ON THE L^1 BEHAVIOR OF EIGENFUNCTION EXPANSIONS AND SINGULAR INTEGRAL OPERATORS

Michael Christ[1] *and Christopher D. Sogge*[2]

1. INTRODUCTION

Let M be a compact, smooth manifold, without boundary, of dimension $n \geq 2$. Suppose that D is a pseudodifferential operator of the class $S^m_{1,0}$ on M, self-adjoint with respect to some measure μ with a smooth, nonvanishing density in local coordinates. Suppose further that either D is an elliptic differential operator whose principal symbol is real and nonnegative, or that $m = 1$ and D is a pseudodifferential operator whose symbol $a(x, \xi)$ has the property that

$$\lim_{s \to \infty} a(x, s\xi)$$

exists and is real and positive for all $\xi \neq 0$. For any such operator, $L^2(M, \mu)$ admits an orthogonal decomposition

$$L^2 = \bigoplus_{j=0}^{\infty} E_j$$

where each E_j is a finite-dimensional eigenspace of D with eigenvalue λ_j. These eigenvalues are distinct and form a discrete sequence which tends to $+\infty$. Denote by π_j the orthogonal projection of L^2 onto E_j. Then

$$S^0_t f = \sum_{\lambda_j \leq t} \pi_j f \to f$$

in L^2 norm as $j \to \infty$, for all $f \in L^2$, and

$$\|f\|^2_{L^2} = \sum \|\pi_j f\|^2_{L^2} .$$

[1] Supported by the Australian National University and the National Science Foundation of the USA.

[2] Supported in part by the National Science Foundation.

We are interested in the convergence of the series for $f \in L^1$ rather than L^2. However since the characteristic function of the unit ball is not a Fourier multiplier of $L^1(\mathbb{R}^n)$ (an elementary calculation), it cannot be true that $S_t^0 f \to f$ in L^1 for all $f \in L^1$, when M is the n- torus and D is the ordinary Laplace operator. In fact general transplantation theorems imply that L^1 convergence fails for all M and D.

But consider the Riesz means

$$S_t^\delta f(x) = \sum_{\lambda_j \le t} (1 - \lambda_j/t)^\delta \pi_j f(x) ,$$

for $\delta \ge 0$. It is well known in the context of Fourier series that $\{S_t^\delta f\}$ has better convergence properties when $\delta > 0$ than when $\delta = 0$, and in fact $S_t^\delta f \to f$ in L^1 norm as $t \to \infty$ for all $f \in L^1$ as soon as $\delta > 0$. Furthermore in $\mathbb{R}^n$, $n \ge 2$, the Bochner-Riesz means

$$S_t^\delta f(x) = \int_{|\xi| \le t} e^{2\pi i x \cdot \xi} (1 - |\xi|^2/t^2) \hat{f}(\xi)\, d\xi$$

are known to become better behaved as δ increases. In this case

$$\|S_t^\delta f - f\|_1 \to 0 \qquad \text{as } t \to \infty \text{ for all } f \in L^1$$

precisely when $\delta > (n-1)/2$, the so-called critical index. Moreover when δ equals the critical index a weaker result remains valid: $S_t^\delta f \to f$ in the weak L^1 norm, for all $f \in L^1$ [7]. This "norm" is

$$\|f\|_{1,\infty} = \sup_{\lambda > 0} \lambda \big| \{x : |f(x)| > \lambda\} \big| ,$$

and $L^{1,\infty}$ is the set of all measurable f with finite "norm". An operator bounded from L^1 to $L^{1,\infty}$ is said to be of weak type $(1,1)$. The Hardy-Littlewood maximal function and the Hilbert transform are fundamental examples of opertators which map L^1 boundedly to $L^{1,\infty}$ but not to L^1.

When $\delta < (n-1)/2$ even $L^{1,\infty}$ convergence fails for S_t^δ in $\mathbb{R}^n$. The delicacy of the situation is further evidenced by the fact that $S_t^\delta f$ need not converge almost everywhere to $f \in L^1$ when $\delta = (n-1)/2$ [24].

In the context of an elliptic operator on a compact manifold it was proved more recently [21] that still $S_t^\delta f \to f$ in L^1 for all $f \in L^1$, for all $\delta > (n-1)/2$, which still is a necessary restriction. Thus one was led to hope for the sharp result:

THEOREM A. *[11] Let M, D be as above and $\delta = (n-1)/2$. Then there exists $C < \infty$ such that for all $f \in L^1$ and $t \geq 0$,*

$$\|S_t^\delta f\|_{1,\infty} \leq C\|f\|_1 \,.$$

Moreover

$$\|S_t^\delta f - f\|_{1,\infty} \to 0 \qquad \text{as } t \to \infty$$

for all $f \in L^1$.

There is a very closely related result for Fourier multipliers in $\mathbb{R}^n$. Let b be a function supported in a compact subset of $\mathbb{R}^n$. Write

$$\mathbb{R}^n = \left\{(z', z_n) \in \mathbb{R}^{n-1} \times \mathbb{R}\right\} .$$

Suppose that $b \in C^\infty$ off of the hyperplane $\{z_n = 0\}$ and that for each β and i

$$\left|\frac{\partial^\beta}{\partial z'^\beta}\frac{\partial^i}{\partial z_n^i} b(z)\right| \leq C_{\beta,i}|z_n|^{\delta - i}$$

where as usual $\delta = (n-1)/2$. Suppose Φ is a C^∞ diffeomorphism of a neighbourhood of the support of b into $\mathbb{R}^n$, and let

$$m(\xi) = b(\Phi^{-1}(\xi))$$

and

$$(Tf)\hat{}(\xi) = \hat{f}(\xi)m(\xi)\,.$$

THEOREM B. *T is of weak type* $(1,1)$.

This was previously known only in the case where the manifold $\Phi(\{z_n = 0\})$ has nonvanishing scalar curvature at every point [7] (and only for a prototypical subclass of b's).

There are two obstacles to be overcome in proving Theorem A. First it is necessary to obtain some more explicit grasp of the operators S_t^δ than is afforded by their abstract definition in terms of eigenfunction expansions. The proof in [11] relies on Hörmander's work [17] on the Weyl formula for the distribution of the eigenvalues, following earlier work of Sogge [20], [21].

The second problem is how to establish weak (1,1) bounds, even for an operator which is described very explicitly. In the present article we discuss only this second issue, on which some progress has been made in a series of comparatively recent works [3], [7], [10], [11]. A fairly flexible method has been developed; we attempt to describe the method and several distinct results which it yields. In the last section of the article we sketch its application to Theorem B.

2. THE CLASSICAL THEORY

There are two prototypes for the various operators which we shall discuss. The maximal function of Hardy and Littlewood is

$$Mf(x) = \sup_{r>0} \int_{|y|\leq 1} |f(x-ry)|\, dy \tag{2.0}$$

for $F \in L^1_{loc}(\mathbb{R}^n)$. The fundamental result is that M is of weak type (1,1), which by interpolation implies L^p boundedness for all $p > 1$. Second is a class of singular integral operators treated by Calderón and Zygmund:

$$\begin{aligned} Tf(x) &= pv f * K(x) \\ &= \lim_{\epsilon\to 0} \int_{|y|\geq\epsilon} f(x-y)K(y)\, dy \end{aligned} \tag{2.1}$$

with three hypotheses:

(2.2) T is bounded on L^2

(2.3) $K(x) = |x|^{-n}\Omega(x/|x|)$

(2.4) $\Omega \in \Lambda_\alpha(S^{n-1})$ for some $\alpha > 0$

where Λ_α denotes the Hölder class. Then again T is of weak type (1,1) and bounded on L^p for all $p > 1$. All the operators discussed in this article are fairly easily seen to be bounded on L^2 or on L^∞, so that the weak (1,1) property is sharper than L^p boundedness. Furthermore they were all known already to be bounded on L^p for all $p \in (1, \infty)$ before weak (1,1) boundedness was proved. Thus the weak (1,1) property was sought as an endpoint result the sharper the existing theory, rather than in order to deduce the L^p boundedness as for (2.0) and (2.1).

The theory is not limited to convolution operators; for more general integral operators

$$Tf(x) = \int K(x,y)\, f(y)\, \mathrm{d}y$$

(where K is associated to T in the sense of [13]) the natural generalization is to retain (2.2) and to replace (2.3) and (2.4) by

$$|K(x,y)| \leq C|x-y|^{-n} \tag{2.5}$$

plus

$$|\nabla K(x,y)| \leq C|x-y|^{-n-1} \tag{2.6}$$

or a weaker version of (2.6) involving Hölder rather than Lipschitz continuity. (2.2) plus (2.5) and (2.6) imply weak type (1,1). These hypotheses are not optimal; for instance (2.4) may be replaced by an L^1 Dini condition. But the "classical" theory as formulated for instance in [16] always required some regularity for the kernel K. In contrast Calderón and Zygmund [2] showed that if $\Omega \in L^1$ is odd, then (2.1) and (2.2) alone imply L^p boundedness for all $p > 1$.

To see where regularity of K comes into play let us recall the method of Calderón and Zygmund. Using L^2 boundedness and a decomposition of an arbitrary L^1 function, they reduce matters to showing that if $\lambda > 0$, if $B = \sum_Q b_Q$ where the Q are distinct dyadic cubes, b_Q is supported on Q, $\|b_Q\|_1 \leq C\lambda|Q|$, $\int b_Q = 0$ and $\sum_Q |Q| \leq C\lambda^{-1}\|B\|_1$, then

$$\big|\{x : |TB(x)| > \lambda\}\big| \leq C\lambda^{-1}\|B\|_1\,. \tag{2.7}$$

They form the exceptional set

$$E = \bigcup_Q 2Q$$

where CQ denote the cube concentric with Q but C times as large. $|E| \leq C^n \sum |Q| \leq C\lambda^{-1}\|B\|_1$, so it is only necessary to consider $x \notin E$ in (2.7). By Chebychev's inequality it suffices to prove that

$$\|TB\|_{L^1(\mathbb{R}^n \setminus E)} \leq C\|B\|_1 \ , \tag{2.8}$$

which by the triangle inequality follows from

$$\|Tb_Q\|_{L^1(\mathbb{R}^n \setminus 2Q)} \leq C\|b_Q\|_1 \qquad \text{for all } Q \ . \tag{2.9}$$

The reduction to a single cube via the triangle inequality is a small but pivotal step; it is made possible by the introduction of the L^1 norm in (2.8). For $L^{1,\infty}$ there is the quasi-norm property

$$\|f+g\|_{1,\infty} \leq 2\|f\|_{1,\infty} + 2\|g\|_{1,\infty}$$

but no equivalent norm satisfies a true triangle inequality. Thus infinite series cannot be summed in $L^{1,\infty}$ in a straightforward way. This is one of the two essential difficulties in dealing with $L^{1,\infty}$.

The advantage of considering a single b_Q is that now the regularity hypothesis $\Omega \in \Lambda_\alpha$ coupled with the condition $\int b_Q = 0$ gives the pointwise bound

$$|Tb_Q(x)| \leq 2^{-nj}(2^{-j}|x - x_Q|)^{-n-\alpha}\|b_Q\|_1 \tag{2.10}$$

where $|Q| = 2^{nj}$, and similarly (with $\alpha = 1$) if $|\nabla K(x,y)| \leq C|x-y|^{-n-1}$. This implies (2.9). In the sequel we shall treat operators for which the best *pointwise* bound is (2.10) with $\alpha = 0$; Tb_Q will belong to $L^{1,\infty}$ but not to L^1, and the method breaks down for lack of a triangle inequality.

3. VARIANTS INVOLVING OSCILLATORY FACTORS

C. Fefferman [15] examined the L^p boundedness of operators $Tf = f * K$ on $\mathbb{R}^n$, with

$$K(x) = e^{i(|x|^a)}|x|^{-c}\chi_{|x|\leq 1} \tag{3.1}$$

where $a < 0$. The fundamental result was that when $c = n$, T is of weak type (1,1), for any $a < 0$; L^p results could then be obtained by an interpolation based on the analytic family of operators obtained by varying the parameters c and a. Let us restrict our attention to the case $c = n$. Then $|K(x)| \leq C|x|^{-n}$, and a calculation shows that $\hat{K} \in L^\infty$. However

$$|\nabla K(x)| \sim |x|^{-n-1+a} \qquad \text{as } x \to 0\,,$$

as (2.6) fails and the "classical" arguments don't apply. Nonetheless Fefferman established the weak (1,1) property, and from his proof may be extracted a rather general principle: In place of (2.8) it suffices to have

$$\|TB\|^2_{L^2(\mathbb{R}^n \setminus E)} \leq C\lambda \|B\|_1 \,. \tag{3.2}$$

For then (2.7) still follows by Chebychev's inequality. (The homogeneity appears to be inconsistent in (3.2), but actually it is correct because λ scales proportionally to B.) The analysis in [15] also involved considerations more closely tied to the nature of the particular kernels (3.1), relying in particular on the Fourier transform, Plancherel's theorem and an explicit bound

$$|\hat{K}(\xi)| \leq C(1 + |\xi|)^{-b}$$

where $b > 0$ is a known function of a whose precise value is needed in the argument.

Subsequently several authors, in particular Chanillo, Kuntz, Miyachi and Sampson, studied operators of the same form but with $0 < a \neq 1$, and with $\chi_{|x|\leq 1}$ replaced by $\chi_{|x|\geq 1}$ as it must be, relying on the basic method of [15]. Later Ricci and Stein [19] were led to consider operators $Tf(x) = \int f(y)K(x,y)\,dy$ with

$$K(x,y) = e^{iP(x,y)}L(x-y) \tag{3.3}$$

where $P : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is any polynomial and L is a classical Calderón-Zygmund kernel, homogeneous of degree $-n$, C^∞ on the unit sphere and having mean value zero there. They proved the L^p boundedness for all $p \in (1,\infty)$. But again the condition $|\nabla K(x,y)| \le C|x-y|^{-n-1}$ fails badly.

How might one obtain the L^p boundedness without first proving the weak (1,1) property and then interpolating? Consider the special case $n=1$, $L(x,y)=(x-y)^{-1}$ and $P(x,y)=(x-y)^2$, a convolution operator with kernel $e^{ix^2}x^{-1}$. Fix an auxillary function $\zeta \in C_0^\infty(\mathbb{R})$, supported in $\{\frac{1}{4} < |x| < 1\}$, satisfying

$$\sum_{-\infty}^{\infty} \zeta(2^{-j}x) = 1 \qquad \text{on } \mathbb{R}^n \backslash \{0\}\ .$$

Set

$$K_j(x) = K(x)\zeta(2^{-j}x) = e^{ix^2}x^{-1}\zeta(2^{-j}x)$$

and

$$T_j f = f * K_j\ .$$

Then

$$\begin{aligned} \|T_j f\|_1 &\le C\|f\|_1 \|K_j\|_1 \\ &\le C\|f\|_1\ . \end{aligned}$$

A calculation gives

$$\|\hat{K}_j\|_\infty \le C2^{-\epsilon j} \tag{3.4}$$

for some $\epsilon > 0$, for $j \ge 0$, so $\|T_j f\|_2 \le C2^{-\epsilon j}\|f\|_2$. Therefore interpolation gives for each $p \in (1,\infty)$

$$\|T_j f\|_p \le C2^{-\delta j}\|f\|_p$$

for all $f \in L^p$, with $\delta = \delta(p) > 0$. Summing the series yields L^p boundedness for $\sum_{j\ge 0} T_j$ (it turns out that the classical theory applies to $\sum_{j<0} T_j$). The key is the strong L^2 bound (3.4). More sophisticated versions of the argument apply to a variety of singular integral operators [4], [5], [14], [12]. Here the fact that L^p is between L^1 and L^2, for $p \in (1,2)$, is exploited; L^1 is not between any two (useful) spaces, and this is the second fundamental difficulty in dealing with weak type (1,1) estimates.

S. Chanillo raised the question of whether operators of the class (3.3) are of weak type (1,1), and it was proved in [3] to be so, using the basic idea of (3.2). However

(3.2) can no longer be verified by means of the Fourier transform, and one is led to work directly with the kernel K. The fact that the L^2 norm in (3.2) is taken over the somewhat irregular set $\mathbb{R}^n \backslash E$ is awkward, for it is reasonable to expect that orthogonality considerations will be useful in proving an L^2 estimate. On the other hand it is too much to hope that any B will be mapped to $L^2(\mathbb{R}^n)$, so the deletion of E must play a significant role. In [3] a truncation T' was constructed so that

$$T'B \equiv TB \quad \text{on} \quad \mathbb{R}^n \backslash E$$

and it was then shown that

$$\|T'B\|^2_{L^2(\mathbb{R}^n)} \leq C\lambda \|B\|_1 .$$

The truncation was of the form $T'B = \sum T'_Q b_Q$ where T'_Q depends on Q; see [3] for details.

Now (3.2) would follow from the pointwise bound

$$\|T'^* T'B\|_\infty \leq C\lambda .$$

Thus whereas the classical theory relied on pointwise bounds for Tb_Q, roughly speaking the new method requires pointwise bounds for T^*TB (disregarding the truncation). Actually the variant (3.8) below is more typical of the applications.

The distribution-kernel for $T'^* T'$ is

$$J(x,y) = \int \overline{K}'(z,x) K'(z,y) \, \mathrm{d}z$$

where K' is the kernel for T'. It turns out that except for certain degenerate P,

$$|J(x,y)| \leq C(1 + |x-y|)^{-n-\epsilon}$$

for some $\epsilon(P) > 0$, for "most" (x,y). Thus the kernel for $T'^* T'$ is significantly better behaved than the kernel for T itself. This happens in a variety of situations and is at the heart of applications of the method. The technical issue in [3] was to make precise the meaning of "most". This necessitated a quantitative analysis of the zero variety of an arbitrary real-valued polynomial on $\mathbb{R}^n$, based on some elementary algebraic geometry.

Historically the next advance was made in [6], but it was only recognized somewhat later [7] that a general technique with a variety of applications was at hand. The least tractable of the oscillatory kernels (3.1) had turned out to be the case $a = 1$,

$$K(x) = e^{i|x|}|x|^{-n}\chi_{|x|\geq 1} \,.$$

This is a case of intrinsic interest. For setting

$$S_t^\delta f(x) = \int_{|\xi|\leq t} e^{2\pi i x\cdot\xi}\hat{f}(\xi)(1-|\xi|)^\delta \, d\xi$$

with $\delta = (n-1)/2$ in $\mathbb{R}^n$, gives $S_1^\delta f = f * K$ where K is C^∞ and radial, and as $x \to \infty$ admits an asymptotic expansion

$$K(x) = \cos(2\pi|x| + \beta)(c_0|x|^{-n} + c_1|x|^{-n-1} + \ldots)$$

with $c_0 \neq 0$ and β a calculable constant. The oscillatory factors $\cos(2\pi|x| + \beta)$ and $e^{i|x|}$ turn out to have equivalent behavior for our purpose, so when $a = 1$ we have returned to a special case of the original question of convergence of eigenfunction expansions.

Let us set out the skeleton of the proof in some generality, then examine its specialization to the particular kernel $e^{i|x|}|x|^{-n}\chi_{|x|\geq 1}$. Suppose K is some distribution with $\hat{K} \in L^\infty$ and $K = \sum_{-\infty}^{\infty} K_j$ with K_j supported where $|x| \sim 2^j$, and make the rather minimal size hypothesis that K_j is a finite measure, possibly singular, with total mass bounded uniformly in j. To prove that $Tf = f * K$ is of weak type (1,1) it suffices to show that if B is as above and E is the union of the dilated cubes C_0Q for some large C_0 then (3.2) holds. Now all the cubes Q are dyadic, so $B = \sum_{-\infty}^{\infty} B_j$ where

$$B_j = \sum_{|Q|=2^{nj}} b_Q \,.$$

On $\mathbb{R}^n \backslash E$

$$\begin{aligned} TB &= \sum_j \sum_{i \geq j} B_j * K_i \\ &= \sum_{s=0}^{\infty} \Big(\sum_j B_j * K_{j+s} \Big) \\ &\equiv \sum_{s=0}^{\infty} T_s B \end{aligned} \tag{3.5}$$

since $b_Q * K_i$ is supported on $C_0 Q \subset E$ if $|Q| = 2^{nj}$, $i < j$, and C_0 is chosen sufficiently large. This last expression will replace the truncation $T'B$.

It now suffices to show that

$$\|T_s B\|^2_{L^2(\mathbb{R}^n)} \leq C 2^{-\epsilon s} \lambda \|B\|_1 \tag{3.6}$$

for some $\epsilon > 0$. It is useful to group the terms $B_j * K_i$ together according to the difference $s = j - i$, which from the geometric point of view has a natural interpretation as the (logarithm of the) ratio of the scales 2^j and 2^i associated to B_j and to K_i, respectively. In the classical case where K satisfies (2.6),

$$\|K_{j+s} * b_Q\|_1 \leq C 2^{-s} \|b_Q\|_1$$

when $|Q| = 2^{nj}$; the nontrivial factor of 2^{-s} is just this ratio of scales. This motivates hoping for the decaying factor of $2^{-\epsilon s}$ in (3.6).

Of course

$$\begin{aligned} \|T_s B\|_2^2 &= \sum_j \langle \tilde{K}_{j+s} * K_{j+s} * B_j, B_j \rangle \\ &\quad + 2 \sum_j \Big(\Big\langle \sum_{i<j} \tilde{K}_{j+s} * K_{i+s} * B_i, B_j \Big\rangle \Big) \end{aligned} \tag{3.7}$$

where $\tilde{K}_j(x) = \overline{K}_j(-x)$. The off-diagonal terms must be reckoned with, but to see the principal thrust of the analysis let us restrict attention to the sum of the diagonal terms, $i = j$. It suffices to show that

$$\|\tilde{K}_{j+s} * K_{j+s} * B_j\|_\infty \leq C 2^{-\epsilon s} \lambda \,. \tag{3.8}$$

The essential properties of B_j involve both cancellation and size restrictions: B_j has mean value zero on every dyadic cube of sidelength 2^j in $\mathbb{R}^n$, and $\int_Q |B_j| \leq C\lambda|Q|$ for every cube of sidelength greater than or equal to 2^j. The latter condition says heuristically that on scales $\geq 2^j$, B_j looks like a *bounded* function with L^∞ norm $C\lambda$. This is clearly relevant to (3.8), where we seek an L^∞ estimate. This property was not exploited in the classical theory but will be crucial for us.

So far the analysis has been purely formal, but the behavior of $\tilde{K}_j * K_j$ will depend on the nature of the particular kernels (or measures) at hand. If $|K(x)| \leq C|x|^{-n}$ then (3.8) holds trivially with $\epsilon = 0$, and the issue is the decaying factor $2^{-\epsilon s}$. There are two senses in which $\tilde{K}_{j+s} * K_{j+s}$ may be better behaved than K_{j+s} itself, namely in terms of smoothness or of size, and either one might potentially be exploited because of the two properties of B_j. In the instance $K = e^{i|x|}|x|^{-n}\chi_{|x|\geq 1}$, a calculation using the method of stationary phase gives

$$|\tilde{K}_i * K_i(x)| \leq C2^{-ni}(1+|x|)^{-(n-1)/2} ,$$

the latter factor capturing the improvement relative to $|K_i(x)| = C2^{-ni}$. (3.8) follows.

4. SMOOTHNESS CONDITIONS

Let F be a distribution supported in a fixed compact region in $\mathbb{R}^n$. There are many senses in which F may be said to possess some degree of smoothness. Several of these notions and the distinctions between them are quite relevant to the applications of the general method just outlined. Let us digress to contemplate some of them. Consider

(A) $F \in \Lambda_\alpha$ for some $\alpha > 0$.

(B) $\int(\sup_{|y-x|\leq r} |F(x) - F(y)|)\, dx \leq Cr^\epsilon$
for all $r \in [0,1]$, for some $\epsilon > 0$.

(C) $\sup_{|h|\leq r} \int |F(x+h) - F(x)|\, dx \leq Cr^\epsilon$
for all $r \in (0,1]$, for some $\epsilon > 0$.

(D) $r^{-n} \displaystyle\iint_{|x-y|\le r} |F(x)-F(y)|\,dx\,dy \le Cr^{\epsilon}$

for some $\epsilon > 0$.

(E) $F \in L^p_\epsilon$ for some $p \ge 1$ and $\epsilon > 0$.

($\mathrm{F_p}$) $\|Q_r(f * F)\|_p \le Cr^{\epsilon}\|f\|_p$

for all $f \in L^p$ and $r \in (0,1]$, for some $\epsilon > 0$, where

$$(Q_t f)\hat{}(\xi) = \hat{\psi}(t\xi)\hat{f}(\xi)\,,$$

and $\hat{\psi} \in C_0^\infty(\mathbb{R}^n)$ is an auxiliary function identically zero near the origin but nonvanishing everywhere on some annulus. $p \in [1,\infty)$ is a parameter; a different condition ($\mathrm{F_p}$) corresponds to each p.

(G) $|\hat{F}(\xi)| \le C(1+|\xi|)^{-\epsilon}$

for some $\epsilon > 0$.

(H) $\int_0^1 \omega_1(r)\frac{dr}{r} < \infty$

where

$$\omega_1(r) = \sup_{|h|\le r} \int |F(x+h)-F(x)|\,dx\,.$$

The following implications are valid:

$$\begin{array}{ccccccccccccc}
 & & & & (\mathrm{D}) & \Leftrightarrow & (\mathrm{E}) & & & & & & \\
 & & & & \Updownarrow & & \Updownarrow & & & & & & \\
(\mathrm{A}) & \Rightarrow & (\mathrm{B}) & \Rightarrow & (\mathrm{C}) & \Leftrightarrow & (\mathrm{F_1}) & \Rightarrow & (\mathrm{F_q}) & \Rightarrow & (\mathrm{F_2}) & \Rightarrow & (\mathrm{G}) \\
 & & & & & & \Downarrow & & & & & & \\
 & & & & & & (\mathrm{H}) & & & & & &
\end{array}$$

where $q \in (1,2)$, and the implication (G)$\Rightarrow$($\mathrm{F_q}$) is valid under the additional assumption that F is a finite measure. Of course it is implicitly assumed in (B), (C) and (D) that F is a function.

To each condition $(\cdot)$ corresponds $(\cdot)'$

$$\tilde{F} * F \quad \text{satisfies} \quad (\cdot)\,.$$

In each case $(\cdot) \Rightarrow (\cdot)'$, and (G)$\Leftrightarrow$(G)$'$.

We consider these conditions with $F = K_0 = \zeta K$, where ζ is as before and K is some distribution with $\hat{K} \in L^\infty$, which we would like to show defines by convolution an operator of weak type (1,1). ($K_j = \zeta(2^{-j}\cdot)K$ should be analyzed by passing to $K_j'(x) = 2^{nj}\zeta(x)K(2^j x)$.) The most relevant conditions are (H), (B)′ and (G). If K is homogeneous of degree $-n$ and if $F = \zeta K$ satisfies (H) then

$$\|b_Q * K\|_{L^1(\mathbb{R}^n \setminus 2Q)} \leq C\|b_Q\|_1$$

for all b_Q of the type described in section 2; (H) is the weakest condition with this property. (See [23], Remark 6.10, page 51.) Thus (H) is really the limit of the range of applicability of the classical analysis.

(B) was first made explicit in this context in [10]; it is actually completely equivalent to the inequality

$$\|F * B_{-s}\|_\infty \leq C2^{-\epsilon s}\lambda \quad \text{for all } s \geq 0, \text{ all } B_{-s} \text{ and some } \epsilon > 0\ ,$$

where B_{-s} satisfies the cancellation and size conditions spelled out at the close of section 3. Thus (in the most typical applications) our method depends on knowing that K_0 satisfies (B)′. However we should emphasize that details of the method vary from one application to another, and in fact these smoothness conditions do not really enter into the proof of Theorem B or into the treatment of $e^{i|x|}x^{-1}\chi_{|x|\geq 1}$, which rely on the size condition on B_{-s} but not its cancellation property.

In the next section we discuss a case in which K_0 satisfies the weak condition (G). Recall that when $n \geq 4$, the normal derivative of surface measure on the unit sphere S^{n-1} satisfies (G) with $\epsilon = 1/2$, and that on $\mathbb{R}^1$, for any $\epsilon < 1/2$ there exists a totally singular measure which satisfies (G) with that value of ϵ, so (G) is really quite weak. In particular it does not imply (B)′.

($\mathrm{F_p}$) is used in [8], and an L^2 variant of (D) arises in [9].

5. MAXIMAL FUNCTIONS

Consider a variant of the Hardy-Littlewood maximal function: In $\mathbb{R}^n$, $n \geq 2$, let S be a measurable set with finite measure, star-shaped about the origin. Let

$$\mathcal{M}_S f(x) = \sup_{r>0} \int_S |f(x - ry)\, dy\ .$$

Write $S = \{(r,\theta) : 0 \le r \le h(\theta)\}$ in polar coordinates and note that S has finite measure if and only if $h \in L^n(S^{n-1})$. A very closely related type of maximal function is

$$M_\Omega f(x) = \sup_{r>0} \int_{|y|\le 1} |f(x-ry)|\Omega(y/|y|)\, dy ,$$

$\Omega \in L^1(S^{n-1})$ nonnegative. It is quite easy to use the method of rotations [2] to show that $\mathcal{M}_S$ and M_Ω are bounded on $L^p(\mathbb{R}^n)$ for all $p > 1$, for all $h \in L^n$ and $\Omega \in L^1$, respectively, but it is an open question for which S and which Ω they are of weak type (1,1), and consequently when differentiation results such as

$$\lim_{r\to 0} |S|^{-1} \int_S f(x-ry)\, dy \to f(x) \quad a.e. \quad \forall f \in L^1$$

are valid. No example of S nor of Ω is known for which the maximal function is not weak type (1,1). The application of the method of rotations breaks down for $L^{1,\infty}$ because it relies on the triangle inequality, in the form of Minkowski's integral inequality.

R. Fefferman and later F. Soria [22] have shown that M_Ω maps L^1 to $L^{1,\infty}$ if Ω satisfies an entropy condition. More recently S. Hudson [18] obtained an interesting positive result in $\mathbb{R}^2$, assuming $\Omega \in L^1(S^1)$ to be monotone with respect to some choice of an origin on S^1, and making an additional hypothesis on the geometric structure of Ω. With only a size hypothesis on Ω, the best that is known [10] is that M_Ω is of weak type (1,1) if $\Omega \in L(\log L)$, that is,

$$\int_{S^{n-1}} \Omega \log^+ \Omega < \infty .$$

An easy corollary [10] is that $\mathcal{M}_S$ is weak (1,1) if

$$\int_{S^{n-1}} h^n \log^+ h < \infty .$$

In order to apply our now-familiar technique to M_Ω, fix an auxiliary function $\zeta \in C_0^\infty(\mathbb{R}^+)$, nonnegative and identically one on [1/2,1]. Then pointwise, for all $f \ge 0$,

$$M_\Omega f(x) \le CTf(x)$$

where

$$Tf(x) = \sup_j |f * K_j(x)|$$

and

$$K_j(x) = \zeta(2^{-j}|x|)\Omega(x/|x|) .$$

Let B be as before. Introduce the quadratic expressions

$$T_s B(x) = \Big(\sum_j |B_{j-s} * K_j(x)|^2\Big)^{1/2} , \tag{5.1}$$

and observe that

$$|TB| \leq \sum_{s=0}^{\infty} |T_s B|$$

pointwise off of E. Therefore it suffices to prove that

$$\|T_s B\|^2_{L^2(\mathbb{R}^n)} \leq C 2^{-\epsilon s} \lambda \|B\|_1 ,$$

which by homogeneity follows from the familiar inequality

$$\|\tilde{K}_0 * K_0 * B_{-s}\|_\infty \leq C 2^{-\epsilon s} \lambda . \tag{5.2}$$

A simple interpolation with crude estimates for the case $\Omega \in L^1$ reduces our task to proving (5.2) for $\Omega \in L^\infty$, with a bound proportional to $\|\Omega\|^2_\infty$; actually (5.2) won't hold unless $\Omega \in L^2$.

In $\mathbb{R}^2$ it is easy to see that $\tilde{K}_0 * K_0$ is Hölder continuous (except at 0). Define measures μ_θ, for $\theta \in S^{n-1}$, by

$$\int f \, d\mu_\theta = \int f(r\theta) r^{n-1} \zeta(r) \, dr .$$

Then

$$\tilde{K}_0 * K_0 = \iint (\tilde{\mu}_\theta * \mu_\omega) \Omega(\theta) \Omega(\omega) \, d\theta \, d\omega .$$

In $\mathbb{R}^2$, $\tilde{\mu}_\theta * \mu_\omega$ is clearly absolutely continuous and has a smooth density when $\theta \neq \pm\omega$. Making this quantitative and integrating with respect to $d\theta \, d\omega$ yields the

Hölder continuity except at 0, where the bounds blow up at a rate which may be estimated [7].

In higher dimensions $\tilde{\mu}_\theta * \mu_\omega$ will be supported on a two-dimensional plane, hence certainly will be singular. The situation is subtle: the *threefold* convolution $K_0 * K_0 * K_0$ is Hölder continuous in any dimension, indeed this holds for any $K_0 \in L^2$ satisfying (G); yet it is possible to construct an example with $\Omega \in L^\infty$ but $\tilde{K}_0 * K_0$ not Hölder continuous, even away from 0, in $\mathbb{R}^3$.

It was proved in [10] that $\tilde{K}_0 * K_0$ satisfies the smoothness condition (B), in all dimensions, for all $\Omega \in L^\infty$; in other words K_0 satisfies (B)'. The proof involved writing

$$(\tilde{K}_0 * K_0 * B_{-s})(0) = \langle L\Omega, \Omega \rangle$$

where $L : L^\infty(S^{n-1}) \to L^1(S^{n-1})$ is a linear operator which depends implicitly on B_{-s} and is rather singular; it involves integrations over curves on S^{n-1}. The L^2-based technique of [4] and [5] can then be applied to L.

Perhaps the most interesting question in this area is whether there is an L^1 theory for a class of more singular maximal functions involving integrations over lower-dimensional sets, of which a typical example is

$$Mf(x) = \sup_j |f * \mu_j(x)|$$

where

$$\int f \, \mathrm{d}\mu_j = \int_{S^{n-1}} f(2^{-j}x) \, \mathrm{d}\sigma(x)$$

and σ is surface measure on the unit sphere. An extension of the basic technique was introduced in [6] and used to show that this M maps the real-variable Hardy space $H^1(\mathbb{R}^n)$ to $L^{1,\infty}$, a result intermediate in strength between L^p boundedness for $p > 1$ and weak type (1,1).

6. PROOF OF THEOREM B (SKETCH)

The obstacle to applying the general method is that unless the singular locus $\mathcal{M} = \Phi(\{z_n = 0\})$ has nonvanishing scalar curvature at every point, one cannot hope to obtain any explicit expression for $K = \check{m}$. In fact its behavior changes dramatically from the case where M is curved to the case where it is a hyperplane. The real complications arise when M is flat to infinite order, but not linear, at some points. Therefore we are led to work less with K and more on the Fourier transform side, bringing Plancherel's theorem into play in a decisive way. However it is too much to hope to get away with Fourier transform side arguments alone, for the properties of B admit no direct re-interpretation in terms of $\hat{B}$.

Let $\eta \in C_0^\infty(\mathbb{R}^+)$ satisfy

$$\sum_{-\infty}^{\infty} \eta(2^{-j}t) \equiv 1 \quad \text{on} \quad \mathbb{R}^+$$

and $\tilde{\eta} \in C_0^\infty(\mathbb{R}^{n-1})$ satisfy

$$\sum_{\nu \in Z^{n-1}} \tilde{\eta}(z' + \nu) \equiv 1 \quad \text{on} \quad \mathbb{R}^{n-1} .$$

Set

$$b_j(z) = b(z)\eta(2^j|z_n|) \quad \text{and} \quad b_j^\nu(z) = b_j(z)\tilde{\eta}(2^{j/2}z' + \nu) ,$$

so that $b = \sum b_j = \sum\sum b_j^\nu$, where j ranges only over Z^+, modulo a C_0^∞ function which may be disregarded. Composing with Φ produces a corresponding decomposition of m. Set

$$(T_j f)\hat{} = \hat{f} m_j \quad \text{and} \quad (T_j^\nu f)\hat{} = \hat{f} m_j^\nu .$$

Let B be as before and decompose it as $\sum B_i$ as in section 3, except that $B_i = 0$ for all $i < 0$, and $B_0 = \sum b_Q$, summed over all Q with $|Q| \leq 1$. It suffices to prove that

$$\left\| \sum_{j>0} T_{j+s} B_j \right\|_{L^2(\mathbb{R}^n)}^2 \leq C 2^{-\epsilon s} \lambda \|B\|_1 \qquad \text{for } s \geq 0 , \tag{6.1}$$

$$\left\| \sum_{j>0} T_{j+s} B_j \right\|_{L^1(\mathbb{R}^n \setminus E)} \leq C 2^{s} \|B\|_1 \qquad \text{for } s < 0 , \tag{6.2}$$

and

$$\|T_s B_0\|_{L^2(\mathbb{R}^n)}^2 \leq C 2^{-\epsilon s} \lambda \|B_0\|_1 \qquad \text{for } s \geq 0 . \tag{6.3}$$

To prove (6.2) we need some crude estimates on the inverse Fourier transforms of the m_j^ν. Fix j and ν, let w be any fixed point in the support of b_j^ν, and rotate coordinates in the ξ-space so that at $\Phi(w)$, $(D\Phi)(\partial/\partial z_n)$ points in the direction of $\partial/\partial\xi_n$. Then a careful integration by parts establishes pointwise bounds for $(m_j^\nu)^\check{}$, in the coordinates (x', x_n) dual to these rotated ξ-coordinates.

LEMMA 1. $|(m_j^\nu)^\check{}(x)| \leq C_N 2^{-jn}(1 + 2^{-j/2}|x'| + 2^{-j}|x_n|)^{-N}$ *for all* N, j, ν.

Since the rotation depends on j and ν it is not possible to sum over j and ν to obtain pointwise bounds for $\check{m}$, without sacrificing essential information. But Lemma 1 does yield

LEMMA 2. $\|(m_j^\nu)^\check{}\|_{L^1\{|x|\geq 2^{s+j}\}} \leq C_N 2^{-Ns}2^{-j(n-1)/2}$ *for all* j, ν, N *and all* $s \geq 0$.

Summing over ν, which ranges over an index set of cardinality comparable to $2^{j(n-1)/2}$, gives a bound of $C_N 2^{-Ns}$ for the L^1 norm of $\check{m}_j$ on the same region. Now (6.2) follows at once.

As for (6.1), there exists $C < \infty$ such that no ξ is contained in the supports of more than C of the m_j^ν, so Plancherel's theorem yields a bound of

$$\sum_{j,\nu} \|T_{j+s}^\nu B_j\|_2^2$$

for the left-hand side of (6.1). This innocuous exploitation of orthogonality is a key step. Since there are about $2^{j(n-1)/2}$ values of ν for each j, it suffices to show that

$$\|T_j^{\nu *}T_j^\nu B_{j-s}\|_\infty \leq C 2^{-j(n-1)/2}\lambda 2^{-\epsilon s} \tag{6.4}$$

for all $j > s \geq 0$ and all ν.

Fix j and ν. The multiplier for $T_j^{\nu *}T_j^\nu$ is $|m_j^\nu|^2$, which has the same size and smoothness properties as $2^{-j(n-1)/2}m_j^\nu$. Thus in the coordinates of Lemma 1 the same integration by parts argument establishes the following.

LEMMA 3. *The convolution kernel for* $T_j^{\nu *}T_j^\nu$ *is majorized pointwise by*

$$C_N 2^{-j(n-1)/2}2^{-jn}(1 + 2^{-j/2}|x'| + 2^{-j}|x|)^{-N}$$

for all N.

It is straightforward to see that m_j^ν is supported in

$$\{\xi : |\xi' - w'| \leq C2^{-j/2} \quad \text{and} \quad |\xi_n - w_n| \leq C2^{-j}\}$$

for some C independent of j and ν, in the rotated coordinates adopted above. Moreover for all i, β

$$\left\| \frac{\partial^\beta}{\partial \xi'^\beta} \frac{\partial^i}{\partial \xi_n^i} m_j^\nu \right\|_\infty \leq C2^{-j(n-1)/2} 2^{j|\beta|/2} 2^{ji} \tag{6.5}$$

where C depends on i, β but not on j, ν. This follows from the chain rule. Lemmas 1 and 3 follow from (6.5) by a straightforward integration by parts.

Roughly speaking, Lemma 3 implies that $(m_j^\nu)^\vee$ is supported essentially on a rectangle of dimensions $C2^{j/2}$ in the x' directions and $C2^j$ in the x_n direction, and satisfies a certain favorable L^∞ bound. We have already remarked that on scales larger than 2^i, B_i behaves like an L^∞ function with norm at most $C\lambda$. When $i \leq j/2$ we may clearly combine these two facts in a straightforward way to obtain an upper bound on $\|T_j^{\nu*} T_j^\nu B_i\|_\infty$, and a slightly more careful analysis yields (6.4) for all $i \leq j$ with $\epsilon = (n-1)/4$. (6.3) follows in the same way. The proof of Theorem B is, in outline, complete. More details may be found in the proof of Theorem A in [11], though the notation there must be unraveled.

The decomposition of m as $\sum m_j^\nu$ is both natural and in a sense optimal for our purpose. The first decomposition as $\sum m_j$ has two motivations: first, in the case $m(\xi) = \text{distance}(\xi, \mathcal{M})^\delta$ which inspires the more general multipliers of the theorem, $\check{m}_j$ is roughly the same as $\check{m}\zeta(2^{-j}x)$, and second, it respects the homogeneity of distance$(\xi, \mathcal{M})$. Unfortunately no satisfactory bounds hold for $\check{m}_j$ and a further decomposition is suggested. We have chosen the coarsest decomposition for which pointwise bounds may be obtained for the m_j^ν without sacrificing essential information. Given that b_j^ν is to be supported on a rectangle with one side of length comparable to 2^{-j}, purely geometric considerations suggest that the other sides should be chosen to have lengths no longer than $C2^{-j/2}$. This choice gives the largest rectangles which, under an arbitrary diffeomorphism, are mapped to sets which in a reasonable sense look like rectangles, with comparable dimensions.

REFERENCES

[1] A.P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Math. 88 (1952), 85–139.

[2] A.P. Calderón and A. Zygmund, *On singular integrals*, Amer. J. Math. 78 (1956), 289–309.

[3] S. Chanillo and M. Christ, *Weak (1,1) bounds for oscillatory singular integrals*, Duke Math. J. 55 (1987), 141–155.

[4] M. Christ, *Hilbert transforms along curves, I: Nilpotent groups*, Annals of Math. 122 (1985), 575–596.

[5] M. Christ, *Hilbert transforms along curves, III: Rotational curvature*, preprint Dec. 1984, unpublished.

[6] M. Christ, *Two singular maximal functions act on Hardy spaces*, preprint, subsumed in [7].

[7] M. Christ, *Weak type (1,1) bounds for rough operators*, to appear, Annals of Math.

[8] M. Christ, *On the regularity of inverses of singular integral operators*, to appear, Duke Math. J.

[9] M. Christ and J.L. Journé, *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. 159 (1987), 51–80.

[10] M. Christ and J.L. Rubio de Francia, *Weak type (1,1) bounds for rough operators, II*, submitted.

[11] M. Christ and C.D. Sogge, *The weak type L^1 convergence of eigenfunction expansions for pseudodifferential operators*, submitted.

[12] M. Christ and E.M. Stein, *A remark on singular Calderón-Zygmund theory*, Proc. AMS 99 (1987), 71–75.

[13] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators*, Annals of Math. 120 (1984), 371–397.

[14] J. Duoandikoetxea and J.L. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates*, Invent. Math. 84 (1986), 541–561.

[15] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. 124 (1970), 9–36.

[16] L. Hörmander, *Estimates for translation invariant operators on L^p spaces*, Acta Math. 104 (1960), 93–139.

[17] L. Hörmander, *The spectral function of an elliptic operator*, Acta Math. 88 (1968), 341–370.

[18] S. Hudson, *A covering lemma for maximal operators with unbounded kernels*, to appear, Michigan Math. J.

[19] F. Ricci and E.M. Stein, *Harmonic analysis on nilpotent groups and singular integrals, I, II*, preprints.

[20] C.D. Sogge, *Concerning the L^p norm of spectral clusters for second order elliptic operators on compact manifolds*, to appear in J. Funct. Anal.

[21] C.D. Sogge, *On the convergence of Riesz means on compact manifolds*, Annals of Math. 126 (1987), 439–447.

[22] F. Soria, *Characterizations of classes of functions generated by blocks and associated Hardy spaces*, Indiana Univ. Math. J. 34 (1985), 463–492.

[23] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, N.J., 1971.

[24] E.M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, N.J., 1971.

M. Christ:
University of California,
Los Angeles,
California 90024.

C.D. Sogge:
University of Chicago,
Chicago,
Illinois 60637.