series converges in H^1.
$\big[$*Hint:* To obtain convergence in L^1 for $j \ge 0$ use the estimate $\big\|\Delta_j(f)\big\|_{L^1} \le 2^{-j}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} 2^{jn}|\Psi(2^j y)|\,|2^j y|\,|(\nabla f)(x-\theta y)|\,dy\,dx$ for some θ in $[0,1]$ and consider the cases $|x| \ge 2|y|$ and $|x| \le 2|y|$. When $j \le 0$ use the simple identity $f * \Psi_{2^{-j}} = (f_{2^j} * \Psi)_{2^{-j}}$ and reverse the roles of f and Ψ. To show convergence in H^1, use that $\big\|\Delta_j(\varphi)\big\|_{H^1} \approx \big\|(\sum_k |\Delta_k\Delta_j(\varphi)|^2)^{\frac{1}{2}}\big\|_{L^1}$ and that only at most three terms in the square function are nonzero.$\big]$

8.4.2. Without appealing to the H^1-BMO duality theorem, prove that there is a dimensional constant C_n such that for all $b \in BMO(\mathbf{R}^n)$ we have

$$\sup_{j\in\mathbf{Z}} \big\|\Delta_j(b)\big\|_{L^\infty} \le C_n \big\|b\big\|_{BMO}.$$

8.4.3. (a) Show that for all $1 < p,q,r < \infty$ with $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$ there is a constant C_{pqr} such that for all Schwartz functions f,g on $\mathbf{R}^n$ we have

$$\big\|P_g(f)\big\|_{L^r} \le C_{pqr}\big\|f\big\|_{L^p}\big\|g\big\|_{L^q}.$$

(b) Obtain the same conclusion for the bilinear operator

$$\widetilde{P}_g(f) = \sum_j \sum_{k\le j} \Delta_j(g)\,\Delta_k(f)\,.$$

$\big[$*Hint:* Part (a): Estimate the L^r norm using duality. Part (b): Use part (a).$\big]$

8.4.4. (a) Let f be a normalized bump (see Definition 8.3.1). Prove that

$$\big\|\Delta_j(f_R)\big\|_{L^\infty} \le C(n,\Psi)\min\big(2^{-j}R^{-(n+1)}, 2^{nj}\big)$$

for all $R>0$. Then interpolate between L^1 and L^∞ to obtain

$$\big\|\Delta_j(f_R)\big\|_{L^2} \le C(n,\Psi)\min\Big(2^{-\frac{j}{2}}R^{-\frac{n+1}{2}}, 2^{\frac{nj}{2}}\Big)\,.$$

(b) Observe that the same result is valid for the operators Q_j as defined in Proposition 8.4.4. Conclude that for some constant C_n we have

$$\sum_{j\in\mathbf{Z}} \big\|Q_j(g_R)\big\|_{L^2} \le C_n R^{-\frac{n}{2}}\,.$$

(c) Show that there is a constant C_n such that for all normalized bumps f and g we have

$$\big|\big\langle P_b(\tau^{x_0}(f_R)),\tau^{x_0}(g_R)\big\rangle\big| \le C_n R^{-n}\big\|b\big\|_{BMO}\,.$$

$\big[$*Hint:* Part (a): Use the cancellation of the functions f and Ψ. Part (c): Write

$$\big\langle P_b(\tau^{x_0}(f_R)),\tau^{x_0}(g_R)\big\rangle = \sum_j \int_{\mathbf{R}^n} S_{j-3}[\Delta_j(\tau^{-x_0}(b))Q_j(g_R)]f_R\,dx\,.$$

Apply the Cauchy–Schwarz inequality, and use the boundedness of S_{j-3} on L^2, Exercise 8.4.2, and part (b).$\big]$

8.4.5. (*Continuous paraproducts*) (a) Let Φ and Ψ be Schwartz functions on $\mathbf{R}^n$ with $\int_{\mathbf{R}^n}\Phi(x)\,dx = 1$ and $\int_{\mathbf{R}^n}\Psi(x)\,dx = 0$. For $t>0$ define operators $P_t(f) = \Phi_t * f$ and $Q_t(f) = \Psi_t * f$. Let $b \in BMO(\mathbf{R}^n)$ and $f \in L^2(\mathbf{R}^n)$. Show that the limit

$$\lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N Q_t\big(Q_t(b)\,P_t(f)\big)\,\frac{dt}{t}$$

converges in $L^2(\mathbf{R}^n)$ and defines an operator $\Pi_b(f)$ that satisfies

$$\big\|\Pi_b\big\|_{L^2\to L^2} \le C_n\big\|b\big\|_{BMO}$$

for some dimensional constant C_n.
(b) Under the additional assumption that

$$\lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N Q_t^2\,\frac{dt}{t} = I\,,$$

identify $\Pi_b(1)$ and $\Pi_b(b)$.
$\big[$*Hint:* Suitably adapt the proofs of Theorem 8.4.3 and Proposition 8.4.4.$\big]$

8.5 An Almost Orthogonality Lemma and Applications

In this section we discuss an important lemma regarding orthogonality of operators and some of its applications.

It is often the case that a linear operator T is given as an infinite sum of other linear operators T_j such that the T_j's are uniformly bounded on L^2. This sole condition is not enough to imply that the sum of the T_j's is also L^2 bounded, although this is often the case. Let us consider, for instance, the linear operators $\{T_j\}_{j\in\mathbf{Z}}$ given by convolution with the smooth functions $e^{2\pi i jt}$ on the circle $\mathbf{T}^1$. Each T_j can be written as $T_j(f) = (\widehat{f}\otimes\delta_j)^\vee$, where $\widehat{f}$ is the sequence of Fourier coefficients of f; here δ_j is the infinite sequence consisting of zeros everywhere except at the jth entry, in which it has 1, and $\otimes$ denotes term-by-term multiplication of infinite sequences. It follows that each operator T_j is bounded on $L^2(\mathbf{T}^1)$ with norm 1. Moreover, the sum of the T_j's is the identity operator, which is also L^2 bounded with norm 1.

It is apparent from the preceding discussion that the crucial property of the T_j's that makes their sum bounded is their orthogonality. In the preceding example we have $T_jT_k = 0$ unless $j = k$. It turns out that this orthogonality condition is a bit too strong, and it can be weakened significantly.

8.5.1 The Cotlar–Knapp–Stein Almost Orthogonality Lemma

The next result provides a sufficient orthogonality criterion for boundedness of sums of linear operators on a Hilbert space.

Lemma 8.5.1. *Let $\{T_j\}_{j\in\mathbf{Z}}$ be a family of operators mapping a Hilbert space H to itself. Assume that there is a a function $\gamma : \mathbf{Z}\to\mathbf{R}^+$ such that*

$$\big\|T_j^*T_k\big\|_{H\to H} + \big\|T_jT_k^*\big\|_{H\to H} \le \gamma(j-k) \tag{8.5.1}$$

for all j,k in $\mathbf{Z}$. Suppose that

$$A = \sum_{j\in\mathbf{Z}} \sqrt{\gamma(j)} < \infty\,.$$

Then the following three conclusions are valid:

(i) For all finite subsets Λ of $\mathbf{Z}$ we have

$$\Big\|\sum_{j\in\Lambda} T_j\Big\|_{H\to H} \le A.$$

(ii) For all $x\in H$ we have

$$\sum_{j\in\mathbf{Z}} \big\|T_j(x)\big\|_H^2 \le A^2\big\|x\big\|_H^2.$$

(iii) For all $x\in H$ the sequence $\sum_{|j|\le N} T_j(x)$ converges to some $T(x)$ as $N\to\infty$ in the norm topology of H. The linear operator T defined in this way is bounded from H to H with norm

$$\big\|T\big\|_{H\to H} \le A.$$

Proof. As usual we denote by S^* the adjoint of a linear operator S. It is a simple fact that any bounded linear operator $S:\ H\to H$ satisfies

$$\big\|S\big\|_{H\to H}^2 = \big\|SS^*\big\|_{H\to H}. \tag{8.5.2}$$

See Exercise 8.5.1. By taking $j=k$ in (8.5.1) and using (8.5.2), we obtain

$$\big\|T_j\big\|_{H\to H}\le\sqrt{\gamma(0)}\tag{8.5.3}$$

for all $j\in\mathbf{Z}$. It also follows from (8.5.2) that if an operator S is self-adjoint, then $\big\|S\big\|^2_{H\to H}=\big\|S^2\big\|_{H\to H}$, and more generally,

$$\big\|S\big\|^m_{H\to H}=\big\|S^m\big\|_{H\to H}\tag{8.5.4}$$

for m that are powers of 2. Now observe that the linear operator

$$\Big(\sum_{j\in\Lambda}T_j\Big)\Big(\sum_{j\in\Lambda}T_j^*\Big)$$

is self-adjoint. Applying (8.5.2) and (8.5.4) to this operator, we obtain

$$\Big\|\sum_{j\in\Lambda}T_j\Big\|^2_{H\to H}=\Big\|\Big[\Big(\sum_{j\in\Lambda}T_j\Big)\Big(\sum_{j\in\Lambda}T_j^*\Big)\Big]^m\Big\|^{\frac{1}{m}}_{H\to H},\tag{8.5.5}$$

where m is a power of 2. We now expand the mth power of the expression in (8.5.5). So we write the right side of the identity in (8.5.5) as

$$\Big\|\sum_{j_1,\cdots,j_{2m}\in\Lambda}T_{j_1}T^*_{j_2}\cdots T_{j_{2m-1}}T^*_{j_{2m}}\Big\|^{\frac{1}{m}}_{H\to H},\tag{8.5.6}$$

which is controlled by

$$\Big(\sum_{j_1,\cdots,j_{2m}\in\Lambda}\big\|T_{j_1}T^*_{j_2}\cdots T_{j_{2m-1}}T^*_{j_{2m}}\big\|_{H\to H}\Big)^{\frac{1}{m}}.\tag{8.5.7}$$

We estimate the expression inside the sum in (8.5.7) in two different ways. First we group j_1 with j_2, j_3 with j_4, ..., j_{2m-1} with j_{2m} and we apply (8.5.3) and (8.5.1) to control this expression by

$$\gamma(j_1-j_2)\gamma(j_3-j_4)\cdots\gamma(j_{2m-1}-j_{2m}).$$

Grouping j_2 with j_3, j_4 with j_5, ..., j_{2m-2} with j_{2m-1} and leaving j_1 and j_{2m} alone, we also control the expression inside the sum in (8.5.7) by

$$\sqrt{\gamma(0)}\gamma(j_2-j_3)\gamma(j_4-j_5)\cdots\gamma(j_{2m-2}-j_{2m-1})\sqrt{\gamma(0)}.$$

Taking the geometric mean of these two estimates, we obtain the following bound for (8.5.7):

$$\Big(\sum_{j_1,\dots,j_{2m}\in\Lambda}\sqrt{\gamma(0)}\sqrt{\gamma(j_1-j_2)}\sqrt{\gamma(j_2-j_3)}\cdots\sqrt{\gamma(j_{2m-1}-j_{2m})}\Big)^{\frac{1}{m}}.$$

Summing first over j_1, then over j_2, and finally over j_{2m-1}, we obtain the estimate

$$\gamma(0)^{\frac{1}{2m}} A^{\frac{2m-1}{m}} \Big(\sum_{j_{2m}\in\Lambda} 1 \Big)^{\frac{1}{m}}$$

for (8.5.7). Using (8.5.5), we conclude that

$$\Big\| \sum_{j\in\Lambda} T_j \Big\|_{H\to H}^2 \le \gamma(0)^{\frac{1}{2m}} A^{\frac{2m-1}{m}} |\Lambda|^{\frac{1}{m}},$$

and letting $m\to\infty$, we obtain conclusion (i) of the proposition.

To prove (ii) we use the Rademacher functions r_j of Appendix C.1. These functions are defined for nonnegative integers j, but we can reindex them so that the subscript j runs through the integers. The fundamental property of these functions is their orthogonality, that is, $\int_0^1 r_j(\omega)r_k(\omega)\,d\omega = 0$ when $j\neq k$. Using the fact that the norm $\|\cdot\|_H$ comes from an inner product, for every finite subset Λ of $\mathbf{Z}$ and x in H we obtain

$$\begin{aligned}
&\int_0^1 \Big\| \sum_{j\in\Lambda} r_j(\omega) T_j(x) \Big\|_H^2 \, d\omega \\
&\quad = \sum_{j\in\Lambda} \big\| T_j(x)\big\|_H^2 + \int_0^1 \sum_{\substack{j,k\in\Lambda\\ j\neq k}} r_j(\omega) r_k(\omega) \big\langle T_j(x), T_k(x)\big\rangle_H \, d\omega \qquad (8.5.8)\\
&\quad = \sum_{j\in\Lambda} \big\| T_j(x)\big\|_H^2 .
\end{aligned}$$

For any fixed $\omega\in[0,1]$ we now use conclusion (i) of the proposition for the operators $r_j(\omega)T_j$, which also satisfy assumption (8.5.1), since $r_j(\omega)=\pm 1$. We obtain that

$$\Big\| \sum_{j\in\Lambda} r_j(\omega) T_j(x) \Big\|_H^2 \le A^2 \big\|x\big\|_H^2 ,$$

which, combined with (8.5.8), gives conclusion (ii).

We now prove (iii). First we show that given $x\in H$ the sequence

$$\Big\{ \sum_{j=-N}^{N} T_j(x) \Big\}_N$$

is Cauchy in H. Suppose that this is not the case. This means that there is some $\varepsilon > 0$ and a subsequence of integers $1\le N_1 < N_2 < N_3 < \cdots$ such that

$$\big\| \widetilde{T}_k(x)\big\|_H \ge \varepsilon , \qquad (8.5.9)$$

where we set

$$\widetilde{T}_k(x) = \sum_{N_k \le |j| < N_{k+1}} T_j(x).$$

For any fixed $\omega \in [0,1]$, apply conclusion (i) to the operators $S_j = r_k(\omega)T_j$ whenever $N_k \le |j| < N_{k+1}$, since these operators clearly satisfy hypothesis (8.5.1). Taking $N_1 \le |j| \le N_{K+1}$, we obtain

$$\Big\| \sum_{k=1}^{K} r_k(\omega) \sum_{N_k \le |j| < N_{k+1}} T_j(x) \Big\|_H = \Big\| \sum_{k=1}^{K} r_k(\omega) \widetilde{T}_k(x) \Big\|_H \le A \left\| x \right\|_H .$$

Squaring and integrating this inequality with respect to ω in $[0,1]$, and using (8.5.8) with $\widetilde{T}_k$ in the place of T_k and $\{1,2,\dots,K\}$ in the place of Λ, we obtain

$$\sum_{k=1}^{K} \big\|\widetilde{T}_k(x)\big\|_H^2 \le A^2 \|x\|_H^2 .$$

But this clearly contradicts (8.5.9) as $K \to \infty$.

We conclude that every sequence $\big\{\sum_{j=-N}^{N} T_j(x)\big\}_N$ is Cauchy in H and thus it converges to Tx for some linear operator T. In view of conclusion (i), it follows that T is a bounded operator on H with norm at most A. □

Remark 8.5.2. At first sight, it appears strange that the norm of the operator T is independent of the norm of every piece T_j and depends only on the quantity A in (8.5.1). But as observed in the proof, if we take $j = k$ in (8.5.1), we obtain

$$\|T_j\|_{H\to H}^2 = \|T_j T_j^*\|_{H\to H} \le \gamma(0) \le A^2 ;$$

thus the norm of each individual T_j is also controlled by the constant A.

We also note that there wasn't anything special about the role of the index set $\mathbf{Z}$ in Lemma 8.5.1. Indeed, the set $\mathbf{Z}$ can be replaced by any countable group, such as $\mathbf{Z}^k$ for some k. For instance, see Theorem 8.5.7, in which the index set is $\mathbf{Z}^{2n}$. See also Exercises 8.5.7 and 8.5.8, in which versions of Lemma 8.5.1 are given with no group structure on the set of indices.

8.5.2 An Application

We now discuss an application of the almost orthogonality lemma just proved concerning sums of nonconvolution operators on $L^2(\mathbf{R}^n)$. We begin with the following version of Theorem 8.3.3, in which it is assumed that $T(1) = T^t(1) = 0$.

Proposition 8.5.3. *Suppose that $K_j(x,y)$ are functions on $\mathbf{R}^n \times \mathbf{R}^n$ indexed by $j \in \mathbf{Z}$ that satisfy*

$$|K_j(x,y)| \le \frac{A 2^{nj}}{(1+2^j|x-y|)^{n+\delta}} , \tag{8.5.10}$$

$$|K_j(x,y) - K_j(x,y')| \le A 2^{\gamma j} 2^{nj} |y-y'|^{\gamma} , \tag{8.5.11}$$

$$|K_j(x,y) - K_j(x',y)| \le A 2^{\gamma j} 2^{nj} |x-x'|^{\gamma} , \tag{8.5.12}$$

for some $0 < A, \gamma, \delta < \infty$ *and all* $x, y, x', y' \in \mathbf{R}^n$. *Suppose also that*

$$\int_{\mathbf{R}^n} K_j(z,y)\,dz = 0 = \int_{\mathbf{R}^n} K_j(x,z)\,dz, \tag{8.5.13}$$

for all $x, y \in \mathbf{R}^n$ *and all* $j \in \mathbf{Z}$. *For* $j \in \mathbf{Z}$ *define integral operators*

$$T_j(f)(x) = \int_{\mathbf{R}^n} K_j(x,y)\, f(y)\,dy$$

for $f \in L^2(\mathbf{R}^n)$. *Then the series*

$$\sum_{j\in\mathbf{Z}} T_j(f)$$

converges in L^2 *to some* $T(f)$ *for all* $f \in L^2(\mathbf{R}^n)$, *and the linear operator* T *defined in this way is* L^2 *bounded.*

Proof. It is a consequence of (8.5.10) that the kernels K_j are in $L^1(dy)$ uniformly in $x \in \mathbf{R}^n$ and $j \in \mathbf{Z}$ and hence the operators T_j map $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$ uniformly in j. Our goal is to show that the sum of the T_j's is also bounded on $L^2(\mathbf{R}^n)$. We achieve this using the orthogonality considerations of Lemma 8.5.1. To be able to use Lemma 8.5.1, we need to prove (8.5.1). Indeed, we show that for all $k, j \in \mathbf{Z}$ we have

$$\big\|T_j T_k^*\big\|_{L^2\to L^2} \le CA^2\, 2^{-\frac{1}{4}\frac{\delta}{n+\delta}\min(\gamma,\delta)|j-k|}, \tag{8.5.14}$$

$$\big\|T_j^* T_k\big\|_{L^2\to L^2} \le CA^2\, 2^{-\frac{1}{4}\frac{\delta}{n+\delta}\min(\gamma,\delta)|j-k|}, \tag{8.5.15}$$

for some $0 < C = C_{n,\gamma,\delta} < \infty$. We prove only (8.5.15), since the proof of (8.5.14) is similar. In fact, since the kernels of T_j and T_j^* satisfy similar size, regularity, and cancellation estimates, (8.5.15) is directly obtained from (8.5.14) when T_j are replaced by T_j^*.

It suffices to prove (8.5.15) under the extra assumption that $k \le j$. Once (8.5.15) is established under this assumption, taking $j \le k$ yields

$$\big\|T_j^* T_k\big\|_{L^2\to L^2} = \big\|(T_k^* T_j)^*\big\|_{L^2\to L^2} = \big\|T_k^* T_j\big\|_{L^2\to L^2} \le CA^2 2^{-\frac{1}{2}\min(\gamma,\delta)|j-k|},$$

thus proving (8.5.15) also under the assumption $j \le k$.

We therefore take $k \le j$ in the proof of (8.5.15). Note that the kernel of $T_j^* T_k$ is

$$L_{jk}(x,y) = \int_{\mathbf{R}^n} \overline{K_j(z,x)} K_k(z,y)\,dz.$$

We prove that

$$\sup_{x\in\mathbf{R}^n} \int_{\mathbf{R}^n} |L_{kj}(x,y)|\,dy \;\le\; CA^2\, 2^{-\frac{1}{4}\frac{\delta}{n+\delta}\min(\gamma,\delta)|k-j|}, \tag{8.5.16}$$

$$\sup_{y\in\mathbf{R}^n} \int_{\mathbf{R}^n} |L_{kj}(x,y)|\,dx \;\le\; CA^2\, 2^{-\frac{1}{4}\frac{\delta}{n+\delta}\min(\gamma,\delta)|k-j|}. \tag{8.5.17}$$

Once (8.5.16) and (8.5.17) are established, (8.5.15) follows directly from the classical Schur lemma in Appendix I.1.

We need to use the following estimate, valid for $k \le j$:

$$\int_{\mathbf{R}^n} \frac{2^{nj}\min(1,(2^k|u|)^\gamma)}{(1+2^j|u|)^{n+\delta}}\,du \le C_{n,\delta}2^{-\frac{1}{2}\min(\gamma,\delta)(j-k)}. \tag{8.5.18}$$

Indeed, to prove (8.5.18), we observe that by changing variables we may assume that $j=0$ and $k\le 0$. Taking $r=k-j\le 0$, we establish (8.5.18) as follows:

$$\begin{aligned}
\int_{\mathbf{R}^n} \frac{\min(1,(2^r|u|)^\gamma)}{(1+|u|)^{n+\delta}}\,du &\le \int_{\mathbf{R}^n} \frac{\min\left(1,(2^r|u|)^{\frac{1}{2}\min(\gamma,\delta)}\right)}{(1+|u|)^{n+\delta}}\,du \\
&\le \int\limits_{|u|\le 2^{-r}} \frac{(2^r|u|)^{\frac{1}{2}\min(\gamma,\delta)}}{(1+|u|)^{n+\delta}}\,du + \int\limits_{|u|\ge 2^{-r}} \frac{1}{(1+|u|)^{n+\delta}}\,du \\
&\le 2^{\frac{1}{2}\min(\gamma,\delta)r}\int\limits_{\mathbf{R}^n} \frac{1}{(1+|u|)^{n+\frac{\delta}{2}}}\,du + \int\limits_{|u|\ge 2^{-r}} \frac{1}{|u|^{n+\delta}}\,du \\
&\le C_{n,\delta}\left[2^{\frac{1}{2}\min(\gamma,\delta)r}+2^{\delta r}\right] \\
&\le C_{n,\delta}\,2^{-\frac{1}{2}\min(\gamma,\delta)|r|},
\end{aligned}$$

We now obtain estimates for L_{jk} in the case $k\le j$. Using (8.5.13), we write

$$\begin{aligned}
|L_{jk}(x,y)| &= \left|\int_{\mathbf{R}^n} K_k(z,y)\overline{K_j(z,x)}\,dz\right| \\
&= \left|\int_{\mathbf{R}^n} \left[K_k(z,y)-K_k(x,y)\right]\overline{K_j(z,x)}\,dz\right| \\
&\le A^2\int_{\mathbf{R}^n} 2^{nk}\min(1,(2^k|x-z|)^\gamma)\frac{2^{nj}}{(1+2^j|z-x|)^{n+\delta}}\,dz \\
&\le CA^2\,2^{kn}\,2^{-\frac{1}{2}\min(\gamma,\delta)(j-k)}
\end{aligned}$$

using estimate (8.5.18). Combining this estimate with

$$|L_{jk}(x,y)| \le \int_{\mathbf{R}^n} |K_j(z,x)|\,|K_k(z,y)|\,dz \le \frac{CA^2 2^{kn}}{(1+2^k|x-y|)^{n+\delta}},$$

which follows from (8.5.10) and the result in Appendix K.1 (since $k\le j$), yields

$$|L_{jk}(x,y)| \le C_{n,\gamma,\delta}A^2\,2^{-\frac{1}{2}\frac{\delta/2}{n+\delta}\min(\gamma,\delta)(j-k)}\frac{2^{kn}}{(1+2^k|x-y|)^{n+\frac{\delta}{2}}},$$

which easily implies (8.5.16) and (8.5.17). This concludes the proof of the proposition. □

8.5.3 Almost Orthogonality and the $T(1)$ Theorem

We now give an important application of the proposition just proved. We re-prove the difficult direction of the $T(1)$ theorem proved in Section 8.3. We have the following:

Theorem 8.5.4. *Let K be in $SK(\delta,A)$ and let T be a continuous linear operator from $\mathscr{S}(\mathbf{R}^n)$ to $\mathscr{S}'(\mathbf{R}^n)$ associated with K. Assume that*

$$\|T(1)\|_{BMO}+\|T^t(1)\|_{BMO}+\|T\|_{WB}=B_4<\infty.$$

Then T extends to bounded linear operator on $L^2(\mathbf{R}^n)$ with norm at most a constant multiple of $A+B_4$.

Proof. Consider the paraproduct operators $P_{T(1)}$ and $P_{T^t(1)}$ introduced in the previous section. Then, as we showed in Proposition 8.4.4, we have

$$\begin{aligned} P_{T(1)}(1) &= T(1), & (P_{T(1)})^t(1) &= 0,\\ P_{T^t(1)}(1) &= T^t(1), & (P_{T^t(1)})^t(1) &= 0. \end{aligned}$$

Let us define an operator

$$L=T-P_{T(1)}-(P_{T^t(1)})^t.$$

Using Proposition 8.4.4, we obtain that

$$L(1)=L^t(1)=0.$$

In view of (8.4.16), we have that L is an operator whose kernel satisfies the estimates (8.1.1), (8.1.2), and (8.1.3) with constants controlled by a dimensional constant multiple of

$$A+\|T(1)\|_{BMO}+\|T^t(1)\|_{BMO}.$$

Both of these numbers are controlled by $A+B_4$. We also have

$$\begin{aligned} \|L\|_{WB} &\le C_n\big(\|T\|_{WB}+\|P_{T(1)}\|_{L^2\to L^2}+\|(P_{T^t(1)})^t\|_{L^2\to L^2}\big)\\ &\le C_n\big(\|T\|_{WB}+\|T(1)\|_{BMO}+\|T^t(1)\|_{BMO}\big)\\ &\le C_n(A+B_4), \end{aligned}$$

which is a very useful fact.

Next we introduce operators Δ_j and S_j; one should be cautious as these are not the operators Δ_j and S_j introduced in Section 8.4 but rather discrete analogues of those introduced in the proof of Theorem 8.3.3. We pick a smooth radial real-valued function Φ with compact support contained in the unit ball $B(0,\frac{1}{2})$ that satisfies $\int_{\mathbf{R}^n}\Phi(x)\,dx=1$ and we define

$$\Psi(x)=\Phi(x)-2^{-n}\Phi(\tfrac{x}{2}). \tag{8.5.19}$$

Notice that Ψ has mean value zero. We define

$$\Phi_{2^{-j}}(x) = 2^{nj}\Phi(2^j x) \qquad \text{and} \qquad \Psi_{2^{-j}}(x) = 2^{nj}\Psi(2^j x)$$

and we observe that both Φ and Ψ are supported in $B(0,1)$ and are multiples of normalized bumps. We then define Δ_j to be the operator given by convolution with the function $\Psi_{2^{-j}}$ and S_j the operator given by convolution with the function $\Phi_{2^{-j}}$. In view of identity (8.5.19) we have that $\Delta_j = S_j - S_{j-1}$. Notice that

$$S_j L S_j = S_{j-1} L S_{j-1} + \Delta_j L S_j + S_{j-1} L \Delta_j\,,$$

which implies that for all integers $N < M$ we have

$$\begin{aligned} S_M L S_M - S_{N-1} L S_{N-1} &= \sum_{j=N}^{M} \big(S_j L S_j - S_{j-1} L S_{j-1}\big) \\ &= \sum_{j=N}^{M} \Delta_j L S_j - \sum_{j=N}^{M} S_{j-1} L \Delta_j. \end{aligned} \tag{8.5.20}$$

Until the end of the proof we fix a Schwartz function f whose Fourier transform vanishes in a neighborhood of the origin; such functions are dense in L^2; see Exercise 5.2.9. We would like to use Proposition 8.5.3 to conclude that

$$\sup_{M\in\mathbf{Z}} \sup_{N<M} \big\| S_M L S_M(f) - S_{N-1} L S_{N-1}(f)\big\|_{L^2} \le C_n (A_2 + B_4) \big\|f\big\|_{L^2} \tag{8.5.21}$$

and that $S_M L S_M(f) - S_{N-1} L S_{N-1}(f) \to \widetilde{L}(f)$ in L^2 as $M \to \infty$ and $N \to -\infty$. Once these statements are proved, we deduce that $\widetilde{L}(f) = L(f)$. To see this, it suffices to prove that $S_M L S_M(f) - S_{N-1} L S_{N-1}(f)$ converges to $L(f)$ weakly in L^2. Indeed, let g be another Schwartz function. Then

$$\begin{aligned} &\big\langle S_M L S_M(f) - S_{N-1} L S_{N-1}(f), g\big\rangle - \big\langle L(f), g\big\rangle \\ &\qquad = \big\langle S_M L S_M(f) - L(f), g\big\rangle - \big\langle S_{N-1} L S_{N-1}(f), g\big\rangle. \end{aligned} \tag{8.5.22}$$

We first prove that the first term in (8.5.22) tends to zero as $M \to \infty$. Indeed,

$$\begin{aligned} \big\langle S_M L S_M(f) - L(f), g\big\rangle &= \big\langle L S_M(f), S_M g\big\rangle - \big\langle L(f), g\big\rangle \\ &= \big\langle L(S_M(f) - f), S_M(g)\big\rangle + \big\langle L(f), S_M(g) - g\big\rangle, \end{aligned}$$

and both terms converge to zero, since $S_M(f) - f \to 0$ and $S_M(g) - g$ tend to zero in $\mathscr{S}$, L is continuous from $\mathscr{S}$ to $\mathscr{S}'$, and all Schwartz seminorms of $S_M(g)$ are bounded uniformly in M; see also Exercise 8.3.1.

The second term in (8.5.22) is $\big\langle S_{N-1} L S_{N-1}(f), g\big\rangle = \big\langle L S_{N-1}(f), S_{N-1}(g)\big\rangle$. Since $\widehat{f}$ is supported away from the origin, $S_N(f) \to 0$ in $\mathscr{S}$ as $N \to -\infty$; see Exercise 8.3.2. By the continuity of L, $L S_{N-1}(f) \to 0$ in $\mathscr{S}'$, and since all Schwartz seminorms of $S_{N-1}(g)$ are bounded uniformly in N, we conclude that the term $\langle LS_{N-1}(f), S_{N-1}(g)\rangle$ tends to zero as $N \to -\infty$. We thus deduce that $\widetilde{L}(f) = L(f)$.

It remains to prove (8.5.21). We now define

$$L_j = \Delta_j L S_j \qquad \text{and} \qquad L'_j = S_{j-1} L \Delta_j$$

for $j \in \mathbf{Z}$. In view of identity (2.3.21) and the convergence of the Riemann sums to the integral defining $f * \Phi_{2^{-j}}$ in the topology of $\mathscr{S}$ (see the proof of Theorem 2.3.20), we have

$$\big(L(f*\Phi_{2^{-j}}) * \Psi_{2^{-j}}\big)(x) = \int_{\mathbf{R}^n} \big\langle L(\tau^y(\Phi_{2^{-j}})), \tau^x(\Psi_{2^{-j}})\big\rangle f(y)\,dy,$$

where $\tau^y(g)(u) = g(u-y)$. Thus the kernel K_j of L_j is

$$K_j(x,y) = \big\langle L(\tau^y(\Phi_{2^{-j}})), \tau^x(\Psi_{2^{-j}})\big\rangle$$

and the kernel K'_j of L'_j is

$$K'_j(x,y) = \big\langle L(\tau^y(\Psi_{2^{-j}})), \tau^x(\Phi_{2^{-(j-1)}})\big\rangle .$$

We plan to prove that

$$|K_j(x,y)| + 2^{-j}|\nabla K_j(x,y)| \le C_n(A+B_4)2^{nj}(1+2^j|x-y|)^{-n-\delta}, \tag{8.5.23}$$

noting that an analogous estimate holds for $K'_j(x,y)$. Once (8.5.23) is established, Exercise 8.5.2 and the conditions

$$L_j(1) = \Delta_j L S_j(1) = \Delta_j L(1) = 0, \qquad L'_j(1) = S_{j-1} L \Delta_j(1) = 0,$$

yield the hypotheses of Proposition 8.5.3. Recalling (8.5.20), the conclusion of this proposition yields (8.5.21).

To prove (8.5.23) we quickly repeat the corresponding argument from the proof of Theorem 8.3.3. We consider the following two cases: If $|x-y| \le 5\cdot 2^{-j}$, then the weak boundedness property gives

$$\begin{aligned}\big|\langle L(\tau^y(\Phi_{2^{-j}})), \tau^x(\Psi_{2^{-j}})\rangle\big| &= \big|\langle L(\tau^x(\tau^{2^j(y-x)}(\Phi)_{2^{-j}})), \tau^x(\Psi_{2^{-j}})\rangle\big| \\ &\le C_n \big\|L\big\|_{WB} 2^{jn},\end{aligned}$$

since Ψ and $\tau^{2^j(y-x)}(\Phi)$, whose support is contained in $B(0,\frac{1}{2}) + B(0,5) \subseteq B(0,10)$, are multiples of normalized bumps. This proves the first of the two estimates in (8.5.23) when $|x-y| \le 5\cdot 2^{-j}$.

We now turn to the case $|x-y| \ge 5\cdot 2^{-j}$. Then the functions $\tau^y(\Phi_{2^{-j}})$ and $\tau^x(\Psi_{2^{-j}})$ have disjoint supports, and so we have the integral representation

$$K_j(x,y) = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_{2^{-j}}(v-y)K(u,v)\Psi_{2^{-j}}(u-x)\,du\,dv .$$

Using that Ψ has mean value zero, we can write the previous expression as

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_{2^{-j}}(v-y)\big(K(u,v)-K(x,v)\big)\Psi_{2^{-j}}(u-x)\,du\,dv\,.$$

We observe that $|u-x|\le 2^{-j}$ and $|v-y|\le 2^{-j}$ in the preceding integral. Since $|x-y|\ge 5\cdot 2^{-j}$, this makes $|u-v|\ge |x-y|-2\cdot 2^{-j}\ge 2\cdot 2^{-j}$, which implies that $|u-x|\le \frac{1}{2}|u-v|$. Using the regularity condition (8.1.2), we deduce

$$|K(u,v)-K(x,v)|\le A\frac{|x-u|^{\delta}}{|u-v|^{n+\delta}}\le C_{n,\delta}A\,\frac{2^{-j\delta}}{|x-y|^{n+\delta}}\,.$$

Inserting this estimate in the preceding double integral, we obtain the first estimate in (8.5.23). The second estimate in (8.5.23) is proved similarly. □

8.5.4 Pseudodifferential Operators

We now turn to another elegant application of Lemma 8.5.1 regarding pseudodifferential operators. We first introduce pseudodifferential operators.

Definition 8.5.5. Let $m\in\mathbf{R}$ and $0<\rho,\delta\le 1$. A $\mathscr{C}^\infty$ function $\sigma(x,\xi)$ on $\mathbf{R}^n\times\mathbf{R}^n$ is called a *symbol of class* $S^m_{\rho,\delta}$ if for all multi-indices α and β there is a constant $B_{\alpha,\beta}$ such that

$$|\partial_x^\alpha\partial_\xi^\beta\sigma(x,\xi)|\le B_{\alpha,\beta}(1+|\xi|)^{m-\rho|\beta|+\delta|\alpha|}\,. \tag{8.5.24}$$

For $\sigma\in S^m_{\rho,\delta}$, the linear operator

$$T_\sigma(f)(x)=\int_{\mathbf{R}^n}\sigma(x,\xi)\widehat{f}(\xi)e^{2\pi i x\cdot\xi}\,d\xi$$

initially defined for f in $\mathscr{S}(\mathbf{R}^n)$ is called a *pseudodifferential operator* with symbol $\sigma(x,\xi)$.

Example 8.5.6. The paraproduct P_b introduced in the previous section is a pseudodifferential operator with symbol

$$\sigma_b(x,\xi)=\sum_{j\in\mathbf{Z}}\Delta_j(b)(x)\widehat{\Psi}(2^{-j}\xi)\,. \tag{8.5.25}$$

It is not hard to see that the symbol σ_b satisfies

$$|\partial_x^\alpha\partial_\xi^\beta\sigma_b(x,\xi)|\le B_{\alpha,\beta}|\xi|^{-|\beta|+|\alpha|} \tag{8.5.26}$$

for all multi-indices α and β. Indeed, every differentiation in x produces a factor of 2^j, while every differentiation in ξ produces a factor of 2^{-j}. But since $\widehat{\Psi}$ is supported in $\frac{1}{2}\cdot 2^j\le|\xi|\le 2\cdot 2^j$, it follows that $|\xi|\approx 2^j$, which yields (8.5.26).

It follows that σ_b is not in any of the classes $S^m_{\rho,\delta}$ introduced in Definition 8.5.5. However, if we restrict the indices of summation in (8.5.25) to $j \geq 0$, then $|\xi| \approx 1+|\xi|$ and we obtain a symbol of class $S^0_{1,1}$. Note that not all symbols in $S^0_{1,1}$ give rise to bounded operators on L^2. See Exercise 8.5.6.

An example of a symbol in $S^m_{0,0}$ is $(1+|\xi|^2)^{\frac{1}{2}(m+it)}$ when $m,t \in \mathbf{R}$.

We do not plan to embark on a systematic study of pseudodifferential operators here, but we would like to study the L^2 boundedness of symbols of class $S^0_{0,0}$.

Theorem 8.5.7. *Suppose that a symbol σ belongs to the class $S^0_{0,0}$. Then the pseudodifferential operator T_σ with symbol σ, initially defined on $\mathscr{S}(\mathbf{R}^n)$, has a bounded extension on $L^2(\mathbf{R}^n)$.*

Proof. In view of Plancherel's theorem, it suffices to obtain the L^2 boundedness of the linear operator

$$\widetilde{T_\sigma}(f)(x) = \int_{\mathbf{R}^n} \sigma(x,\xi) f(\xi) e^{2\pi i x\cdot\xi}\, d\xi\,. \tag{8.5.27}$$

We fix a nonnegative smooth function $\varphi(\xi)$ supported in a small multiple of the unit cube $Q_0=[0,1]^n$ (say in $[-\frac{1}{9},\frac{10}{9}]^n$) that satisfies

$$\sum_{j\in\mathbf{Z}^n} \varphi(x-j) = 1\,, \qquad x\in\mathbf{R}^n\,. \tag{8.5.28}$$

For $j,k\in\mathbf{Z}^n$ we define symbols

$$\sigma_{j,k}(x,\xi) = \varphi(x-j)\sigma(x,\xi)\varphi(\xi-k)$$

and corresponding operators $T_{j,k}$ given by (8.5.27) in which $\sigma(x,\xi)$ is replaced by $\sigma_{j,k}(x,\xi)$. Using (8.5.28), we obtain that

$$\widetilde{T_\sigma} = \sum_{j,k\in\mathbf{Z}^n} T_{j,k}\,,$$

where the double sum is easily shown to converge in the topology of $\mathscr{S}(\mathbf{R}^n)$. Our goal is to show that for all $N\in\mathbf{Z}^+$ we have

$$\big\|T^*_{j,k}T_{j',k'}\big\|_{L^2\to L^2} \leq C_N(1+|j-j'|+|k-k'|)^{-2N}\,, \tag{8.5.29}$$

$$\big\|T_{j,k}T^*_{j',k'}\big\|_{L^2\to L^2} \leq C_N(1+|j-j'|+|k-k'|)^{-2N}\,, \tag{8.5.30}$$

where C_N depends on N and n but is independent of j,j',k,k'.

We note that

$$T^*_{j,k}T_{j',k'}(f)(x) = \int_{\mathbf{R}^n} K_{j,k,j',k'}(x,y) f(y)\, dy\,,$$

where

$$K_{j,k,j',k'}(x,y) = \int_{\mathbf{R}^n} \overline{\sigma_{j,k}(z,x)}\sigma_{j',k'}(z,y)e^{2\pi i(y-x)\cdot z}\,dz. \tag{8.5.31}$$

We integrate by parts in (8.5.31) using the identity

$$e^{2\pi i z\cdot(y-x)} = \frac{(I-\Delta_z)^N(e^{2\pi i z\cdot(y-x)})}{(1+4\pi^2|x-y|^2)^N},$$

and we obtain the pointwise estimate

$$\frac{\varphi(x-k)\varphi(y-k')}{(1+4\pi^2|x-y|^2)^N}\left|(I-\Delta_z)^N(\varphi(z-j)\overline{\sigma(z,x)}\sigma(z,y)\varphi(z-j'))\right|$$

for the integrand in (8.5.31). The support property of φ forces $|j-j'| \le c_n$ for some dimensional constant c_n; indeed, $c_n = 2\sqrt{n}$ suffices. Moreover, all derivatives of σ and φ are controlled by constants, and φ is supported in a cube of finite measure. We also have $1+|x-y| \approx 1+|k-k'|$. It follows that

$$|K_{j,k,j',k'}(x,y)| \le \begin{cases} \dfrac{C_N\varphi(x-k)\varphi(y-k')}{(1+|k-k'|)^{2N}} & \text{when } |j-j'| \le c_n, \\ 0 & \text{otherwise.} \end{cases}$$

We can rewrite the preceding estimates in a more compact (and symmetric) form as

$$|K_{j,k,j',k'}(x,y)| \le \frac{C_{n,N}\varphi(x-k)\varphi(y-k')}{(1+|j-j'|+|k-k'|)^{2N}},$$

from which we easily obtain that

$$\sup_{x\in\mathbf{R}^n}\int_{\mathbf{R}^n}|K_{j,k,j',k'}(x,y)|\,dy \le \frac{C_{n,N}}{(1+|j-j'|+|k-k'|)^{2N}}, \tag{8.5.32}$$

$$\sup_{y\in\mathbf{R}^n}\int_{\mathbf{R}^n}|K_{j,k,j',k'}(x,y)|\,dx \le \frac{C_{n,N}}{(1+|j-j'|+|k-k'|)^{2N}}. \tag{8.5.33}$$

Using the classical Schur lemma in Appendix I.1, we obtain that

$$\left\|T_{j,k}^*T_{j',k'}\right\|_{L^2\to L^2} \le \frac{C_{n,N}}{(1+|j-j'|+|k-k'|)^{2N}},$$

which proves (8.5.29). Since $\rho = \delta = 0$, the roles of the variables x and ξ are symmetric, and (8.5.30) can be proved in exactly the same way as (8.5.29). The almost orthogonality Lemma 8.5.1 now applies, since

$$\sum_{j,k\in\mathbf{Z}^n}\sqrt{\frac{1}{(1+|j|+|k|)^{2N}}} \le \sum_{j\in\mathbf{Z}^n}\sum_{k\in\mathbf{Z}^n}\frac{1}{(1+|j|)^{\frac{N}{2}}}\frac{1}{(1+|k|)^{\frac{N}{2}}} < \infty$$

for $N \ge 2n+2$, and the boundedness of $\widetilde{T_\sigma}$ on L^2 follows. □

Remark 8.5.8. The reader may want to check that the argument in Theorem 8.5.7 is also valid for symbols of the class $S^0_{\rho,\rho}$ whenever $0<\rho<1$.

Exercises

8.5.1. Prove that any bounded linear operator $S:\ H\to H$ satisfies

$$\big\|S\big\|^2_{H\to H}=\big\|SS^*\big\|_{H\to H}.$$

8.5.2. Show that if a family of kernels K_j satisfy (8.5.10) and

$$|\nabla_x K_j(x,y)|+|\nabla_y K_j(x,y)|\le \frac{A2^{(n+1)j}}{(1+2^j|x-y|)^{n+\delta}}$$

for all $x,y\in\mathbf{R}^n$, then conditions (8.5.11) and (8.5.12) hold with $\gamma=1$.

8.5.3. Prove the boundedness of the Hilbert transform using Lemma 8.5.1 and without using the Fourier transform.
[*Hint:* Pick a smooth function η supported in $[1/2,2]$ such that $\sum_{j\in\mathbf{Z}}\eta(2^{-j}x)=1$ for $x\neq 0$ and set $K_j(x)=x^{-1}\eta(2^{-j}|x|)$ and $H_j(f)=f*K_j$. Note that $H_j^*=-H_j$. Estimate $\big\|H_kH_j\big\|_{L^2\to L^2}$ by $\big\|K_k*K_j\big\|_{L^1}\le\big\|K_k*K_j\big\|_{L^\infty}|\operatorname{supp}(K_k*K_j)|$. When $j<k$, use the mean value property of K_j and that $\big\|K_k'\big\|_{L^\infty}\le C2^{-2k}$ to obtain that $\big\|K_k*K_j\big\|_{L^\infty}\le C2^{-2k+j}$. Conclude that $\big\|H_kH_j\big\|_{L^2\to L^2}\le C2^{-|j-k|}$.]

8.5.4. For a symbol $\sigma(x,\xi)$ in $S^0_{1,0}$, let $k(x,z)$ denote the inverse Fourier transform (evaluated at z) of the function $\sigma(x,\cdot)$ with x fixed. Show that for all $x\in\mathbf{R}^n$, the distribution $k(x,\cdot)$ coincides with a smooth function away from the origin in $\mathbf{R}^n$ that satisfies the estimates

$$|\partial_x^\alpha\partial_z^\beta k(x,z)|\le C_{\alpha,\beta}|z|^{-n-|\beta|},$$

and conclude that the kernels $K(x,y)=k(x,x-y)$ are well defined and smooth functions away from the diagonal in $\mathbf{R}^{2n}$ that belong to $SK(1,A)$ for some $A>0$. Conclude that pseudodifferential operators with symbols in $S^0_{1,0}$ are associated with standard kernels.
[*Hint:* Consider the distribution $(\partial^\gamma\sigma(x,\cdot))^\vee=(-2\pi iz)^\gamma k(x,\cdot)$. Since $\partial_\xi^\gamma\sigma(x,\xi)$ is integrable in ξ when $|\gamma|\ge n+1$, it follows that $k(x,\cdot)$ coincides with a smooth function on $\mathbf{R}^n\setminus\{0\}$. Next, set $\sigma_j(x,\xi)=\sigma(x,\xi)\widehat{\Psi}(2^{-j}\xi)$, where Ψ is as in Section 8.4 and k_j the inverse Fourier transform of σ_j in z. For $|\gamma|=M$ use that

$$(-2\pi iz)^\gamma\partial_x^\alpha\partial_\xi^\beta k_j(x,z)=\int_{\mathbf{R}^n}\partial_\xi^\gamma\big((2\pi i\xi)^\beta\partial_x^\alpha\sigma_j(x,\xi)\big)2^{2\pi i\xi\cdot z}\,d\xi$$

to obtain $|\partial_x^\alpha\partial_\xi^\beta k_j(x,z)|\le B_{M,\alpha,\beta}2^{jn}2^{j|\alpha|}(2^jn|z|)^{-M}$ and sum over $j\in\mathbf{Z}$.]

8.5.5. Prove that pseudodifferential operators with symbols in $S^0_{1,0}$ that have compact support in x are elements of $CZO(1,A,B)$ for some $A,B>0$.
[*Hint:* Write

$$T_\sigma(f)(x) = \int_{\mathbf{R}^n}\left(\int_{\mathbf{R}^n}\widehat{\sigma}(a,\xi)\widehat{f}(\xi)e^{2\pi i x\cdot\xi}\,d\xi\right)e^{2\pi i x\cdot a}\,da\,,$$

where $\widehat{\sigma}(a,\xi)$ denotes the Fourier transform of $\sigma(x,\xi)$ in the variable x. Use integration by parts to obtain $\sup_\xi|\widehat{\sigma}(a,\xi)|\le C_N(1+|a|)^{-N}$ and pass the L^2 norm inside the integral in a to obtain the required conclusion using the translation-invariant case.]

8.5.6. Let $\widehat{\eta}(\xi)$ be a smooth bump on $\mathbf{R}$ that is supported in $2^{-\frac{1}{2}}\le|\xi|\le 2^{\frac{1}{2}}$ and is equal to 1 on $2^{-\frac{1}{4}}\le|\xi|\le 2^{\frac{1}{4}}$. Let

$$\sigma(x,\xi) = \sum_{k=1}^{\infty} e^{-2\pi i 2^k x}\widehat{\eta}(2^{-k}\xi)\,.$$

Show that σ is an element of $S^0_{1,1}$ on the line but the corresponding pseudodifferential operator T_σ is not L^2 bounded.
[*Hint:* To see the latter statement, consider the sequence of functions $f_N(x) = \sum_{k=5}^{N}\frac{1}{k}e^{2\pi i 2^k x}h(x)$, where $h(x)$ is a Schwartz function whose Fourier transform is supported in the set $|\xi|\le\frac{1}{4}$. Show that $\|f_N\|_{L^2}\le C\|h\|_{L^2}$ but $\|T_\sigma(f_N)\|_{L^2}\ge c\log N\|h\|_{L^2}$ for some positive constants c,C.]

8.5.7. Prove conclusions (i) and (ii) of Lemma 8.5.1 if hypothesis (8.5.1) is replaced by

$$\|T_j^*T_k\|_{H\to H} + \|T_jT_k^*\|_{H\to H}\le\Gamma(j,k)\,,$$

where Γ is a nonnegative function on $\mathbf{Z}\times\mathbf{Z}$ such that

$$\sup_j\sum_{k\in\mathbf{Z}}\sqrt{\Gamma(j,k)} = A<\infty\,.$$

8.5.8. Let $\{T_t\}_{t\in\mathbf{R}^+}$ be a family of operators mapping a Hilbert space H to itself. Assume that there is a function $\gamma:\mathbf{R}^+\times\mathbf{R}^+\to\mathbf{R}^+\cup\{0\}$ satisfying

$$A_\gamma = \sup_{t>0}\int_0^\infty\sqrt{\gamma(t,s)}\,\frac{ds}{s}<\infty$$

such that

$$\|T_t^*T_s\|_{H\to H} + \|T_tT_s^*\|_{H\to H}\le\gamma(t,s)$$

for all t,s in $\mathbf{R}^+$. [An example of a function with $A_\gamma<\infty$ is $\gamma(t,s)=\min\left(\frac{s}{t},\frac{t}{s}\right)^\varepsilon$ for some $\varepsilon>0$.] Then prove that for all $0<\varepsilon<N$ we have

$$\left\|\int_\varepsilon^N T_t\,\frac{dt}{t}\right\|_{H\to H}\le A_\gamma\,.$$

8.6 The Cauchy Integral of Calderón and the $T(b)$ Theorem

The Cauchy integral is almost as old as complex analysis itself. In the classical theory of complex analysis, if Γ is a curve in $\mathbf{C}$ and f is a function on the curve, the Cauchy integral of f is given by

$$\frac{1}{2\pi i}\int_{\Gamma}\frac{f(\zeta)}{\zeta-z}\,d\zeta\,.$$

One situation in which this operator appears is the following: If Γ is a closed simple curve (i.e., a Jordan curve), Ω_+ is the interior connected component of $\mathbf{C}\setminus\Gamma$, Ω_- is the exterior connected component of $\mathbf{C}\setminus\Gamma$, and f is a smooth complex function on Γ, is it possible to find analytic functions F_+ on Ω_+ and F_- on Ω_-, respectively, that have continuous extensions on Γ such that their difference is equal to the given f on Γ? It turns out that a solution of this problem is given by

$$F_+(w)=\frac{1}{2\pi i}\int_{\Gamma}\frac{f(\zeta)}{\zeta-w}\,d\zeta\,,\qquad w\in\Omega_+\,,$$

and

$$F_-(w)=\frac{1}{2\pi i}\int_{\Gamma}\frac{f(\zeta)}{\zeta-w}\,d\zeta\,,\qquad w\in\Omega_-\,.$$

We are would like to study the case in which the Jordan curve Γ passes through infinity, in particular, when it is the graph of a Lipschitz function on $\mathbf{R}$. In this case we compute the boundary limits of F_+ and F_- and we see that they give rise to a very interesting operator on the curve Γ. To fix notation we let

$$A:\ \mathbf{R}\to\mathbf{R}$$

be a Lipschitz function. This means that there is a constant $L>0$ such that for all $x,y\in\mathbf{R}$ we have $|A(x)-A(y)|\le L|x-y|$. We define a curve

$$\gamma:\ \mathbf{R}\to\mathbf{C}$$

by setting

$$\gamma(x)=x+iA(x)$$

and we denote by

$$\Gamma=\{\gamma(x):\ x\in\mathbf{R}\}\tag{8.6.1}$$

the graph of γ. Given a smooth function f on Γ we set

$$F(w)=\frac{1}{2\pi i}\int_{\Gamma}\frac{f(\zeta)}{\zeta-w}\,d\zeta\,,\qquad w\in\mathbf{C}\setminus\Gamma\,.\tag{8.6.2}$$

We now show that for $z\in\Gamma$, both $F(z+i\delta)$ and $F(z-i\delta)$ have limits as $\delta\downarrow 0$, and these limits give rise to an operator on the curve Γ that we would like to study.

8.6.1 Introduction of the Cauchy Integral Operator along a Lipschitz Curve

For a smooth function f on the curve Γ and $z \in \Gamma$ we define the *Cauchy integral of f at z* as

$$\mathfrak{C}_\Gamma(f)(z) = \lim_{\varepsilon \to 0+} \frac{1}{\pi i} \int\limits_{\substack{\zeta \in \Gamma \\ |\operatorname{Re}\zeta - \operatorname{Re} z| > \varepsilon}} \frac{f(\zeta)}{\zeta - z}\, d\zeta\,, \tag{8.6.3}$$

assuming that $f(\zeta)$ has some decay as $|\zeta| \to \infty$. The latter assumption makes the integral in (8.6.3) converge when $|\operatorname{Re}\zeta - \operatorname{Re} z| \geq 1$. The fact that the limit in (8.6.3) exists as $\varepsilon \to 0$ for almost all $z \in \Gamma$ is shown in the next proposition.

Proposition 8.6.1. *Let Γ be as in (8.6.1). Let $f(\zeta)$ be a smooth function on Γ that has decay as $|\zeta| \to \infty$. Given f, we define a function F as in (8.6.2) related to f. Then the limit in (8.6.3) exists as $\varepsilon \to 0$ for almost all $z \in \Gamma$ and gives rise to a well defined operator $\mathfrak{C}_\Gamma(f)$ acting on such functions f. Moreover, for almost all $z \in \Gamma$ we have that*

$$\lim_{\delta \downarrow 0} F(z + i\delta) = \frac{1}{2}\mathfrak{C}_\Gamma(f)(z) - \frac{1}{2} f(z)\,, \tag{8.6.4}$$

$$\lim_{\delta \downarrow 0} F(z - i\delta) = \frac{1}{2}\mathfrak{C}_\Gamma(f)(z) + \frac{1}{2} f(z)\,. \tag{8.6.5}$$

Proof. We show first that the limit in (8.6.3) exists as $\varepsilon \to 0$. For $z \in \Gamma$ and $0 < \varepsilon < 1$ we write

$$\begin{aligned}\frac{1}{\pi i} \int\limits_{\substack{\zeta \in \Gamma \\ |\operatorname{Re}\zeta - \operatorname{Re} z| > \varepsilon}} \frac{f(\zeta)\, d\zeta}{\zeta - z} = {} & \frac{1}{\pi i} \int\limits_{\substack{\zeta \in \Gamma \\ |\operatorname{Re}\zeta - \operatorname{Re} z| > 1}} \frac{f(\zeta)\, d\zeta}{\zeta - z} \\ & + \frac{1}{\pi i} \int\limits_{\substack{\zeta \in \Gamma \\ \varepsilon \leq |\operatorname{Re}\zeta - \operatorname{Re} z| \leq 1}} \frac{(f(\zeta) - f(z))\, d\zeta}{\zeta - z} \\ & + \frac{f(z)}{\pi i} \int\limits_{\substack{\zeta \in \Gamma \\ \varepsilon \leq |\operatorname{Re}\zeta - \operatorname{Re} z| \leq 1}} \frac{d\zeta}{\zeta - z}\,.\end{aligned} \tag{8.6.6}$$

By the smoothness of f, the middle term of the sum in (8.6.6) has a limit as $\varepsilon \to 0$. We therefore study the third (last) term of this sum.

We consider two branches of the complex logarithm: first $\log_{upper}(z)$ defined for z in $\mathbf{C} \setminus \{0\}$ minus the negative imaginary axis normalized so that $\log_{upper}(1) = 0$; this logarithm satisfies $\log_{upper}(i) = \frac{\pi i}{2}$ and $\log_{upper}(-1) = \pi i$. Second, $\log_{lower}(z)$ defined for z in $\mathbf{C} \setminus \{0\}$ minus the positive imaginary axis normalized so that $\log_{lower}(1) = 0$; this logarithm satisfies $\log_{lower}(-i) = -\frac{\pi i}{2}$ and $\log_{lower}(-1) = -\pi i$.

Let $\tau = \operatorname{Re} z$ and $t = \operatorname{Re}\zeta$; then $z = \gamma(\tau) = \tau + iA(\tau)$ and $\zeta = \gamma(t)$. The function A is Lipschitz and thus differentiable almost everywhere; consequently, the function $\gamma(\tau) = \tau + iA(\tau)$ is differentiable a.e. in $\tau \in \mathbf{R}$. Moreover, $\gamma'(\tau) = 1 + iA'(\tau) \neq 0$ whenever γ is differentiable at τ. Fix a $\tau = \operatorname{Re} z$ at which γ is differentiable.

We rewrite the last term in the sum in (8.6.6) as

$$\int_{\varepsilon}^{1} \frac{\gamma'(t)}{\gamma(t+\tau)-\gamma(\tau)}\,dt + \int_{-1}^{-\varepsilon} \frac{\gamma'(t)}{\gamma(t+\tau)-\gamma(\tau)}\,dt\,. \tag{8.6.7}$$

The curve $t \mapsto \gamma(t+\tau)-\gamma(\tau) = t + i(A(t+\tau)-A(\tau))$ lies in the complex plane minus a small angle centered at the origin that does not contain the negative imaginary axis. Using the upper branch of the logarithm, we evaluate (8.6.7) as

$$\begin{aligned}
&\frac{f(z)}{\pi i}\Big[\log_{upper}\big(\gamma(1+\tau)-\gamma(\tau)\big) - \log_{upper}\big(\gamma(\varepsilon+\tau)-\gamma(\tau)\big)\\
&\qquad\qquad - \log_{upper}\big(\gamma(-1+\tau)-\gamma(\tau)\big) + \log_{upper}\big(\gamma(-\varepsilon+\tau)-\gamma(\tau)\big)\Big]\\
&\quad = \log_{upper}\big(\gamma(\tau-\varepsilon)-\gamma(\tau)\big) - \log_{upper}\big(\gamma(\varepsilon+\tau)-\gamma(\tau)\big)\\
&\quad = \log_{upper} \frac{\dfrac{\gamma(\tau-\varepsilon)-\gamma(\tau)}{\varepsilon}}{\dfrac{\gamma(\varepsilon+\tau)-\gamma(\tau)}{\varepsilon}}\,.
\end{aligned}$$

This expression converges to $\log_{upper}\big(-\frac{\gamma'(\tau)}{\gamma'(\tau)}\big) = \log_{upper}(-1) = i\pi$ as $\varepsilon \to 0$. Thus the limit in (8.6.6), and hence in (8.6.3), exists as $\varepsilon \to 0$ for almost all z on the curve. Hence $\mathfrak{C}_\Gamma(f)$ is a well defined operator whenever f is a smooth function with decay at infinity.

We proceed with the proof of (8.6.4). For fixed $\delta > 0$ and $0 < \varepsilon < 1$ we write

$$\begin{aligned}
F(z+i\delta) &= \frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma\\ |\operatorname{Re}\zeta-\operatorname{Re}z|>\varepsilon}} \frac{f(\zeta)}{\zeta-z-i\delta}\,d\zeta\\
&\qquad + \frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma\\ |\operatorname{Re}\zeta-\operatorname{Re}z|\le\varepsilon}} \frac{f(\zeta)-f(z)}{\zeta-z-i\delta}\,d\zeta\\
&\qquad + f(z)\frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma\\ |\operatorname{Re}\zeta-\operatorname{Re}z|\le\varepsilon}} \frac{1}{\zeta-z-i\delta}\,d\zeta\,.
\end{aligned} \tag{8.6.8}$$

With $\tau = \operatorname{Re} z$, the last term in the sum in (8.6.8) is equal to

$$\int_{\varepsilon}^{1} \frac{\gamma'(t)}{\gamma(t+\tau)-(\gamma(\tau)+i\delta)}\,dt + \int_{-1}^{-\varepsilon} \frac{\gamma'(t)}{\gamma(t+\tau)-(\gamma(\tau)+i\delta)}\,dt\,. \tag{8.6.9}$$

Since $\delta > 0$, the curve $\gamma(t+\tau) - (\gamma(\tau) + i\delta)$ lies below the curve $t \mapsto \gamma(t+\tau) - \gamma(\tau)$ and therefore outside a small angle centered at the origin that does not contain the positive imaginary axis. In this region, $\log_{lower}$ is an analytic branch of the logarithm. Evaluation of (8.6.9) yields

$$\frac{f(z)}{2\pi i} \log_{lower} \frac{\gamma(\varepsilon+\tau) - \gamma(\tau) - i\delta}{\gamma(-\varepsilon+\tau) - \gamma(\tau) - i\delta}.$$

So, taking limits as $\delta \downarrow 0$ in (8.6.8), we obtain that

$$\begin{aligned} \lim_{\delta\downarrow 0} F(z+i\delta) &= \frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma \\ |\mathrm{Re}\,\zeta - \mathrm{Re}\,z|>\varepsilon}} \frac{f(\zeta)}{\zeta - z}\, d\zeta \\ &+ \frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma \\ |\mathrm{Re}\,\zeta - \mathrm{Re}\,z|\le\varepsilon}} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta + \frac{f(z)}{2\pi i} \log_{lower} \frac{\gamma(\tau+\varepsilon) - \gamma(\tau)}{\gamma(\tau-\varepsilon) - \gamma(\tau)}, \end{aligned} \tag{8.6.10}$$

in which $z = \gamma(\tau) = \tau + iA(\tau)$ and both integrals converge absolutely.

Up until this point, $\varepsilon \in (0,1)$ was arbitrary and we may let it tend to zero. In doing so we first observe that the middle integral in (8.6.10) tends to zero because of the smoothness of f. But for almost all $\tau \in \mathbf{R}$, the limit as $\varepsilon \to 0$ of the logarithm in (8.6.10) is equal to $\log_{lower}(-\frac{\gamma'(\tau)}{\gamma'(\tau)}) = \log_{lower}(-1) = -\pi i$. From this we conclude that for almost all $z \in \Gamma$ we have

$$\lim_{\delta\downarrow 0} F(z+i\delta) = \lim_{\varepsilon\to 0} \frac{1}{2\pi i} \int\limits_{\substack{\zeta\in\Gamma \\ |\mathrm{Re}\,\zeta - \mathrm{Re}\,z|>\varepsilon}} \frac{f(\zeta)}{\zeta - z}\, d\zeta + f(z) \frac{1}{2\pi i}(-\pi i), \tag{8.6.11}$$

which proves (8.6.4).

The only difference in the proof of (8.6.5) is that $\log_{upper}$ is replaced by $\log_{lower}$, and for this reason $(-\pi i)$ should be replaced by πi in (8.6.11). □

Remark 8.6.2. If we let F_+ be the restriction of F on the region above the graph Γ and let F_- be the restriction of F on the region below the graph Γ, we have that F_+ and F_- have continuous extensions on Γ, and moreover,

$$F_+ - F_- = -f,$$

where f is the given smooth function on the curve. We also note that the argument given in Proposition 8.6.1 does not require f to be smoother than $\mathscr{C}^1$.

8.6.2 Resolution of the Cauchy Integral and Reduction of Its L^2 Boundedness to a Quadratic Estimate

Having introduced the Cauchy integral $\mathfrak{C}_\Gamma$ as an operator defined on smooth functions on the graph Γ of a Lipschitz function A, we turn to some of its properties. We are mostly interested in obtaining an a priori L^2 estimate for $\mathfrak{C}_\Gamma$. Before we achieve this goal, we make some observations. First we can write $\mathfrak{C}_\Gamma$ as

$$\mathfrak{C}_\Gamma(H)(x+iA(x)) = \lim_{\varepsilon\to 0}\frac{1}{\pi i}\int\limits_{|x-y|>\varepsilon}\frac{H(y+iA(y))(1+iA'(y))}{y+iA(y)-x-iA(x)}\,dy\,, \tag{8.6.12}$$

where the integral is over the real line and H is a function on the curve Γ. (Recall that Lipschitz functions are differentiable almost everywhere.) To any function H on Γ we can associate a function h on the line $\mathbf{R}$ by setting

$$h(y) = H(y+iA(y))\,.$$

We have that

$$\int_\Gamma |H(y)|^2\,dy = \int_{\mathbf{R}} |h(y)|^2(1+|A'(y)|^2)^{\frac{1}{2}}\,dy \approx \int_{\mathbf{R}} |h(y)|^2\,dy$$

for some constants that depend on the Lipschitz constant L of A. Therefore, the boundedness of the operator in (8.6.12) is equivalent to that of the operator

$$\mathcal{C}_\Gamma(h)(x) = \lim_{\varepsilon\to 0}\frac{1}{\pi i}\int\limits_{|x-y|>\varepsilon}\frac{h(y)(1+iA'(y))}{y-x+i(A(y)-A(x))}\,dy \tag{8.6.13}$$

acting on Schwartz functions h on the line. It is this operator that we concentrate on in the remainder of this section. We recall that (see Example 8.1.6) the function

$$\frac{1}{y-x+i(A(y)-A(x))}$$

defined on $\mathbf{R}\times\mathbf{R}\setminus\{(x,x):\ x\in\mathbf{R}\}$ is a standard kernel in $SK(1,cL)$ for some $c>0$. We note that this is not the case with the kernel

$$\frac{1+iA'(y)}{y-x+i(A(y)-A(x))}\,, \tag{8.6.14}$$

for conditions (8.1.2) and (8.1.3) fail for this kernel, since the function $1+iA'$ does not possess any smoothness. [Condition (8.1.1) trivially holds for the function in (8.6.14).] We note, however, that the L^p boundedness of the operator in (8.6.13) is equivalent to that of

$$\widetilde{\mathcal{C}}_\Gamma(h)(x) = \lim_{\varepsilon\to 0} \frac{1}{\pi i} \int\limits_{|x-y|>\varepsilon} \frac{h(y)}{y-x+i(A(y)-A(x))}\,dy\,, \tag{8.6.15}$$

since the function $1+iA'$ is bounded above and below and can be absorbed in h. Therefore, the L^2 boundedness of $\mathcal{C}_\Gamma$ is equivalent to that of $\widetilde{\mathcal{C}}_\Gamma$, which has a kernel that satisfies standard estimates. This equivalence, however, is not as useful in the approach we take in the sequel. We choose to work with the operator $\mathcal{C}_\Gamma$, in which the appearance of the term $1+iA'(y)$ plays a crucial cancellation role.

In the proof of Theorem 8.3.3 we used a *resolution* of an operator T with standard kernel of the form

$$\int_0^\infty P_s T_s Q_s \,\frac{ds}{s}\,,$$

where P_s and Q_s are nice averaging operators that approximate the identity and the zero operator, respectively. Our goal is to achieve a similar resolution for the operator $\mathcal{C}_\Gamma$ defined in (8.6.13). To achieve this, for every $s>0$ we introduce the auxiliary operator

$$\mathcal{C}_\Gamma(h)(x;s) = \frac{1}{\pi i} \int\limits_{\mathbf{R}} \frac{h(y)(1+iA'(y))}{y-x+i(A(y)-A(x))+is}\,dy \tag{8.6.16}$$

defined for Schwartz functions h on the line. We make two preliminary observations regarding this operator: For almost all $x\in\mathbf{R}$ we have

$$\lim_{s\to\infty} \mathcal{C}_\Gamma(h)(x;s) \;=\; 0\,, \tag{8.6.17}$$

$$\lim_{s\to 0} \mathcal{C}_\Gamma(h)(x;s) \;=\; \mathcal{C}_\Gamma(h)(x)+h(x)\,. \tag{8.6.18}$$

Identity (8.6.17) is trivial. To obtain (8.6.18), for a fixed $\varepsilon>0$ we write

$$\begin{aligned}\mathcal{C}_\Gamma(h)(x;s) \;&= \frac{1}{\pi i} \int\limits_{|x-y|>\varepsilon} \frac{h(y)(1+iA'(y))}{y-x+i(A(y)-A(x))+is}\,dy \\ &\qquad + \frac{1}{\pi i} \int\limits_{|x-y|\le\varepsilon} \frac{(h(y)-h(x))(1+iA'(y))}{y-x+i(A(y)-A(x))+is}\,dy \\ &\qquad + h(x)\,\frac{1}{\pi i}\log_{upper} \frac{\varepsilon+i(A(x+\varepsilon)-A(x))+is}{-\varepsilon+i(A(x-\varepsilon)-A(x))+is}\,,\end{aligned} \tag{8.6.19}$$

where $\log_{upper}$ denotes the analytic branch of the complex logarithm defined in the proof of Proposition 8.6.1. We used this branch of the logarithm, since for $s>0$, the graph of the function $y\mapsto y+i(A(y+x)-A(x))+is$ lies outside a small angle centered at the origin that contains the negative imaginary axis.

We now take successive limits first as $s\to 0$ and then as $\varepsilon\to 0$ in (8.6.19). We obtain that

$$\lim_{s\to 0}\mathcal{C}_\Gamma(h)(x;s) = \lim_{\varepsilon\to 0}\frac{1}{\pi i}\int\limits_{|x-y|>\varepsilon}\frac{h(y)(1+iA'(y))}{y-x+i(A(y)-A(x))}\,dy + h(x)\lim_{\varepsilon\to 0}\frac{1}{\pi i}\log_{upper}\frac{\varepsilon+i(A(x+\varepsilon)-A(x))}{-\varepsilon+i(A(x-\varepsilon)-A(x))}\,.$$

Since this expression inside the logarithm tends to -1 as $\varepsilon\to 0$, this logarithm tends to πi, and this concludes the proof of (8.6.18).

We now consider the second derivative in s of the auxiliary operator $\mathcal{C}_\Gamma(h)(x;s)$.

$$\begin{aligned}
&\int_0^\infty s^2\frac{d^2}{ds^2}\mathcal{C}_\Gamma(h)(x;s)\,\frac{ds}{s}\\
&= \int_0^\infty s\frac{d^2}{ds^2}\mathcal{C}_\Gamma(h)(x;s)\,ds\\
&= \lim_{s\to\infty} s\frac{d}{ds}\mathcal{C}_\Gamma(h)(x;s) - \lim_{s\to 0} s\frac{d}{ds}\mathcal{C}_\Gamma(h)(x;s) - \int_0^\infty \frac{d}{ds}\mathcal{C}_\Gamma(h)(x;s)\,ds\\
&= 0-0+\lim_{s\to 0}\mathcal{C}_\Gamma(h)(x;s) - \lim_{s\to\infty}\mathcal{C}_\Gamma(h)(x;s)\\
&= \mathcal{C}_\Gamma(h)(x)+h(x)\,,
\end{aligned}$$

where we used integration by parts, the fact that for almost all $x\in\mathbf{R}$ we have

$$\lim_{s\to\infty} s\frac{d}{ds}\mathcal{C}_\Gamma(h)(x;s) = \lim_{s\to 0} s\frac{d}{ds}\mathcal{C}_\Gamma(h)(x;s) = 0\,, \tag{8.6.20}$$

and identities (8.6.17) and (8.6.18) whenever h is a Schwartz function. One may consult Exercise 8.6.2 for a proof of the identities in (8.6.20). So we have succeeded in writing the operator $\mathcal{C}_\Gamma(h)+h$ as an average of smoother operators. Precisely, we have shown that for $h\in\mathscr{S}(\mathbf{R})$ we have

$$\mathcal{C}_\Gamma(h)(x)+h(x) = \int_0^\infty s^2\frac{d^2}{ds^2}\mathcal{C}_\Gamma(h)(x;s)\,\frac{ds}{s}\,, \tag{8.6.21}$$

and it remains to understand what the operator

$$\frac{d^2}{ds^2}\mathcal{C}_\Gamma(h)(x;s) = \mathcal{C}_\Gamma(h)''(x;s)$$

really is. Differentiating (8.6.16) twice, we obtain

$$\begin{aligned}
\mathcal{C}_\Gamma(h)(x)+h(x) &= \int_0^\infty s^2\mathcal{C}_\Gamma(h)''(x;s)\,\frac{ds}{s}\\
&= 4\int_0^\infty s^2\mathcal{C}_\Gamma(h)''(x;2s)\,\frac{ds}{s}\\
&= -\frac{8}{\pi i}\int_0^\infty\int_{\mathbf{R}}\frac{s^2h(y)(1+iA'(y))}{(y-x+i(A(y)-A(x))+2is)^3}\,dy\,\frac{ds}{s}\\
&= -\frac{8}{\pi i}\int_0^\infty\int_\Gamma\frac{s^2H(\zeta)}{(\zeta-z+2is)^3}\,d\zeta\,\frac{ds}{s}\,,
\end{aligned}$$

where in the last step we set $z = x + iA(x)$, $H(z) = h(x)$, and we switched to complex integration over the curve Γ. We now use the following identity from complex analysis. For $\zeta, z \in \Gamma$ we have

$$\frac{1}{(\zeta - z + 2is)^3} = -\frac{1}{4\pi i}\int_\Gamma \frac{1}{(\zeta - w + is)^2}\frac{1}{(w - z + is)^2}\,dw, \tag{8.6.22}$$

for which we refer to Exercise 8.6.3. Inserting this identity in the preceding expression for $\mathcal{C}_\Gamma(h)(x) + h(x)$, we obtain

$$\mathcal{C}_\Gamma(h)(x) + h(x) = -\frac{2}{\pi^2}\int_0^\infty \left[\int_\Gamma \frac{s}{(w - z + is)^2}\left(\int_\Gamma \frac{s\,H(\zeta)}{(\zeta - w + is)^2}\,d\zeta\right)dw\right]\frac{ds}{s},$$

recalling that $z = x + iA(x)$. Introducing the linear operator

$$\Theta_s(h)(x) = \int_{\mathbf{R}} \theta_s(x,y)\,h(y)\,dy, \tag{8.6.23}$$

where

$$\theta_s(x,y) = \frac{s}{(y - x + i(A(y) - A(x)) + is)^2}, \tag{8.6.24}$$

we may therefore write

$$\mathcal{C}_\Gamma(h)(x) + h(x) = -\frac{2}{\pi^2}\int_0^\infty \Theta_s\big((1 + iA')\Theta_s\big((1 + iA')h\big)\big)(x)\,\frac{ds}{s}. \tag{8.6.25}$$

We also introduce the multiplication operator

$$M_b(h) = b\,h,$$

which enables us to write (8.6.25) in a more compact form as

$$\mathcal{C}_\Gamma(h) = -h - \frac{2}{\pi^2}\int_0^\infty \Theta_s M_{1+iA'}\Theta_s M_{1+iA'}(h)\,\frac{ds}{s}. \tag{8.6.26}$$

This gives us the desired resolution of the operator $\mathcal{C}_\Gamma$. It suffices to obtain an L^2 estimate for the integral expression in (8.6.26). Using duality, we write

$$\left\langle \int_0^\infty \Theta_s M_{1+iA'}\Theta_s M_{1+iA'}(h)\,\frac{ds}{s}, g\right\rangle = \int_0^\infty \left\langle M_{1+iA'}\Theta_s M_{1+iA'}(h), \Theta_s^t(g)\right\rangle \frac{ds}{s},$$

which is easily bounded by

$$\begin{aligned}
&\sqrt{1+L^2}\int_0^\infty \big\|\Theta_s M_{1+iA'}(h)\big\|_{L^2}\big\|\Theta_s^t(g)\big\|_{L^2}\frac{ds}{s}\\
&\qquad\le \sqrt{1+L^2}\left(\int_0^\infty \big\|\Theta_s M_{1+iA'}(h)\big\|_{L^2}^2\frac{ds}{s}\right)^{\frac12}\left(\int_0^\infty \big\|\Theta_s(g)\big\|_{L^2}^2\frac{ds}{s}\right)^{\frac12}.
\end{aligned}$$

We have now reduced matters to the following estimate:

$$\left(\int_0^\infty \|\Theta_s(h)\|_{L^2}^2 \frac{ds}{s}\right)^{\frac{1}{2}} \le C\|h\|_{L^2}. \tag{8.6.27}$$

We derive (8.6.27) as a consequence of Theorem 8.6.6 discussed in Section 8.6.4.

8.6.3 A Quadratic $T(1)$ Type Theorem

We review what we have achieved so far and we introduce definitions that place matters into a new framework.

For the purposes of the subsequent exposition we can switch to $\mathbf{R}^n$, since there are no differences from the one-dimensional argument. Suppose that for all $s > 0$, there is a family of functions θ_s defined on $\mathbf{R}^n \times \mathbf{R}^n$ such that

$$|\theta_s(x,y)| \le \frac{1}{s^n} \frac{A}{\left(1+\frac{|x-y|}{s}\right)^{n+\delta}} \tag{8.6.28}$$

and

$$|\theta_s(x,y) - \theta_s(x,y')| \le \frac{A}{s^n} \frac{|y-y'|^\gamma}{s^\gamma} \tag{8.6.29}$$

for all $x,y,y' \in \mathbf{R}^n$ and some $0 < \gamma, \delta, A < \infty$. Let Θ_s be the operator with kernel θ_s, that is,

$$\Theta_s(h)(x) = \int_{\mathbf{R}^n} \theta_s(x,y)\, h(y)\, dy\,, \tag{8.6.30}$$

which is well defined for all h in $\bigcup_{1\le p\le\infty} L^p(\mathbf{R}^n)$ in view of (8.6.28).

At this point we observe that both (8.6.28) and (8.6.29) hold for the θ_s defined in (8.6.24) with $\gamma = \delta = 1$ and A a constant multiple of L. We leave the details of this calculation to the reader but we note that (8.6.29) can be obtained quickly using the mean value theorem. Our goal is to figure out under what additional conditions on Θ_s the quadratic estimate (8.6.27) holds. If we can find such a condition that is easily verifiable for the Θ_s associated with the Cauchy integral, this will conclude the proof of its L^2 boundedness.

We first consider a simple condition that implies the quadratic estimate (8.6.27).

Theorem 8.6.3. *For $s > 0$, let θ_s be a family of kernels satisfying (8.6.28) and (8.6.29) and let Θ_s be the linear operator whose kernel is θ_s. Suppose that for all $s > 0$ we have*

$$\Theta_s(1) = 0\,. \tag{8.6.31}$$

Then there is a constant $C_{n,\delta}$ such that for all $f \in L^2$ we have

$$\left(\int_0^\infty \|\Theta_s(f)\|_{L^2}^2 \frac{ds}{s}\right)^{\frac{1}{2}} \le C_{n,\delta} A\|f\|_{L^2}. \tag{8.6.32}$$

We note that condition (8.6.31) is not satisfied for the operators Θ_s associated with the Cauchy integral as defined in (8.6.23). However, Theorem 8.6.3 gives us an idea of what we are looking for, something like the action of Θ_s on a specific function. We also observe that condition (8.6.31) is "basically" saying that $\Theta(1) = 0$, where

$$\Theta = \int_0^\infty \Theta_s \frac{ds}{s}.$$

Proof. We introduce Littlewood–Paley operators Q_s given by convolution with a smooth function $\Psi_s = \frac{1}{s^n}\Psi(\frac{\cdot}{s})$ whose Fourier transform is supported in the annulus $s/2 \le |\xi| \le 2s$ that satisfies

$$\int_0^\infty Q_s^2 \frac{ds}{s} = \lim_{\substack{\varepsilon\to 0\\ N\to\infty}} \int_\varepsilon^N Q_s^2 \frac{ds}{s} = I, \tag{8.6.33}$$

where the limit is taken in the sense of distributions and the identity holds in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}$. This identity and properties of Θ_t imply the operator identity

$$\Theta_t = \Theta_t \int_0^\infty Q_s^2 \frac{ds}{s} = \int_0^\infty \Theta_t Q_s^2 \frac{ds}{s}.$$

The key fact is the following estimate:

$$\big\|\Theta_t Q_s\big\|_{L^2\to L^2} \le A\, C_{n,\Psi} \min\left(\frac{s}{t},\frac{t}{s}\right)^{\varepsilon}, \tag{8.6.34}$$

which holds for some $\varepsilon = \varepsilon(\gamma,\delta,n) > 0$. [Recall that A, γ, and δ are as in (8.6.28) and (8.6.29).] Assuming momentarily estimate (8.6.34), we can quickly prove Theorem 8.6.3 using duality. Indeed, let us take a function $G(x,t)$ such that

$$\int_0^\infty \int_{\mathbf{R}^n} |G(x,t)|^2\, dx \frac{dt}{t} \le 1. \tag{8.6.35}$$

Then we have

$$\begin{aligned}
&\int_0^\infty \int_{\mathbf{R}^n} G(x,t)\, \Theta_t(f)(x)\, dx \frac{dt}{t} \\
&\qquad = \int_0^\infty \int_{\mathbf{R}^n} G(x,t) \int_0^\infty \Theta_t Q_s^2(f)(x) \frac{ds}{s}\, dx \frac{dt}{t} \\
&\qquad = \int_0^\infty \int_0^\infty \int_{\mathbf{R}^n} G(x,t)\, \Theta_t Q_s^2(f)(x)\, dx \frac{dt}{t}\frac{ds}{s} \\
&\qquad \le \left(\int_0^\infty \int_0^\infty \int_{\mathbf{R}^n} |G(x,t)|^2 dx \min\left(\frac{s}{t},\frac{t}{s}\right)^{\varepsilon} \frac{dt}{t}\frac{ds}{s}\right)^{\frac{1}{2}} \\
&\qquad\qquad \times \left(\int_0^\infty \int_0^\infty \int_{\mathbf{R}^n} |\Theta_t Q_s(Q_s(f))(x)|^2\, dx \min\left(\frac{s}{t},\frac{t}{s}\right)^{-\varepsilon} \frac{dt}{t}\frac{ds}{s}\right)^{\frac{1}{2}}.
\end{aligned}$$

But we have the estimate

$$\sup_{t>0}\int_0^\infty \min\Big(\frac{s}{t},\frac{t}{s}\Big)^{\varepsilon}\frac{ds}{s}\le C_\varepsilon\,,$$

which, combined with (8.6.35), yields that the first term in the product of the two preceding square functions is controlled by $\sqrt{C_\varepsilon}$. Using this fact and (8.6.34), we write

$$\begin{aligned}
&\int_0^\infty\int_{\mathbf{R}^n} G(x,t)\,\Theta_t(f)(x)\,dx\,\frac{dt}{t}\\
&\quad\le \sqrt{C_\varepsilon}\bigg(\int_0^\infty\int_0^\infty\int_{\mathbf{R}^n}|\Theta_t Q_s(Q_s(f))(x)|^2\,dx\,\min\Big(\frac{s}{t},\frac{t}{s}\Big)^{-\varepsilon}\frac{dt}{t}\frac{ds}{s}\bigg)^{\frac12}\\
&\quad\le A\sqrt{C_\varepsilon}\bigg(\int_0^\infty\int_0^\infty\int_{\mathbf{R}^n}|Q_s(f)(x)|^2\,dx\,\min\Big(\frac{s}{t},\frac{t}{s}\Big)^{2\varepsilon}\min\Big(\frac{s}{t},\frac{t}{s}\Big)^{-\varepsilon}\frac{dt}{t}\frac{ds}{s}\bigg)^{\frac12}\\
&\quad\le A\sqrt{C_\varepsilon}\bigg(\int_0^\infty\int_0^\infty\int_{\mathbf{R}^n}|Q_s(f)(x)|^2\,dx\,\min\Big(\frac{s}{t},\frac{t}{s}\Big)^{\varepsilon}\frac{dt}{t}\frac{ds}{s}\bigg)^{\frac12}\\
&\quad\le C_\varepsilon A\bigg(\int_0^\infty\int_{\mathbf{R}^n}|Q_s(f)(x)|^2\,dx\,\frac{ds}{s}\bigg)^{\frac12}\\
&\quad\le C_{n,\varepsilon}A\big\|f\big\|_{L^2}\,,
\end{aligned}$$

where in the last step we used the continuous version of Theorem 5.1.2 (cf. Exercise 5.1.4). Taking the supremum over all functions $G(x,t)$ that satisfy (8.6.35) yields estimate (8.6.32).

It remains to prove (8.6.34). What is crucial here is that both Θ_t and Q_s satisfy the cancellation conditions $\Theta_t(1)=0$ and $Q_s(1)=0$. The proof of estimate (8.6.34) is similar to that of estimates (8.5.14) and (8.5.15) in Proposition 8.5.3. Using ideas from the proof of Proposition 8.5.3, we quickly dispose of the proof of (8.6.34).

The kernel of $\Theta_t Q_s$ is seen easily to be

$$L_{t,s}(x,y)=\int_{\mathbf{R}^n}\theta_t(x,z)\Psi_s(z-y)\,dz\,.$$

Notice that the function $(y,z)\mapsto\Psi_s(z-y)$ satisfies (8.6.28) with $\delta=1$ and $A=C_\Psi$ and satisfies

$$|\Psi_s(z-y)-\Psi_s(z'-y)|\le\frac{C_\Psi}{s^n}\frac{|z-z'|}{s}$$

for all $z,z',y\in\mathbf{R}^n$ for some $C_\Psi<\infty$. We prove that

$$\sup_{x\in\mathbf{R}^n}\int_{\mathbf{R}^n}|L_{t,s}(x,y)|\,dy\;\le\;C_\Psi A\min\Big(\frac{t}{s},\frac{s}{t}\Big)^{\frac14\frac{\min(\delta,1)}{n+\min(\delta,1)}\min(\gamma,\delta,1)}\,,\tag{8.6.36}$$

$$\sup_{y\in\mathbf{R}^n}\int_{\mathbf{R}^n}|L_{t,s}(x,y)|\,dx\;\le\;C_\Psi A\min\Big(\frac{t}{s},\frac{s}{t}\Big)^{\frac14\frac{\min(\delta,1)}{n+\min(\delta,1)}\min(\gamma,\delta,1)}\,.\tag{8.6.37}$$

Once (8.6.36) and (8.6.37) are established, (8.6.34) follows directly from the lemma in Appendix I.1 with $\varepsilon = \frac{1}{4}\frac{\min(\delta,1)}{n+\min(\delta,1)}\min(\gamma,\delta,1)$.

We begin by observing that when $s \le t$ we have the estimate

$$\int_{\mathbf{R}^n} \frac{s^{-n}\min(2,(t^{-1}|u|)^{\gamma})}{(1+s^{-1}|u|)^{n+1}}\,du \le C_n\Big(\frac{s}{t}\Big)^{\frac{1}{2}\min(\gamma,1)}. \tag{8.6.38}$$

Also when $t \le s$ we have the analogous estimate

$$\int_{\mathbf{R}^n} \frac{t^{-n}\min(2,s^{-1}|u|)}{(1+t^{-1}|u|)^{n+\delta}}\,du \le C_n\Big(\frac{t}{s}\Big)^{\frac{1}{2}\min(\delta,1)}. \tag{8.6.39}$$

Both (8.6.38) and (8.6.39) are trivial reformulations or consequences of (8.5.18).

We now take $s \le t$ and we use that $Q_s(1) = 0$ for all $s > 0$ to obtain

$$\begin{aligned}
|L_{t,s}(x,y)| &= \bigg|\int_{\mathbf{R}^n} \theta_t(x,z)\Psi_s(z-y)\,dz\bigg| \\
&= \bigg|\int_{\mathbf{R}^n} \big[\theta_t(x,z)-\theta_t(x,y)\big]\Psi_s(z-y)\,dz\bigg| \\
&\le CA\int_{\mathbf{R}^n} \frac{\min(2,(t^{-1}|z-y|)^{\gamma})}{t^n}\,\frac{s^{-n}}{(1+s^{-1}|z-y|)^{n+1}}\,dz \\
&\le C_n' A\,\frac{1}{t^n}\Big(\frac{s}{t}\Big)^{\frac{1}{2}\min(\gamma,1)} \\
&\le C_n' A\,\min\Big(\frac{1}{t},\frac{1}{s}\Big)^n \min\Big(\frac{t}{s},\frac{s}{t}\Big)^{\frac{1}{2}\min(\gamma,\delta,1)}
\end{aligned}$$

using estimate (8.6.38). Similarly, using (8.6.39) and the hypothesis that $\Theta_t(1) = 0$ for all $t > 0$, we obtain for $t \le s$,

$$\begin{aligned}
|L_{t,s}(x,y)| &= \bigg|\int_{\mathbf{R}^n} \theta_t(x,z)\Psi_s(z-y)\,dz\bigg| \\
&= \bigg|\int_{\mathbf{R}^n} \theta_t(x,z)\big[\Psi_s(z-y)-\Psi_s(x-y)\big]\,dz\bigg| \\
&\le xCA\int_{\mathbf{R}^n} \frac{t^{-n}}{(1+t^{-1}|x-z|)^{n+\delta}}\,\frac{\min(2,s^{-1}|x-z|)}{s^n}\,dz \\
&\le C_n' A\,\frac{1}{s^n}\Big(\frac{t}{s}\Big)^{\frac{1}{2}\min(\delta,1)} \\
&\le C_n' A\,\min\Big(\frac{1}{t},\frac{1}{s}\Big)^n \min\Big(\frac{t}{s},\frac{s}{t}\Big)^{\frac{1}{2}\min(\gamma,\delta,1)}.
\end{aligned}$$

Combining the estimates for $|L_{t,s}(x,y)|$ in the preceding cases $t \le s$ and $s \le t$ with the estimate

$$|L_{t,s}(x,y)| \le \int_{\mathbf{R}^n} |\Theta_t(x,z)|\,|\Psi_s(z-y)|\,dz \le \frac{CA\min(\frac{1}{t},\frac{1}{s})^n}{\big(1+\min(\frac{1}{t},\frac{1}{s})|x-y|\big)^{n+\min(\delta,1)}},$$

which is a consequence of the result in Appendix K.1, gives

$$|L_{t,s}(x,y)| \le \frac{C\min(\frac{t}{s},\frac{s}{t})^{\frac{1}{2}\min(\gamma,\delta,1)(1-\beta)} A\min(\frac{1}{t},\frac{1}{s})^n}{\Big(\big(1+\min(\frac{1}{t},\frac{1}{s})|x-y|\big)^{n+\min(\delta,1)}\Big)^{\beta}}$$

for any $0<\beta<1$. Choosing $\beta=(n+\frac{1}{2}\min(\delta,1))(n+\min(\delta,1))^{-1}$ and integrating over x or y yields (8.6.36) and (8.6.37), respectively, and thus concludes the proof of estimate (8.6.34). □

We end this subsection with a small generalization of the previous theorem that follows by an examination of its proof. The simple details are left to the reader.

Corollary 8.6.4. *For $s>0$ let Θ_s be linear operators that are uniformly bounded on $L^2(\mathbf{R}^n)$ by a constant B. Let Ψ be a Schwartz function whose Fourier transform is supported in the annulus $1/2\le|x|\le 2$ such that the Littlewood–Paley operator Q_s given by convolution with $\Psi_s(x)=s^{-n}\Psi(s^{-1}x)$ satisfies* (8.6.33)*. Suppose that for some $C_{n,\Psi},A,\varepsilon<\infty$,*

$$\big\|\Theta_t Q_s\big\|_{L^2\to L^2} \le A\,C_{n,\Psi}\min\left(\frac{s}{t},\frac{t}{s}\right)^{\varepsilon} \tag{8.6.40}$$

is satisfied for all $t,s>0$. Then there is a constant $C_{n,\Psi,\varepsilon}$ such that for all $f\in L^2(\mathbf{R}^n)$ we have

$$\left(\int_0^\infty \big\|\Theta_s(f)\big\|_{L^2}^2\frac{ds}{s}\right)^{\frac{1}{2}} \le C_{n,\Psi,\varepsilon}(A+B)\big\|f\big\|_{L^2}.$$

8.6.4 A $T(b)$ Theorem and the L^2 Boundedness of the Cauchy Integral

The operators Θ_s defined in (8.6.23) and (8.6.24) that appear in the resolution of the Cauchy integral operator $\mathcal{C}_\Gamma$ do not satisfy the condition $\Theta_s(1)=0$ of Theorem 8.6.3. It turns out that a certain variant of this theorem is needed for the purposes of the application we have in mind, the L^2 boundedness of the Cauchy integral operator. This variant is a quadratic type $T(b)$ theorem discussed in this subsection. Before we state the main theorem, we need a definition.

Definition 8.6.5. A bounded complex-valued function b on $\mathbf{R}^n$ is said to be *accretive* if there is a constant $c_0>0$ such that $\operatorname{Re} b(x)\ge c_0$ for almost all $x\in\mathbf{R}^n$.

The following theorem is the main result of this section.

Theorem 8.6.6. *Let θ_s be a complex-valued function on $\mathbf{R}^n \times \mathbf{R}^n$ that satisfies (8.6.28) and (8.6.29), and let Θ_s be the linear operator in (8.6.30) whose kernel is θ_s. If there is an accretive function b such that*

$$\Theta_s(b) = 0 \tag{8.6.41}$$

for all $s > 0$, then there is a constant $C_n(b)$ such that the estimate

$$\left(\int_0^\infty \|\Theta_s(f)\|_{L^2}^2 \frac{ds}{s}\right)^{\frac{1}{2}} \le C_n(b)\|f\|_{L^2} \tag{8.6.42}$$

holds for all $f \in L^2$.

Corollary 8.6.7. *The Cauchy integral operator $\mathcal{C}_\Gamma$ maps $L^2(\mathbf{R})$ to itself.*

The corollary is a consequence of Theorem 8.6.6. Indeed, the crucial and important cancellation property

$$\Theta_s(1+iA') = 0 \tag{8.6.43}$$

is valid for the accretive function $1+iA'$, when Θ_s and θ_s are as in (8.6.23) and (8.6.24). To prove (8.6.43) we simply note that

$$\begin{aligned}
\Theta_s(1+iA')(x) &= \int_{\mathbf{R}} \frac{s\,(1+iA'(y))\,dy}{(y-x+i(A(y)-A(x))+is)^2} \\
&= \left[\frac{-s}{y-x+i(A(y)-A(x))+is}\right]_{y=-\infty}^{y=+\infty} \\
&= 0-0=0\,.
\end{aligned}$$

This condition plays exactly the role of (8.6.31), which may fail in general. The necessary "internal cancellation" of the family of operators Θ_s is exactly captured by the single condition (8.6.43).

It remains to prove Theorem 8.6.6.

Proof. We fix an approximation of the identity operator, such as

$$P_s(f)(x) = \int_{\mathbf{R}^n} \Phi_s(x-y)\,f(y)\,dy\,,$$

where $\Phi_s(x) = s^{-n}\Phi(s^{-1}x)$, and Φ is a nonnegative Schwartz function with integral 1. Then P_s is a nice positive averaging operator that satisfies $P_s(1) = 1$ for all $s > 0$. The key idea is to decompose the operator Θ_s as

$$\Theta_s = \big(\Theta_s - M_{\Theta_s(1)}P_s\big) + M_{\Theta_s(1)}P_s\,, \tag{8.6.44}$$

where $M_{\Theta_s(1)}$ is the operator given by multiplication by $\Theta_s(1)$. We begin with the first term in (8.6.44), which is essentially an error term. We simply observe that

$$\big(\Theta_s - M_{\Theta_s(1)}P_s\big)(1) = \Theta_s(1) - \Theta_s(1)\,P_s(1) = \Theta_s(1) - \Theta_s(1) = 0\,.$$

Therefore, Theorem 8.6.3 is applicable once we check that the kernel of the operator $\Theta_s - M_{\Theta_s(1)}P_s$ satisfies (8.6.28) and (8.6.29). But these are verified easily, since the kernels of both Θ_s and P_s satisfy these estimates and $\Theta_s(1)$ is a bounded function uniformly in s. The latter statement is a consequence of condition (8.6.28).

We now need to obtain the required quadratic estimate for the term $M_{\Theta_s(1)}P_s$. With the use of Theorem 7.3.7, this follows once we prove that the measure

$$\big|\Theta_s(1)(x)\big|^2 \frac{dx\,ds}{s}$$

is Carleson. It is here that we use condition (8.6.41). Since $\Theta_s(b) = 0$ we have

$$P_s(b)\,\Theta_s(1) = \big(P_s(b)\,\Theta_s(1) - \Theta_s P_s(b)\big) + \big(\Theta_s P_s(b) - \Theta_s(b)\big)\,. \tag{8.6.45}$$

Suppose we could show that the measures

$$\big|\Theta_s(b)(x) - \Theta_s P_s(b)(x)\big|^2 \frac{dx\,ds}{s}\,, \tag{8.6.46}$$

$$\big|\Theta_s P_s(b)(x) - P_s(b)(x)\,\Theta_s(1)(x)\big|^2 \frac{dx\,ds}{s}\,, \tag{8.6.47}$$

are Carleson. Then it would follow from (8.6.45) that the measure

$$\big|P_s(b)(x)\,\Theta_s(1)(x)\big|^2 \frac{dx\,ds}{s}$$

is also Carleson. Using the accretivity condition on b and the positivity of P_s we obtain

$$\big|P_s(b)\big| \ge \operatorname{Re} P_s(b) = P_s(\operatorname{Re} b) \ge P_s(c_0) = c_0,$$

from which it follows that $|\Theta_s(1)(x)|^2 \le c_0^{-2}|P_s(b)(x)\,\Theta_s(1)(x)|^2$. Thus the measure $|\Theta_s(1)(x)|^2 dx\,ds/s$ must be Carleson.

Therefore, the proof will be complete if we can show that both measures (8.6.46) and (8.6.47) are Carleson. Theorem 7.3.8 plays a key role here.

We begin with the measure in (8.6.46). First we observe that the kernel

$$L_s(x,y) = \int_{\mathbf{R}^n} \theta_s(x,z)\Phi_s(z-y)\,dz$$

of $\Theta_s P_s$ satisfies (8.6.28) and (8.6.29). The verification of (8.6.28) is a straightforward consequence of the estimate in Appendix K.1, while (8.6.29) follows easily from the mean value theorem. It follows that the kernel of

$$R_s = \Theta_s - \Theta_s P_s$$

satisfies the same estimates. Moreover, it is easy to see that $R_s(1) = 0$ and thus the quadratic estimate (8.6.32) holds for R_s in view of Theorem 8.6.3. Therefore, the hypotheses of Theorem 7.3.8(c) are satisfied, and this gives that the measure in (8.6.46) is Carleson.

We now continue with the measure in (8.6.47). Here we set

$$T_s(f)(x) = \Theta_s P_s(f)(x) - P_s(f)(x)\Theta_s(1)(x).$$

The kernel of T_s is $L_s(x,y) - \Theta_s(1)(x)\Phi_s(x-y)$, which clearly satisfies (8.6.28) and (8.6.29), since $\Theta_s(1)(x)$ is a bounded function uniformly in $s>0$. We also observe that $T_s(1)=0$. Using Theorem 8.6.3, we conclude that the quadratic estimate (8.6.32) holds for T_s. Therefore, the hypotheses of Theorem 7.3.8(c) are satisfied; hence the measure in (8.6.46) is Carleson. □

We conclude by observing that if we attempt to replace Θ_s with $\widetilde{\Theta}_s = \Theta_s M_{1+iA'}$ in the resolution identity (8.6.26), then $\widetilde{\Theta}_s(1)=0$ would hold, but the kernel of $\widetilde{\Theta}_s$ would not satisfy the regularity estimate (8.6.29). The whole purpose of Theorem 8.6.6 was to find a certain balance between regularity and cancellation.

Exercises

8.6.1. Given a function H on a Lipschitz graph Γ, we associate a function h on the line by setting $h(t) = H(t+iA(t))$. Prove that for all $0<p<\infty$ we have

$$\|h\|_{L^p(\mathbf{R})}^p \le \|H\|_{L^p(\Gamma)}^p \le \sqrt{1+L^2}\,\|h\|_{L^p(\mathbf{R})}^p,$$

where L is the Lipschitz constant of the defining function A of the graph Γ.

8.6.2. Let $A:\ \mathbf{R}\to\mathbf{R}$ satisfy $|A(x)-A(y)|\le L|x-y|$ for all $x,y\in\mathbf{R}$ for some $L>0$. Also, let h be a Schwartz function on $\mathbf{R}$.
(a) Show that for all $s>0$ and $x,y\in\mathbf{R}$ we have

$$\frac{s^2+|x-y|^2}{|x-y|^2+|A(x)-A(y)+s|^2} \le 4L^2+2.$$

(b) Use the Lebesgue dominated convergence theorem to prove that

$$\int_{|x-y|>\sqrt{s}} \frac{s(1+iA'(y))h(y)}{(y-x+i(A(y)-A(x))+is)^2}\,dy \to 0$$

as $s\to 0$.
(c) Integrate directly to show that as $s\to 0$,

$$\int_{|x-y|\le\sqrt{s}} \frac{s(1+iA'(y))}{(y-x+i(A(y)-A(x))+is)^2}\,dy \to 0$$

for every point x at which A is differentiable.
(d) Use part (a) and the Lebesgue dominated convergence theorem to show that as

$s \to 0$,

$$\int_{|x-y|\le\sqrt{s}} \frac{s(1+iA'(y))(h(y)-h(x))}{(y-x+i(A(y)-A(x))+is)^2}\,dy \to 0\,.$$

(e) Use part (a) and the Lebesgue dominated convergence theorem to show that as $s \to \infty$,

$$\int_{\mathbf{R}} \frac{s(1+iA'(y))h(y)}{(y-x+i(A(y)-A(x))+is)^2}\,dy \to 0\,.$$

Conclude the validity of the statements in (8.6.20) for almost all $x \in \mathbf{R}$.

8.6.3. Prove identity (8.6.22).
[*Hint:* Write the identity in (8.6.22) as

$$\frac{-2}{((\zeta+is)-(z-is))^3} = \frac{1}{2\pi i}\int_{\Gamma} \frac{\frac{1}{(w-(z-is))^2}}{(w-(\zeta+is))^2}\,dw$$

and interpret it as Cauchy's integral formula for the derivative of the analytic function $w \mapsto (w-(z-is))^{-2}$ defined on the region above Γ. If Γ were a closed curve containing $\zeta+is$ but not $z-is$, then the previous assertion would be immediate. In general, consider a circle of radius R centered at the point $\zeta+is$ and the region U_R inside this circle and above Γ. See Figure 8.1. Integrate over the boundary of U_R and let $R \to \infty$.]

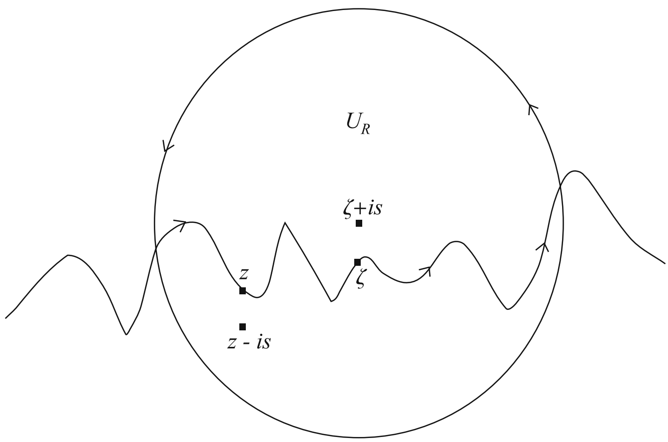

Fig. 8.1 The region U_R inside the circle and above the curve.

8.6.4. Given an accretive function b, define a pseudo-inner product

$$\langle f,g\rangle_b = \int_{\mathbf{R}^n} f(x)\,g(x)\,b(x)\,dx$$

on L^2. For an interval I, set $b_I = \int_I b(x)\,dx$. Let I_L denote the left half of a dyadic interval I and let I_R denote its right half. For a complex number z, let $z^{\frac{1}{2}} = e^{\frac{1}{2}\log_{right} z}$, where $\log_{right}$ is the branch of the logarithm defined on the complex plane minus the negative real axis normalized so that $\log_{right} 1 = 0$ [and $\log_{right}(\pm i) = \pm\frac{\pi}{2} i$]. Show that the family of functions

$$h_I = \frac{-1}{b(I)^{\frac{1}{2}}}\left(\frac{b(I_R)^{\frac{1}{2}}}{b(I_L)^{\frac{1}{2}}}\chi_{I_L} - \frac{b(I_L)^{\frac{1}{2}}}{b(I_R)^{\frac{1}{2}}}\chi_{I_R}\right),$$

where I runs over all dyadic intervals, is an orthonormal family on $L^2(\mathbf{R})$ with respect to the preceding inner product. (This family of functions is called a *pseudo-Haar basis associated with b*.)

8.6.5. Let $I = (a,b)$ be a dyadic interval and let $3I$ be its triple. For a given $x \in \mathbf{R}$, let

$$d_I(x) = \min\left(|x-a|, |x-b|, |x - \tfrac{a+b}{2}|\right).$$

Show that there exists a constant C such that

$$\left|\mathcal{C}_\Gamma(h_I)(x)\right| \le C|I|^{-\frac{1}{2}} \log\frac{10|I|}{|x-d_I(x)|}$$

whenever $x \in 3I$ and also

$$\left|\mathcal{C}_\Gamma(h_I)(x)\right| \le \frac{C|I|^{\frac{3}{2}}}{|x-d_I(x)|^2}$$

for $x \notin 3I$. In the latter case, $d_I(x)$ can be any of $a, b, \frac{a+b}{2}$.

8.6.6. (*Semmes [281]*) We say that a bounded function b is *para-accretive* if for all $s > 0$ there is a linear operator R_s with kernel satisfying (8.6.28) and (8.6.29) such that $|R_s(b)| \ge c_0$ for all $s > 0$. Let Θ_s and P_s be as in Theorem 8.6.6.
(a) Prove that

$$\left|R_s(b)(x) - R_s(1)(x)P_s(b)(x)\right|^2 \frac{dx\,ds}{s}$$

is a Carleson measure.
(b) Use the result in part (a) and the fact that $\sup_{s>0}|R_s(1)| \le C$ to obtain that $\chi_\Omega(x,s)\,dx\,ds/s$ is a Carleson measure, where

$$\Omega = \left\{(x,s) : \ |P_s(b)(x)| \le \frac{c_0}{2}\left(\sup_{s>0}|R_s(1)|\right)^{-1}\right\}.$$

(c) Conclude that the measure $\left|\Theta_s(1)(x)\right|^2 dx\,ds/s$ is Carleson, thus obtaining a generalization of Theorem 8.6.6 for para-accretive functions.

8.6.7. Using the operator $\widetilde{\mathcal{C}}_\gamma$ defined in (8.6.15), obtain that $\mathcal{C}_\Gamma$ is of weak type $(1,1)$ and bounded on $L^p(\mathbf{R})$ for all $1 < p < \infty$.

8.7 Square Roots of Elliptic Operators

In this section we prove an L^2 estimate for the square root of a divergence form second-order elliptic operator on $\mathbf{R}^n$. This estimate is based on an approach in the spirit of the $T(b)$ theorem discussed in the previous section. However, matters here are significantly more complicated for two main reasons: the roughness of the variable coefficients of the aforementioned elliptic operator and the higher-dimensional nature of the problem.

8.7.1 Preliminaries and Statement of the Main Result

For $\xi = (\xi_1, \dots, \xi_n) \in \mathbf{C}^n$ we denote its complex conjugate $(\overline{\xi_1}, \dots, \overline{\xi_n})$ by $\overline{\xi}$. Moreover, for $\xi, \zeta \in \mathbf{C}^n$ we use the inner product notation

$$\xi \cdot \zeta = \sum_{k=1}^{n} \xi_k \zeta_k .$$

Throughout this section, $A = A(x)$ is an $n \times n$ matrix of complex-valued L^∞ functions, defined on $\mathbf{R}^n$, that satisfies the *ellipticity* (or *accretivity*) conditions for some $0 < \lambda \le \Lambda < \infty$, that

$$\begin{aligned} \lambda |\xi|^2 &\le \operatorname{Re}\big(A(x)\,\xi \cdot \overline{\xi}\big), \\ |A(x)\,\xi \cdot \overline{\zeta}| &\le \Lambda |\xi|\,|\zeta|, \end{aligned} \tag{8.7.1}$$

for all $x \in \mathbf{R}^n$ and $\xi, \zeta \in \mathbf{C}^n$. We interpret an element ξ of $\mathbf{C}^n$ as a column vector in $\mathbf{C}^n$ when the matrix A acts on it.

Associated with such a matrix A, we define a second-order *divergence form operator*

$$L(f) = -\operatorname{div}(A \nabla f) = \sum_{j=1}^{n} \partial_j \big((A \nabla f)_j \big) , \tag{8.7.2}$$

which we interpret in the weak sense whenever f is a distribution.

The accretivity condition (8.7.1) enables us to define a square root operator $L^{1/2} = \sqrt{L}$ so that the operator identity $L = \sqrt{L}\sqrt{L}$ holds. The *square root operator* can be written in several ways, one of which is

$$\sqrt{L}(f) = \frac{16}{\pi} \int_0^{+\infty} (I + t^2 L)^{-3} t^3 L^2(f) \, \frac{dt}{t} . \tag{8.7.3}$$

We refer the reader to Exercise 8.7.3 for the existence of the square root operator and the validity of identity (8.7.3).

An important problem in the subject is to determine whether the estimate

$$\big\| \sqrt{L}(f) \big\|_{L^2} \le C_{n,\lambda,\Lambda} \big\| \nabla f \big\|_{L^2} \tag{8.7.4}$$

holds for functions f in a dense subspace of the homogeneous Sobolev space $\dot{L}^2_1(\mathbf{R}^n)$, where $C_{n,\lambda,\Lambda}$ is a constant depending only on n, λ, and Λ. Once (8.7.4) is known for a dense subspace of $\dot{L}^2_1(\mathbf{R}^n)$, then it can be extended to the entire space by density. The main purpose of this section is to discuss a detailed proof of the following result.

Theorem 8.7.1. *Let L be as in (8.7.2). Then there is a constant $C_{n,\lambda,\Lambda}$ such that for all smooth functions f with compact support, estimate (8.7.4) is valid.*

The proof of this theorem requires certain estimates concerning elliptic operators. These are presented in the next subsection, while the proof of the theorem follows in the remaining four subsections.

8.7.2 Estimates for Elliptic Operators on $\mathbf{R}^n$

The following lemma provides a quantitative expression for the mean decay of the resolvent kernel.

Lemma 8.7.2. *Let E and F be two closed sets of $\mathbf{R}^n$ and set*

$$d = \operatorname{dist}(E,F),$$

the distance between E and F. Then for all complex-valued functions f supported in E and all vector-valued functions $\vec{f}$ supported in E, we have

$$\int_F |(I+t^2L)^{-1}(f)(x)|^2\,dx \le Ce^{-c\frac{d}{t}} \int_E |f(x)|^2\,dx, \tag{8.7.5}$$

$$\int_F |t\nabla(I+t^2L)^{-1}(f)(x)|^2\,dx \le Ce^{-c\frac{d}{t}} \int_E |f(x)|^2\,dx, \tag{8.7.6}$$

$$\int_F |(I+t^2L)^{-1}(t\operatorname{div}\vec{f}\,)(x)|^2\,dx \le Ce^{-c\frac{d}{t}} \int_E |\vec{f}(x)|^2\,dx, \tag{8.7.7}$$

where $c = c(\lambda,\Lambda)$, $C = C(n,\lambda,\Lambda)$ are finite constants.

Proof. It suffices to obtain these inequalities whenever $d \ge t > 0$. Let us set $u_t = (I+t^2L)^{-1}(f)$. For all $v \in L^2_1(\mathbf{R}^n)$ we have

$$\int_{\mathbf{R}^n} u_t v\,dx + t^2 \int_{\mathbf{R}^n} A\nabla u_t \cdot \nabla v\,dx = \int_{\mathbf{R}^n} f v\,dx.$$

Let η be a nonnegative smooth function with compact support that does not meet E and that satisfies $\|\eta\|_{L^\infty} = 1$. Taking $v = \overline{u_t}\,\eta^2$ and using that f is supported in E, we obtain

$$\int_{\mathbf{R}^n} |u_t|^2\eta^2\,dx + t^2 \int_{\mathbf{R}^n} A\nabla u_t \cdot \overline{\nabla u_t}\;\eta^2\,dx = -2t^2 \int_{\mathbf{R}^n} A(\eta\nabla u_t)\cdot \overline{u_t\nabla\eta}\,dx.$$

Using (8.7.1) and the inequality $2ab \le \varepsilon|a|^2 + \varepsilon^{-1}|b|^2$, we obtain for all $\varepsilon > 0$,

$$\int_{\mathbf{R}^n} |u_t|^2 \eta^2\,dx + \lambda t^2 \int_{\mathbf{R}^n} |\nabla u_t|^2\ \eta^2\,dx \\ \le \Lambda \varepsilon t^2 \int_{\mathbf{R}^n} |\nabla u_t|^2\ \eta^2\,dx + \Lambda \varepsilon^{-1} t^2 \int_{\mathbf{R}^n} |u_t|^2 |\nabla \eta|^2\,dx,$$

and this reduces to

$$\int_{\mathbf{R}^n} |u_t|^2 |\eta|^2\,dx \le \frac{\Lambda^2 t^2}{\lambda} \int_{\mathbf{R}^n} |u_t|^2 |\nabla \eta|^2\,dx \tag{8.7.8}$$

by choosing $\varepsilon = \frac{\lambda}{\Lambda}$. Replacing η by $e^{k\eta} - 1$ in (8.7.8), where

$$k = \frac{\sqrt{\lambda}}{2\Lambda t \|\nabla \eta\|_{L^\infty}},$$

yields

$$\int_{\mathbf{R}^n} |u_t|^2 |e^{k\eta} - 1|^2\,dx \le \frac{1}{4} \int_{\mathbf{R}^n} |u_t|^2 |e^{k\eta}|^2\,dx. \tag{8.7.9}$$

Using that $|e^{k\eta} - 1|^2 \ge \frac{1}{2}|e^{k\eta}|^2 - 1$, we obtain

$$\int_{\mathbf{R}^n} |u_t|^2 |e^{k\eta}|^2\,dx \le 4 \int_{\mathbf{R}^n} |u_t|^2\,dx \le 4C \int_E |f|^2\,dx,$$

where in the last estimate we use the uniform boundedness of $(I + t^2 L)^{-1}$ on $L^2(\mathbf{R}^n)$ (Exercise 8.7.2). If, in addition, we have $\eta = 1$ on F, then

$$|e^k|^2 \int_F |u_t|^2\,dx \le \int_{\mathbf{R}^n} |u_t|^2 |e^{k\eta}|^2\,dx,$$

and picking η so that $\|\nabla \eta\|_{L^\infty} \approx 1/d$, we conclude (8.7.5).

Next, choose $\varepsilon = \lambda/2\Lambda$ and η as before to obtain

$$\begin{aligned} \int_F |t \nabla u_t|^2\,dx &\le \int_{\mathbf{R}^n} |t \nabla u_t|^2 \eta^2\,dx \\ &\le \frac{2\Lambda^2 t^2}{\lambda} \int_{\mathbf{R}^n} |u_t|^2 |\nabla \eta|^2\,dx \\ &\le C t^2 d^{-2} e^{-c\frac{d}{t}} \int_E |f|^2\,dx, \end{aligned}$$

which gives (8.7.6). Finally, (8.7.7) is obtained by duality from (8.7.6) applied to $L^* = -\mathrm{div}\,(A^* \nabla)$ when the roles of E and F are interchanged. □

Lemma 8.7.3. *Let M_f be the operator given by multiplication by a Lipschitz function f. Then there is a constant C that depends only on n, λ, and Λ such that*

$$\big\| \big[(I + t^2 L)^{-1}, M_f \big] \big\|_{L^2 \to L^2} \le C t \,\big\| \nabla f \big\|_{L^\infty} \tag{8.7.10}$$

and

$$\big\|\nabla\big[(I+t^2L)^{-1},M_f\big]\big\|_{L^2\to L^2}\le C\big\|\nabla f\big\|_{L^\infty} \tag{8.7.11}$$

for all $t>0$. *Here* $[T,S]=TS-ST$ *is the* commutator *of the operators* T *and* S.

Proof. Set $\vec{b}=A\nabla f$, $\vec{d}=A^t\nabla f$ and note that the operators given by pointwise multiplication by these vectors are L^2 bounded with norms at most a multiple of $C\big\|\nabla f\big\|_{L^\infty}$. Write

$$\begin{aligned}\big[(I+t^2L)^{-1},M_f\big] &= -(I+t^2L)^{-1}\big[(I+t^2L),M_f\big](I+t^2L)^{-1}\\ &= -(I+t^2L)^{-1}t^2(\operatorname{div}\vec{b}+\vec{d}\cdot\nabla)(1+t^2L)^{-1}.\end{aligned}$$

The uniform L^2 boundedness of $(I+t^2L)^{-1}\,t\nabla(I+t^2L)^{-1}$ and $(I+t^2L)^{-1}t\operatorname{div}$ on L^2 (see Exercise 8.7.2) implies (8.7.10). Finally, using the L^2 boundedness of the operator $t^2\nabla(I+t^2L)^{-1}\operatorname{div}$ yields (8.7.11). □

Next we have a technical lemma concerning the mean square deviation of f from $(I+t^2L)^{-1}$.

Lemma 8.7.4. *There exists a constant* C *depending only on* n, λ, *and* Λ *such that for all* Q *cubes in* $\mathbf{R}^n$ *with sides parallel to the axes, for all* $t\le\ell(Q)$, *and all Lipschitz functions* f *on* $\mathbf{R}^n$ *we have*

$$\frac{1}{|Q|}\int_Q|(I+t^2L)^{-1}(f)-f|^2\,dx\;\le\;Ct^2\big\|\nabla f\big\|_{L^\infty}^2, \tag{8.7.12}$$

$$\frac{1}{|Q|}\int_Q|\nabla((I+t^2L)^{-1}(f)-f)|^2\,dxx\;\le\;C\big\|\nabla f\big\|_{L^\infty}^2. \tag{8.7.13}$$

Proof. We begin by proving (8.7.12), while we omit the proof of (8.7.13), since it is similar. By a simple rescaling, we may assume that $\ell(Q)=1$ and that $\|\nabla f\|_{L^\infty}=1$. Set $Q_0=2Q$ (i.e., the cube with the same center as Q with twice its side length) and write $\mathbf{R}^n$ as a union of cubes Q_k of side length 2 with disjoint interiors and sides parallel to the axes. Lemma 8.7.2 implies that

$$(I+t^2L)^{-1}(1)=1$$

in the sense that

$$\lim_{R\to\infty}(I+t^2L)^{-1}(\eta_R)=1$$

in $L^2_{\text{loc}}(\mathbf{R}^n)$, where $\eta_R(x)=\eta(x/R)$ and η is a smooth bump function with $\eta\equiv 1$ near 0. Hence, we may write

$$(I+t^2L)^{-1}(f)(x)-f(x)=\sum_{k\in\mathbf{Z}^n}(I+t^2L)^{-1}((f-f(x))\chi_{Q_k})(x)=\sum_{k\in\mathbf{Z}^n}g_k(x).$$

The term for $k=0$ in the sum is $[(I+t^2L)^{-1},M_f](\chi_{Q_0})(x)$. Hence, its $L^2(Q)$ norm is controlled by $Ct\big\|\chi_{Q_0}\big\|_{L^2}$ by (8.7.10). The terms for $k\neq 0$ are dealt with using the further decomposition

$$g_k(x) = (I+t^2L)^{-1}((f-f(x_k))\chi_{Q_k})(x) + (f(x_k)-f(x))(I+t^2L)^{-1}(\chi_{Q_k})(x),$$

where x_k is the center of Q_k. Applying Lemma 8.7.2 for $(I+t^2L)^{-1}$ on the sets $E = Q_k$ and $F = Q$ and using that f is a Lipschitz function, we obtain

$$\int_Q |g_k|^2\,dx \le Ct^2 e^{-c\frac{|x_k|}{t}} \|\chi_{Q_k}\|_{L^2}^2 = Ct^2 e^{-c\frac{|x_k|}{t}} 2^n |Q|.$$

The desired bound on the $L^2(Q)$ norm of $(I+t^2L)^{-1}(f) - f$ follows from these estimates, Minkowski's inequality, and the fact that $t \le 1 = \ell(Q)$. □

8.7.3 Reduction to a Quadratic Estimate

We are given a divergence form elliptic operator as in (8.7.2) with ellipticity constants λ and Λ in (8.7.1). Our goal is to obtain the a priori estimate (8.7.4) for functions f in some dense subspace of $\dot{L}_1^2(\mathbf{R}^n)$.

To obtain this estimate we need to resolve the operator $\sqrt{L}$ as an average of simpler operators that are uniformly bounded from $\dot{L}_1^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$. In the sequel we use the following resolution of the square root:

$$\sqrt{L}(f) = \frac{16}{\pi}\int_0^\infty (I+t^2L)^{-3} t^3 L^2(f)\,\frac{dt}{t},$$

in which the integral converges in $L^2(\mathbf{R}^n)$ for $f \in \mathscr{C}_0^\infty(\mathbf{R}^n)$. Take $g \in \mathscr{C}_0^\infty(\mathbf{R}^n)$ with $\|g\|_{L^2} = 1$. Using duality and the Cauchy–Schwarz inequality, we can control the quantity $\big|\big\langle \sqrt{L}(f)\,|\,g\big\rangle\big|^2$ by

$$\frac{256}{\pi^2}\left(\int_0^\infty \big\|(I+t^2L)^{-1} tL(f)\big\|_2^2\,\frac{dt}{t}\right)\left(\int_0^\infty \big\|V_t(g)\big\|_{L^2}^2\,\frac{dt}{t}\right), \tag{8.7.14}$$

where we set

$$V_t = t^2L^*(I+t^2L^*)^{-2}.$$

Here L^* is the adjoint operator to L and note that the matrix corresponding to L^* is the conjugate-transpose matrix A^* of A (i.e., the transpose of the matrix whose entries are the complex conjugates of the matrix A). We explain why the estimate

$$\int_0^\infty \big\|V_t(g)\big\|_{L^2}^2\,\frac{dt}{t} \le C\big\|g\big\|_{L^2}^2 \tag{8.7.15}$$

is valid. Fix a real-valued function $\Psi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$ with mean value zero normalized so that

$$\int_0^\infty |\widehat{\Psi}(s\xi)|^2\,\frac{ds}{s} = 1$$

for all $\xi \in \mathbf{R}^n$ and define $\Psi_s(x) = \frac{1}{s^n}\Psi(\frac{x}{s})$. Throughout the proof, Q_s denotes the operator

$$Q_s(h) = h * \Psi_s. \tag{8.7.16}$$

Obviously we have

$$\int_0^\infty \|Q_s(g)\|_{L^2}^2 \frac{ds}{s} = \|g\|_{L^2}^2$$

for all L^2 functions g.

We obtain estimate (8.7.15) as a consequence of Corollary 8.6.4 applied to the operators V_t that have uniform (in t) bounded extensions on $L^2(\mathbf{R}^n)$. To apply Corollary 8.6.4, we need to check that condition (8.6.40) holds for $\Theta_t = V_t$. Since

$$V_t Q_s = -(I + t^2 L^*)^{-2} t^2 \operatorname{div} A^* \nabla Q_s\,,$$

we have

$$\|V_t Q_s\|_{L^2 \to L^2} \le \|(I + t^2 L^*)^{-2} t^2 \operatorname{div} A^*\|_{L^2 \to L^2} \|\nabla Q_s\|_{L^2 \to L^2} \le c\,\frac{t}{s}\,, \tag{8.7.17}$$

with C depending only on n, λ, and Λ. Choose $\Psi = \Delta\varphi$ with $\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$ radial so that in particular, $\Psi = \operatorname{div}\vec{h}$. This yields $Q_s = s\operatorname{div}\vec{R}_s$ with $\vec{R}_s$ uniformly bounded; hence

$$\|V_t Q_s\|_{L^2 \to L^2} \le \|t^2 L^* (I + t^2 L^*)^{-2} \operatorname{div}\|_{L^2 \to L^2} \|s\vec{R}_s\|_{L^2 \to L^2} \le c\,\frac{s}{t}\,, \tag{8.7.18}$$

with C depending only on n, λ, and Λ.

Combining (8.7.17) and (8.7.18) proves (8.6.40) with $\Theta_t = V_t$. Hence Corollary 8.6.4 is applicable and (8.7.15) follows.

Therefore, the second integral on the right-hand side of (8.7.14) is bounded, and estimate (8.7.4) is reduced to proving

$$\int_0^\infty \|(I + t^2 L)^{-1} t L(f)\|_2^2 \frac{dt}{t} \le C \int_{\mathbf{R}^n} |\nabla f|^2\, dx \tag{8.7.19}$$

for all $f \in \mathscr{C}_0^\infty(\mathbf{R}^n)$.

8.7.4 Reduction to a Carleson Measure Estimate

Our next goal is to reduce matters to a Carleson measure estimate. We first introduce some notation to be used throughout. For $\mathbf{C}^n$-valued functions $\vec{f} = (f_1, \dots, f_n)$ define

$$Z_t(\vec{f}) = -\sum_{k=1}^n \sum_{j=1}^n (I + t^2 L)^{-1} t \partial_j (a_{j,k} f_k)\,.$$

In short, we write $Z_t = -(I+t^2L)^{-1}t\,\mathrm{div}A$. With this notation, we reformulate (8.7.19) as

$$\int_0^\infty \big\|Z_t(\nabla f)\big\|_2^2 \frac{dt}{t} \le C\int_{\mathbf{R}^n} |\nabla f|^2\,dx. \tag{8.7.20}$$

Also, define

$$\gamma_t(x) = Z_t(\mathbf{1})(x) = \Big(-\sum_{j=1}^n (I+t^2L)^{-1}t\,\partial_j(a_{j,k})(x)\Big)_{1\le k\le n},$$

where $\mathbf{1}$ is the $n\times n$ identity matrix and the action of Z_t on $\mathbf{1}$ is columnwise.

The reduction to a Carleson measure estimate and to a $T(b)$ argument requires the following inequality:

$$\int_{\mathbf{R}^n}\int_0^\infty |\gamma_t(x)\cdot P_t^2(\nabla g)(x) - Z_t(\nabla g)(x)|^2\,\frac{dx\,dt}{t} \le C\int_{\mathbf{R}^n} |\nabla g|^2\,dx, \tag{8.7.21}$$

where C depends only on n, λ, and Λ. Here, P_t denotes the operator

$$P_t(h) = h * p_t\,, \tag{8.7.22}$$

where $p_t(x) = t^{-n}p(t^{-1}x)$ and p denotes a nonnegative smooth function supported in the unit ball of $\mathbf{R}^n$ with integral equal to 1. To prove this, we need to handle Littlewood–Paley theory in a setting a bit more general than the one encountered in the previous section.

Lemma 8.7.5. *For $t>0$, let U_t be integral operators defined on $L^2(\mathbf{R}^n)$ with measurable kernels $L_t(x,y)$. Suppose that for some $m>n$ and for all $y\in\mathbf{R}^n$ and $t>0$ we have*

$$\int_{\mathbf{R}^n}\Big(1+\frac{|x-y|}{t}\Big)^{2m} |L_t(x,y)|^2\,dx \le t^{-n}. \tag{8.7.23}$$

Assume that for any ball $B(y,t)$, U_t has a bounded extension from $L^\infty(\mathbf{R}^n)$ to $L^2(B(y,t))$ such that for all f in $L^\infty(\mathbf{R}^n)$ and $y\in\mathbf{R}^n$ we have

$$\frac{1}{t^n}\int_{B(y,t)} |U_t(f)(x)|^2\,dx \le \big\|f\big\|_{L^\infty}^2. \tag{8.7.24}$$

Finally, assume that $U_t(1)=0$ in the sense that

$$U_t(\chi_{B(0,R)}) \to 0 \qquad \textit{in} \qquad L^2(B(y,t)) \tag{8.7.25}$$

as $R\to\infty$ for all $y\in\mathbf{R}^n$ and $t>0$.

Let Q_s and P_t be as in (8.7.16) and (8.7.22), respectively. Then for some $\alpha>0$ and C depending on n and m we have

$$\big\|U_tP_tQ_s\big\|_{L^2\to L^2} \le C\min\Big(\frac{t}{s},\frac{s}{t}\Big)^\alpha \tag{8.7.26}$$

and also

$$\left\| U_t Q_s \right\|_{L^2 \to L^2} \le C \left(\frac{t}{s} \right)^{\alpha}, \qquad t \le s. \tag{8.7.27}$$

Proof. We begin by observing that $U_t^* U_t$ has a kernel $K_t(x,y)$ given by

$$K_t(x,y) = \int_{\mathbf{R}^n} \overline{L_t(z,x)} L_t(z,y)\, dz.$$

The simple inequality $(1+a+b) \le (1+a)(1+b)$ for $a,b>0$ combined with the Cauchy–Schwarz inequality and (8.7.23) yield that $\left(1+\frac{|x-y|}{t}\right)^m |K_t(x,y)|$ is bounded by

$$\int_{\mathbf{R}^n} \left(1+\frac{|x-z|}{t}\right)^m |L_t(z,x)|\, |L_t(z,y)| \left(1+\frac{|z-y|}{t}\right)^m dy \le t^{-n}.$$

We conclude that

$$|K_t(x,y)| \le \frac{1}{t^n} \left(1+\frac{|x-y|}{t}\right)^{-m}. \tag{8.7.28}$$

Hence $U_t^* U_t$ is bounded on all L^p, $1 \le p \le +\infty$, and in particular, for $p=2$. Since L^2 is a Hilbert space, it follows that U_t is bounded on $L^2(\mathbf{R}^n)$ uniformly in $t>0$.

For $s \le t$ we use that $\left\| U_t \right\|_{L^2 \to L^2} \le B < \infty$ and basic estimates to deduce that

$$\left\| U_t P_t Q_s \right\|_{L^2 \to L^2} \le B \left\| P_t Q_s \right\|_{L^2 \to L^2} \le CB \left(\frac{s}{t} \right)^{\alpha}.$$

Next, we consider the case $t \le s$. Since P_t has an integrable kernel, and the kernel of $U_t^* U_t$ satisfies (8.7.28), it follows that $W_t = U_t^* U_t P_t$ has a kernel that satisfies a similar estimate. If we prove that $W_t(1) = 0$, then we can deduce from standard arguments that when $t \le s$ we have

$$\left\| W_t Q_s \right\|_{L^2 \to L^2} \le C \left(\frac{t}{s} \right)^{2\alpha} \tag{8.7.29}$$

for $0 < \alpha < m-n$. This would imply the required estimate (8.7.26), since

$$\left\| U_t P_t Q_s \right\|_{L^2 \to L^2}^2 = \left\| Q_s^* P_t U_t^* U_t P_t Q_s \right\|_{L^2 \to L^2} \le C \left\| U_t^* U_t P_t Q_s \right\|_{L^2 \to L^2}.$$

We have that $W_t(1) = U_t^* U_t(1)$. Suppose that a function φ in $L^2(\mathbf{R}^n)$ is compactly supported. Then φ is integrable over $\mathbf{R}^n$ and we have

$$\left\langle U_t^* U_t(1) \,|\, \varphi \right\rangle = \lim_{R\to\infty} \left\langle U_t^* U_t(\chi_{B(0,R)}) \,|\, \varphi \right\rangle = \lim_{R\to\infty} \left\langle U_t(\chi_{B(0,R)}) \,|\, U_t(\varphi) \right\rangle.$$

We have

$$\left\langle U_t(\chi_{B(0,R)}) \,|\, U_t(\varphi) \right\rangle = \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} U_t(\chi_{B(0,R)})(x) \overline{U_t(x,y)\varphi(y)}\, dy\, dx,$$

and this is in absolute value at most a constant multiple of

$$\left(t^{-n}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\left(1+\frac{|x-y|}{t}\right)^{-2m}|U_t(\chi_{B(0,R)})(x)|^2|\varphi(y)|\,dy\,dx\right)^{\frac{1}{2}}\|\varphi\|_{L^1}^{\frac{1}{2}}$$

by (8.7.23) and the Cauchy–Schwarz inequality for the measure $|\varphi(y)|\,dy\,dx$. Using a covering in the x variable by a family of balls $B(y+ckt,t)$, $k\in\mathbf{Z}^n$, we deduce easily that the last displayed expression is at most

$$C_\varphi\left(\sum_{k\in\mathbf{Z}^n}\int_{\mathbf{R}^n}(1+|k|)^{-2m}c_R(y,k)|\varphi(y)|\,dy\right)^{\frac{1}{2}},$$

where C_φ is a constant that depends on φ and

$$c_R(y,k)=t^{-n}\int_{B(y+ckt,t)}|U_t(\chi_{B(0,R)})(x)|^2\,dx.$$

Applying the dominated convergence theorem and invoking (8.7.24) and (8.7.25) as $R\to\infty$, we conclude that $\langle U_t^*U_t(1)\,|\,\varphi\rangle=0$. The latter implies that $U_t^*U_t(1)=0$. The same conclusion follows for W_t, since $P_t(1)=1$.

To prove (8.7.27) when $t\le s$ we repeat the previous argument with $W_t=U_t^*U_t$. Since $W_t(1)=0$ and W_t has a nice kernel, it follows that (8.7.29) holds. Thus

$$\big\|U_tQ_s\big\|_{L^2\to L^2}^2=\big\|Q_s^*U_t^*U_tQ_s\big\|_{L^2\to L^2}\le C\big\|U_t^*U_tQ_s\big\|_{L^2\to L^2}\le C\left(\frac{t}{s}\right)^{2\alpha}.$$

This concludes the proof of the lemma. □

Lemma 8.7.6. *Let P_t be as in Lemma 8.7.5. Then the operator U_t defined by $U_t(\vec f)(x)=\gamma_t(x)\cdot P_t(\vec f)(x)-Z_tP_t(\vec f)(x)$ satisfies*

$$\int_0^\infty\big\|U_tP_t(\vec f)\big\|_{L^2}^2\frac{dt}{t}\le C\big\|\vec f\big\|_{L^2}^2,$$

where C depends only on n, λ, and Λ. Here the action of P_t on $\vec f$ is componentwise.

Proof. By the off-diagonal estimates of Lemma 8.7.2 for Z_t and the fact that p has support in the unit ball, it is simple to show that there is a constant C depending on n, λ, and Λ such that for all $y\in\mathbf{R}^n$,

$$\frac{1}{t^n}\int_{B(y,t)}|\gamma_t(x)|^2\,dx\le C \tag{8.7.30}$$

and that the kernel of $C^{-1}U_t$ satisfies the hypotheses in Lemma 8.7.5. The conclusion follows from Corollary 8.6.4 applied to U_tP_t. □

We now return to (8.7.21). We begin by writing

$$\gamma_t(x)\cdot P_t^2(\nabla g)(x)-Z_t(\nabla g)(x)=U_tP_t(\nabla g)(x)+Z_t(P_t^2-I)(\nabla g)(x),$$

and we prove (8.7.21) for each term that appears on the right. For the first term we apply Lemma 8.7.6. Since P_t commutes with partial derivatives, we may use that

$$\|Z_t\nabla\|_{L^2\to L^2} = \|(I+t^2L)^{-1}tL\|_{L^2\to L^2} \le Ct^{-1},$$

and therefore we obtain for the second term

$$\begin{aligned}\int_{\mathbf{R}^n}\int_0^\infty |Z_t(P_t^2-I)(\nabla g)(x)|^2\,\frac{dx\,dt}{t} &\le C^2\int_{\mathbf{R}^n}\int_0^\infty |(P_t^2-I)(g)(x)|^2\,\frac{dt}{t^3}\,dx\\ &\le C^2c(p)\|\nabla g\|_2^2\end{aligned}$$

by Plancherel's theorem, where C depends only on n, λ, and Λ. This concludes the proof of (8.7.21).

Lemma 8.7.7. *The required estimate (8.7.4) follows from the Carleson measure estimate*

$$\sup_Q \frac{1}{|Q|}\int_Q\int_0^{\ell(Q)} |\gamma_t(x)|^2\,\frac{dx\,dt}{t} < \infty, \tag{8.7.31}$$

where the supremum is taken over all cubes in $\mathbf{R}^n$ *with sides parallel to the axes.*

Proof. Indeed, (8.7.31) and Theorem 7.3.7 imply

$$\int_{\mathbf{R}^n}\int_0^\infty |P_t^2(\nabla g)(x)\cdot\gamma_t(x)|^2\,\frac{dx\,dt}{t} \le C\int_{\mathbf{R}^n}|\nabla g|^2\,dx,$$

and together with (8.7.21) we deduce that (8.7.20) holds. □

Next we introduce an auxiliary averaging operator. We define a dyadic averaging operator S_t^Q as follows:

$$S_t^Q(\vec{f})(x) = \left(\frac{1}{|Q'_x|}\int_{Q'_x}\vec{f}(y)\,dy\right)\chi_{Q'_x}(x),$$

where Q'_x is the unique dyadic cube contained in Q that contains x and satisfies $\frac{1}{2}\ell(Q'_x) < t \le \ell(Q'_x)$. Notice that S_t^Q is a projection, i.e., it satisfies $S_t^Q S_t^Q = S_t^Q$. We have the following technical lemma concerning S_t^Q.

Lemma 8.7.8. *For some C depending only on n, λ, and Λ, we have*

$$\int_Q\int_0^{\ell(Q)} |\gamma_t(x)\cdot(S_t^Q-P_t^2)(\vec{f})(x)|^2\,\frac{dx\,dt}{t} \le C\int_{\mathbf{R}^n}|\vec{f}|^2\,dx. \tag{8.7.32}$$

Proof. We actually obtain a stronger version of (8.7.32) in which the t-integration on the left is taken over $(0,+\infty)$. Let Q_s be as in (8.7.16). Set $\Theta_t = \gamma_t\cdot(S_t^Q-P_t^2)$. The proof of (8.7.32) is based on Corollary 8.6.4 provided we show that for some $\alpha>0$,

$$\|\Theta_t Q_s\|_{L^2\to L^2} \le C\min\left(\frac{t}{s},\frac{s}{t}\right)^\alpha.$$

Suppose first that $t \le s$. Notice that $\Theta_t(1) = 0$, and thus (8.7.25) holds. With the aid of (8.7.30), we observe that Θ_t satisfies the hypotheses (8.7.23) and (8.7.24) of Lemma 8.7.5. Conclusion (8.7.27) of this lemma yields that for some $\alpha > 0$ we have

$$\big\|\Theta_t Q_s\big\|_{L^2\to L^2} \le C\left(\frac{t}{s}\right)^{\alpha}.$$

We now turn to the case $s \le t$. Since the kernel of P_t is bounded by $ct^{-n}\chi_{|x-y|\le t}$, condition (8.7.30) yields that $\gamma_t P_t$ is uniformly bounded on L^2 and thus

$$\big\|\gamma_t P_t^2 Q_s\big\|_{L^2\to L^2} \le C\big\|P_t Q_s\big\|_{L^2\to L^2} \le C'\frac{s}{t}.$$

It remains to consider the case $s \le t$ for the operator $U_t = \gamma_t \cdot S_t^Q$. We begin by observing that U_t is L^2 bounded uniformly in $t > 0$; this follows from a standard $U_t^* U_t$ argument using condition (8.7.23). Secondly, as already observed, S_t^Q is an orthogonal projection. Therefore, we have

$$\begin{aligned}
\big\|(\gamma_t \cdot S_t^Q)Q_s\big\|_{L^2\to L^2} &\le \big\|(\gamma_t \cdot S_t^Q)S_t^Q Q_s\big\|_{L^2\to L^2}\\
&\le \big\|S_t^Q Q_s\big\|_{L^2\to L^2}\\
&\le \big\|S_t^Q\big\|_{L^2\to \dot L^2_\alpha}\big\|Q_s\big\|_{\dot L^2_\alpha\to L^2}\\
&\le C s^{\alpha} t^{-\alpha}.
\end{aligned}$$

The last inequality follows from the facts that for any α in $(0,\frac12)$, Q_s maps the homogeneous Sobolev space $\dot L^2_\alpha$ to L^2 with norm at most a multiple of Cs^α and that the dyadic averaging operator S_t^Q maps $L^2(\mathbf{R}^n)$ to $\dot L^2_\alpha(\mathbf{R}^n)$ with norm $Ct^{-\alpha}$. The former of these statements is trivially verified by taking the Fourier transform, while the latter statement requires some explanation.

Fix an $\alpha \in (0,\frac12)$ and take $h, g \in L^2(\mathbf{R}^n)$. Also fix $j \in \mathbf{Z}$ such that $2^{-j-1} \le t < 2^{-j}$. We then have

$$\big\langle S_t^Q(-\Delta)^{\frac{\alpha}{2}}(h), g\big\rangle = \sum_{J_{j,k}\subseteq Q}\Big\langle(-\Delta)^{\frac{\alpha}{2}}(h), \chi_{J_{j,k}}(x)(\operatorname*{Avg}_{J_{j,k}} \overline{g})\Big\rangle,$$

where $J_{j,k} = \prod_{r=1}^n [2^{-j}k_r, 2^{-j}(k_r+1))$ and $k = (k_1,\dots,k_n)$. It follows that

$$\begin{aligned}
\big\langle S_t^Q(-\Delta)^{\frac{\alpha}{2}}(h), g\big\rangle &= \sum_{J_{j,k}\subseteq Q}\Big\langle h, \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big)(-\Delta)^{\frac{\alpha}{2}}(\chi_{J_{j,k}})(x)\Big\rangle\\
&= \Big\langle h, \sum_{J_{j,k}\subseteq Q} 2^{\alpha j}\big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big)(-\Delta)^{\frac{\alpha}{2}}(\chi_{[0,1)^n})(2^j(\cdot)-k)\Big\rangle.
\end{aligned}$$

Set $\chi_\alpha = (-\Delta)^{\frac{\alpha}{2}}(\chi_{[0,1)^n})$. We estimate the L^2 norm of the preceding sum. We have

$$
\begin{aligned}
&\int_{\mathbf{R}^n} \Big| \sum_{J_{j,k} \subseteqq Q} 2^{\alpha j} \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big)\, \chi_\alpha(2^j x - k) \Big|^2 dx \\
&\qquad = 2^{2\alpha j - nj} \int_{\mathbf{R}^n} \Big| \sum_{J_{j,k} \subseteqq Q} \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big)\, \chi_\alpha(x - k) \Big|^2 dx \\
&\qquad = 2^{2\alpha j - nj} \int_{\mathbf{R}^n} \Big| \sum_{J_{j,k} \subseteqq Q} e^{-2\pi i k\cdot\xi} \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big) \Big|^2 |\widehat{\chi_\alpha}(\xi)|^2 \, d\xi \\
&\qquad = 2^{2\alpha j - nj} \int_{[0,1]^n} \Big| \sum_{J_{j,k} \subseteqq Q} e^{-2\pi i k\cdot\xi} \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big) \Big|^2 \sum_{l \in \mathbf{Z}^n} |\widehat{\chi_\alpha}(\xi + l)|^2 \, d\xi \\
&\qquad \le 2^{2\alpha j - nj} \int_{[0,1]^n} \Big| \sum_{J_{j,k} \subseteqq Q} e^{-2\pi i k\cdot\xi} \big(\operatorname*{Avg}_{J_{j,k}} \overline{g}\big) \Big|^2 d\xi \sup_{\xi \in [0,1]^n} \sum_{l \in \mathbf{Z}^n} |\widehat{\chi_\alpha}(\xi + l)|^2 \\
&\qquad = 2^{2\alpha j - nj} \sum_{k \in \mathbf{Z}^n} \big| \operatorname*{Avg}_{J_{j,k}} \overline{g} \big|^2 C(n,\alpha)^2 ,
\end{aligned}
$$

where we used Plancherel's identity on the torus (Proposition 3.1.16) and we set

$$
C(n,\alpha)^2 = \sup_{\xi \in [0,1]^n} \sum_{l \in \mathbf{Z}^n} |\widehat{\chi_\alpha}(\xi + l)|^2 .
$$

Since

$$
\widehat{\chi_\alpha}(\xi) = |\xi|^\alpha \prod_{r=1}^{n} \frac{1 - e^{-2\pi i \xi_r}}{2\pi i \xi_r} ,
$$

it follows that $C(n,\alpha) < \infty$ when $0 < \alpha < \frac{1}{2}$. In this case we conclude that

$$
\begin{aligned}
\big| \big\langle S_t^Q (-\Delta)^{\frac{\alpha}{2}}(h), g \big\rangle \big| \;&\le\; C(n,\alpha) \big\| h \big\|_{L^2} 2^{j\alpha} \Big(2^{-nj} \sum_{k \in \mathbf{Z}^n} \big| \operatorname*{Avg}_{J_{j,k}} \overline{g} \big|^2 \Big)^{\frac{1}{2}} \\
&\le\; C' \big\| h \big\|_{L^2} t^{-\alpha} \big\| g \big\|_{L^2} ,
\end{aligned}
$$

and this implies that $\big\| S_t^Q \big\|_{L^2 \to \dot{L}^2_\alpha} \le C t^{-\alpha}$ and hence the required conclusion. □

8.7.5 The $T(b)$ Argument

To obtain (8.7.31), we adapt the $T(b)$ theorem of the previous section for square roots of divergence form elliptic operators. We fix a cube Q with center c_Q, an $\varepsilon \in (0,1)$, and a unit vector w in $\mathbf{C}^n$. We define a scalar-valued function

$$
f^{\varepsilon}_{Q,w} = (1 + (\varepsilon \ell(Q))^2 L)^{-1} (\Phi_Q \cdot \overline{w}) , \tag{8.7.33}
$$

where

$$
\Phi_Q(x) = x - c_Q .
$$

We begin by observing that the following estimates are consequences of Lemma 8.7.4:

$$\int_{5Q} |f^{\varepsilon}_{Q,w} - \Phi_Q \cdot \overline{w}|^2 \, dx \le C_1 \varepsilon^2 \ell(Q)^2 |Q| \tag{8.7.34}$$

and

$$\int_{5Q} |\nabla (f^{\varepsilon}_{Q,w} - \Phi_Q \cdot \overline{w})|^2 \, dx \le C_2 |Q| \,, \tag{8.7.35}$$

where C_1, C_2 depend on n, λ, Λ and not on ε, Q, and w. It is important to observe that the constants C_1, C_2 are independent of ε.

The proof of (8.7.31) follows by combining the next two lemmas. The rest of this section is devoted to their proofs.

Lemma 8.7.9. *There exists an $\varepsilon > 0$ depending on n, λ, Λ, and a finite set $\mathscr{F}$ of unit vectors in $\mathbf{C}^n$ whose cardinality depends on ε and n, such that*

$$\begin{aligned} &\sup_Q \frac{1}{|Q|} \int_Q \int_0^{\ell(Q)} |\gamma_t(x)|^2 \, \frac{dx\,dt}{t} \\ &\qquad \le C \sum_{w \in \mathscr{F}} \sup_Q \frac{1}{|Q|} \int_Q \int_0^{\ell(Q)} |\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)|^2 \, \frac{dx\,dt}{t} \,, \end{aligned}$$

where C depends only on ε, n, λ, and Λ. The suprema are taken over all cubes Q in $\mathbf{R}^n$ with sides parallel to the axes.

Lemma 8.7.10. *For C depending only on n, λ, Λ, and $\varepsilon > 0$, we have*

$$\int_Q \int_0^{\ell(Q)} \big|\gamma_t(x) \cdot (S^Q_t \nabla f^{\varepsilon}_{Q,w})(x)\big|^2 \, \frac{dx\,dt}{t} \le C|Q|. \tag{8.7.36}$$

We begin with the proof of Lemma 8.7.10, which is the easiest of the two.

Proof of Lemma 8.7.10. Pick a smooth bump function $\mathscr{X}_Q$ localized on $4Q$ and equal to 1 on $2Q$ with $\|\mathscr{X}_Q\|_{L^\infty} + \ell(Q)\|\nabla \mathscr{X}_Q\|_{L^\infty} \le c_n$. By Lemma 8.7.5 and estimate (8.7.21), the left-hand side of (8.7.36) is bounded by

$$\begin{aligned} &C \int_{\mathbf{R}^n} \big|\nabla (\mathscr{X}_Q f^{\varepsilon}_{Q,w})\big|^2 \, dx + 2 \int_Q \int_0^{\ell(Q)} \big|\gamma_t(x) \cdot (P_t^2 \nabla (\mathscr{X}_Q f^{\varepsilon}_{Q,w}))(x)\big|^2 \, \frac{dx\,dt}{t} \\ &\qquad \le C \int_{\mathbf{R}^n} \big|\nabla (\mathscr{X}_Q f^{\varepsilon}_{Q,w})\big|^2 \, dx + 4 \int_Q \int_0^{\ell(Q)} \big|(Z_t \nabla (\mathscr{X}_Q f^{\varepsilon}_{Q,w}))(x)\big|^2 \, \frac{dx\,dt}{t} \,. \end{aligned}$$

It remains to control the last displayed expression by $C|Q|$.

First, it follows easily from (8.7.34) and (8.7.35) that

$$\int_{\mathbf{R}^n} |\nabla (\mathscr{X}_Q f^{\varepsilon}_{Q,w})|^2 \, dx \le C|Q| \,,$$

where C is independent of Q and w (but it may depend on ε). Next, we write

$$Z_t \nabla(\mathscr{X}_Q f^{\varepsilon}_{Q,w}) = W^1_t + W^2_t + W^3_t\,,$$

where

$$\begin{aligned} W^1_t &= (I+t^2L)^{-1} t\left(\mathscr{X}_Q L(f^{\varepsilon}_{Q,w})\right), \\ W^2_t &= -(I+t^2L)^{-1} t\left(\operatorname{div}(A f^{\varepsilon}_{Q,w} \nabla \mathscr{X}_Q)\right), \\ W^3_t &= -(I+t^2L)^{-1} t\left(A\nabla f^{\varepsilon}_{Q,w} \cdot \nabla \mathscr{X}_Q\right), \end{aligned}$$

and we use different arguments to treat each term W^j_t.

To handle W^1_t, observe that

$$L(f^{\varepsilon}_{Q,w}) = \frac{f^{\varepsilon}_{Q,w} - \Phi_Q \cdot \overline{w}}{\varepsilon^2 \ell(Q)^2}\,,$$

and therefore it follows from (8.7.34) that

$$\int_{\mathbf{R}^n} |\mathscr{X}_Q L(f^{\varepsilon}_{Q,w})|^2 \le C|Q|(\varepsilon\ell(Q))^{-2}\,,$$

where C is independent of Q and w. Using the (uniform in t) boundedness of the operator $(I+t^2L)^{-1}$ on $L^2(\mathbf{R}^n)$, we obtain

$$\int_Q \int_0^{\ell(Q)} \left|W^1_t(x)\right|^2 \frac{dx\,dt}{t} \le \int_0^{\ell(Q)} \frac{C|Q|t^2}{(\varepsilon\ell(Q))^2}\,\frac{dt}{t} \le \frac{C|Q|}{\varepsilon^2}\,,$$

which establishes the required quadratic estimate for W^1_t.

To obtain a similar quadratic estimate for W^2_t, we apply Lemma 8.7.2 for the operator $(I+t^2L)^{-1}t\,\mathrm{div}$ with sets $F = Q$ and $E = \operatorname{supp}(f^{\varepsilon}_{Q,w}\nabla\mathscr{X}_Q) \subseteq 4Q \setminus 2Q$. We obtain that

$$\int_Q \int_0^{\ell(Q)} |W^2_t(x)|^2 \frac{dx\,dt}{t} \le C \int_0^{\ell(Q)} e^{-\frac{\ell(Q)}{ct}} \frac{dt}{t} \int_{4Q\setminus 2Q} |A f^{\varepsilon}_{Q,w} \nabla \mathscr{X}_Q|^2\,dx\,.$$

The first integral on the right provides at most a constant factor, while we handle the second integral by writing

$$f^{\varepsilon}_{Q,w} = (f^{\varepsilon}_{Q,w} - \Phi_Q \cdot \overline{w}) + \Phi_Q \cdot \overline{w}\,.$$

Using (8.7.34) and the facts that $\left\|\nabla \mathscr{X}_Q\right\|_{L^\infty} \le c_n \ell(Q)^{-1}$ and that $|\Phi_Q| \le c_n \ell(Q)$ on the support of $\mathscr{X}_Q$, we obtain that

$$\int_{4Q\setminus 2Q} |A f^{\varepsilon}_{Q,w} \nabla \mathscr{X}_Q|^2\,dx \le C\,|Q|\,,$$

where C depends only on n, λ, and Λ. This yields the required result for W^2_t.

To obtain a similar estimate for W^3_t, we use the (uniform in t) boundedness of $(I+t^2L)^{-1}$ on $L^2(\mathbf{R}^n)$ (Exercise 8.7.2) to obtain that

$$\int_Q \int_0^{\ell(Q)} |W_t^3(x)|^2 \frac{dx\,dt}{t} \leq C \int_0^{\ell(Q)} t^2 \frac{dt}{t} \int_{4Q\setminus 2Q} |A\nabla f_{Q,w}^{\varepsilon} \cdot \nabla \mathscr{X}_Q|^2\,dx.$$

But the last integral is shown easily to be bounded by $C|Q|$ by writing $f_{Q,w}^{\varepsilon}$, as in the previous case, and using (8.7.35) and the properties of $\mathscr{X}_Q$ and Φ_Q. Note that C here depends only on n, λ, and Λ. This concludes the proof of Lemma 8.7.10. □

8.7.6 The Proof of Lemma 8.7.9

It remains to prove Lemma 8.7.9. The main ingredient in the proof of Lemma 8.7.9 is the following proposition, which we state and prove first.

Proposition 8.7.11. *There exists an $\varepsilon > 0$ depending on n, λ, and Λ, and $\eta = \eta(\varepsilon) > 0$ such that for each unit vector w in $\mathbf{C}^n$ and each cube Q with sides parallel to the axes, there exists a collection $\mathscr{S}_w' = \{Q'\}$ of nonoverlapping dyadic subcubes of Q such that*

$$\Big|\bigcup_{Q'\in\mathscr{S}_w'} Q'\Big| \leq (1-\eta)|Q|, \tag{8.7.37}$$

and moreover, if $\mathscr{S}_w''$ is the collection of all dyadic subcubes of Q not contained in any $Q' \in \mathscr{S}_w'$, then for any $Q'' \in \mathscr{S}_w''$ we have

$$\frac{1}{|Q''|}\int_{Q''} \mathrm{Re}\,(\nabla f_{Q,w}^{\varepsilon}(y)\cdot w)\,dy \geq \frac{3}{4} \tag{8.7.38}$$

and

$$\frac{1}{|Q''|}\int_{Q''} |\nabla f_{Q,w}^{\varepsilon}(y)|^2\,dy \leq (4\varepsilon)^{-2}. \tag{8.7.39}$$

Proof. We begin by proving the following crucial estimate:

$$\left|\int_Q (1-\nabla f_{Q,w}^{\varepsilon}(x)\cdot w)\,dx\right| \leq C\varepsilon^{\frac{1}{2}}|Q|, \tag{8.7.40}$$

where C depends on n, λ, and Λ, but not on ε, Q, and w. Indeed, we observe that

$$\nabla(\Phi_Q\cdot\overline{w})(x)\cdot w = |w|^2 = 1\,,$$

so that

$$1-\nabla f_{Q,w}^{\varepsilon}(x)\cdot w = \nabla g_{Q,w}^{\varepsilon}(x)\cdot w\,,$$

where we set

$$g_{Q,w}^{\varepsilon}(x) = \Phi_Q(x)\cdot\overline{w} - f_{Q,w}^{\varepsilon}(x)\,.$$

Next we state another useful lemma, whose proof is postponed until the end of this subsection.

Lemma 8.7.12. *There exists a constant $C = C_n$ such that for all $h \in \dot{L}^2_1$ we have*

$$\left| \int_Q \nabla h(x)\, dx \right| \le C \ell(Q)^{\frac{n-1}{2}} \left(\int_Q |h(x)|^2\, dx \right)^{\frac{1}{4}} \left(\int_Q |\nabla h(x)|^2\, dx \right)^{\frac{1}{4}} .$$

Applying Lemma 8.7.12 to the function $g^{\varepsilon}_{Q,w}$, we deduce (8.7.40) as a consequence of (8.7.34) and (8.7.35).

We now proceed with the proof of Proposition 8.7.11. First we deduce from (8.7.40) that

$$\frac{1}{|Q|} \int_Q \operatorname{Re}\left(\nabla f^{\varepsilon}_{Q,w}(x) \cdot w\right) dx \ge \frac{7}{8},$$

provided that ε is small enough. We also observe that as a consequence of (8.7.35) we have

$$\frac{1}{|Q|} \int_Q |\nabla f^{\varepsilon}_{Q,w}(x)|^2\, dx \le C_3 ,$$

where C_3 is independent of ε. Now we perform a stopping-time decomposition to select a collection $\mathscr{S}'_w$ of dyadic subcubes of Q that are maximal with respect to either one of the following conditions:

$$\frac{1}{|Q'|} \int_{Q'} \operatorname{Re}\left(\nabla f^{\varepsilon}_{Q,w}(x) \cdot w\right) dx, \;\le\; \frac{3}{4} \tag{8.7.41}$$

$$\frac{1}{|Q'|} \int_{Q'} |\nabla f^{\varepsilon}_{Q,w}(x)|^2\, dx \;\ge\; (4\varepsilon)^{-2} . \tag{8.7.42}$$

This is achieved by subdividing Q dyadically and by selecting those cubes Q' for which either (8.7.41) or (8.7.42) holds, subdividing all the nonselected cubes, and repeating the procedure. The validity of (8.7.38) and (8.7.39) now follows from the construction and (8.7.41) and (8.7.42).

It remains to establish (8.7.37). Let B_1 be the union of the cubes in $\mathscr{S}'_w$ for which (8.7.41) holds. Also, let B_2 be the union of those cubes in $\mathscr{S}'_w$ for which (8.7.42) holds. We then have

$$\Big| \bigcup_{Q' \in \mathscr{S}'_w} Q' \Big| \le |B_1| + |B_2| .$$

The fact that the cubes in $\mathscr{S}'_w$ do not overlap yields

$$|B_2| \le (4\varepsilon)^2 \int_Q |\nabla f^{\varepsilon}_{Q,w}(x)|^2\, dx \le (4\varepsilon)^2 C_3 |Q|.$$

Setting $b^{\varepsilon}_{Q,w}(x) = 1 - \operatorname{Re}\left(\nabla f^{\varepsilon}_{Q,w}(x) \cdot w\right)$, we also have

$$|B_1| \le 4 \sum \int_{Q'} b^{\varepsilon}_{Q,w}\, dx = 4 \int_Q b^{\varepsilon}_{Q,w}\, dx - 4 \int_{Q \setminus B_1} b^{\varepsilon}_{Q,w}\, dx , \tag{8.7.43}$$

where the sum is taken over all cubes Q' that comprise B_1. The first term on the right in (8.7.43) is bounded above by $C\varepsilon^{\frac{1}{2}} |Q|$ in view of (8.7.40). The second term on the right in (8.7.43) is controlled in absolute value by

$$4|Q\setminus B_1|+4|Q\setminus B_1|^{\frac{1}{2}}(C_3|Q|)^{\frac{1}{2}}\le 4|Q\setminus B_1|+4C_3\varepsilon^{\frac{1}{2}}|Q|+\varepsilon^{-\frac{1}{2}}|Q\setminus B_1|.$$

Since $|Q\setminus B_1|=|Q|-|B_1|$, we obtain

$$(5+\varepsilon^{-\frac{1}{2}})|B_1|\le(4+C\varepsilon^{\frac{1}{2}}+\varepsilon^{-\frac{1}{2}})|Q|,$$

which yields $|B_1|\le(1-\varepsilon^{\frac{1}{2}}+o(\varepsilon^{\frac{1}{2}}))|Q|$ if ε is small enough. Hence

$$|B|\le(1-\eta(\varepsilon))|Q|$$

with $\eta(\varepsilon)\approx\varepsilon^{\frac{1}{2}}$ for small ε. This concludes the proof of Proposition 8.7.11. □

Next, we need the following simple geometric fact.

Lemma 8.7.13. *Let w,u,v be in $\mathbf{C}^n$ such that $|w|=1$ and let $0<\varepsilon\le 1$ be such that*

$$|u-(u\cdot\overline{w})w| \le \varepsilon\,|u\cdot\overline{w}|, \tag{8.7.44}$$

$$\mathrm{Re}\,(v\cdot w) \ge \frac{3}{4}, \tag{8.7.45}$$

$$|v| \le (4\varepsilon)^{-1}. \tag{8.7.46}$$

Then we have $|u|\le 4\,|u\cdot v|$.

Proof. It follows from (8.7.45) that

$$\tfrac{3}{4}\,|u\cdot\overline{w}|\le|(u\cdot\overline{w})(v\cdot w)|. \tag{8.7.47}$$

Moreover, (8.7.44) and the triangle inequality imply that

$$|u|\le(1+\varepsilon)|u\cdot\overline{w}|\le 2\,|u\cdot\overline{w}|. \tag{8.7.48}$$

Also, as a consequence of (8.7.44) and (8.7.46), we obtain

$$|(u-(u\cdot\overline{w})w)\cdot v|\le\tfrac{1}{4}\,|u\cdot\overline{w}|. \tag{8.7.49}$$

Finally, using (8.7.47) and (8.7.49) together with the triangle inequality, we deduce that

$$|u\cdot v|\ge|(u\cdot\overline{w})(v\cdot w)|-|(u-(u\cdot\overline{w})w)\cdot v|\ge(\tfrac{3}{4}-\tfrac{1}{4})\,|u\cdot\overline{w}|\ge\tfrac{1}{4}\,|u|,$$

where in the last inequality we used (8.7.48). □

We now proceed with the proof of Lemma 8.7.9. We fix an $\varepsilon>0$ to be chosen later and we choose a finite number of cones $\mathscr{C}_w$ indexed by a finite set $\mathscr{F}$ of unit vectors w in $\mathbf{C}^n$ defined by

$$\mathscr{C}_w=\left\{u\in\mathbf{C}^n:\ |u-(u\cdot\overline{w})w|\le\varepsilon\,|u\cdot\overline{w}|\right\}, \tag{8.7.50}$$

so that

$$\mathbf{C}^n = \bigcup_{w\in\mathscr{F}} \mathscr{C}_w .$$

Note that the size of the set $\mathscr{F}$ can be chosen to depend only on ε and the dimension n.

It suffices to show that for each fixed $w \in \mathscr{F}$ we have a Carleson measure estimate for $\gamma_{t,w}(x) \equiv \chi_{\mathscr{C}_w}(\gamma_t(x))\gamma_t(x)$, where $\chi_{\mathscr{C}_w}$ denotes the characteristic function of $\mathscr{C}_w$. To achieve this we define

$$A_w \equiv \sup_Q \frac{1}{|Q|}\int_Q\int_0^{\ell(Q)} |\gamma_{t,w}(x)|^2 \frac{dx\,dt}{t}, \tag{8.7.51}$$

where the supremum is taken over all cubes Q in $\mathbf{R}^n$ with sides parallel to the axes. By truncating $\gamma_{t,w}(x)$ for t small and t large, we may assume that this quantity is finite. Once an a priori bound independent of these truncations is obtained, we can pass to the limit by monotone convergence to deduce the same bound for $\gamma_{t,w}(x)$.

We now fix a cube Q and let $\mathscr{S}''_w$ be as in Proposition 8.7.11. We pick Q'' in $\mathscr{S}''_w$ and we set

$$v = \frac{1}{|Q''|}\int_{Q''} \nabla f^{\varepsilon}_{Q,w}(y)\,dy \in \mathbf{C}^n.$$

It is obvious that statements (8.7.38) and (8.7.39) in Proposition 8.7.11 yield conditions (8.7.45) and (8.7.46) of Lemma 8.7.13. Set $u = \gamma_{t,w}(x)$ and note that if $x \in Q''$ and $\frac{1}{2}\ell(Q'') < t \le \ell(Q'')$, then $v = S^Q_t(\nabla f^{\varepsilon}_{Q,w})(x)$; hence

$$\big|\gamma_{t,w}(x)\big| \le 4\,\big|\gamma_{t,w}(x)\cdot S^Q_t(\nabla f^{\varepsilon}_{Q,w})(x)\big| \le 4\,\big|\gamma_t(x)\cdot S^Q_t(\nabla f^{\varepsilon}_{Q,w})(x)\big| \tag{8.7.52}$$

from Lemma 8.7.13 and the definition of $\gamma_{t,w}(x)$.

We partition the Carleson region $Q\times(0,\ell(Q)]$ as a union of boxes $Q'\times(0,\ell(Q')]$ for Q' in $\mathscr{S}'_w$ and Whitney rectangles $Q''\times(\frac{1}{2}\ell(Q''),\ell(Q'')]$ for Q'' in $\mathscr{S}''_w$. This allows us to write

$$\begin{aligned}\int_Q\int_0^{\ell(Q)} |\gamma_{t,w}(x)|^2 \frac{dx\,dt}{t} &= \sum_{Q'\in\mathscr{S}'_w}\int_{Q'}\int_0^{\ell(Q')} |\gamma_{t,w}(x)|^2 \frac{dx\,dt}{t} \\ &\quad + \sum_{Q''\in\mathscr{S}''_w}\int_{Q''}\int_{\frac{1}{2}\ell(Q'')}^{\ell(Q'')} |\gamma_{t,w}(x)|^2 \frac{dx\,dt}{t}.\end{aligned}$$

First observe that

$$\sum_{Q'\in\mathscr{S}'_w}\int_{Q'}\int_0^{\ell(Q')} |\gamma_{t,w}(x)|^2 \frac{dx\,dt}{t} \le \sum_{Q'\in\mathscr{S}'_w} A_w|Q'|A_w(1-\eta)|Q|.$$

Second, using (8.7.52), we obtain

$$\sum_{Q''\in\mathscr{S}''_w}\int_{Q''}\int_{\frac{1}{2}\ell(Q'')}^{\ell(Q'')}|\gamma_{t,w}(x)|^2\,\frac{dx\,dt}{t}$$
$$\leq 16\sum_{Q''\in\mathscr{S}''_w}\int_{Q''}\int_{\frac{1}{2}\ell(Q'')}^{\ell(Q'')}|\gamma_t(x)\cdot S_t^Q(\nabla f_{Q,w}^\varepsilon)(x)|^2\,\frac{dx\,dt}{t}$$
$$\leq 16\int_Q\int_0^{\ell(Q)}|\gamma_t(x)\cdot S_t^Q(\nabla f_{Q,w}^\varepsilon)(x)|^2\,\frac{dx\,dt}{t}\,.$$

Altogether, we obtain the bound

$$\int_Q\int_0^{\ell(Q)}|\gamma_{t,w}(x)|^2\,\frac{dx\,dt}{t}$$
$$\leq A_w(1-\eta)|Q|+16\int_Q\int_0^{\ell(Q)}|\gamma_t(x)\cdot S_t^Q(\nabla f_{Q,w}^\varepsilon)(x)|^2\,\frac{dx\,dt}{t}\,.$$

We divide by $|Q|$, we take the supremum over all cubes Q with sides parallel to the axes, and we use the definition and the finiteness of A_w to obtain the required estimate

$$A_w\leq 16\,\eta^{-1}\sup_Q\frac{1}{|Q|}\int_Q\int_0^{\ell(Q)}|\gamma_t(x)\cdot S_t^Q(\nabla f_{Q,w}^\varepsilon)(x)|^2\,\frac{dx\,dt}{t}\,,$$

thus concluding the proof of the lemma. □

We end by verifying the validity of Lemma 8.7.12 used earlier.

Proof of Lemma 8.7.12. For simplicity we may take Q to be the cube $[-1,1]^n$. Once this case is established, the case of a general cube follows by translation and rescaling. Set

$$M=\left(\int_Q|h(x)|^2\,dx\right)^{\frac{1}{2}},\qquad M'=\left(\int_Q|\nabla h(x)|^2\,dx\right)^{\frac{1}{2}}.$$

If $M\geq M'$, there is nothing to prove, so we may assume that $M<M'$. Take $t\in(0,1)$ and $\varphi\in\mathscr{C}_0^\infty(Q)$ with $\varphi(x)=1$ when $\operatorname{dist}(x,\partial Q)\geq t$ and $0\leq\varphi\leq 1$, $\|\nabla\varphi\|_{L^\infty}\leq C/t$, $C=C(n)$; here the distance is taken in the L^∞ norm of $\mathbf{R}^n$. Then

$$\int_Q\nabla h(x)\,dx=\int_Q(1-\varphi(x))\nabla h(x)\,dx-\int_Q h(x)\nabla\varphi(x)\,dx\,,$$

and the Cauchy–Schwarz inequality yields

$$\left|\int_Q\nabla h(x)\,dx\right|\leq C(M'\,t^{\frac{1}{2}}+M\,t^{-\frac{1}{2}})\,.$$

Choosing $t=M/M'$, we conclude the proof of the lemma. □

The proof of Theorem 8.7.1 is now complete. □

Exercises

8.7.1. Let A and L be as in the statement of Theorem 8.7.1.
(a) Consider the generalized heat equation

$$\frac{\partial u}{\partial t} - \operatorname{div}(A\nabla u) = 0$$

on $\mathbf{R}^{n+1}_+$ with initial condition $u(0,x) = u_0$. Assume a uniqueness theorem for solutions of these equations to obtain that the solution of the equation in part (a) is

$$u(t,x) = e^{-tL}(u_0).$$

(b) Take $u_0 = 1$ to deduce the identity

$$e^{-tL}(1) = 1$$

for all $t > 0$. Conclude that the family of $\{e^{-tL}\}_{t>0}$ is an approximate identity, in the sense that

$$\lim_{t\to 0} e^{-tL} = I.$$

8.7.2. Let L be as in (8.7.2). Show that the operators

$$\begin{aligned} L_1 &= (I+t^2L)^{-1}, \\ L_2 &= t\nabla(I+t^2L)^{-1}, \\ L_3 &= (I+t^2L)^{-1}t\operatorname{div} \end{aligned}$$

are bounded on $L^2(\mathbf{R}^n)$ uniformly in t with bounds depending only on n, λ, and Λ. $\big[$*Hint:* The L^2 boundedness of L_3 follows from that of L_2 via duality and integration by parts. To prove the L^2 boundedness of L_1 and L_2, let $u_t = (I+t^2L)^{-1}(f)$. Then $u_t + t^2L(u_t) = f$, which implies $\int_{\mathbf{R}^n} |u_t|^2\,dx + t^2\int_{\mathbf{R}^n} \overline{u_t}\,L(u_t)\,dx = \int_{\mathbf{R}^n} \overline{u_t}\,f\,dx$. The definition of L and integration by parts yield $\int_{\mathbf{R}^n} |u_t|^2\,dx + t^2\int_{\mathbf{R}^n} A\nabla u_t\,\nabla\overline{u_t}\,dx = \int_{\mathbf{R}^n} \overline{u_t}\,f\,dx$. Apply the ellipticity condition to bound the left side of this identity from below by $\int_{\mathbf{R}^n} |u_t|^2\,dx + \lambda\int_{\mathbf{R}^n} |t\nabla u_t|^2\,dx$. Also $\int_{\mathbf{R}^n} \overline{u_t}\,f\,dx$ is at most $\varepsilon^{-1}\int_{\mathbf{R}^n} |f|^2\,dx + \varepsilon\int_{\mathbf{R}^n} |u_t|^2\,dx$ by the Cauchy–Schwarz inequality. Choose ε small enough to complete the proof when $\|u_t\|_{L^2} < \infty$. In the case $\|u_t\|_{L^2} = \infty$, multiply the identity $u_t + t^2L(u_t) = f$ by $\overline{u_t}\eta_R$, where η_R is a suitable cutoff localized in a ball $B(0,R)$, and use the idea of Lemma 8.7.2. Then let $R\to\infty$.$\big]$

8.7.3. Let L be as in the proof of Theorem 8.7.1.
(a) Show that for all $t > 0$ we have

$$(I+t^2L^2)^{-2} = \int_0^\infty e^{-u(I+t^2L)}u\,du$$

by checking the identities

$$\int_0^\infty (I+t^2L)^2 e^{-u(I+t^2L)}u\,du = \int_0^\infty e^{-u(I+t^2L)}(I+t^2L)^2 u\,du = I\,.$$

(b) Prove that the operator

$$T = \frac{4}{\pi}\int_0^\infty L(I+t^2L)^{-2}\,dt$$

satisfies $TT = L$.
(c) Conclude that the operator

$$S = \frac{16}{\pi}\int_0^{+\infty} t^3L^2(I+t^2L)^{-3}\,\frac{dt}{t}$$

satisfies $SS = L$, that is, S is the square root of L. Moreover, all the integrals converge in $L^2(\mathbf{R}^n)$ when restricted to functions in $f \in \mathscr{C}_0^\infty(\mathbf{R}^n)$.
$\big[$*Hint:* Part (a): Write $(I+t^2L)e^{-u(I+t^2L)} = -\frac{d}{du}(e^{-u(I+t^2L)})$, apply integration by parts twice, and use Exercise 8.7.1. Part (b): Write the integrand as in part (a) and use the identity

$$\int_0^\infty\int_0^\infty e^{-(ut^2+vs^2)L}L^2\,dt\,ds = \frac{\pi}{4}(uv)^{-\frac{1}{2}}\int_0^\infty e^{-r^2L}L^2\,2r\,dr.$$

Set $\rho = r^2$ and use $e^{-\rho L}L = \frac{d}{d\rho}(e^{-\rho L})$. Part (c): Show that $T = S$ using an integration by parts starting with the identity $L = \frac{d}{dt}(tL)$.$\big]$

8.7.4. Suppose that μ is a measure on $\mathbf{R}^{n+1}_+$. For a cube Q in $\mathbf{R}^n$ we define the tent $T(Q)$ of Q as the set $Q\times(0,\ell(Q))$. Suppose that there exist two positive constants $\alpha < 1$ and β such that for all cubes Q in $\mathbf{R}^n$ there exist subcubes Q_j of Q with disjoint interiors such that

1. $\left|Q\setminus\bigcup_j Q_j\right| > \alpha\,|Q|$,
2. $\mu\Big(T(Q)\setminus\bigcup_j T(Q_j)\Big) \le \beta\,|Q|$.

Then μ is a Carleson measure with constant

$$\big\|\mu\big\|_{\mathscr{C}} \le \frac{\beta}{\alpha}\,.$$

$\big[$*Hint:* We have

$$\begin{aligned}\mu(T(Q)) &\le \mu\Big(T(Q)\setminus\bigcup_j T(Q_j)\Big) + \sum_j \mu(T(Q_j))\\ &\le \beta|Q| + \big\|\mu\big\|_{\mathscr{C}}\sum_j |Q_j|\,,\end{aligned}$$

and the last expression is at most $(\beta + (1-\alpha)\|\mu\|_{\mathscr{C}})|Q|$. Assuming that $\|\mu\|_{\mathscr{C}} < \infty$, we obtain the required conclusion. In general, approximate the measure by a sequence of truncated measures.]

HISTORICAL NOTES

Most of the material in Sections 8.1 and 8.2 has been in the literature since the early development of the subject. Theorem 8.2.7 was independently obtained by Peetre [254], Spanne [286], and Stein [290].

The original proof of the $T(1)$ theorem obtained by David and Journé [103] stated that if $T(1)$, $T^t(1)$ are in BMO and T satisfies the weak boundedness property, then T is L^2 bounded. This proof is based on the boundedness of paraproducts and is given in Theorem 8.5.4. Paraproducts were first exploited by Bony [28] and Coifman and Meyer [81]. The proof of L^2 boundedness using condition (iv) given in the proof of Theorem 8.3.3 was later obtained by Coifman and Meyer [82]. The equivalent conditions (ii), (iii), and (vi) first appeared in Stein [292], while condition (iv) is also due to David and Journé [103]. Condition (i) appears in the article of Nazarov, Volberg, and Treil [245] in the context of nondoubling measures. The same authors [246] obtained a proof of Theorems 8.2.1 and 8.2.3 for Calderón–Zygmund operators on nonhomogeneous spaces. Multilinear versions of the $T(1)$ theorem were obtained by Christ and Journé [70], Grafakos and Torres [154], and Bényi, Demeter, Nahmod, Thiele, Torres, and Villaroya [20]. The article [70] also contains a proof of the quadratic $T(1)$ type Theorem 8.6.3. Smooth paraproducts viewed as bilinear operators have been studied by Bényi, Maldonado, Nahmod, and Torres [21] and Dini-continuous versions of them by Maldonado and Naibo [225].

The orthogonality Lemma 8.5.1 was first proved by Cotlar [94] for self-adjoint and mutually commuting operators T_j. The case of general noncommuting operators was obtained by Knapp and Stein [190]. Theorem 8.5.7 is due to Calderón and Vaillancourt [49] and is also valid for symbols of class $S^0_{\rho,\rho}$ when $0 \le \rho < 1$. For additional topics on pseudodifferential operators we refer to the books of Coifman and Meyer [81], Journé [180], Stein [292], Taylor [309], Torres [315], and the references therein. The last reference presents a careful study of the action of linear operators with standard kernels on general function spaces. The continuous version of the orthogonality Lemma 8.5.1 given in Exercise 8.5.8 is due to Calderón and Vaillancourt [49]. Conclusion (iii) in the orthogonality Lemma 8.5.1 follows from a general principle saying that if $\sum x_j$ is a series in a Hilbert space such that $\|\sum_{j\in F} x_j\| \le M$ for all finite sets F, then the series $\sum x_j$ converges in norm. This is a consequence of the Orlicz–Pettis theorem, which states that in any Banach space, if $\sum x_{n_j}$ converges weakly for every subsequence of integers n_j, then $\sum x_j$ converges in norm.

A nice exposition on the Cauchy integral that presents several historical aspects of its study is the book of Muskhelishvili [243]. See also the book of Journé [180]. Proposition 8.6.1 is due to Plemelj [265] when Γ is a closed Jordan curve. The L^2 boundedness of the first commutator $\mathscr{C}_1$ in Example 8.3.8 is due to Calderón [42]. The L^2 boundedness of the remaining commutators $\mathscr{C}_m$, $m \ge 2$, is due to Coifman and Meyer [80], but with bounds of order $m!\,\|A'\|^m_{L^\infty}$. These bounds are not as good as those obtained in Example 8.3.8 and do not suffice in obtaining the boundedness of the Cauchy integral by summing the series of commutators. The L^2 boundedness of the Cauchy integral when $\|A'\|_{L^\infty}$ is small enough is due to Calderón [43]. The first proof of the boundedness of the Cauchy integral with arbitrary $\|A'\|_{L^\infty}$ was obtained by Coifman, M[c]Intosh, and Meyer [79]. This proof is based on an improved operator norm for the commutators $\|\mathscr{C}_m\|_{L^2\to L^2} \le C_0 m^4 \|A'\|^m_{L^\infty}$. The quantity m^4 was improved by Christ and Journé [70] to $m^{1+\delta}$ for any $\delta > 0$; it is announced in Verdera [326] that Mateu and Verdera have improved this result by taking $\delta = 0$. Another proof of the L^2 boundedness of the Cauchy integral was given by David [102] by employing the following bootstrapping argument: If the Cauchy integral is L^2 bounded whenever $\|A'\|_{L^\infty} \le \varepsilon$, then it is also L^2 bounded whenever $\|A'\|_{L^\infty} \le \frac{10}{9}\varepsilon$. A refinement of this bootstrapping technique was independently obtained by Murai [241], who was also able to obtain the best possible bound for the operator norm $\|\widetilde{\mathcal{C}}_\Gamma\|_{L^2\to L^2} \le C\big(1+\|A'\|_{L^\infty}\big)^{1/2}$ in terms of $\|A'\|_{L^\infty}$. Here $\widetilde{\mathcal{C}}_\Gamma$ is the operator defined in (8.6.15). Note that the corresponding estimate for $\mathcal{C}_\Gamma$ involves the power $3/2$ instead of $1/2$. See the book of Murai [242] for this result and a variety of topics related to the commutators and the Cauchy integral. Two elementary proofs of the L^2 boundedness of the Cauchy integral were given by Coifman, Jones, and Semmes [77]. The first of these proofs uses complex variables and the second a pseudo-Haar basis of L^2 adapted to the accretive function $1+iA'$. A geometric proof was given by Melnikov and Verdera [231]. Other proofs were obtained by Verdera [326] and Tchamitchian [310]. The proof of boundedness of the Cauchy integral given in Section 8.6 is taken from Semmes [281]. The book of Christ [67] contains an insightful exposition of many of the preceding results and discusses connections between the Cauchy integral and analytic capacity. The book of David and Semmes [105] presents several extensions of the results in this chapter to singular integrals along higher-dimensional surfaces.

The $T(1)$ theorem is applicable to many problems only after a considerable amount of work; see, for instance, Christ [67] for the case of the Cauchy integral. A more direct approach to many problems was given by McIntosh and Meyer [224], who replaced the function 1 by an accretive function b and showed that any operator T with standard kernel that satisfies $T(b) = T^t(b) = 0$ and $\|M_bTM_b\|_{WB} < \infty$ must be L^2 bounded. (M_b here is the operator given by multiplication by b.) This theorem easily implies the boundedness of the Cauchy integral. David, Journé, and Semmes [104] generalized this theorem even further as follows: If b_1 and b_2 are para-accretive functions such that T maps $b_1\mathscr{C}_0^\infty \to (b_2\mathscr{C}_0^\infty)'$ and is associated with a standard kernel, then T is L^2 bounded if and only if $T(b_1) \in BMO$, $T^t(b_2) \in BMO$, and $\|M_{b_1}TM_{b_2}\|_{WB} < \infty$. This is called the $T(b)$ theorem. The article of Semmes [281] contains a different proof of this theorem in the special case $T(b) = 0$ and $T^t(1) = 0$ (Exercise 8.6.6). Our proof of Theorem 8.6.6 is based on ideas from [281]. An alternative proof of the $T(b)$ theorem was given by Fabes, Mitrea, and Mitrea [121] based on a lemma due to Krein [200]. Another version of the $T(b)$ theorem that is applicable to spaces with no Euclidean structure was obtained by Christ [66].

Theorem 8.7.1 was posed as a problem by Kato [181] for maximal accretive operators and reformulated by McIntosh [222], [223] for square roots of elliptic operators. The reformulation was motivated by counterexamples found to Kato's original abstract formulation, first by Lions [215] for maximal accretive operators, and later by McIntosh [220] for regularly accretive ones. The one-dimensional Kato problem and the boundeness of the Cauchy integral along Lipschitz curves are equivalent problems as shown by Kenig and Meyer [188]. See also Auscher, McIntosh, and Nahmod [8]. Coifman, Deng, and Meyer [73] and independently Fabes, Jerison, and Kenig [119], [120] solved the square root problem for small peturbations of the identity matrix. This method used multilinear expansions and can be extended to operators with smooth coefficients. McIntosh [221] considered coefficients in Sobolev spaces, Escauriaza in VMO (unpublished), and Alexopoulos [3] real Hölder coefficients using homogenization techniques. Peturbations of real symmetric matrices with L^∞ coefficients were treated in Auscher, Hofmann, Lewis, and Tchamitchian [10]. The solution of the two-dimensional Kato problem was obtained by Hofmann and McIntosh [164] using a previously derived $T(b)$ type reduction due to Auscher and Tchamitchian [9]. Hofmann, Lacey, and McIntosh [165] extended this theorem to the case in which the heat kernel of e^{-tL} satisfies Gaussian bounds. Theorem 8.7.1 was obtained by Auscher, Hofmann, Lacey, McIntosh, and Tchamitchian [11]; the exposition in the text is based on this reference. Combining Theorem 8.7.1 with a theorem of Lions [215], it follows that the domain of $\sqrt{L}$ is $\dot{L}_1^2(\mathbf{R}^n)$ and that for functions f in this space the equivalence of norms $\|\sqrt{L}(f)\|_{L^2} \approx \|\nabla f\|_{L^2}$ is valid.