Hilbert integrals, singular integrals, and Radon transforms I

by

D. H. PHONG(1) and E. M. STEIN(2)

Columbia University
New York, NY, U.S.A.

Princeton University
Princeton, NJ, U.S.A

Table of contents

Introduction

The purpose of this series of papers is to introduce two new classes of operators, clarify their connection, and then show how they may be applied to the study of boundary-value problems. The first class we shall consider has as its geometric setting the Radon transform but it is combined here with basic features of singular integrals. While this class of operators, the singular Radon transforms, has an intrinsic interest in harmonic analysis and real-variable theory, it will also be important for us because of its applicability to the second class of operators: the Hilbert integral operators. The latter class arises in boundary-value problems, and is of particular interest in the non-coercive case, such as in the $\bar{\partial}$-Neumann problem for strongly pseudo-convex domains. This class will be the subject of a second paper in this series.

(1) Alfred P. Sloan Fellow, also partially supported by NSF Grant No. MCS-78-27119 and No. DMS-84-02710.

(2) Partially supported by NSF Grant No. MCS-80-03072.

Singular Radon transforms. Let Ω be a smooth manifold, and suppose that for each $P\in\Omega$ we are given a co-dimension one smooth sub-manifold Ω_P so that $P\in\Omega_P$, and a singular integral density $K(P,\cdot)$ concentrated on Ω_P with its singularity at P. Then (if the mappings $P\to\Omega_P$ and $P\to K(P,\cdot)$ are smooth) we define the singular Radon integral by

$$R(f)(P)=\langle K(P,\cdot),f|_{\Omega_P}\rangle, \quad \text{whenever } f\in C_0^\infty(\Omega). \tag{0.1}$$

Our first main task then is to prove the boundedness of the operator R on L^p, when $1<p<\infty$. There is also a closely related maximal function for which we might expect similar results. To define it fix a Riemannian metric ds^2 on Ω, with ds_P^2 the resulting induced measure on Ω_P and $d\sigma_P$ the corresponding volume element of Ω_P. Let $B(P,\delta)$ denote the geodesic ball in Ω_P centered at P of radius δ, and denote by $|B(P,\delta)|$ its σ_P measure. Then the maximal function is defined by

$$M(f)(P)=\sup_{0<\delta<1}\frac{1}{|B(P,\delta)|}\int_{B(P,\delta)}|f(Q)|\,d\sigma_P(Q).$$

It turns out that in order to prove the desired results for R and M some geometric properties related to the family $\{\Omega_P\}$ must be assumed, and to a formulation of such conditions we now turn.

Rotational curvature. There are several ways of stating the curvature condition we use (see Section 1). One is manifestly invariant, and reverts to a condition formulated by Guillemin and Sternberg [21] in their generalization of the invertibility of the Radon transform. Thus when the singular density $K(P,\cdot)$ is replaced by a C^∞ function these conditions imply that the transform R is a Fourier integral operator in the sense of Hörmander [26], whose Lagrangian manifold is the normal bundle of

$$\mathscr{C}=\{(P,Q)\,|\,Q\in\Omega_P\}\quad\text{in } \Omega\times\Omega.$$

Our analysis requires another formulation of the curvature condition: assuming $\dim\Omega=m+1$, one can cover Ω by coordinate systems (t,x), with $t\in\mathbf{R}$, $x\in\mathbf{R}^m$, so that if $P=(t,x)$ then

$$\Omega_P=\{(s,y)|s=t+S(t,x,y)\}\quad\text{with}\quad S(t,x,x)=0 \text{ and } \det\left(\frac{\partial^2S(t,x,y)}{\partial x_j\partial y_k}\right)_{x=y}\neq 0. \tag{0.2}$$

Some examples. We describe briefly several examples of the above structures $\{\Omega_P\}$, and their corresponding singular Radon transforms.

(i) Suppose $\Omega=\mathbf{R}^{m+1}$, Ω_0 is a hypersurface passing through the origin, and $\Omega_P=\Omega_0+P$ is the translate of Ω_0 by P. Then the condition is equivalent to the non-vanishing of all the principal curvatures of Ω_0 at the origin. In this translation-invariant setting our results for the singular Radon transform and corresponding maximal function are closely related to earlier work of Nagel, Riviere, Wainger and one of the authors (see e.g. [49], and [35]).

(ii) Even when the Ω_P are flat the non-zero curvature condition may hold when the Ω_P "rotate" in a suitable manner as P varies. An enlightening example of this occurs when Ω is the Heisenberg group $H^n=\{(z,t)\}$. Then we can take Ω_P to be the left *group* translate of the hyperplane $\Omega_P=\{(z,0)\}$, i.e. $\Omega_P=P\cdot\Omega_0$ (see Examples 2 and 3 in Section 1). In this case the singular Radon transforms were studied by Geller and Stein [16], [17].

(iii) For us the most fundamental example will be the following generalization of the previous one, where Ω is the boundary of a domain $\mathscr{D}$ in $\mathbf{C}^{n+1}$. Suppose r is a defining function for $\mathscr{D}$ i.e. $\mathscr{D}=\{z\in\mathbf{C}^{n+1}|r(z)<0\}$, and let $\psi(z,w)$ be an extension of r which is almost analytic in z, almost anti-analytic in w, and so that $\psi(z,z)=r(z)$. (ψ and its variants already appear in the formula for the Bergman kernel obtained by Fefferman [11], and Boutet de Monvel-Sjöstrand [4].) If we take $\Omega_P=\{z\in\Omega|\operatorname{Im}\psi(z,w)=0\}$, with $w=P$, then our non-vanishing curvature condition is equivalent with the non-degeneracy of the Levi form of $\mathscr{D}$. In this case the singular Radon transform plays a crucial role in the $\bar{\partial}$-Neumann problem.

Oscillatory integrals. We study the operator R (given by (0.1)), by expressing it as a pseudo-differential operator in one variable, once we have chosen coordinates as in (0.2). That is we write

$$R(f)(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda t}\,a(t,\lambda)\,\hat{f}(\lambda)\,d\lambda. \tag{0.3}$$

Here $\hat{f}(\lambda)=\hat{f}(\lambda,x)$ is a function which takes its values in $L^2(\mathbf{R}^m_x)$, and for each (t,λ) the symbol $a(t,\lambda)$ is the oscillatory operator given by

$$a(t,\lambda)(f)(x)=\int_{\mathbf{R}^m} e^{i\lambda S(t,x,y)}\,K(t,x;x-y)\,f(y)\,dy, \tag{0.4}$$

where $K(t,x;\cdot)$ are a smooth family of singular integral kernels. Thus the study of R, at least for L^2, is reduced to the properties of the oscillatory operators (0.4). For these one can prove $S_{1/2,\,1/2}$ estimates, i.e.

$$\left\| \left(\frac{\partial}{\partial t}\right)^k \left(\frac{\partial}{\partial \lambda}\right)^l a(t,\lambda) \right\|_{\mathrm{op}} \leq A(1+|\lambda|)^{k/2-l/2} \tag{0.5}$$

for appropriate k and l.

Our approach to the estimates for R draws on three sources:

(i) The idea of "twisted convolution" for the Heisenberg group which in that case makes the passage to the pseudo-differential representation (0.3) entirely natural. The notion of twisted convolution for the Heisenberg group goes back to Segal, and was later exploited by many authors (the papers of Grossman, Loupias and Stein [20], Howe [28], Mauceri, Picardello and Ricci [30], and Ricci [42] being the most germane here).

(ii) The suggestive results of Hörmander [27] for oscillatory integrals like (0.4), where C^∞ functions replace the singular kernel K.

(iii) The construction of appropriate analytic families of operators in order to get L^p estimates. (This idea was used systematically in [49].)

We should stress an important fact about the $S_{1/2,\,1/2}$ estimates (0.5): these cannot hold for $k+l>m$. We are therefore limited in the degree of smoothness of the symbol at our disposal in trying to apply the Calderón-Vaillancourt theorem or its variants. Fortunately one can use methods developed by Coifman and Meyer [7] for this purpose, and adapt them to our situation where the symbol is operator-valued (for this, see the appendix).

The ideas we have alluded to allow us to prove the L^p estimate for the singular Radon transform and maximal function, with the understanding that we always take $\dim \Omega \geq 3$. (The case $\dim \Omega = 2$ has been considered previously in [35] by methods which do not use the pseudo-differential realization (0.3).)

Model case. The prototype of the oscillatory integral (0.4) is the operator

$$f \to (Tf)(x) = \int_{\mathbf{R}^n} e^{i\langle Bx,y\rangle} K(x-y)\, f(y)\, dy, \tag{0.6}$$

where $\langle Bx, y\rangle$ is a real bilinear form, and $K(x)$ is a singular kernel. When B is anti-symmetric and non-degenerate we are dealing essentially with twisted convolutions, and operators like (0.6) were studied in [30], and [17]. Other special cases had been considered also by Sampson [43] and Sjölin [45].

We make a brief study of these model operators in Section 2. We do this partly to motivate the considerations of the more general form (0.4), but also because these model operators lead us to a suggestive generalization of some notions of Hardy space theory, such as BMO, and "sharp-functions".

Hilbert integrals, coercive case. As we have already stated one of our motivations for considering the singular Radon transforms is that they allow us to deal with the Hilbert integral operators which are the second class of operators we intend to study. Let us briefly describe this application which will be carried out in the succeeding paper of this series. The archetype of the Hilbert integral operator is the classical example

$$u \to \int_0^\infty \frac{u(y)\,dy}{x+y}, \quad x>0,\ u \in L^p(\mathbf{R}^+). \tag{0.7}$$

Another example arises in the usual Dirichlet problem for Laplace's equation. Thus in the upper half space $\mathbf{R}_+^{n+1}=\{(x,\varrho); x\in\mathbf{R}^n, \varrho\in\mathbf{R}^+\}$, it is a classical fact that the solution to the problem

$$\Delta f \equiv \left(\frac{\partial^2}{\partial\varrho^2}+\sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}\right) f(x,\varrho) = u(x,\varrho)$$

$$f(x,0)=0$$

can be written as

$$f = N(u)+H(u) \tag{0.8}$$

where $N(u)$ is the Newtonian potential

$$(Nu)(x,\varrho) = c_n \int_{\mathbf{R}_+^{n+1}} (|x-y|^2+(\varrho-\mu)^2)^{-(n-1)/2}\, u(y,\mu)\, dy\, d\mu$$

and

$$(Hu)(x,\varrho) = -c_n \int_{\mathbf{R}_+^{n+1}} (|x-y|^2+(\varrho+\mu)^2)^{-(n-1)/2}\, u(y,\mu)\, dy\, d\mu.$$

One has $\Delta N(u)=u$, so $H(u)$ is the compensating term that takes into account the Dirichlet boundary condition. The estimates for N are well-known (N is a standard singular integral operator of order -2), but H is essentially an example of a Hilbert integral operator of order -2. We remark that in the coercive case sharp estimates for the Hilbert integrals are easy and are quickly reducible to simple inequalities for absolutely convergent integrals, and in effect to the original example (0.7).

We now turn to non-coercive problems and in particular the $\bar{\partial}$-Neumann problem on a strongly pseudo-convex domain. For this question the approach in Kohn (see [13]), Greiner and Stein [19] yielded sharp estimates, and the work of Phong [36], Lieb and Range [20], Harvey and Polking [23] and Stanton [46] (see also the survey paper of Beals, Fefferman and Grossman [2]), have provided us with a pretty clear picture of the nature of the singularity of the kernel of the Neumann operator. The problem that arises is how to make sharp estimates for the general class of operators whose kernels display this kind of singularity. It will be natural to consider such operators as further variants of the Hilbert integrals described above, but with significant differences which make their study substantially more difficult. In particular the estimates for these operators are essentially dependent on delicate cancellation properties, and unlike the coercive case are not reducible to easy estimates or (0.7).

Instead, one can obtain the estimates for the Hilbert integrals by writing them as integrals of families of singular Radon transforms. The integral decomposition corresponds to a two-fold fibration. First, the domain $\mathscr{D}$ is fibered into a one-parameter family of copies of its boundary Ω, i.e. in terms of $S_\varrho=\{z\in\mathscr{D};\varrho(z)=\varrho\}$, with ϱ a defining function. Then a second integration corresponds to a one-parameter flow in each S_ϱ, determined by an appropriate vector-field conjugate to $\partial/\partial\varrho$.

Concluding remarks. A preliminary announcement of our results was made in [40]; earlier work having some bearing on the present paper is in [38], [39]. We shall now describe some further areas of research suggested in part by the above.

(a) In view of Corollary 1 of Theorem 1 it seems highly likely that the analogues of our results for singular Radon transforms and Hilbert integrals for domains $\mathscr{D}$ will still hold if the Levi form of $\mathscr{D}$ has merely one non-vanishing eigenvalue.

(b) It is possible to define a maximal singular Radon transform, via the truncations implicit in the definition (0.1), and it is expected that it should satisfy properties similar to R itself. For the case corresponding to the Heisenberg group this is carried out in Greenleaf [18].

(c) Our methods carry over when the codimension of Ω_P in Ω is small, but the case of general co-dimension raises interesting questions. Thus when codimension $\Omega_P>\frac{1}{2}\dim\Omega$, the analogue of the non-zero curvature condition formulated above can never hold. In that case the work [49] indicates that appropriate non-vanishing "higher curvature" or even real-analyticity would be a suitable substitute. Thus when $\dim\Omega=3$, and Ω_P is the translation by P of a curve Ω_0 passing through the origin, a sufficient condition is the non-vanishing of the curvature and torsion of the curve Ω_0, or that Ω_0 be real-analytic. Other results for higher codimension, in the setting of nilpotent groups, are in Muller [33], [34], and Christ [9].

(d) G. Uhlmann has called our attention to the possible connection of the present paper with joint work done with Melrose and Guillemin (see [22], [31] and [32]) on Fourier integral operators with singular symbols associated to pairs of Lagrangians. This point merits further exploration.

1. Singular Radon transforms

We now give a precise description of the setting for singular Radon transforms referred to in the introduction.

Let Ω be a C^∞ manifold without boundary. Denote by Δ the diagonal in $\Omega\times\Omega$, i.e., $\Delta=\{(P,Q)\in\Omega\times\Omega; P=Q\}$, and by π_1 and π_2 respectively the projections from $\Omega\times\Omega$ on the first and second factor. Then the submanifolds Ω_P are defined to be

$$\Omega_P=\pi_2(\mathscr{C}\cap\pi_1^{-1}(P))$$

where $\mathscr{C}$ is a given C^∞ hypersurface in $\Omega\times\Omega$ satisfying the conditions:

(1) $\mathscr{C}$ contains the diagonal Δ.

(2) The projections π_1 and π_2 are submersions near Δ.

(3) Let $N(\mathscr{C})\subset T^*(\Omega\times\Omega)$ be the normal bundle of $\mathscr{C}$ in $\Omega\times\Omega$, and denote by ϱ_1 and ϱ_2 the restrictions to $N(\mathscr{C})$ of the projections of $T^*(\Omega\times\Omega)$ on the first and second factor. Then the mappings

$$d\varrho_j\colon T_\lambda(N(\mathscr{C}))\to T_{\varrho_j(\lambda)}(T^*(\Omega)),\quad j=1,2$$

are isomorphisms at every point $\lambda\in N(\mathscr{C})\setminus 0$ lying above Δ.

Conditions such as (3) are due to Guillemin and Sternberg [21] who introduced them in their approach to Fourier integral operators, and related them to the Radon transform. Observe that they are symmetric with respect to P and Q.

Often only a neighborhood of the diagonal in $\mathscr{C}$ is relevant to our purposes. We shall thus restrict our attention to an open subset $\mathscr{C}'$ of $\mathscr{C}$ containing Δ and having compact closure in $\mathscr{C}$, for which condition (3) holds at all $\lambda\in N(\mathscr{C}')\setminus 0$, and assume that Ω_P and $\Omega'_P=\pi_2(\mathscr{C}'\cap\pi_1^{-1}(P))$ are C^∞ hypersurfaces in Ω.

For future reference (when discussing adjoints) we set

$$\Omega_Q^*=\pi_1(\mathscr{C}\cap\pi_2^{-1}(Q)),\quad \Omega_Q^{*\prime}=\pi_1(\mathscr{C}'\cap\pi_2^{-1}(Q)).$$

We shall consider two classes of densities. The densities $K(P,Q)$ in the first class are smooth in $\mathscr{C}\setminus\Delta$, with principal value type singularities for Q near P in Ω_P, and the main task will be to establish L^p boundedness of the corresponding operators. In the second class, whose study is closely related to and actually implies the results of the first class, the densities K will be C^∞ everywhere on $\mathscr{C}$; the L^p boundedness is then easy, and of interest is rather the exact dependence of the operator norms on a suitable family of semi-norms for K.

To define the first class, observe that on each Ω_P there is a well-defined class $K^0(\Omega_P)$ of generalized densities, namely the class of all linear functionals on $C_0^\infty(\Omega_P)$ of the form

$$C_0^\infty(\Omega_P)\ni\varphi\to(L\varphi)(P)$$

for some $L\in OPS^0_{1,0}(\Omega_P)$, the class of pseudo-differential operators of order 0 in Ω_P. A family of distributions $(K(P,\cdot))_{P\in\Omega}$, $K(P,\cdot)\in K^0(\Omega_P)$ will be said to be a smooth family in P if $\operatorname{supp}K(P,\cdot)\subset\subset\Omega'_P$ and

$$\langle K(P,\cdot),\varphi|_{\Omega_P}\rangle\in C^\infty(\Omega)\quad\text{for each }\varphi\in C_0^\infty(\Omega).$$

Definition 1. A singular density K is a smooth family in P of distributions $(K(P,\cdot))_{P\in\Omega}$ with $K(P,\cdot)\in K^0(\Omega_P)$.

The singular Radon transform R associated to K is then the operator given by

$$(Rf)(P)=\langle K(P,\cdot),f|_{\Omega_P}\rangle,\quad f\in C_0^\infty(\Omega). \tag{1.1}$$

The second class of densities is simply the space of C^∞ functions on $\mathscr{C}$, supported in $\mathscr{C}'$, with semi-norms defined as follows. An admissible coordinate system ι is a covering of a neighborhood of $\mathscr{C}'$ by open sets $(\mathscr{C}_j)$, with a C^∞ function ι_j on each $(\mathscr{C}_j)$ satisfying

(a) $\iota_j|_{\mathscr{C}_j\cap\Delta}=0$;

(b) $\iota_j(P,Q)\in\mathbf{R}^{\dim\Omega_P}$ is for each fixed P a coordinate system (y) for Ω'_P;

(c) $\pi_1(\mathscr{C}_j)$ is included for each j in a coordinate patch Ω_j, with coordinates (x).

We assume that $\mathscr{C}'$ is initially taken to be small enough for existence of at least one such admissible coordinate system.

Given $K\in C_0^\infty(\mathscr{C}')$, an admissible coordinate system ι, and integers $N,M\geq 0$, rewrite K as a function of (x,y) and set

$$\|K\|_{N,M}^{\iota,j} = \sup \sum_{\substack{|\alpha|\leq N \\ |\beta|\leq M}} |y|^{\dim \Omega_P + |\alpha|} |\partial_x^\beta \partial_y^\alpha K(x,y)| \tag{1.2}$$

where the sup is taken over $x \in \pi_1(\mathscr{C}_j)$, $y \in \iota_j(P, \Omega_P)$ (x=coordinates of P), and

$$\|K\|_M^{\iota,j} = \sup \sum_{|\beta|\leq M} \left| \int_{\varepsilon\leq|y|\leq 1} \partial_x^\beta K(x,y)\, dy \right| \tag{1.3}$$

with the sup taken over $0<\varepsilon\leq 1$, and $x \in \pi_1(\mathscr{C}_j)$.

It is readily seen that the sets of semi-norms corresponding to two different admissible coordinate systems are equivalent.

Definition 2. An admissible density is a function $K \in C_0^\infty(\mathscr{C}')$ with the above semi-norms. Fix a C^∞ density dv on Ω and a C^∞ density $d\sigma$ on $\mathscr{C}$. The densities dv and $d\sigma$ induce a density $d\sigma_P$ on each fiber of the submersion π_1: $\mathscr{C} \to \Omega$, and the singular Radon transform associated to K is defined to be

$$(Rf)(P) = \int_{\Omega_P} K(P,Q) f(Q)\, d\sigma_P(Q), \quad f \in C_0^\infty(\Omega). \tag{1.4}$$

The manifold $\mathscr{C}$ will be referred to as the Lagrangian support of the singular Radon transform.

We shall also study a maximal operator naturally related to the Radon transform. Fix a Riemannian metric ds^2 on Ω, and let ds_P^2 denote the induced metric on Ω_P, with $d\sigma_P$ the resulting measure on Ω_P. Denote by $B(P,\delta)$ the ball of radius δ with respect to ds_P^2 in Ω_P which is centered at P. Set

$$(Mf)(P) = \sup_{0<\delta<1} \frac{1}{|B(P,\delta)|} \int_{B(P,\delta)} |f(Q)|\, d\sigma_P(Q)$$

where $|B(P,\delta)|$ denotes the σ_P measure of $B(P,\delta)$.

We can now state the main theorems about singular Radon transforms and maximal operators.

Assume throughout that $\dim \Omega \geq 3$, $\mathscr{C} \subset \Omega \times \Omega$ is a C^∞ hypersurface satisfying conditions (1), (2), (3) listed at the beginning of this section, and let $\Omega_1, \Omega_2 \subseteq \Omega$ be two open subsets with compact closures. All L^p norms appearing below are taken with respect to a fixed positive C^∞ density on Ω.

THEOREM A. (a) *Let R be a singular Radon transform defined by a singular density K on $\mathscr{C}$. Then for any p, $1<p<\infty$ we have*

$$\|Rf\|_{L^p(\Omega_2)} \leq C_{p,\Omega_1,\Omega_2,K}\|f\|_{L^p(\Omega_1)}$$

for all $f \in C_0^\infty(\Omega_1)$.

(b) *If K is instead an admissible density with the semi-norms* (1.2) *and* (1.3), *then the same inequality holds for $1<p<\infty$ with the constant $C_{p,\Omega_1,\Omega_2,K}$ depending only on finitely many of the seminorms.*

THEOREM B. *If $1<p\leq\infty$ we have*

$$\|Mf\|_{L^p(\Omega_2)} \leq C_{p,\Omega_1,\Omega_2}\|f\|_{L^p(\Omega_1)}$$

for all $f \in C_0^\infty(\Omega_1)$.

We conclude this section with some observations.

Observation 1. It is of course possible to define Mf in terms of metrics on each manifold Ω_P varying smoothly with P, in analogy with the introduction of admissible coordinate systems used to define seminorms of admissible densities.

Observation 2. The norms for $K(P,Q)$ are equivalent to the norms for $K(Q,P)$ viewed as a density on $\mathscr{C}^*=\{(Q,P);P\in\Omega_Q^*\}$.

Observation 3. To establish L^p bounds for certain ranges of p, we shall have to consider formal adjoints of singular Radon transforms, i.e., operators R^* satisfying

$$\int_\Omega (Rf)(P)\,\overline{g(P)}\,dv(P) = \int_\Omega f(P)\,\overline{(R^*g)(P)}\,dv(P)$$

for all $u, v \in C_0^\infty(\Omega)$. Since estimates for the first class can be reduced to similar ones for the second class, it suffices to determine formal adjoints when R is given by (1.4). In this case, however, it is evident that

$$(R^*g)(Q) = \int_{\Omega_Q^*} \overline{K(Q,P)}g(P)\,d\sigma_Q^*(P)$$

where $d\sigma_Q^*$ is the density induced on Ω_Q^* by the densities dv, $d\sigma$, and the submersion π_2. Thus R^* is also a singular Radon transform, its Lagrangian support $\mathscr{C}^*$ also has non-vanishing curvature (i.e. condition (3) above) since this condition is symmetric with respect to P and Q, and its density has equivalent seminorms to the density of R (Observation 1). Thus whichever L^q estimates already established for R will hold for R^* as well.

Main examples. To treat the main examples in this paper, it is convenient to provide several different reformulations of condition (3) which may have some interest in their own right. Let $\mathscr{C}\subset\Omega\times\Omega$ be a C^∞ hypersurface satisfying conditions (1) and (2). The following are then equivalent.

(i) $d\sigma_1$ and $d\varrho_2$ are isomorphisms from $T_\lambda(N(\mathscr{C}))$ to $T_{\varrho_j(\lambda)}(T^*(\Omega))$ for each $\lambda\in N_\Delta(\mathscr{C})\setminus 0$.

(ii) Let $N(\mathscr{C})'=\{(P,\xi;Q,-\eta)\in T^*(\Omega\times\Omega);\ (P,\xi;Q,\eta)\in N(\mathscr{C})\}$. Then locally near every point above Δ, $N(\mathscr{C})'$ is the graph of a canonical transformation from $T^*(\Omega)$ to $T^*(\Omega)$.

(iii) Let $\mathscr{C}$ be defined near Δ by an equation $\Phi(P,Q)=0$, with $\Phi\in C^\infty(\Omega\times\Omega)$, $d_P\Phi(P,Q)\neq 0$, $d_Q\Phi(P,Q)\neq 0$ (this is no loss of generality since we may assume that $d_{P,Q}\Phi\neq 0$, and on the diagonal $d_P\Phi=-d_Q\Phi$), and let the Hessian $d^2_{QP}\Phi$ be the differential of the map

$$\Omega\ni Q\to(d_P\Phi)(P,Q)\in T^*_P(\Omega)$$

which is a linear mapping from $T_Q(\Omega)$ to $T^*_P(\Omega)$.

Define the rotational curvature form L_Φ to be the bilinear form

$$L_\Phi\colon T_P(\Omega_P)\times T_P(\Omega_P)\to\mathbf{R}$$
$$\langle L_\Phi v_1, v_2\rangle=\langle (d^2_{QP}\Phi)|_{P=Q}v_1, v_2\rangle.$$

Then L_Φ is nondegenerate. (Observe that $L_{\lambda\Phi}=\lambda L_\Phi$ for $\lambda\in C^\infty(\Omega\times\Omega)$, so that the nondegeneracy of L_Φ is directly seen to be independent of the choice of Φ.)

(iv) Let x denote the coordinates of P, and y the coordinates of Q in a coordinate patch in Ω, and let $\Phi(x,y)=0$ be a defining equation with $d_x\Phi(x,y)\neq 0$, $d_y\Phi(x,y)\neq 0$ near the diagonal. Then the Monge-Ampere determinant

$$J(\Phi)=\det\left[\begin{array}{c|c} 0 & \partial\Phi/\partial x_j \\ \hline \dfrac{\partial\Phi}{\partial y_k} & \dfrac{\partial^2\Phi}{\partial y_k\,\partial x_j}\end{array}\right] \tag{1.5}$$

does not vanish when $\Phi=0$.

When any of these conditions is satisfied, we shall say that $\mathscr{C}$ has rotational curvature. Note that the equivalence of (ii) and (iv) can already be found in [26], Section 4.1.

In fact (i) means that $N(\mathscr{C})'$ is locally the graph of a C^∞ invertible mapping. As the normal bundle of a submanifold, $N(\mathscr{C})$ is automatically lagrangian, and hence the mapping is a canonical transformation. The equivalence with (ii) follows.

Next to see the equivalence of (i) with (iv) write $N(\mathscr{C})$ and ϱ_1 as

$$N(\mathscr{C}) = \{(P, td_P\Phi; Q, td_Q\Phi);\ \Phi(P,Q)=0,\ t\in\mathbf{R}\}$$
$$\varrho_1(P, td_P\Phi; Q, td_Q\Phi) = (P, td_P\Phi).$$

That ϱ_1 be a diffeomorphism near $(P, td_P\Phi)$, $t\neq 0$, thus means that given $(P',\mu')\in T^*_{P'}(\Omega)$, (P',μ') near $(P, td_P\Phi)$ we can find $Q\in\Omega$, $s\in\mathbf{R}\setminus 0$ smoothly satisfying the system

$$\Phi(P',Q)=0$$
$$sd_P\Phi(P',Q)=\mu.$$

This in turn means that $\{0\}\times\mathbf{R}^n$ is in the range of the Jacobian at (t,P) of the mapping $(s,Q)\to(\Phi(P,Q), sd_P\Phi(P,Q))$. As the Jacobian is given by

$$J=\left[\begin{array}{c|c} 0 & d'_P\Phi \\ \hline d_Q\Phi & td^2_{QP}\Phi \end{array}\right] = t^{n-2}\left[\begin{array}{c|c} 0 & d'_P\Phi \\ \hline d_Q\Phi & d^2_{QP}\Phi \end{array}\right]$$

and $d_P\Phi\neq 0$, the projection of the first component is surjective, and thus the previous statement is equivalent to the Jacobian having maximal rank, that is, the non-vanishing of (1.5).

By symmetry, the equivalence with ϱ_2 being a local diffeomorphism also follows.

Finally choose local coordinates $(x',t)\in\mathbf{R}^{n-1}\times\mathbf{R}$ near P so that $d_P\Phi(P,P)=(0,1)$, and $T_P(\Omega_P)=\{(x',0); x'\in\mathbf{R}^{n-1}\}$. The matrix of L_Φ is then $(\partial^2\Phi/\partial x'_j\,\partial y'_k)_{i\leq j,k\leq n-1}$, while

$$J(\Phi)=\det\begin{vmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & & & & \\ \vdots & & \dfrac{\partial^2\Phi}{\partial x_j'\,\partial y_k'} & & * \\ 0 & & & & \\ 1 & & * & & * \end{vmatrix}$$

$$=\det\left(\frac{\partial^2\Phi}{\partial x_j'\,\partial y_k'}\right)$$

which shows that (iii)$\Leftrightarrow$(iv).

A consequence of nonvanishing rotational curvature is the following property which will play a key role in the sequel.

COROLLARY. *Assume that $\mathscr{C}\subset\Omega\times\Omega$ satisfies any of the conditions* (i)–(iv) *near the diagonal Δ. Then Ω admits a covering by coordinate patches, on each of which there is a coordinate system*

$$P\to(t,x)\in\mathbf{R}\times\mathbf{R}^{\dim\Omega-1}$$

so that the hypersurface Ω_P can be parametrized by

$$\mathbf{R}^{\dim\Omega-1}\ni z\to(t+S(t,x;z);z) \tag{1.6}$$

with $S(t,x;z)$ a C^∞ function near $(t,x;x)$ satisfying

$$S(t,x;x)=0 \tag{1.7}$$

$$\det\left(\frac{\partial^2 S}{\partial x_j\partial z_k}(t;x;z)\right)\neq 0 \tag{1.8}$$

Proof of the corollary. Since the surfaces Ω_P vary smoothly with P, for each fixed P we can choose a curve γ in Ω which passes through P and is transversal to $\Omega_{P'}$ for any P' on γ near P. Parametrize γ by $t\to\gamma(t)$ with $\gamma(0)=P$, and choose for each $\Omega_{\gamma(t)}$ a coordinate system $\Omega_{\gamma(t)}\ni P'\to x\in\mathbf{R}^{\dim\Omega-1}$ centered at $\gamma(t)$ and varying smoothly with t. We thus obtain a coordinate system for a neighborhood V of P in Ω by letting

$$V\ni P'\to(t,x)$$

if $P'\in\Omega_{\gamma(t)}$, and x are the coordinates of P' in $\Omega_{\gamma(t)}$. If V is small enough, this is well defined, and $\Omega_{P'}$ for P' in V can be parametrized by (1.6) for some C^∞ function $S(t,x;z)$. Property (1.7) just says that $\Omega_{P'}$ passes through P'. As for (1.8), observe that

$$S(t,0;z)=0 \quad \text{for any } (t,z) \tag{1.9}$$

and that

$$J(\Phi)=\det\begin{vmatrix} 0 & -d'_x S & -1-d'_t S \\ d_z S & d^2_{xz} S & -d^2_{tz} S \\ 1 & 0 & 0 \end{vmatrix} \tag{1.10}$$

if we choose the function Φ of (1.5) to be

$$\Phi(t,x;t',z)=t'-t-S(t,x;z). \tag{1.11}$$

For $x=0$, $J(\Phi)$ reduces to $(-1)^{\text{power}}\det(d^2_{xz}S)$ in view of (1.9), and thus

$$\det(d^2_{xz}S)\,(t,0;z)\neq 0.$$

Shrinking V further if necessary, we obtain the desired statement by continuity. Q.E.D.

Some of the examples discussed below can be more readily understood if we modify the parametrization of Ω_P in (1.6) to be

$$y\to(t+B(t,x;y);x-y) \tag{1.12}$$

for y near 0 in $\mathbf{R}^{\dim\Omega-1}$, and the condition (1.8) becomes

$$\det(d^2_{yy}B+d^2_{xy}B)\neq 0. \tag{1.13}$$

In some sense the term $d^2_{yy}B$ represents the curvature of each hypersurface Ω_P, while $d^2_{xy}B$ measures the rate of change with respect to P of the normal to Ω_P.

Example 1. In $\mathbf{R}^n$ let H be a hypersurface passing through 0, and let $\varphi(P)=0$ be a defining function for H, with $|d\varphi(P)|=1$ on H. Define Ω_P as the translate to P of H. Then the function Φ of condition (iii) may be taken to be

$$\Phi(P,Q)=\varphi(Q-P),$$

the mapping $\Omega_P\ni Q\to d_P(\Phi(P,Q))=-d\varphi(Q-P)\in S^{n-1}$ is the Gauss map of Ω_P viewed as an embedded hypersurface in $\mathbf{R}^n$, and

$$L_\Phi=d^2_{QP}\Phi|_{T_P(\Omega_P)\times T_P(\Omega_P)}$$

is just the second fundamental form of H at 0. Thus the condition of nondegeneracy of L_Φ is equivalent to the nonvanishing of the gaussian curvature of H at 0.

Example 2. In $\mathbf{R}^n$ let Ω_P be for each P a hyperplane, and let $\nu(P)$ be the unit normal to Ω_P. Denote by $d\nu(P)$ the differential at P of the map $P\to\nu(P)$. That the distribution of hyperplanes Ω_P have rotational curvature is now equivalent to the nonvanishing of the $(n-1)$ symmetric function of the eigenvalues of $d\nu(P)$.

(Note that the fact that $|\nu(P)|=1$ and simple rank considerations imply that the nth symmetric function, i.e., the determinant of $d\nu(P)$, is always 0.)

To establish this we use (iv) with $\Phi(P,Q)=\langle \nu(P), Q-P\rangle$. The Monge-Ampere determinant $J(\Phi)$ at (P,P) is then given by

$$J(\Phi) = -\det \begin{vmatrix} 0 & \nu(P) \\ \nu(P) & d\nu(P) \end{vmatrix}. \tag{1.14}$$

Observe that the first row of the above matrix is orthogonal to all the other rows since $|\nu(P)|^2=1$ for all P. This fact together with the value of the determinant are invariant under conjugation by matrices of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix}$$

where U is any matrix in $O(n-1)$. Choosing U so that $\nu U=(0,\ldots,0,1)$ we get

$$\begin{aligned} J(\Phi) &= -\det \begin{vmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & & & & 0 \\ \vdots & & \Lambda & & \vdots \\ 0 & & & & 0 \\ 1 & & * & & 0 \end{vmatrix} \\ &= \det\Lambda \end{aligned} \tag{1.15}$$

with Λ an $(n-1)\times(n-1)$ matrix determined by

$$U^t\, d\nu(P)\, U = \begin{bmatrix} \Lambda & 0 \\ * & 0 \end{bmatrix}.$$

On the other hand we have

$$\begin{aligned} &(n-1) \text{ symmetric function of eigenvalues of } d\nu(P) \\ &\quad = (n-1) \text{ symmetric function of eigenvalues of } U^t\, d\nu(P)\, U \\ &\quad = (-1)^{n-1}\, [\text{coefficient of } \lambda \text{ in } \det(\lambda I_n - U^t\, d\nu(P)\, U)] \end{aligned} \tag{1.16}$$

$$= \text{Coefficient of } \lambda \text{ in } [\lambda \det(\lambda I_{n-1} - \Lambda)]$$
$$= (-1)^{n-1}(\det \Lambda).$$

The desired assertion follows from (1.15) and (1.16).

Example 3. Let $\mathbf{H}^n = \{(z,t) \in \mathbf{C}^n \times \mathbf{R}; (z,t)\cdot(z',t') = (z+z', t+t'+2\operatorname{Im} z\cdot \bar{z}')\}$ be the Heisenberg group. Then there is a natural invariant distribution of hyperplanes

$$\Omega_{(z,t)} = \{(z',t') \in \mathbf{C}^n \times \mathbf{R};\ t'-t-2\operatorname{Im} z\cdot \bar{z}' = 0\}. \tag{1.17}$$

Identify $T_{(z,t)}(\Omega_{(z,t)})$ with $\mathbf{C}^n$, and set

$$\Phi(z,t;z',t') = t'-t-2\operatorname{Im} z\cdot \bar{z}'.$$

It is then readily seen that the rotational curvature L_Φ reduces to the standard symplectic form on $\mathbf{R}^{2n}$:

$$\langle L_\Phi v_1, v_2\rangle = \operatorname{Im}(v_1\cdot \bar{v}_2) = \sigma(v_1, v_2), \quad v_1, v_2 \in \mathbf{C}^n. \tag{1.18}$$

Example 4. This example is basically a generalization of the previous one. In $\mathbf{C}^{n+1}$, let Ω be a hypersurface defined by $r(z)=0$, with $r \in C^\infty(\mathbf{C}^{n+1})$ and $dr \neq 0$ when $r=0$. Let $\psi(z,w)$ be an almost-analytic extension of r, i.e., $\psi(z,w)$ is a function having the following Taylor expansion along the diagonal

$$\psi(\xi+z, \eta+z) \sim \sum_{\alpha,\beta} \frac{1}{\alpha!\beta!} \frac{\partial^{\alpha+\beta} r}{\partial z^\alpha \partial \bar{z}^\beta}(z)\, \xi^\alpha \bar{\eta}^\beta \tag{1.19}$$

(see [4]). At each point $z \in \Omega$, set

$$\Omega_z = \{w \in \Omega; \operatorname{Im} \psi(z,w) = 0\}. \tag{1.20}$$

To relate the rotational curvature form $L_{\operatorname{Im}\psi}$ to the complex structure of Ω, rewrite vectors

$$v = \sum_{j=1}^{n+1} \alpha_j \frac{\partial}{\partial x_j} + \beta_j \frac{\partial}{\partial y_j} \quad \text{as} \quad v = \sum_{j=1}^{n+1} v_j \frac{\partial}{\partial z_j} + \bar{v}_j \frac{\partial}{\partial \bar{z}_j}$$

with $v_j = \alpha_j + i\beta_j$, and observe that the real tangent space $T_z^{\mathbf{R}}(\Omega)$ consists of vectors v satisfying

$$\operatorname{Re}\left(\sum_{j=1}^{n+1} v_j \frac{\partial r}{\partial z_j}\right) = 0. \tag{1.21}$$

In particular $T_z^{\mathbf{R}}(\Omega)$ contains the subspace $T_z^{1,0}(\Omega)$ defined by

$$T_z^{1,0}(\Omega)=\left\{v=\sum_{j=1}^{n+1} v_j\frac{\partial}{\partial z_j}+\bar{v}_j\frac{\partial}{\partial \bar{z}_j};\ \sum_{j=1}^{n+1} v_j\frac{\partial r}{\partial z_j}=0\right\},$$

which evidently is a complex vector space of dimension n. We shall often identify

$$v=\sum_{j=1}^{n+1} v_j\frac{\partial}{\partial z_j}+\bar{v}_j\frac{\partial}{\partial \bar{z}_j}\quad\text{with}\quad v=\sum_{j=1}^{n+1} v_j\frac{\partial}{\partial z_j}$$

and view $T_z^{1,0}(\Omega)$ as the space of "tangential holomorphic vectors" given by

$$T_z^{1,0}(\Omega)=\left\{v=\sum_{j=1}^{n+1} v_j\frac{\partial}{\partial z_j};\ \sum_{j=1}^{n+1} v_j\frac{\partial r}{\partial z_j}=0\right\}. \tag{1.22}$$

The Levi form $\mathscr{L}$ is the sesquilinear form on $T_z^{1,0}(\Omega)\times T_z^{1,0}(\Omega)$ given by

$$\mathscr{L}(v_1,v_2)=\sum_{j,k=1}^{n+1}\frac{\partial^2 r(z)}{\partial z_j\partial\bar{z}_k}v_{1,j}\bar{v}_{2,k}\quad\text{where}\quad v_i=\sum_{j=1}^{n+1} v_{i,j}\frac{\partial}{\partial z_j}.$$

Then

(a) the tangent space of Ω_z at z is just $T_z^{1,0}(\Omega)$;

(b) the rotational curvature form $L_{\operatorname{Im}\psi}$ coincides with the imaginary part of the Levi form;

(c) the nondegeneracy of $L_{\operatorname{Im}\psi}$ is equivalent to the nondegeneracy of the Levi form, and thus is satisfied when Ω is strongly pseudo-convex.

To verify the first assertion, observe that up to terms of second order in v

$$\begin{aligned}\operatorname{Im}\psi(z,z+v)&\sim\operatorname{Im}\left(\sum_{j=1}^{n+1}\frac{\partial r}{\partial\bar{z}_j}(z)\,\bar{v}_j\right)\\&=-\operatorname{Im}\left(\sum_{j=1}^{n+1}\frac{\partial r}{\partial z_j}(z)\,v_j\right).\end{aligned}$$

In particular $\langle d(\operatorname{Im}\psi)|_{(z,z)},v\rangle=-\operatorname{Im}(\sum_{j=1}^{n+1}\partial r/\partial z_j(z)\,v_j)$ and thus vectors v tangent to Ω_z are characterized by the two conditions

$$\operatorname{Im}\left(\sum_{j=1}^{n+1}\frac{\partial r}{\partial z_j}(z)\,v_j\right)=0,\quad \operatorname{Re}\left(\sum_{j=1}^{n+1}\frac{\partial r}{\partial z_j}(z)\,v_j\right)=0$$

which are exactly the ones defining $T_z^{1,0}(\Omega)$. The second assertion is an immediate consequence of the following two facts:

$\langle L_{\mathrm{Im}\,\psi} v, v' \rangle =$ Mixed terms of second order in $(\mathrm{Im}\,\psi)(z+v, z+v')$.

Mixed terms of second order in $\psi(z+v, z+v') = \sum_{j,k=1}^{n+1} \partial^2 r/\partial z_j \partial \bar{z}_k(z)\, v_j \bar{v}'_k = \mathscr{L}(v, v')$.

Finally assume that $L_{\mathrm{Im}\,\psi}(v, v')=0$ for all $v' \in T_z^{1,0}(\Omega)$. Then $\mathrm{Im}\,\mathscr{L}(v, v')=0$ and $\mathrm{Re}\,\mathscr{L}(v, v')=-\mathrm{Im}\,\mathscr{L}(v, iv')=-L_{\mathrm{Im}\,\psi}(v, iv')=0$ for all v', which would imply that $v=0$ if $\mathscr{L}$ were nondegenerate. The converse being obvious, the third assertion is proved.

2. The model case: motivation

The main ingredients in our approach are bounds for a class of "oscillatory singular integrals" whose study may be interesting in its own right. To see how they arise we consider the case where Ω is $\mathbf{R}^{n+1}=\{(x, t) \in \mathbf{R}^n \times \mathbf{R}\}$, and the hypersurface Ω_P through $P=(x, t)$ is given as the image of the linear mapping

$$\mathbf{R}^n \ni y \to (x+y, t+\langle Bx, y \rangle) \in \mathbf{R}^{n+1}$$

where $B=(b_{jk})$ is a fixed bilinear form. Let $K(y)$ be a kernel on $\mathbf{R}^n$, and define a singular density K_0 on $\mathscr{C}=\{(P, Q); Q \in \Omega_P\}$ by pushing forth to each manifold Ω_P through the above mapping the density $K(y)\, dy$.

The singular Radon transform R associated to $\mathscr{C}$ and K_0 is then given by the formula

$$(Ru)(x, t) = \int_{\mathbf{R}^n} u(x+y, t+\langle Bx, y \rangle)\, K(y)\, dy.$$

If we denote by $\hat{u}(x, \lambda)$ the Fourier transform of $u(x, t)$ with respect to t, R can be rewritten as

$$\begin{aligned}
(Ru)(x, t) &= \frac{1}{2\pi} \int e^{i\lambda t} \left[\int e^{i\lambda \langle Bx, y \rangle} K(y)\, \hat{u}(x+y, \lambda)\, dy \right] d\lambda \\
&= \frac{1}{2\pi} \left\{ \int e^{i\lambda t} \left[\int e^{i\lambda \langle Bx, y \rangle} K(y-x)\, \hat{u}(y, \lambda)\, dy \right] e^{-i\lambda \langle Bx, x \rangle}\, d\lambda \right\} \\
&= \int e^{i\lambda t} e^{-i\lambda \langle Bx, x \rangle} (T_\lambda \hat{u}(\cdot, \lambda))(x)\, d\lambda
\end{aligned}$$

with the operators $T_\lambda\colon C_0^\infty(\mathbf{R}^n)\to C^\infty(\mathbf{R}^n)$ defined by

$$(T_\lambda \varphi)(x) = \int e^{i\lambda\langle Bx,y\rangle} K(y-x)\,\varphi(y)\,dy, \quad \varphi\in C_0^\infty(\mathbf{R}^n).$$

We may now write using Plancherel's formula

$$\begin{aligned}\iint |Ru(x,t)|^2\,dx\,dt &= \frac{1}{4\pi^2}\iint\left|\int e^{i\lambda t}(T_\lambda \hat{u}(\cdot,\lambda))(x)\,d\lambda\right|^2 dt\,dx\\ &= \int_{\mathbf{R}^n}\left(\int_{-\infty}^{\infty}|(T_\lambda\hat{u}(\cdot,\lambda))(x)|^2\,d\lambda\right)dx\\ &= \int_{-\infty}^{\infty}\|T_\lambda\hat{u}(\cdot,\lambda)\|^2_{L^2(\mathbf{R}^n)}\,d\lambda\\ &\le \int_{-\infty}^{\infty}\|T_\lambda\|^2_{\mathrm{op}}\,\|\hat{u}(\cdot,\lambda)\|^2_{L^2(\mathbf{R}^n)}\,d\lambda.\end{aligned}$$

Here $\|T_\lambda\|_{\mathrm{op}}$ denotes the norm of T_λ as an operator on $L^2(\mathbf{R}^n)$. Thus L^2 bounds for R reduce to bounds for T_λ. In particular if $\|T_\lambda\|_{\mathrm{op}}$ is finite and bounded independently of λ by a constant A we may conclude that

$$\begin{aligned}\|Ru\|^2_{L^2(\mathbf{R}^{n+1})} &\le A^2\int_{-\infty}^{\infty}\|\hat{u}(\cdot,\lambda)\|^2_{L^2(\mathbf{R}^n)}\,d\lambda\\ &= A^2\|u\|^2_{L^2(\mathbf{R}^{n+1})}.\end{aligned}$$

In general, to insure the boundedness of the T_λ's we need an appropriate combination of conditions on the bilinear form B and the kernel K. For example K may be homogeneous of degree $-n$ and B may be non-degenerate (when it is also antisymmetric we get back the Heisenberg group of Example 3; however, in many cases of interest B will be neither symmetric nor antisymmetric). For our purposes, it is necessary to go further in two directions: first, replace K by more general homogeneous kernels (to fulfill conditions for later interpolation on L^p spaces; the ideas involved here will be taken up in the model case); second, replace the bilinear form $\langle Bx,y\rangle$ by more general functions $S(t,x,y)$ depending on a parameter t.

In the special case where K is homogeneous of degree $-n$ and we have a bilinear form as phase, the operators can be treated (see Corollary 2 of Theorem 1 and Corollary 2 of Theorem 2) as a consequence of Sjölin's n-dimensional version of Carleson's theorem on pointwise convergence of Fourier series (see [5], [44]). Since this case does not suffice for our needs, we require an independent approach.

The model case: L^p estimates

We shall now study the following operators

$$(Tf)(x) = \text{P.V.} \int e^{i\langle Bx,y\rangle} K(x-y)\, f(y)\, dy, \quad f \in C_0^\infty(\mathbf{R}^n),$$

where P.V. stands for principal value, $K(x)$ is C^∞ outside the origin, coincides with a homogeneous function of degree $-\mu$ for large $|x|$, with a homogeneous function of degree $-n$ for small $|x|$, and satisfies the cancellation property $\int_{|x|=\varepsilon} K(x)\, d\sigma(x)=0$, for ε small. Finally $\langle Bx, y\rangle$ is a bilinear form, given by the $n\times n$ matrix B.

THEOREM 1. *Assume B is nondegenerate. Then T can be extended as a bounded operator from $L^2(\mathbf{R}^n)$ to itself, if $0\leq\mu$.*

Proof. Step 1. We begin by proving the boundedness of T under the assumption that K vanishes for $|x|\geq 1$. In this case we show first that

$$\int_{B_1} |Tf(x)|^2\, dx \leq c \int_{B_2} |f(x)|^2\, dx, \tag{2.1}$$

where B_1 and B_2 are respectively the balls of radius 1 and 2 centered at the origin. To see this note that since the support of $K(x-y)$ is in the set where $|x-y|\leq 1$, in estimating $Tf(x)$ for $x\in B_1$ we may as well assume that f vanishes outside B_2. Now let T' be the operator defined by $T'(f)(x)=\int K(x-y)\, e^{i\langle By,y\rangle} f(y)\, dy$. For it we have an estimate like (2.1), in view of the standard theory of singular integrals. However $(T-T')(f)(x)=\int K(x-y)\,[e^{i\langle Bx,y\rangle}-e^{i\langle By,y\rangle}]\, f(y)\, dy$, and in absolute value this difference is bounded by

$$c\int_{|x-y|\leq 1} \frac{|e^{i\langle Bx,y\rangle}-e^{i\langle By,y\rangle}|}{|x-y|^n} |f(y)|\, dy \leq c \int_{|x-y|\leq 1} \frac{|f(y)|\, dy}{|x-y|^{n-1}}, \quad \text{if } x\in B_1.$$

From this (2.1) follows.

We next remark that while operators like T do not commute with translations, they do satisfy the identity

$$(\tau_{-h} T \tau_h)(f)(x) = e^{i\langle Bh,h\rangle}\, e^{i\langle Bx,h\rangle}\, T(e^{i\langle Bh,\cdot\rangle} f(\cdot))(x), \tag{2.2}$$

with $\tau_h(f)(x)=f(x-h)$, as a simple change of variables shows. With this we get as an immediate extension of (2.1) the inequality

$$\int \chi_1(x-h)\,|Tf(x)|^2\,dx \leq c \int \chi_2(x-h)\,|f(x)|^2\,dx$$

where χ_1 and χ_2 are respectively the characteristic functions of B_1 and B_2. If we integrate both sides of the inequality with respect to h we obtain

$$\int |Tf(x)|^2\,dx \leq 2^n \cdot c \int |f(x)|^2\,dx$$

which establishes the boundedness of T (under the assumption that $K(x)$ is supported in $|x|\leq 1$).

Step 2. We now turn to the case when $K(x)$ is supported in $|x|\geq 1/2$, and $\mu>0$. It will be convenient to put our assumptions on K in the following more general form

$$|(\partial_x)^\alpha K(x)| \leq A_\alpha(1+|x|)^{-\mu-|\alpha|}. \tag{2.3}$$

In this setting we can always replace our original K by a family K_ε with $K_\varepsilon(x)=K(x)\,\varphi(\varepsilon x)$, where φ is a fixed C_0^∞ function, where $\varphi=1$ near the origin, and $0<\varepsilon\leq 1$. The kernels K_ε then satisfy (2.3) uniformly in ε, and for the operators $T^{(\varepsilon)}f=\int e^{i\langle Bx,y\rangle}\,K_\varepsilon(x-y)\,f(y)\,dy$ there will be no difficulty in justifying the operations carried out below. Once the estimates are obtained for $T^{(\varepsilon)}$, we then let $\varepsilon\to 0$ to get our desired conclusions.

Having made these preparations we remark that the boundedness of T follows from that of T^*T. A straightforward calculation shows that the operator T^*T has as its kernel

$$L(x,y) = \int_{\mathbf{R}^n} e^{-i\langle Bz,x-y\rangle}\,\bar{K}(z-x)\,K(z-y)\,dz. \tag{2.4}$$

The main point will be the following estimate for L:

$$|L(x,y)| \leq C_N|x-y|^{-N}, \quad \text{whenever } N\geq 0, \text{ and } N>n-2\mu. \tag{2.5}$$

We proceed as follows. We have $(a,\nabla_z)\,\langle Bz,x-y\rangle=\langle Ba,x-y\rangle$. Therefore

$$(a,\nabla_z)\,e^{-i\langle Bz,x-y\rangle}=i\langle Ba,x-y\rangle\,e^{-i\langle Bz,x-y\rangle}.$$

So if we set $\mathscr{L}_z=i(a,\nabla_z)/\langle Ba,x-y\rangle$, then

$$(\mathscr{L}_z)^N\,e^{-i\langle Bz,x-y\rangle} = e^{-i\langle Bz,x-y\rangle}.$$

Inserting this in (2.4) with

$$a = B^{-1}\left(\frac{x-y}{|x-y|}\right)$$

and integrating by parts N times gives us

$$C_N|x-y|^{-N} \sum_{|\alpha|+|\beta|=N} \int |\partial_z^\alpha K(z-x)|\,|\partial_\beta^z K(z-y)|\,dz$$

as an estimate for $L(x,y)$. It is now convenient to use the remark that

$$(1+|x|)^{-\sigma} * (1+|x|)^{-\tau} \leq c \quad \text{if } 0 \leq \sigma, \tau, \text{ and } \sigma+\tau > n. \tag{2.6}$$

In fact,

$$\begin{aligned}\int (1+|x-y|)^{-\sigma}(1+|y|)^{-\tau}\,dy &= \int_{|y|\leq|x|/2} + \int_{|x|/2\leq|y|\leq 2|x|} + \int_{|y|\geq 2|x|} \\ &= O\Bigg((1+|x|)^{-\sigma}\int_{|y|\leq|x|/2}(1+|y|)^{-\tau}\,dy \\ &\quad +(1+|x|)^{-\tau}\int_{|y|\leq 3|x|}(1+|y|)^{-\sigma}\,dy \\ &\quad +\int_{|y|\geq 2|x|}(1+|y|)^{-\sigma-\tau}\,dy\Bigg),\end{aligned}$$

which proves (2.6). For further reference we set down two other inequalities proved by this argument.

$$(1+|x|)^{-\sigma} * (1+|x|)^{-\tau} \leq c(1+|x|)^{n-\sigma-\tau}, \quad \text{if } 0 \leq \sigma, \tau < n, \text{ and } \sigma+\tau > n \tag{2.6a}$$

$$(1+|x|)^{-\sigma} * (1+|x|)^{-n} \leq c(1+|x|)^{-\sigma}\log(2+|x|), \quad \text{if } 0 < \sigma. \tag{2.6b}$$

Now observe that if $N > n-2\mu$, $\sigma = |\alpha|+\mu$, $\tau = |\beta|+\mu$, with $N = |\alpha|+|\beta|$, then $\sigma+\tau > n$. Thus by (2.6) we can conclude that $|L(x,y)| \leq C_N|x-y|^{-N}$; that is we have proved (2.5) when N is an integer. To treat the general case, let $N_0 < N \leq N_0+1$, with N_0 an integer, and consider the analytic family L_s given by

$$L_s(x,y) = e^{s^2}\int_{\mathbf{R}^n} e^{-i\langle Bz, x-y\rangle}\,\bar{K}(z-x)\,K(z-y)\,(1+|z-y|^2)^{-s/2}\,dz$$

in the strip $N_0 - N \leq \operatorname{Re}(s) \leq N_0 - N + 1$. Observe that $L_0 = L$ while the argument we used in the case when N is integral gives

$$|L_s(x, y)| \le c|x-y|^{-N_0}, \quad \text{for } \operatorname{Re}(s) = N_0 - N$$

and

$$|L_s(x, y)| \le c|x-y|^{-N_0-1}, \quad \text{for } \operatorname{Re}(s) = N_0 - N + 1.$$

Thus by the three lines theorem $|L_0(x, y)| \le C_N |x-y|^{-N}$ and (2.5) is proved.

Next using (2.5) with $N<n$ (since $\mu>0$), for $|x-y|\le 1$, and with $N>n$, when $|x-y|\ge 1$ shows that

$$\sup_x \int_{\mathbf{R}^n} |L(x, y)|\, dy \quad \text{and} \quad \sup_y \int_{\mathbf{R}^n} |L(x, y)|\, dx$$

are both finite, concluding the proof of the boundedness of T^*T and thus T, when $\mu>0$.

The case $\mu=0$ remains. Let us temporarily introduce the notation $T=T_K^B$ to make explicit the dependence of T on the bilinear form B and the kernel K. We also introduce the Fourier transform $\mathcal{F}$, defined by

$$\mathcal{F}(f)(z) = \int_{\mathbf{R}^n} e^{-i\langle Bx, z\rangle} f(x)\, dx.$$

Finally M will denote the multiplication operator given by $M(f)(y) = e^{i\langle By, y\rangle} f(y)$. Then for C_0^∞ we have the identity

$$\mathcal{F} T_K^B = T_{\mathcal{F}(K)}^{-B^*} \cdot M, \tag{2.7}$$

where B^* is the adjoint to B. This identity is proved by writing the left-side as

$$\int e^{-i\langle Bx, z\rangle} \left\{ \int e^{i\langle Bx, y\rangle} K(x-y) f(y)\, dy \right\} dx,$$

interchanging the order of integration, and noting that

$$\int e^{i\langle Bx, y-z\rangle} K(x-y)\, dx = e^{i\langle By, y-z\rangle} \mathcal{F}(K)(z-y).\text{(}^1\text{)}$$

These formal manipulations are justified when we interpret K and $\mathcal{F}(K)$ as tempered distributions, and restrict f to say C_0^∞. The identity (2.7) makes clear that to prove the L^2 boundedness of T_K^B it suffices to do the same for $T_{\mathcal{F}(K)}^{-B^*}$. From (2.3) the following

(¹) For such identities in the special case of "twisted convolution", see e.g. [20], [28].

properties of $\mathscr{F}(K)$ may be proved without difficulty (here ^ denotes the usual Fourier transform).

(a) $\mathscr{F}(K)^{\wedge}$ which equals $(2\pi)^n K(-B^{-1}\cdot)/|\det B|$ is bounded.

(b) $\mathscr{F}(K)$ coincides with a C^∞ function in $\mathbf{R}^n \setminus \{0\}$ which is rapidly decreasing at ∞.

(c) $|\mathscr{F}(K)(x)| \leq A/|x|^n$.

Now write $\mathscr{F}(K)=K_1+K_\infty$, where $K_1=\varphi\mathscr{F}(K)\, K_\infty=(1-\varphi)\,\mathscr{F}(K)$, with $\varphi\in C_0^\infty$, and $\varphi=1$ near the origin. Then $T^{-B^*}_{\mathscr{F}(K)}=T_1+T_\infty$. Here $T_\infty=T^{-B^*}_{K_\infty}$ is trivially bounded, because of property (b). Finally to prove that T_1 is bounded on L^2 is merely a reprise of step 1, carried out before. In fact if we set

$$T_1'(f)(x)=\int K_1(x-y)\, e^{-i\langle By,y\rangle} f(y)\, dy,$$

then T_1' is bounded on L^2 because of (a), while T_1-T_1' is bounded because of (c). We can then continue as in the argument in step 1. This concludes the proof of Theorem 1.

COROLLARY 1. *Suppose T is defined as in Theorem* 1, *except now we assume only that rank of $B=k$. If $\mu>n-k$, then T extends to a bounded operator on $L^2(\mathbf{R}^n)$ to itself.*

Proof. As in the proof of Theorem 1 we divide consideration in two cases, first when $K(x)$ is supported in $|x|\leq 1$. We remark that the proof given above for that case works also in the present situation since it did not depend on the nondegeneracy of B. Thus we turn to the case where K is supported in $|x|\geq 1/2$.

Let P denote the orthogonal projection on the range of B. We may assume that the rank of B is ≥ 1.(1)

LEMMA 1. *Suppose B has rank k, $k\geq 1$, $\mu>n-k$, and K satisfies* (2.3). *Then the kernel $L(x,y)$ (given by* (2.4)), *satisfies*

$$|L(x,y)|\leq c|P(x-y)|^{-k+a}(1+|x-y|)^{-n+k-b} \tag{2.8 i}$$

for some $a>0$, and $b>0$.

$$|L(x,y)|\leq c|P(x-y)|^{-N}(1+|x-y|)^{-n+k-b}, \tag{2.8 ii}$$

for some $b>0$, and all N sufficiently large.

(1) Notice that if rank $B=0$, the proof becomes trivial in view of the fact that $\mu>n$.

Proof of the lemma. We can always find a matrix B_1 so that $BB_1=P$, and $B_1P=B_1$. Recall that with $\mathscr{L}_z=i(a,\nabla_z)/\langle B(a),x-y\rangle$,

$$(\mathscr{L}_z)^N e^{-i\langle Bz,x-y\rangle}=e^{-i\langle Bz,x-y\rangle}.$$

Let $a=B_1(P(x-y)/|P(x-y)|)$. Then $B(a)=BB_1P(x-y)/|P(x-y)|=P(x-y)/|P(x-y)|$. Thus $\langle B(a),x-y\rangle=|P(x-y)|$, and

$$|L(x,y)|\leqslant c|P(x-y)|^{-N}\sum_{|\alpha|+|\beta|=N}\int|\partial_z^\alpha K(z-x)|\,|\partial_z^\beta K(z-y)|\,dz.$$

Conclusion (ii) now follows from (2.6b) if $N\geqslant 2n$, by (2.3), because either $\mu+|\alpha|$ or $\mu+|\beta|$ must be at least n, and $\mu>n-k$.

To prove (i) we may assume that $n-k<\mu\leqslant n-k+1$, since the case $\mu>n-k+1$ is a consequence of the case when $\mu\leqslant n-k+1$. Next repeat the same argument with $N=k$; thus we invoke (2.6a) or (2.6b) with $\sigma=\mu+|\alpha|$, $\tau=\mu+|\beta|$. We have $\sigma+\tau-n=2\mu+k-n>\mu$; also $\sigma\geqslant\mu$, and $\tau\geqslant\mu$. Therefore we get

$$|L(x,y)|\leqslant C|P(x-y)|^{-k}(1+|x-y|)^{-\mu}\log(2+|x-y|). \tag{2.9}$$

Next define $L_s(x,y)$ by

$$L_s(x,y)=e^{s^2}\int e^{i\langle Bz,x-y\rangle}\bar{K}(z-x)\,K(z-y)\,(1+|z-y|^2)^{-s/2}\,dz.$$

Applying the same reasoning gives

$$|L_s(x,y)|\leqslant c|P(x-y)|^{-k}(1+|x-y|)^{-n+k}\log(2+|x-y|), \tag{2.10}$$

when $\operatorname{Re}(s)=\sigma_0=n-k-\mu$. Similarly

$$|L_s(x,y)|\leqslant c|P(x-y)|^{-k+1}(1+|x-y|)^{-n+k}\log(2+|x-y|), \tag{2.11}$$

when $\operatorname{Re}(s)=\sigma_0+1=n-k-\mu+1$.

Now we have $0=(1-\theta)\sigma_0+\theta(\sigma_0+1)$, with $\theta=-\sigma_0$, and so by the three-lines theorem (note that $L_0=L$)

$$|L(x,y)|\leqslant c|P(x,y)|^{-k-\sigma_0}(1+|x-y|)^{-n+k}\log(2+|x-y|);$$

however $-k-\sigma_0=-n+\mu$. Hence,

$$|L(x,y)|\leqslant c|P(x-y)|^{-n+\mu}(1+|x-y|)^{-n+k}\log(2+|x-y|). \tag{2.12}$$

Therefore conclusion (ii) of the lemma follows if we take the geometric mean of (2.9) and (2.12).

We are dealing with the case when $K(x)$ is supported in $|x|\geq 1/2$, so the assertion of Corollary 1 is trivial when $k=0$; hence we assume now that $\operatorname{rank} B\geq 1$. Now following the argument of Theorem 1 we will show that T is bounded, by demonstrating that the kernel L of T^*T can be estimated as follows: $|L(x,y)|\leq M(x-y)$, with $M(x)\in L^1(\mathbf{R}^n)$. To do this write $\mathbf{R}^n=\mathbf{R}^k\times\mathbf{R}^{n-k}$, with $\mathbf{R}^k$ identified with the range of B, and $\mathbf{R}^{n-k}$ with its orthogonal complement; write accordingly $x=(x',x'')$, with $x'=P(x)$, and $x''=(I-P)(x)$. Then by (2.8) we can take

$$M(x)=c\min\{|x'|^{-k+a},|x'|^{-N}\}\times(1+|x|)^{-n+k-b},\quad \text{with } a>0,\ b>0.$$

Therefore, $\int_{\mathbf{R}^n} M(x)\,dx<\infty$, since

$$\int_{|x'|\leq 1}|x'|^{-k+a}\,dx'+\int_{|x'|\geq 1}|x'|^{-N}\,dx'<\infty,$$

and $\int(1+|x''|)^{-n+k-b}\,dx''<\infty$.

The proof of Corollary 1 is therefore concluded.

Remark. When $0<\operatorname{rank} B<n$, and $\mu=n-\operatorname{rank} B$, it would be interesting to find the additional conditions on K that guarantee the boundedness of T on L^2. Of course when $\operatorname{rank} B=n$, Theorem 1 shows that no additional conditions are needed; and when $\operatorname{rank} B=0$ the boundedness holds when K has vanishing mean-value on large spheres, by the standard results in singular integrals.

COROLLARY 2. *Suppose K is homogeneous of degree $-n$, smooth away from the origin, and has vanishing mean-value. Let $\langle Bx,y\rangle$ be any real bilinear form. Then the operator T defined by*

$$(Tf)(x)=\text{P.V.}\int e^{i\langle Bx,y\rangle}K(x-y)f(y)\,dy,\quad f\in C_0^\infty,$$

extends to a bounded operator on $L^2(\mathbf{R}^n)$, with bound independent of B.

Proof. Observe that if we replace the operator T by $\eta_{\delta^{-1}}T\eta_\delta$, where $\eta_\delta(f)(x)=f(\delta^{-1}x)$, then we get an operator having the same norm and of the same type, with K unchanged (because of its homogeneity of degree $-n$), but with $\langle Bx,y\rangle$ replaced by $\delta^2\langle Bx,y\rangle$. Thus we may assume that either $B=0$, in which case the assertion holds by the usual theory of singular integrals, or that $\|B\|=1$. In the later case, $\operatorname{rank} B\geq 1$; we also observe that the bounds arising in the proof of Corollary 1 depend in a uniform way on the entires of B, and this proves the corollary.

We now turn to the L^p theory of these operators.

THEOREM 2. *Let T be the kind of operator considered in Theorem* 1, *with B nondegenerate. Then T extends to a bounded operator from $L^p(\mathbf{R}^n)$ to itself under the restrictions that* $1<p<\infty$, *and* $|1/2-1/p|\leq\mu/2n$.

The proof of Theorem 2 requires the introduction of the appropriate variants of the Hardy spaces, BMO, and the "sharp function" in this context. These notions are here adapted so as to exploit the particular way our operators behave with respect to translations, (see (2.2)). We shall begin by doing this in a more general setting.

Let us assume we are given a family $E=\{e_Q\}$ of functions, $e_Q\colon \mathbf{R}^n\to\mathbf{C}$, as Q ranges over the cubes of $\mathbf{R}^n$. That is, to each such cube Q we associate a function, e_Q, in our family. The assumption we shall make is that $|e_Q|=\chi_Q$, where χ_Q is the characteristic function of Q. We define an (E) *atom*, associated to Q, to be a function a supported in Q, so that $|a(x)|\leq 1/|Q|$, and $\int a(x)\,\bar{e}_Q(x)\,dx=0$. We then define H^1_E to consist of the subspace of L^1 of functions f which can be written as $f=\Sigma\,\lambda_j a_j$, where a_j are (E) atoms, and $\lambda_j\in\mathbf{C}$, with $\Sigma\,|\lambda_j|<\infty$. Then the infimum of $\Sigma\,|\lambda_j|$, ranging over all such representations of f will be the H^1_E norm. Similarly we define the *sharp function*, $f^\#_E$, by

$$(f^\#_E)(x)=\sup_{x\in Q}\frac{1}{|Q|}\int_Q |f(x)-f^E_Q(x)|\,dx, \tag{2.13}$$

where $f^E_Q(x)=e_Q(x)\cdot\int f(y)\,\bar{e}_Q(y)\,dy\cdot|Q|^{-1}$. We define BMO_E to be the space of locally integrable f, for which $f^\#_E\in L^\infty$, and take $\|f^\#_E\|_{L^\infty}$ to be the norm.

Observe that if $e_Q=\chi_Q$ for each Q, then H^1_E, BMO_E, and $f^\#_E$ are the usual H^1, BMO, and $f^\#$ (see e.g. [12], [15], [30]). In the context of the operators we shall be dealing with we shall take the family $E=\{e_Q\}$ to be given by

$$e_Q(x)=\chi_Q(x)\,e^{-i\langle Bx,x_Q\rangle}, \tag{2.14}$$

with x_Q the center of Q. We shall be dealing with the operator $T=T^B_K$ given by $T^B_K(f)(x)=\mathrm{P.V.}\int e^{i\langle Bx,y\rangle}K(x-y)\,f(y)\,dy$, where we shall assume that K is C^∞ outside the origin, coincides with a homogeneous function of degree $-n$ near the origin, and satisfies $\int_{|x|=\varepsilon} K(x)\,d\sigma(x)=0$ for small $\varepsilon>0$. For large x we assume that K coincides with a function that satisfies the differential inequalities

$$|\partial_x^\alpha K(x)| \leq A_\alpha |x|^{-n-|\alpha|}. \tag{2.15}$$

LEMMA 2. *Let $T=T_K^B$ be as above, with $T^*=T_{K^*}^{-B^*}$ where $K^*(x)=\bar{K}(-x)$. Then*

(a) *T^* extends as a bounded operator from H_E^1 to L^1.*

(b) *T extends as a bounded operator from L^∞ to* BMO_E.

Here the family E is given by (2.14).

Proof of the lemma. Statements (a) and (b) are dual statements, and in fact the proof of (a) is very similar to that of (b). Since we shall not be using (a) below, we shall restrict ourselves to proving (b).

Write $F=T(f)$, and assume that $\|f\|_{L^\infty}=1$. In estimating $F_E^\#$, we shall first make the estimates for cubes Q *centered at the origin*. Now fix such a cube Q, let δ denote the diameter of Q, and write $Q=Q_\delta$, and ${}^cQ$ for the complement of Q in $\mathbf{R}^n$. We decompose f as $f=f_1+f_2+f_3$, where $f_1=f$ in $Q_{2\delta}$ but $f_1=0$ otherwise; $f_2=f$ in $({}^cQ_{2\delta})\cap Q_{\delta^{-1}}$, $f_2=0$ otherwise; $f_3=f$ in $({}^cQ_{2\delta})\cap({}^cQ_{\delta^{-1}})$, $f_3=0$ otherwise. Notice that $f_2\equiv0$, if $\delta\geq\sqrt{2}/2$. Write $F_j=T(f_j)$, $j=1,2,3$. Now by the L^2 theory (Theorem 1)

$$\int_Q |F_1|^2\,dx \leq \int_{\mathbf{R}^n} |F_1|^2\,dx \leq c\int |f_1|^2\,dy \leq c|Q_{2\delta}|,$$

therefore

$$\frac{1}{|Q|}\int_Q |F_1|\,dx \leq \left(\frac{1}{|Q|}\int_Q |F_1|^2\,dx\right)^{1/2} \leq c. \tag{2.16}$$

Now

$$F_2(x)=T(f_2)=\int e^{i\langle Bx,y\rangle} K(x-y)\,f_2(y)\,dy;$$

define the constant c_Q by $c_Q=\int K(-y)\,f_2(y)\,dy$. Then

$$F_2(x)-c_Q=\int \left(e^{i\langle Bx,y\rangle} K(x-y)-K(-y)\right) f_2(y)\,dy.$$

However,

$$e^{i\langle Bx,y\rangle} K(x-y)-K(-y) = (e^{i\langle Bx,y\rangle}-1)\,K(x-y)+\{K(x-y)-K(-y)\},$$

which is $O(\delta|y|^{-n+1}+\delta|y|^{-n-1})$, if $x\in Q_\delta$, and $y\in{}^cQ_{2\delta}$. Hence if $x\in Q=Q_\delta$,

$$|F_2(x)-c_Q| \leqslant C\delta\left\{\int \frac{|f_2(y)|\,dy}{|y|^{n-1}} + \int \frac{|f_2(y)|\,dy}{|y|^{n+1}}\right\} \leqslant C'\delta\left\{\int_{|y|\leqslant c\delta^{-1}} \frac{dy}{|y|^{n-1}} + \int_{|y|\geqslant c\delta} \frac{dy}{|y|^{n+1}}\right\} \leqslant C'',$$

because f_2 is supported in $({}^cQ_{2\delta}) \cap Q_{\delta^{-1}}$. So

$$\frac{1}{Q}\int_Q |F_2(x)-c_Q|\,dx \leqslant c. \tag{2.17}$$

Next,

$$\begin{aligned} F_3(x) &= \int e^{i\langle Bx,y\rangle} K(x-y)\, f_3(y)\,dy \\ &= \int e^{i\langle Bx,y\rangle} \{K(x-y)-K(-y)\}\, f_3(y)\,dy + \int e^{i\langle Bx,y\rangle} K(-y)\, f_3(y)\,dy \\ &= F_3^1(x)+F_3^2(x). \end{aligned}$$

However $|K(x-y)-K(-y)| \leqslant c|x|/|y|^{n+1}$ if $x \in Q_\delta$ and $y \in {}^cQ_{2\delta}$, and therefore

$$|F_3^1(x)| \leqslant c\delta \int_{|y|\geqslant c\delta} \frac{dy}{|y|^{n+1}} \leqslant c,$$

which gives

$$\frac{1}{|Q|}\int_Q |F_3^1(x)|\,dx \leqslant c. \tag{2.18}$$

Finally, by Plancherel's theorem

$$\begin{aligned} \int_Q |F_3^2(x)|^2\,dx &\leqslant \int_{\mathbf{R}^n} |F_3^2(x)|^2\,dx = c\int |K(-y)|^2\,|f_3(y)|^2\,dy \\ &\leqslant \int_{|y|\geqslant c\delta^{-1}} |y|^{-2n}\,dy = c'\delta^n. \end{aligned}$$

Thus $(1/|Q|)\int_Q |F_3^2(x)|\,dx \leqslant c$. Altogether then $(1/|Q|)\int_Q |F(x)-c_Q|\,dx \leqslant c$, and as a result

$$\frac{1}{|Q|}\int_Q |F(x)-F_Q|\,dx \leqslant 2c, \tag{2.19}$$

where F_Q is the mean-value of F over Q. We can now use the translation formula (2.2) to drop the assumption that Q is centered at the origin. The result is

$$\frac{1}{|Q|}\int_Q |F(x)-F_Q^E(x)|\,dx \leq c, \tag{2.20}$$

since

$$F_Q^E(x)=\chi_Q(x)\,e^{-i\langle Bx,x_Q\rangle}\,(F(\cdot)\,e^{i\langle B\cdot,x_Q\rangle})_Q$$

and so (b) of the lemma is proved.

We shall also need the following:

LEMMA 3. *Suppose* $F\in L^2(\mathbf{R}^n)$, *and* $2\leq p<\infty$. *If* $F_E^\#\in L^p(\mathbf{R}^n)$, *then* $F\in L^p(\mathbf{R}^n)$, *and*

$$\|F\|_{L^p(\mathbf{R}^n)}\leq c_p\|F_E^\#\|_{L^p(\mathbf{R}^n)}. \tag{2.21}$$

This lemma is an immediate consequence of the special case for the standard sharp function (see [12], § 4). In fact let $G(x)=|F(x)|$. Observe that

$$\frac{1}{|Q|}\int_Q \big||F(y)|-|F_Q^E(y)|\big|\,dy\leq F_E^\#(x)\quad \text{if } x\in Q;$$

but $|F_Q^E(y)|$ is constant in Q. Thus

$$\frac{1}{|Q|}\int_Q |G(y)-G_Q|\,dy\leq 2F_E^\#(x),$$

and as a result $G^\#(x)\leq 2F_E^\#(x)$, for all x. The known inequality $\|G\|_{L^p}\leq c_p\|G^\#\|_{L^p}$ then implies our result.

Proof of Theorem 2. We can now prove the theorem by using the complex interpolation method of [12]. To do this we consider first the case $\mu<n$, and break up our kernel K as K_0+K_∞, where K_0 is supported in $|x|\leq 1$, and $K_\infty(x)$ is supported in $|x|\geq 1/2$. Define the analytic family of operators T_s by

$$(T_sf)(x)=e^{s^2}\left\{\text{P.V.}\int e^{i\langle Bx,y\rangle}K_0(x-y)\,f(y)\,dy+\int e^{i\langle Bx,y\rangle}K_\infty(x-y)\,|x-y|^{\mu-ns}f(y)\,dy\right\}.$$

When $\mathrm{Re}(s)=0$ we get by Theorem 1 (where one uses only estimates like (2.3) for the second term)

$$\|T_s(f)\|_{L^2}\leq c\|f\|_{L^2},$$

which implies

$$\|(T_s(f))_E^\#\|_{L^2} \leq c\|f\|_{L^2}, \quad \mathrm{Re}(s) = 0. \tag{2.22}$$

Also by Lemma 2, (b), we have

$$\|(T_s(f))_E^\#\|_{L^\infty} \leq c\|f\|_{L^\infty}, \quad \mathrm{Re}(s) = 1.$$

Thus by complex interpolation we obtain

$$\|(T_\theta(f))_E^\#\|_{L^p} \leq c\|f\|_{L^p}$$

where $0<\theta<1$, $1/p=(1-\theta)/2+\theta\cdot 0$. But $T_\theta=cT$, $c\neq 0$, if $\mu=n\theta$, and this gives (via Lemma 3) the result when $1/2-1/p=\mu/2n$. The result when $1/p-1/2=\mu/2n$ follows by duality, and then the rest of the range is filled in by the M. Riesz convexity theorem. For $\mu=n$ the argument is similar with $T_s=T$ for all s. Finally the case $\mu>n$ is trivial since the kernel of T is then integrable near infinity.

COROLLARY 1. *Let T be an operator as in Theorem 2, except that we now assume only that* rank $B=k$. *Then T extends to a bounded operator on $L^p(\mathbf{R}^n)$ for $1<p<\infty$ under the condition that $|1/2-1/p|<(\mu-n+k)/2k$ when $k\geq 1$, and $\mu=n$ when $k=0$.*

To prove the corollary we may assume rank $B>0$, for otherwise it is a simple consequence of standard facts about singular integrals. Now the case rank $B\geq 1$ is very much the same as that of Theorem 2, except that the assertion (b) of Lemma 2 needs to be reexamined. We write $F=T(f)$ as before and estimate $F=F_1+F_2+F_3$. The estimates for F_1 and F_2 are unchanged; next $F_3=F_3^1+F_3^2$, and the estimate for F_3^1 is also unchanged. We come therefore to $F_3^2(x)$ which equals $\int e^{i\langle Bx,y\rangle} K(-y) f_3(y)\,dy$.

Now the matrix B can be written as $O_1 \Delta O_2$, where O_1 and O_2 are orthogonal matrices and Δ is a diagonal matrix, with entries $\alpha_1, \alpha_2, \ldots, \alpha_n$. Because of our assumption that rank $(B)\geq 1$, we may also assume that $\alpha_1\neq 0$. Now write $x=(x_1, x')$, with $x_1\in\mathbf{R}^1$ and $x'\in\mathbf{R}^{n-1}$. Then if we set $\tilde{F}_3^2(x)=F_3^2(O_1 x)$, $\tilde{K}(y)=K(O_2^{-1}y)$, $\tilde{f}_3(y)=f_3(O_2^{-1}y)$, we have

$$\tilde{F}_3^2(x) = \int_{\mathbf{R}^1} e^{i\alpha_1 x_1 y_1} g(y_1, x')\,dy_1 \tag{2.23}$$

where

$$g(y_1, x') = \int_{\mathbf{R}^{n-1}} e^{i\langle B'x', y'\rangle} \tilde{K}(-y) \tilde{f}_3(y)\,dy'.$$

We make the following estimate for $g(y_1, x')$

$$|g(y_1, x')| \leq c \min(\delta, |y_1|^{-1}). \tag{2.24}$$

In fact first take $|y_1| \leq c_1 \delta^{-1}$, where c_1 is sufficiently small. Then since $\tilde{f}_3(y)$ is supported in the set where $|y| \geq c\delta^{-1}$, we have

$$|g(y_1, x')| \leq c \int_{|y'| \geq c_2 \delta^{-1}} |y'|^{-n} \, dy' \leq c\delta.$$

Next if $|y_1| \geq c_1 \delta^{-1}$, use the estimate that

$$|g(y_1, x')| \leq \int |\tilde{K}_3(-y)| \, dy' \leq c \int_{\mathbf{R}^{n-1}} (|y_1|^2 + |y'|^2)^{-n/2} \, dy' \leq c|y_1|^{-1}.$$

Therefore (2.24) is proved. Hence by Plancherel's theorem in $\mathbf{R}$

$$\int |\tilde{F}_3^2(x_1, x')|^2 \, dx_1 \leq c \left\{ \int_{|y_1| \geq \delta^{-1}} \frac{dy_1}{|y_1|^2} + \delta^2 \int_{|y_1| \leq \delta^{-1}} dy_1 \right\} = c\delta.$$

Thus an extra integration in x' gives

$$\int_{O_1(Q_\delta)} |\tilde{F}_3^2(x)|^2 \, dx = \int_{Q_\delta} |F_3^2|^2 \, dx \leq c\delta^n,$$

and finally

$$\frac{1}{|Q_\delta|} \int_{Q_\delta} |F_3^2(x)| \, dx \leq c.$$

This concludes the proof of Corollary 1.

In the same way as we showed Corollary 2 of Theorem 1 we get the corresponding result for L^p.

COROLLARY 2. *Let T be the operator considered in Corollary 2 of Theorem 1. Then T extends to a bounded operator on $L^p(\mathbf{R}^n)$ to itself, if $1<p<\infty$, with bounds independent of B.*

3. Estimates for the singular Radon transforms

The bulk of the proof of Theorem A is contained in the following localized version.

Set $\mathbf{R}^{n+1} = \mathbf{R} \times \mathbf{R}^n$; a point in $\mathbf{R}^{n+1}$ will be written as (t, x), $t \in \mathbf{R}$, $x \in \mathbf{R}$ or (s, y), $s \in \mathbf{R}$, $y \in \mathbf{R}^n$. we will *always take* $n \geq 2$, and work in a fixed compact neighborhood of $\mathbf{R}^{n+1}$.

We assume that through any point (t_0, x_0) in our compact neighborhood, there passes a distinguished hypersurface, given as a graph by the equation $s=t_0+S(t_0, x_0, y)$ where S is a smooth function. Thus $S(t_0, x_0, x_0)=0$. Now our singular integrals will be defined by giving for each (t_0, x_0) a kernel concentrated in the hypersurface assigned to (t_0, x_0). The formal expression of this kernel will be $\delta(s-t-S(t,x,y))K(t,x;x-y)$, where K will be specified below. There will be several ways of writing out our operator R, which maps functions on $\mathbf{R}^{n+1}$ to functions on $\mathbf{R}^{n+1}$. If $f(s,y)$ is a function on $\mathbf{R}^{n+1}$, $\hat{f}(\lambda, y)$ will denote its Fourier transform in the s variable. Then we shall write R as

$$(Rf)(t,x)=\frac{1}{2\pi}\int e^{i\lambda t}\int e^{i\lambda S(t,x,y)}K(t,x;x-y)\hat{f}(\lambda,y)\,dy\,d\lambda \tag{*}$$

or more compactly as a pseudo differential operator

$$(Rf)(t)=\frac{1}{2\pi}\int e^{i\lambda t}a(t,\lambda)\hat{f}(\lambda)\,d\lambda. \tag{**}$$

Here $\hat{f}(\lambda)$ is a function which for each λ takes its values in the Hilbert space $L^2(\mathbf{R}^n)$, with $\hat{f}(\lambda)=\hat{f}(\lambda, y)$. Also $a(t,\lambda)$ is for each (t,λ) a bounded operator from $L^2(\mathbf{R}^n)$ to itself; $a(t,\lambda)$ has as its kernel representation

$$(a(t,\lambda)f(x))=\int e^{i\lambda S(t,x,y)}K(t,x;x-y)f(y)\,dy. \tag{***}$$

Our assumptions on S and K are as follows: S is a real C^∞ function such that

$$S(t,x,x)=0 \quad \text{for all } (t,x). \tag{3.1}$$

For each t, the Hessian $\left\{\frac{\partial^2}{\partial x_j \partial y_k}S(t,x,y)\right\}$ is a nonsingular $n\times n$ matrix. (3.2)

$K(t,x;z)$ is C^∞ with fixed compact support and satisfies

$$\left|\left(\frac{\partial}{\partial t}\right)^\alpha\left(\frac{\partial}{\partial x}\right)^\beta\left(\frac{\partial}{\partial z}\right)^\lambda K(t,x;z)\right| \leqslant A_{\alpha,\beta,\gamma}|z|^{-n-|\gamma|} \tag{3.3}$$

and

$$\left|\int_{\varepsilon<|z|<1}\left(\frac{\partial}{\partial t}\right)^\alpha\left(\frac{\partial}{\partial x}\right)^\beta K(t,x;z)\,dz\right| \leqslant A_{\alpha,\beta}, \quad \text{for } 0<\varepsilon\leqslant 1. \tag{3.4}$$

Our main theorem is as follows.

THEOREM 3. *Assume $n\geq 2$, and S and K satisfy* (3.1) *to* (3.4). *Then $R(f)$ initially defined for $f\in C_0^\infty$ extends to a bounded operator on $L^p\in(\mathbf{R}^{n+1})$ to itself,* $1<p<\infty$,

$$\|R(f)\|_p\leq A_p\,\|f_p\|.$$

With S given, the bound A_p depends only on finitely many of $A_{\alpha,\beta,\gamma}$, and $A_{\alpha,\beta}$ in (3.3) *and* (3.4).

For the proof we shall embed the operator R in an analytic family T_γ, $\gamma=\alpha+i\beta$, so that $\tau_0=R$, and when α is negative we can make L^2 estimates, while for α positive, the situation becomes more akin to the usual singular operators. To define T_γ we choose a fixed C^∞ function φ on $[0,\infty]$ which is $=1$ near the origin, and has compact support. We then write $T_\gamma=T^1+T_\gamma^2$, where T^1 has as its symbol (see (**)) $a^1(t,\lambda)$, with $a^1(t,\lambda)$ having as its kernel

$$\varphi(|x-y|^2\lambda)\,e^{i\lambda S(t,x,y)}K(t,x;x-y). \tag{3.5}$$

T_γ^2 has as its symbol the operator $a_\gamma^2(t,\lambda)$, whose kernel in turn is

$$(1-\varphi(|x-y|^2\lambda))\frac{|x-y|^{-2\gamma}}{|\lambda|^\gamma}e^{i\lambda S(t,x,y)}K(t,x;x-y). \tag{3.6}$$

LEMMA 1. *The operator $a^1(t,\lambda)$: $L^2(\mathbf{R}^n)\to L^2(\mathbf{R}^n)$ whose kernel is given by* (3.5) *satisfies the estimates*

$$\|a^1(t,\lambda)\|_{\mathrm{op}}\leq A, \tag{3.7}$$

and more generally

$$\left\|\left(\frac{\partial}{\partial t}\right)^k\left(\frac{\partial}{\partial\lambda}\right)^l a^1(t,\lambda)\right\|_{\mathrm{op}}\leq A(1+|\lambda|)^{k/2-l/2}. \tag{3.7$'$}$$

Thus $a^1(t,\lambda)$ is a symbol of type $S^0_{1/2,1/2}$.

Proof. Let us show first the estimates for the norm of $a^1(t,\lambda)$.

We break up $\mathbf{R}^n$ into a disjoint mesh of cubes $\{Q_j^\lambda\}$, so that diameter $Q_j^\lambda=\lambda^{-1/2}$. Since the kernel of $a^1(t,\lambda)$ is supported in the set $|x-y|\leq c\lambda^{-1/2}$ we see that if f is supported in Q_j^λ, then $a^1(t,\lambda)f$ is supported in $*Q_j^\lambda$, where $*Q_j^\lambda$ is a ball having the same center as $*Q_j^\lambda$, but whose diameter is $(c+1)\lambda^{-1/2}$. However the balls $*Q_j^\lambda$ have the property that every point is contained in at most a bounded number of such $*Q_j^\lambda$'s. Thus it suffices to prove that

$$\|a^1(t,\lambda)f\|_{L^2} \leqslant A\|f\|_{L^2}$$

for each f supported in Q_j^λ, where the bound A is of course independent of λ and j. To do this let y_j denote the center of Q_j^λ. Now $S(t,x,y)=S(t,x,x)+s_1(t,x)\cdot(y-x)+O|x-y|^2$. But $S(t,x,x)=0$, while $S_1(t,x)=S_1(t,y_j)+O(|x-y|)$. Thus since $|x-y_j| \leqslant (c+1)\lambda^{-1/2}$, and $|x-y| \leqslant (c+1)\lambda^{-1/2}$ we get

$$S(t,x,y)=S_1(t,y_j)\cdot(y-x)+O(|\lambda|^{-1/2}|x-y|)$$

so

$$e^{i\lambda s(t,x,y)}=e^{i\lambda S_1(t,y_j)\cdot y}\, e^{-i\lambda S_1(t,y_j)\cdot x}+O|\lambda|^{1/2}|x-y|.$$

Hence for each f supported on Q_j we can write

$$a^1(t,\lambda)f=Ma^3(t,\lambda)\bar{M}f+O\left(|\lambda|^{1/2}\int_{|x-y|\leqslant c|\lambda|^{-1/2}}|x-y|^{-n+1}|f(y)|dy\right)$$

where M is the multiplication operator

$$Mf(x)=e^{-i\lambda S_1(t,y_j)\cdot x}f(x),$$

and $a^3(t,\lambda)$ is the operator with kernel

$$\varphi(|x-y|^2\lambda)\,K(t,x;x-y).$$

Now $a^3(t,\lambda)$ can be handled by the usual theory of singular integrals. In view of the smoothness of K in the first two variables and the compact support we can write

$$K(t,x,x-y)=\int \tilde{K}(\tau,\sigma,x-y)\,e^{i(t\tau+x\cdot\sigma)}\,d\tau\,d\sigma$$

with each $\tilde{K}(\tau,\sigma,z)$ satisfying

$$\left|\left(\frac{\partial}{\partial z}\right)^\lambda \tilde{K}(\tau,\sigma,z)\right| \leqslant A_\gamma(\tau,\sigma)\,|z|^{-n-\gamma}$$

and

$$\left|\int_{\varepsilon<|z|\leqslant 1}\tilde{K}(\tau,\sigma,z)\,dz\right| \leqslant A(\tau,\sigma), \quad \text{independent of } \varepsilon,\ 0<\varepsilon\leqslant 1$$

with $A(\tau,\sigma)$ rapidly decreasing as $|\tau|+|\sigma|\to\infty$. The boundedness on $L^2(\mathbf{R}^n)$ of the operators with kernels

$$\varphi(|x-y|^2\lambda)\tilde{K}(\sigma,\tau,x-y),$$

with bounds rapidly decreasing in σ and τ, and uniform in λ then follows by a known argument (see [47] pp. 35 and 51). This proves the boundedneses of $a^3(t,\lambda)$ and hence $a^1(t,\lambda)$ on Q_j^λ completing the proof of (3.7).

The proof of (3.7′) proceeds by noting that the kernel of $(\partial/\partial t)^k(\partial/\partial\lambda)^l a^1(t,\lambda)$ can be written as the sum of two kinds of terms. One kind occurs only when $l=0$, and all the $k\partial/\partial t$ derivatives fall on $K(t,\lambda,x-y)$. This type of term is, because of (3.3) and (3.4), similar to $a^1(t,\lambda)$ and hence the bound (3.7) already proved takes care of it. The other kind is, in view of the fact that $S(t,x,x)=0$, given by kernels majorized by

$$A|\lambda|^{k'}|x-y|^{k'+l}|x-y|^{-n}, \quad \text{where } k'+l\geq 1,$$

and these kernels are supported on the set where $|x-y|\leq c\lambda^{-1/2}$. This immediately leads to the estimate

$$A|\lambda|^{k'}\int_{|x|\leq c\lambda^{-1/2}}|x|^{k'+l-n}\,dx=A|\lambda|^{k'/2-l/2}\leq A(1+|\lambda|)^{k/2-l/2}$$

for the norms of these terms, when $|\lambda|\geq 1$, and thus Lemma 1 is completely proved.

LEMMA 2. *The operators $a_\gamma^2(t,\lambda)$ whose kernels are given by* (3.6) *satisfy*

$$\left\|\left(\frac{\partial}{\partial t}\right)^k\left(\frac{\partial}{\partial\lambda}\right)^l a_\gamma^2(t,\lambda)\right\|_{\mathrm{op}}\leq A(1+|\lambda|)^{k/2-l/2}, \tag{3.8}$$

as long as $k+l<n+2\mathrm{Re}(\gamma)$.

The basic facts we shall need are contained in the following.

PROPOSITION 1. *Suppose $\Phi=\Phi(x,y)$ is a smooth real function whose Hessian $\{\partial^2\Phi/\partial x_j\partial y_k\}$ is nondegenerate. Suppose $\varphi_1(x)$ is a C^∞ function, vanishing near $x=0$, and $=1$ for large x. Suppose $\psi(x,y)\in C_0^\infty(\mathbf{R}^n\times\mathbf{R}^n)$, and $K(z)$ satisfies*

$$\left|\partial_z^\alpha K(z)\right|\leq A_\alpha|z|^{-n+m-|\alpha|}, \quad z\neq 0.$$

Define the operator $B(\lambda)$ by

$$(B(\lambda)f)(x)=\int\varphi_1(|x-y|^2\lambda)\,\psi(x,y)\,e^{i\lambda\Phi(x,y)}K(x-y)\,f(y)\,dy.$$

Then

$$\|B(\lambda)\|_{\mathrm{op}} \leqslant A(1+|\lambda|)^{-m/2} \tag{3.9}$$

as long as $m<n$. Moreover the bound A depends only on finitely many of the A_α, a lower bound for the determinant of the Hessian of Φ, and upper bounds on finitely many derivatives of Φ, φ_1, and ψ.

Proof. The operator $B(\lambda)$ is clearly bounded for each λ. It suffices to compute the operator norm $\|B(\lambda)B^*(\lambda)\|$ since this equals $\|B(\lambda)\|^2$. Now the operator $B(\lambda)B^*(\lambda)$ has kernel $L_\lambda(x,y)$ given by

$$L_\lambda(x,y) = \int \varphi_\lambda(x,y,z)\,\psi(x,y,z)\,e^{i\lambda(\Phi(x,z)-\Phi(y,z))}\,K(x-z)\,\bar{K}(z-y)\,dz \tag{3.10}$$

with $\varphi_\lambda=\varphi_1(|x-z|^2\lambda)\,\varphi_1(|z-y|^2\lambda)$; $\psi(x,y,z)=\psi(x,z)\,\psi(z,y)$. Note that φ_λ is supported where $|x-z|\geqslant c\lambda^{-1/2}$ and $|z-y|\geqslant c\lambda^{-1/2}$, while $\psi(x,y,z)$ has compact support; thus $B(\lambda)=0$ for small λ, and we need to prove (3.9) only when $|\lambda|\geqslant c_1>0$. For the kernel $L_\lambda(x,y)$ we shall make the following estimate

$$|L_\lambda(x,y)| \leqslant A_N \frac{|\lambda|^{n/2-m}}{(|\lambda|^{1/2}\,|x-y|)^N} \tag{3.11}$$

whenever $N\geqslant 0$, and $N>2m-n$.

We prove (3.11) first when N is an integer $N\geqslant 0$, $N>2m-n$. For this purpose we introduce the differential operator $D=\Sigma_{j=1}^n a_j\,\partial/\partial z_j$, and observe that

$$D(\Phi(x,z)-\Phi(y,z)) = \sum_{j,k} a_j\left[\partial^2/\partial x_k\,\partial z_j\,\Phi(y,z)\right](x_k-y_k)+O|x-y|^2.$$

Thus for appropriate a_j smooth in z, with $|a_j|\leqslant 1$

$$D(\Phi(x,z)-\Phi(y,z)) = \Delta(x,y,z), \quad \text{with } |\Delta(x,y,z)| \geqslant A|x-y|$$

for $|x-y|$ sufficiently small.([1]) Since

$$[(c\lambda\Delta)^{-1}D]^N\,e^{i\lambda(\Phi(x,z)-\Phi(y,z))} = e^{i\lambda(\Phi(x,z)-\Phi(y,z))},$$

the N corresponding integrations by parts in (3.10) give the bound

([1]) If we assume, as we may, that the support of ψ is sufficiently small.

$$|L_\lambda(x,y)| \leqslant A(|\lambda|\,|x-y|)^{-N} \sum_{0\leqslant k+l\leqslant N} \int_{\substack{c_1\geqslant|x-z|\geqslant c|\lambda|^{-1/2}\\ c_1\geqslant|z-y|\geqslant c|\lambda|^{-1/2}}} |x-z|^{-n+m-k}|z-y|^{-n+m-l}dz \tag{3.12}$$

However each integral in the sum above is actually majorized by

$$\int_{c|\lambda|^{-1/2}\leqslant|z|\leqslant c_2} |z|^{-2n+2m-k-l}\,dz,$$

and each of these integrals is in turn majorized by

$$A(|\lambda|^{-1/2})^{-n+2m-N}, \quad \text{as long as } -n+2m-N<0.$$

Substituting this in (3.12) gives (3.11) when N is an integer satisfying $N\geqslant 0$, and $N>2m-n$. To drop the integrality condition on N we can use a simple convexity argument. More precisely, we first establish the analogue of (3.11) for $m=m_j$, N_j integers satisfying $N_j\geqslant 0$, $N_j>2mj-n$, $j=0,1$, and $K(z)$ replaced by $|z|^{-m+m_j}K(z)$. Then whenever $0<\theta<1$, $N=(1-\theta)N_0+\theta N_1$, $m=(1-\theta)m_0+\theta m_1$, we have that (3.11) holds as a consequence for N, and m. Therefore (3.11) is completely proved. From (3.11) we get as a result that the norm of $B(\lambda)B^*(\lambda)$ is majorized by

$$A_{N_1}|\lambda|^{n/2-m}\int_{|\lambda|^{1/2}|x|\leqslant 1}(|\lambda|^{1/2}|x|)^{-N_1}dx + A_{N_2}|\lambda|^{n/2-m}\int_{|\lambda|^{1/2}|x|\geqslant 1}(|\lambda|^{1/2}|x|)^{-N_2}dx$$

where we take $2m-n<N_1<n$ (which is permissible since $m<n$), and $N_2>n$. It follows that $\|BB^*\|\leqslant A|\lambda|^{-m}$, and (3.9) is proved.

We turn to the proof of Lemma 2 and the inequality (3.8). Looking back at the formula (3.6) for the kernel of the operator $a^2_\gamma(t,\lambda)$ we see that the required estimates (3.8) can be reduced to those of the operators of the type $B(\lambda)$, using the same device as was used in the proof of Lemma 1 when writing $K(x,t,x-y)$ as a Fourier transform. The relevant operators $B(\lambda)$ that occur have $m=k+l-2\mathrm{Re}(\gamma)$, thus the restriction $m<n$ yields the restriction $k+l<n+2\mathrm{Re}(\gamma)$ of Proposition 1, and its conclusions give the proof of (3.8).

We shall need the following consequence of Lemma 2.

COROLLARY 1. *Suppose $a^2_\gamma(t,\lambda)$ is as in Lemma 2; assume that k and l are integers with $k+l<n+2$ Re (γ). Fix two functions φ, $\psi\in C_0^\infty(\mathbf{R})$, such that ψ is 1 for $|x|\leqslant 1$, while φ vanishes near the origin and equals 1 near $|x|=1$. For any u, $0<u\leqslant 1$, write*

$$C^1_\gamma(t,\lambda)=\left(\frac{\partial}{\partial t}\right)^k\left(\frac{\partial}{\partial\lambda}\right)^l(a^2_\gamma(ut,u^{-1}\lambda)\,\varphi(u\lambda)). \tag{3.13 a}$$

$$C^2_\gamma(t,\lambda)=\left(\frac{\partial}{\partial\lambda}\right)^l a^2_\gamma(t,\lambda)\,\psi(\lambda). \tag{3.13 b}$$

Then

$$\|C^j_\gamma(t,\lambda)\|_{\mathrm{op}}\leqslant A,\quad j=1,2 \tag{3.14}$$

with the constant A independent of u, $0<u\leqslant 1$.

This is merely a simple rewording of inequality (3.8). Note that C^2_γ does not depend on μ; its only role is to handle small frequencies.

COROLLARY 2. *Suppose* $a^2_\gamma(t,\lambda)$ *and* $C_\gamma(t,\lambda)$ *are as in the above corollary (where the upper index j has been dropped for convenience). Assume we are given* m_1, m_2, *with* $0<m_1\leqslant 1$, $0<m_2\leqslant 1$, *and so that* $k+l+m_1+m_2<n+2\mathrm{Re}(\gamma)$. *Then in addition to* (3.14) *we have*

$$\|C_\gamma(t+h_1,\lambda)-C_\gamma(t,\lambda)\|_{\mathrm{op}}\leqslant A|h_1|^{m_1} \tag{3.15}$$

$$\|C_\gamma(t,\lambda+h_2)-C_\gamma(t,\lambda)\|_{\mathrm{op}}\leqslant A|h_2|^{m_2} \tag{3.16}$$

$$\|C_\gamma(t+h_1,\lambda+h_2)-C_\gamma(t+h_1,\lambda)-C_\gamma(t,\lambda+h_2)+C_\gamma(t,\lambda)\|_{\mathrm{op}}\leqslant A|h_1|^{m_1}|h_2|^{m_2} \tag{3.17}$$

with A independent of u, $0<u\leqslant 1$.

Proof. Let γ_0, γ_1, γ_2, be such that

$$\mathrm{Re}(\gamma_0)=\mathrm{Re}(\gamma)-(m_1+m_2)/2,$$

$$\mathrm{Re}(\gamma_1)=\mathrm{Re}(\gamma)-(m_1+m_2)/2+\tfrac{1}{2},$$

$$\mathrm{Re}(\gamma_2)=\mathrm{Re}(\gamma)-(m_1+m_2)/2+1.$$

Then $\|C_{\gamma_0}(t,\lambda)\|_{\mathrm{op}}\leqslant A$, by the previous corollary. Similarly

$$\left\|\frac{\partial}{\partial t}\frac{\partial}{\partial\lambda}C_{\gamma_2}(t,\lambda)\right\|_{\mathrm{op}}\leqslant A.$$

The first inequality gives

$$\left\|C_{\gamma_0}(t+h_1,\lambda+h_2)-C_{\gamma_0}(t+h_1,\lambda)-C_{\gamma_0}(t,\lambda+h_2)+C_{\gamma_0}(t,\lambda)\right\|_{\mathrm{op}}\leqslant A$$

while the second yields

$$\left\|C_{\gamma_2}(t+h_1,\lambda+h_2)-C_{\gamma_2}(t+h_1,\lambda)-C_{\gamma_2}(t,\lambda+h_2)+C_{\gamma_2}(t,\lambda)\right\|_{\mathrm{op}}\leqslant A|h_1|\,|h_2|.$$

By similar reasoning one can show that

$$\left\|C_{\gamma_1}(t+h_1,\lambda)-C_{\gamma_1}(t,\lambda)\right\|_{\mathrm{op}}\leqslant A|h_1|,$$

$$\|C_{\gamma_1}(t,\lambda+h_2)-C_{\gamma_1}(t,\lambda)\|_{\mathrm{op}}\leqslant A|h_2|,$$

and thus

$$\|C_{\gamma_1}(t+h_1,\lambda+h_2)-C_{\gamma_1}(t+h_1,\lambda)-C_{\gamma_1}(t,\lambda+h_2)+C_{\gamma_1}(t,\lambda)\|_{\mathrm{op}}\leqslant\begin{cases}A|h_1|\\ A|h_2|\end{cases}.$$

A combination of these inequalities via complex interpolation then gives (3.15) to (3.17), proving the corollary.

We now invoke a version of the Calderón-Vaillancourt theorem for boundedness of pseudo-differential operators with symbols of the class $S^0_{1/2,1/2}$. What is important here is that there is a version where the symbol $a(t,\lambda)$ is operator-valued, ($a(t,\lambda)$ takes its values as bounded operators from $\mathscr{H}_1$ to $\mathscr{H}_2$); moreover we need to be careful about the degree of smoothness required for the symbol. Notice that here the variables t and λ range over $\mathbf{R}^1$.

PROPOSITION. *Suppose* $a(t,\lambda)$ *is given, and write*

$$C^1(t,\lambda)=\left(\frac{\partial}{\partial\lambda}\right)^l(a(ut,u^{-1}\lambda)\,\varphi(u\lambda)),\quad 0<u\leqslant 1,\quad C^2(t,\lambda)=\left(\frac{\partial}{\partial\lambda}\right)^l a(t,\lambda)\,\psi(\lambda)$$

with φ *and* ψ *as in Corollary* 1. *Suppose* $l=0$, *or* 1; $m_1>1/2$, $m_2>0$ *and* $C^j(t,\lambda)$ *satisfies the conditions* (3.15) *to* (3.17), *uniformly in* u, $0<u\leqslant 1$, $j=1,2$. *Then the operator*

$$T(f)(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\lambda t}\,a(t,\lambda)\hat{f}(\lambda)\,d\lambda$$

extends to a bounded operator from $L^2(\mathbf{R},\mathscr{H}_1)$ *to* $L^2(\mathbf{R},\mathscr{H}_2)$.

(For a proof see the Appendix.)

In applying the proposition note that $k+l+m_1+m_2=1+m_1+m_2$ and so it applies whenever $(3/2-n)/2<\mathrm{Re}(\gamma)$; and in particular when $n\geqslant 2$ we get a strip which includes the origin in its interior. The result is

$$\|T_\gamma(f)\|_2 \leqslant A(\gamma)\|f\|_2, \tag{3.18}$$

when

$$\mathrm{Re}(\gamma) > (3/2-n)/2, \tag{3.19}$$

where the constants $A(\gamma)$ depend on only finitely many of the bounds $A_{\alpha,\beta,\gamma}$ and $A_{\alpha,\beta}$ appearing in (3.3) and (3.4), and are of at most polynomial growth in γ for γ in any strip of the form $\gamma_1\leqslant\mathrm{Re}(\gamma)\leqslant\gamma_2$, with $\gamma_1>(3/2-n)/2$. These conclusions are arrived at by combining Lemma 1, Lemma 2 and its second corollary.

We shall now consider the operator T_γ, $\gamma=\alpha+i\beta$, with $\alpha=\mathrm{Re}(\gamma)>0$. We shall write T_γ in its kernel expression with

$$T_\gamma(f)(P) = \int K_\gamma(P,Q)f(Q)dQ$$

where K_γ is a singular kernel on $\mathbf{R}^{n+1}\times\mathbf{R}^{n+1}$, with dQ the Euclidean measure on $\mathbf{R}^{n+1}$. Going back to the definitions (see (*), (**), (3.5) and (3.6)) we see that we can write

$$K_\gamma(P,Q) = |x-y|^{-2}\Phi_\gamma\left(\frac{s+S-t}{|x-y|^2}\right)K(t,x;x-y)$$

where

$$\Phi_\gamma(u) = \left(\frac{1}{2\pi}\right)\int_0^\infty e^{-iu\lambda}\lambda^{-\gamma}\,d\lambda. \tag{3.20}$$

Here we have used the notation $\int_0^\infty$ to indicate that $\lambda^{-\gamma}$ has been modified near $\lambda=0$ so as to be smooth here; also $P=(t,x)$, $Q=(s,y)$.

It will be natural to take P as the center of a coordinate system, with Q the variable point. Thus we define this coordinate system by assigning Q the coordinates $[\sigma, z]$, with $\sigma\in\mathbf{R}$, $z\in\mathbf{R}^n$, where

$$\begin{aligned}\sigma &= s+S(t,x,y)-t\\ z &= y-x.\end{aligned} \tag{3.21}$$

This assignment of coordinates for Q varies smoothly with P. Also it is to be noted that we can think of Q as the fixed point (i.e coordinate center), with P varying near Q;

then $[\sigma, z]$ also give coordinates for P near Q. When we do this, and integrate with respect to P, it will be useful to observe that $dP \approx dz\, d\sigma$.

Now the coordinate systems we have introduced define a family of "balls" (and a resulting quasi-distance) which will be controlling in what follows. Thus with $P=(t,x)$, $Q=(s,y)$ we write $d(P,Q)<\delta$ (with $\delta \leq 1$), if

$$|x-y|<\delta, \text{ i.e } |z|<\delta \quad \text{and} \quad |t-s-S(t,x,y)|<\delta^2, \text{ i.e. } |\sigma|<\delta^2. \tag{3.22}$$

Keep P fixed, and let Q_1, Q_2 be two points with coordinates (centered at P) given by $[\sigma_1, z_1]$ and $[\sigma_2, z_2]$ respectively. Then it is not difficult to prove (see the analogous argument e.g. in Folland-Stein [14], p. 475–6) that

$$\begin{gathered} |z_1-z_2| \leq c d(Q_1, Q_2) \\ |\sigma_1-\sigma_2| \leq c\{d(Q_1,Q_2)^2 + d(P,Q_1)\, d(Q_1,Q_2)\}. \end{gathered} \tag{3.23}$$

From this it can be seen without difficulty that $d(Q_1,Q_2) \approx d(Q_2,Q_1)$, and the quasi-triangle inequality holds.

We shall temporarily write

$$K_\gamma(P,Q) = M(P;\sigma,z)$$

where $[\sigma, z]$ are the coordinates of Q with respect to P. We observe the following differential estimates for M:

LEMMA 3. *Let* $\alpha = \operatorname{Re}(\gamma) > 0$. *Then*
(i) *for* $|\sigma| \leq |z|^2$

$$\begin{aligned} &\text{(a)} \quad |M(P;\sigma,z)| \leq c|z|^{-n-2\alpha}|\sigma|^{-1+\alpha} \\ &\text{(b)} \quad \left|\frac{\partial M}{\partial z_j}(P;\sigma,z)\right| \leq c|z|^{-n-1-2\alpha}|\sigma|^{-1+\alpha}, \quad j=1,\dots,n. \\ &\text{(c)} \quad \left|\frac{\partial M}{\partial \sigma}(P;\sigma,z)\right| \leq c|z|^{-n-2\alpha}|\sigma|^{-2+\alpha} \end{aligned} \tag{3.24}$$

(ii) *for* $|\sigma| \geq |z|^2$

$$\begin{aligned} &\text{(a)} \quad |M(P;\sigma,z)| \leq c|\sigma|^{-n/2-1} \\ &\text{(b)} \quad \left|\frac{\partial M}{\partial z_j}(P;\sigma,z)\right| \leq c|\sigma|^{-n/2-3/2} \\ &\text{(c)} \quad \left|\frac{\partial M}{\partial \sigma}(P;\sigma,z)\right| \leq c|\sigma|^{-n/2-2}. \end{aligned} \tag{3.25}$$

Proof. These estimates follow directly from the definition (3.20) viz.

$$K_\gamma(P, Q) = M(P, \sigma, z) = |z|^{-2}\Phi_\gamma(\sigma/|z|^2)\, K(P, z),$$

the property (3.3) of $K(P, z)$, and the fact that

$$\Phi_\gamma(u) = \textstyle\int_0^\infty e^{-iu\gamma}\lambda^{-\gamma}\, d\lambda$$

is $O(|u|^{-1+\alpha})$, and $\Phi'_\gamma(u)$ is $O(|u|^{-2+\alpha})$ as $u \to 0$, while Φ_γ is rapidly decreasing with its derivatives as $|u| \to \infty$.

At this stage the basic fact about the kernel $K_\gamma(P, Q)$ will be contained in the following lemma:

LEMMA 4. *Assume $K_\gamma(P, Q) = M(P, \sigma, z)$ satisfies the conclusions* (3.29) *and* (3.25) *of the previous lemma. Then*

$$\int |K_\gamma(P, Q_1) - K_\gamma(P, Q_2)|\, dP \leq A \tag{3.26}$$

where the integral is taken over the region where $d(P, Q_1) \geq \bar{c}d(Q_1, Q_2)$, and $\bar{c} \gg 1$.

To prove the lemma we begin by considering $\Sigma_{j=0}^\infty \psi_j(u) = 1$, a standard partition of unity of $\mathbf{R}$, with $\psi_0(u) = 1$ for $|u| \geq 2$, and $\psi_j(u)$ supported where $|u| \approx 2^{-j}$, with $|\psi'_j(u)| \leq c2^j$. Write $K^j(P, Q) = K(P, Q)\psi_j(\sigma/|z|^2) = M(P; \sigma, z)\psi_j(\sigma/|z^2|)$. We are going to estimate $K^j(P, Q_1) - K^j(P, Q_2)$, when $j \geq 1$ first. We fix Q_1, Q_2 so that $d(Q_1, Q_2) \leq a$, and P will vary where $d(P, Q_2) \geq \bar{c}a$. Here $\bar{c}$ will be a constant (which will be fixed later as large) and a is to take all values in $(0,1]$.

We shall write $\int_{d(P, Q_1) \geq \bar{c}a} |K^j(P, Q_1) - K^j(P, Q_2)|\, dP$ as $\int_{\mathrm{I}} + \int_{\mathrm{II}}$, where the region I is defined (for each j) to be the set of P where $d(Q_1, Q_2) = a > \underline{c}2^{-j}d(P, Q_1)$, and II the set of P where $d(Q_1, Q_2) \leq \underline{c}2^{-j}d(P, Q_1)$. Here $\underline{c}$ is a positive constant, which will be fixed later to be small.

Let us consider first $\int_{\mathrm{I}}$. We shall write

$$\int_{\mathrm{I}} |K^j(P, Q_2)|\, dP \leq \int_{\mathrm{I}} |K^j(P, Q_1)| + \int_{\mathrm{I}} |K^j(P, Q_2)|\, dP.$$

Now $d(P, Q_1) \approx |z_1| + |\sigma_1|^{1/2}$, but $|\sigma_1| 2^j \approx |z_1|^2$ on the support of K^j (because of the cut-off function $\psi_j(\sigma/|z|^2)$) thus $|z_1|\approx d(P, Q_1)\leq c2^j d(Q_1, Q_2)=c2^j a$, since we are in the region I. Hence by (3.24) (a),

$$\int_{\begin{cases} I \\ d(P,Q_1)\geq ca \end{cases}} |K^j(P, Q_1)|\, dP \leq c \int_{\begin{cases} c_1 a<|z_1|<c_2 2^j a \\ |z_1|^2 \approx |\sigma_1| 2^j \end{cases}} |z_1|^{-n-2} dz_1\, d\sigma_1 \cdot 2^{(1-\alpha)}.$$

But

$$c \int_{\begin{cases} c_0 a<|z_1|<c_2 2^j a \\ |z_1|^2 \approx |\sigma_1| 2^j \end{cases}} |z_1|^{-n-2} dz_1\, d\sigma \leq c \int_{c_1 a<|z_1|<c_2 2^j a} |z_1|^{-n} dz_1 \cdot 2^{-j}$$

$$\leq c_0(\log j+c_3)2^{-j}.$$

Thus,

$$\int_{\begin{cases} I \\ d(P,Q_1)\geq \bar{c}a \end{cases}} |K(P, Q_1)|\, dP \leq c(\log j+c)2^{-\alpha j}. \tag{3.27}$$

The integral of $|K(P, Q_2)|$ taken over the same region is handled similarly. In fact since $d(P, Q_1)\geq \bar{c}d=d(Q_1, Q_2)$ where $\bar{c}$ is *large*, then by the triangle inequality $d(P, Q_2)\geq ca$. Also $d(P, Q_2)\leq c\{d(P, Q_1)+d(Q_1, Q_2)\}\leq c''2^j a$ in region *I*. Thus we can apply the same argument that worked for $K(P, Q_1)$ giving

$$\int_{\begin{cases} I \\ d(P,Q_1)\geq ca \end{cases}} |K(P, Q_2)|\, dP \leq c(\log j+c)2^{-\alpha j}. \tag{3.27'}$$

In making the estimates for the region II, we shall consider Q_1 as fixed; then we can think of $[\sigma_1, z_1]$ as coordinates specifying the point P. Also $[\sigma_2, z_2]$ (which were originally the coordinates of Q_2 in the coordinate system centered at P), can be thought of as functions of P, and hence of $[\sigma_1, z_1]$.

Let us observe that if we are in the region II, then whenever $\sigma_1/|z_1|^2\approx 2^{-j}$, then also $\sigma_2/|z_2|^2\approx 2^{-j}$. In fact $|z_1-z_2|\leq cd(Q_1, Q_2)$ (see (3.23)), so

$$|z_1-z_2|\leq ca\leq c\underline{c}2^{-j}d(P, Q_1)\leq c'\underline{c}2^{-j}|z_1|.$$

Thus if $\underline{c}$ is sufficiently small, then $|z_2|\approx|z_1|$. Next, again by (3.23), $|\sigma_1-\sigma_2|\leq c(a^2+|z_1|a)$, and since as we have seen $a\leq c\underline{c}2^{-j}|z_1|$, with $\underline{c}$ small, we get $|\sigma_2|\approx 2^{-j}|z_2|$. The same argument shows that if $[\sigma, z]$ are the points in the line segments joining $[\sigma_1, z_1]$ to $[\sigma_2, z_1]$ and $[\sigma_2, z_1]$ to $[\sigma_2, z_2]$ then $|\sigma|/|z|^2\approx 2^{-j}$ throughout.

Now

$$K^j(P, Q_1)-K^j(P, Q_2) = M(P, \sigma_1, z_1)\psi_j(\sigma_1/|z_1|^2)-M(P, \sigma_2, z_2)\psi_j(\sigma_2/|z_2|^2) = A+B$$

where

$$A = M(P, \sigma_1, z_1)\, \psi_j(\sigma_1/|z_1|^2)-M(P, \sigma_2, z_1)\, \psi_j(\sigma_2/|z_1|^2)$$

and B has a similar definition.

However (3.24) (c) allows us to make the following estimate:

$$|A| \leqslant c|z_1|^{-n-4}|\sigma_1-\sigma_2|2^{(2-\alpha)j}\tilde{\psi}_j(|\sigma_1|/|z_1|^2) \tag{3.28}$$

when $\tilde{\psi}_j$ is the characteristic function of the set $c_1 2^{-j}<|\sigma_1|/|z_1|^2<c_2 2^{-j}$ for two appropriate constants c_1 and c_2.

Also by (3.24)(a) we can say that

$$|A| \leqslant c|z_1|^{-n-2}\cdot 2^{(1-\alpha)j}\tilde{\psi}_j(|\sigma_1|/|z_1|^2). \tag{3.28$'$}$$

Combining these two yields

$$|A| \leqslant c|z_1|^{-n-2-2\varepsilon}(|\sigma_1-\sigma_2|2^j)^\varepsilon 2^{(1-\alpha)j}\tilde{\psi}_j(|\sigma_1|/|z_1|^2), \quad 0 \leqslant \varepsilon \leqslant 1. \tag{3.29}$$

However $|\sigma_1-\sigma_2| \leqslant c(a^2+a|z_1|)$, by (3.23) as we have already remarked, therefore, since $|z_1| \geqslant ca$, we get

$$|A| \leqslant c|z_1|^{-n-\varepsilon}a^\varepsilon 2^{\varepsilon j}2^{(1-\alpha)j}\tilde{\psi}_j(|\sigma_1|/|z_1|^2). \tag{3.30}$$

Thus

$$\int_{\begin{cases}\mathrm{II} \\ d(P, Q_1)\geqslant \bar{c}a\end{cases}} |A|dP \approx \int|A|dz_1\, d\sigma_1 \leqslant c2^{\varepsilon j}2^{-\alpha j}, \quad \text{if } \varepsilon > 0.$$

The same estimate holds for the contribution of B (here we use (3.24)(b) instead of (3.24)(c), and things are even a littler simpler). Altogether then (taking into account (3.27)), we have

$$\int_{d(P, Q_1)\geqslant \bar{c}d(Q_1, Q_2)} |K^j(P, Q_1)-K^j(P, Q_2)|\, dP \leqslant c(\log j\cdot 2^{-\alpha j}+2^{\varepsilon j}2^{-\alpha j}) \tag{3.31}$$

whenever $0<\varepsilon\leqslant 1$. Now we merely need to take $0<\varepsilon<\alpha$ and sum in j (the term $j=0$ is dealt with separately via inequalities (3.25); the argument is the same as before, but is now in its simplest form). The conclusion is (3.26) and Lemma 4 is proved.

We can now invoke the theory of singular integrals in the setting for which there is a quasi-metric $d(P, Q)$ of the type we have used (see e.g. Coifman and Weiss [8], Chapter III). We observe that our proof of (3.26) showed that $A=A_\gamma$ grows only linearly in $\beta=\operatorname{Im}(\gamma)$, as $\beta\to\infty$, when $\operatorname{Re}(\gamma)>0$, and by (3.18), (3.19) the operator T_γ is bounded on L^2 for $\operatorname{Re}(\gamma)>0$, with bounds growing at most polynomially in β. The conclusion is that T_γ is bounded in L^p, $1<p\leq 2$, $\operatorname{Re}(\gamma)>0$, again with bounds growing at most polynomially with β. Invoking (3.18) (3.19) this time with $\operatorname{Re}(\gamma)<0$ we get that $R=T_0$ is bounded on L^p, $1<p\leq 2$. The case $2\leq p<\infty$ is handled by duality. That this can be done is an immediate consequence of the Observations 2 and 3 of Section 1. The proof of Theorem 3 is complete.

We can now give the

Proof of Theorem A (*case* (b)): Let $\mathscr{C}_1\subset\subset\mathscr{C}_2$ be small neighborhoods of the diagnonal in $\mathscr{C}$, and $\chi(P, Q)$ a C^∞ function which is 1 in $\mathscr{C}_1$ and 0 outside $\mathscr{C}_2$. Then Rf may be decomposed as

$$(Rf)(P)=\int_{\Omega_P}\chi(P,Q)K(P,Q)f(Q)\,d\sigma_P(Q)+\int_{\Omega_P}[1-\chi(P,Q)]\,K(P,Q)f(Q)\,d\sigma_P(Q)$$

$$\equiv(R_1f)(P)+(R_2f)(P).$$

The density $(1-\chi(P, Q))K(P, Q)$ is smooth with C^0 norms bounded by $c\|K\|^{0,0}$. Thus we may write using the fact that the support of u is included in the compact set $\bar{\Omega}_1$

$$\begin{aligned}\int_{\Omega_2}|R_2f(P)|^p dv &\leq C_K\int_{\Omega_2}\left[\int_{\Omega_P}|f(Q)|d\sigma_P(Q)\right]^p dv(P)\\ &\leq C_K\int_{\Omega_2}\left[\int_{\Omega_P\cap(\operatorname{supp}u)}d\sigma_P(Q)\right]^{p/q}\left(\int_{\Omega_P}|f(Q)|^p d\sigma_P(Q)\right)dv(P)\\ &\leq C_{\Omega_1,K}\int_{\Omega_2}\int_{\Omega_P}|f(Q)|^p d\sigma_P(Q)\,dv(P)\\ &= C_{\Omega_1,K}\int_{\mathscr{C}\cap(\Omega_2\times\Omega)}|f(Q)|^p d_\sigma(P,Q)\\ &= C_{\Omega_1,K}\int_{\Omega}|f(Q)|^p\left(\int_{\Omega_Q^*\cap\Omega_2}d\sigma_Q(P)\right)dQ\\ &\leq C_{\Omega_1,\Omega_2,K}\|f\|^p_{L^p(\Omega_1)}.\end{aligned}$$

Turning to R_1, we may assume that $\Omega_1=\Omega_2$, and by use of a partition of unity, that they are contained in coordinate patches. If the support of χ is narrow enough, the coordinate patches may be taken as in the Corollary in Section 1, with condition (1.8) holding uniformly in the sense that $\det(\partial^2 S/\partial x_j \partial y_k)$ is bounded away from 0 by a fixed constant. In terms of these coordinates, the density K on $\mathscr{C}$ can be written as a C^∞ function $K(t,x;z)$ and its seminorms as an admissible density are routinely checked to be equivalent to the best constants $A_{\alpha,\beta,\gamma}$ satisfying (3.3) and (3.4). Thus R_1 reduces to an operator of the form (*) defined at the beginning of this section, the desired hypotheses hold, and Theorem 3 applies. This proves part (b) of Theorem A. Part (a) follows from part (b) by standard approximation procedurs. Q.E.D.

4. The maximal function

Our setting is as before in § 2, but now at each P we consider the ball of radius $\varepsilon>0$, on M_P centered at P. We let $A_\varepsilon(f)$ denote the "average" of f over this ball and we are concerned with $\lim_{\varepsilon\to 0} A_\varepsilon(f)$ and $\sup_{\varepsilon>0}|A_\varepsilon(f)|$.

Following the localization we have used in § 3, we define (on $\mathbf{R}^{n+1}$), the operator A_ε as follows

$$A_\varepsilon(f)(t,x)=\int \delta(t-s-S(t,x,y))\psi_\varepsilon(x-y)\,\eta(t,x)\,f(s,y)\,ds\,dy. \tag{4.1}$$

Here $\psi_\varepsilon(u)=\psi(u/\varepsilon)\varepsilon^{-n}$, where ψ is a fixed C_0^∞ function on $\mathbf{R}^n$ supported in $|u|\geqslant 1$, $\psi\geqslant 0, \int \psi\,dx=1$; and η is a fixed cut-off function of compact support.

We shall have to compare the averages A_ε, with others which can be handled by more standard methods. Thus we define

$$B_\varepsilon(f)(t,x)=\int \varphi_\varepsilon(t-s-S(t,x,y))\psi_\varepsilon(x,y)\eta(t,x)f(s,y)ds dy \tag{4.2}$$

where $\varphi_\varepsilon(u)=\varepsilon^{-2}\varphi(u/\varepsilon^2)$, with φ a fixed C^∞ function in $\mathbf{R}^1$, $\varphi\geqslant 0$, $\int_{-\infty}^{\infty}\varphi(u)du=1$.

The averages (4.2) correspond essentially to mean-values taken over the balls used in § 3 above and defined by (3.22).

The basic estimate we shall make will be in terms of a square function G, defined by

$$G(f)(\varepsilon,x)=\left(\int_0^1 |A_\varepsilon(f)-B_\varepsilon(f)|^2\frac{d\varepsilon}{\varepsilon}\right)^{1/2}. \tag{4.3}$$

THEOREM 4. *Assume $n \geqslant 2$, then*

$$\|G(f)\|_p \leqslant A_p \|f\|_p, \quad 1 < p \leqslant 2.$$

The way to deal with the function G is to consider a closely related linear operator T from $L^p(\mathbf{R}^{n+1})$, to $L^p(\mathbf{R}^{n+1}, \mathscr{H})$ where $\mathscr{H}$ is the Hilbert space $L^2(d\varepsilon/\varepsilon, (0,1))$. T is defined by

$$(Tf)(t, x, \varepsilon) = \int (\delta(t-s-S(t,x,y)) - \varphi_\varepsilon(t-s-S(t,x,y)))\psi_\varepsilon(x-y)\eta(t,x)f(s,y)\,ds\,dy \quad (4.4)$$

where we consider the right side of (4.4) as a function on the (t, x) space (i.e. $\mathbf{R}^{n+1}$) with values in $L^2(d\varepsilon/\varepsilon, (0,1))$.

We shall also use the "pseudo-differntial" version of (4.4), in analogy with § 3, where

$$(Tf)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda t} a(t, \lambda) \hat{f}(\lambda)\, d\lambda \quad (4.5)$$

where $\hat{f}(\lambda)$ is a function which for each λ takes its values in the Hilbert space $L^2(\mathbf{R}^n)$, with $\hat{f}(\lambda) = \hat{f}(\lambda, y)$. Also $a(t, \lambda)$ is for each (t, λ) an operator from $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n) \times L^2(d\varepsilon/\varepsilon, (0,1))$; $a(t, \lambda)$ has as its kernel representation:

$$a(t, \lambda) f(x) = (1 - \hat{\varphi}(\varepsilon^2 \lambda)) \int_{\mathbf{R}^n} e^{i\lambda S(t,x,y)} \psi_\varepsilon(x, y) \eta(t, x) f(y)\, dy. \quad (4.6)$$

Together with the operator T, we shall consider an analytic family of operators T_γ, such that $T_0 = T$. The operator T_γ will be defined in analogy with T (see (4.5)). However now its symbol $a_\gamma(t, \lambda)$ will be the operator from $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n) \times L^2(d\varepsilon/\varepsilon)$, $(0,1))$ which has as its kernel representation

$$a_\gamma(t, \lambda) = (1 + \varepsilon^4 \lambda^2)^{-\gamma/2} a(t, \lambda). \quad (4.7)$$

In analogy with Section 3 and in particular with (3.8) we shall prove for $a_\gamma(t, \lambda)$ the estimate

$$\left\| \left(\frac{\partial}{\partial t}\right)^k \left(\frac{\partial}{\partial \lambda}\right)^l a_\gamma(t, \lambda) \right\|_{\mathrm{op}} \leqslant A_\gamma (1 + |\lambda|)^{k/2 - l/2} \quad (4.8)$$

when $k + l < n + 2\mathrm{Re}(\gamma)$.

For this it will suffice (using the notation of Proposition 1) to consider the operators $\tilde{B}_\gamma(\lambda)$ (from $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n) \times L^2(d\varepsilon/\varepsilon, (0,1))$ given by

$$(\tilde{B}_\gamma(\lambda))(f)(x,\varepsilon) = (1-(\hat{\varphi})(\varepsilon^2\lambda))(1+\varepsilon^4\lambda^2)^{-\gamma/2}\int_{\mathbf{R}^n} e^{i\lambda\Phi(x,y)}\psi_\varepsilon(x,y)f(y)dy$$

where $\psi_\varepsilon(x,y)$ is supported in $|x-y|\leqslant\varepsilon$, with $0<\varepsilon\leqslant 1, \psi_\varepsilon$ satisfies the estimates

$$\left|\left(\frac{\partial}{\partial x}\right)^\alpha\left(\frac{\partial}{\partial y}\right)^\beta \psi_\varepsilon(x,y)\right| \leqslant A_{\alpha,\beta}\varepsilon^{-n-|\alpha|-|\beta|+\delta}.$$

Under these assumptions we shall show that

$$\|\tilde{B}_\gamma(\lambda)\|_{\mathrm{op}} \leqslant A(1+|\lambda|)^{-\delta/2}, \tag{4.9}$$

when

$$-1<\delta<n+2\,\mathrm{Re}(\gamma).$$

Writing out

$$\int_0^1\int_{\mathbf{R}^n} |\tilde{B}_\gamma(f)(x,\varepsilon)|^2 dx\frac{d\varepsilon}{\varepsilon}$$

shows that it equals

$$\int\int_{\mathbf{R}^n\times\mathbf{R}^n} L_\lambda^\gamma(x,y)f(x)\bar{f}(y)\,dx\,dy,$$

where the kernel L_λ^γ is given by

$$L_\lambda^\gamma(x,y) = \int_0^1 |1-\hat{\varphi}(\varepsilon^2\lambda)|^2(1+\varepsilon^4\lambda^2)^{-\mathrm{Re}(\gamma)}I_\lambda^\gamma(x,y,\varepsilon)\frac{d\varepsilon}{\varepsilon}$$

where

$$I_\lambda^\gamma(x,y,\varepsilon) = \int_{\mathbf{R}^n} e^{i\lambda\{\Phi(z,x)-\Phi(z,y)\}}\psi_\varepsilon(z,x)\bar{\psi}_\varepsilon(z,y)\,dz.$$

Now in view of our assumptions on ψ_ε it is obvious that

$$|I_\lambda^\gamma| \leqslant A\varepsilon^{-n+2\delta}\chi_\varepsilon(x-y)$$

where $\chi_\varepsilon(u)$ is the characteristic function of the ball $|u|\leqslant 2\varepsilon$. Invoking the integration-by-parts argument preceeding (3.12) then shows that

$$|I_\lambda^\gamma(x,y,\varepsilon)| \leqslant A_N\varepsilon^{-n+2\delta}(1+\varepsilon|\lambda|\,|x-y|)^{-N}\chi_\varepsilon(x-y) \tag{4.10}$$

for every $N \geqslant 0$. Thus since

$$|1-\hat{\varphi}(\varepsilon^2\lambda)|^2 \leqslant A\frac{\varepsilon^4\lambda^2}{1+\varepsilon^4\lambda^2},$$

(recall that $\varphi \in C_0^\infty$, and $\hat{\varphi}(0)=1$), we have

$$|L_\lambda^\gamma(x,y)| \leqslant A_N \int_{|x-y|/2}^{1} \varepsilon^{-n+3+2\delta}\lambda^2(1+\varepsilon^4\lambda^2)^{-\mathrm{Re}\,\gamma-1}(1+\varepsilon|\lambda|\,|x-y|)^{-N}\,d\varepsilon. \tag{4.11}$$

Observe that if $|\lambda| \leqslant 1$, this shows that $|L_\lambda^\gamma(x,y)| \leqslant A|x-y|^{-n+2+2\delta}$ since the function $L(u)$ which is $O(|u|^{-n+2+2\delta})$, $|u| \leqslant 1$ and vanishes when $|u|>2$ is integrable when $\delta > -1$, we get $\|\tilde{B}_\gamma(\lambda)\|_{\mathrm{op}} \leqslant A$, when $|\lambda| \leqslant 1$, and $\delta > -1$.

To consider what happens when $|\lambda| \geqslant 1$, make the change of variables replacing ε by $\varepsilon|\lambda|^{-1/2}$. Then the right side if (4.11) is majorized by $A_N\lambda^{n/2-\delta}L^\lambda(|\lambda|^{1/2}(x-y))$, with

$$L^\gamma(u) = \int_{|u|/2}^{\infty} \varepsilon^{3-n+2\delta}(1+\varepsilon^4)^{-\mathrm{Re}(\gamma)-1}(1+\varepsilon|u|)^{-N}\,d\varepsilon,$$

and $L^\gamma(u)$ is $O(|u|^{n-2\delta+4\mathrm{Re}(\gamma)})$ when $|u| \to 0$ and is $O(|u|^{-n-1})$ if N is sufficiently large, and this gives an integrable function over $\mathbf{R}^n$ when $-1 < \delta < n+2\,\mathrm{Re}(\gamma)$. This proves $\|\tilde{B}_\gamma(\lambda)\|_{\mathrm{op}} \leqslant A|\lambda|^{-\delta/2}$if $|\lambda| \geqslant 1$ under those conditions, and therefore (4.9) is proved.

As a consequence of this and in parallel with the argument leading to (3.18) (using pseudo-differential operators of class $S_{1/2,1/2}$) we obtain the L^2 estimates for our operator T_γ, namely

$$\|T_\gamma(f)\|_2 \leqslant A(\gamma)\,\|f\|_2 \tag{4.12}$$

whenever

$$\mathrm{Re}(\gamma) > (1-n)/2.$$

We now pass to L^p estimates for T_γ, when $\mathrm{Re}\,(\gamma) > 0$, $\gamma = \alpha + i\beta$. To do this we write T_γ in its kernel expression

$$T_\gamma(f)\,(P,\varepsilon) = \int K_\gamma(P,Q,\varepsilon)f(Q)dQ,$$

where K_γ is a function on $\mathbf{R}^{n+1} \times \mathbf{R}^{n+1} \times (0,1)$, and dQ euclidean measure on $\mathbf{R}^{n+1}$. (Recall we are viewing T_γ as an operator from scalar-valued functions on $\mathbf{R}^{n+1}$ to L^2 $(d\varepsilon/\varepsilon,\ (0,1))$-valued functions on $\mathbf{R}^{n+1}$.) Going back to the definitions ((4.4), (4.5), (4.7)) used in (3.21) and the notation $\sigma=s+S(t,x,y)-t$, $z=y-x$, we see that we can write

$$K_\gamma(P,Q,\varepsilon)=\varepsilon^{-2}K_\gamma(\sigma/\varepsilon^2)\,\psi_\varepsilon(z)\cdot\tilde{\eta}(\sigma,z). \tag{4.13}$$

Here

$$K_\gamma(u)=\int_{-\infty}^{\infty} e^{-iu\lambda}(1+\lambda^2)^{-\gamma/2}(1-\varphi(\lambda))\,d\lambda,$$

with $\psi_\varepsilon(z)=\varepsilon^{-n}\psi(z/\varepsilon)$, and ψ a C^∞ function with support in $|z|\le 1$, and $\tilde{\eta}(\sigma,z)=\eta(t,x)$.

Now for $K_\gamma(u)$ we make the following estimates

$$|K_\gamma(u)|\le A|u|^{-1+\mathrm{Re}(\gamma)},\quad u\to 0,\quad K_\gamma(u)\text{ rapidly decreasing as }|u|\to\infty \tag{4.14}$$

$$\left|\frac{\partial K_\gamma(u)}{\partial u}\right|\le A|u|^{-2+\mathrm{Re}(\gamma)},\quad u\to 0,\quad \frac{\partial K_\gamma(u)}{\partial u}\text{ rapidly decreasing as }|u|\to\infty. \tag{4.14$'$}$$

In fact K_γ equals a function in the space $\mathscr{S}$ plus a Bessel function for which the above estimates are well known (see e.g. [47], pp. 130–134).

Following closely the argument in Section 3 and using the parallel notation $M(P;\sigma,z)$ for $K_\gamma(P,Q,\varepsilon)$, we claim

LEMMA. *The analogues of* (3.24) *and* (3.25) *hold for $M(P;\sigma,z)$. It is understood that in the present context $|M(P;\sigma,z)|$ stands for*

$$\left(\int_0^1 |K_\gamma(P,Q,\varepsilon)|^2\frac{d\varepsilon}{\varepsilon}\right)^{1/2},$$

with similar definitions for $|\partial M/\partial z_j\,(P;\sigma,z)|$ and $|\partial M/\partial\sigma(P;\sigma,z)|$.

Let us consider

$$\int_0^1 |K_\gamma(P,Q,\varepsilon)|^2\frac{d\varepsilon}{\varepsilon}.$$

By (4.13), the fact that $\psi_\varepsilon(z)$ is supported in $|z|<\varepsilon$, and using (4.14) we see that

$$\int_0^1 |K_\gamma(P,Q,\varepsilon)|^2\frac{d\varepsilon}{\varepsilon}\le c\int_{|z|}^{\infty}\varepsilon^{-2n-4}|K_\gamma(\sigma/\varepsilon^2)|^2\frac{d\varepsilon}{\varepsilon}$$
$$=c|\sigma|^{-n-2}J(|z|/\sigma^{1/2}),$$

with

$$J(u) = \int_0^{1/u} \varepsilon^{2n+4} |K_\gamma(\varepsilon^2)|^2 \frac{d\varepsilon}{\varepsilon}.$$

Now $J(u)$ remains bounded as $u \to 0$, since $K_\gamma(u)$ is rapidly decreasing as $u \to \infty$. Moreover $J(u) \leqslant cu^{-2n-4\alpha}$, $\alpha = \mathrm{Re}(\gamma)$, since $|K_\gamma(u)| \leqslant |u|^{-1+\alpha}$. Thus the analogues of (3.24) (a), and (3.25) (a) are proved. The other inequalities are shown similarly, if one uses also (4.14′), and the lemma is proved.

If we then apply Lemma 4 of Section 3, and the theory of singular integrals as used above, one proves that T_γ is bounded on L^p, $1 < p \leqslant 2$, when $\mathrm{Re}(\gamma) > 0$.(¹) Arguing as near the end of Section 3, one can then prove using complex interpolation that $T_0 = T$ is bounded on L^p, $1 < p \leqslant 2$.

This completes the proof of Theorem 4, giving the L^p inequality for the square function.

Corollary 1. *Suppose* $1 < p \leqslant \infty$, *and* $M(f)(t, x) = \sup_{0<\varepsilon<1} |A_\varepsilon(f)(t, x)|$. *Then* $\|M(f)\|_p \leqslant A_p \|f\|_p$.

Proof. Since this result is clear when $p = \infty$, it suffices to prove it for $1 < p \leqslant 2$. We may also assume that $f \geqslant 0$.

Let

$$\tilde{M}(f) = \sup_{1 \geqslant \delta > 0} \frac{1}{\delta} \int_0^\delta A_\varepsilon(f) d\varepsilon.$$

But

$$\frac{1}{\delta} \int_0^\delta A_\varepsilon(f) \, d\varepsilon = \frac{1}{\delta} \int_0^\delta (A_\varepsilon(f) - B_\varepsilon(f)) d\varepsilon + \frac{1}{\delta} \int_0^\delta B_\varepsilon(f) d\varepsilon.$$

The second term on the right is majorized by the usual maximal function associated to the balls corresponding to the metric d given by (3.22). For it the Vitali covering arguments hold (see [8]), therefore

$$\sup_{1 \geqslant \delta > 0} \left| \frac{1}{\delta} \int_0^\delta B_\varepsilon(f) d\varepsilon \right|$$

(¹) Notice that all constants depending on γ are at most of polynomial growth in $\mathrm{Im}(\gamma)$ as long as $\mathrm{Re}(\gamma)$ is restricted to compact subintervals.

is bounded in L^p, $1<p$. However if $0<\delta<1$

$$\left|\frac{1}{\delta}\int_0^\delta A_\varepsilon(f)-B_\varepsilon(f)d\varepsilon\right| \leqslant \left(\int_0^1 |A_\varepsilon(f)-B_\varepsilon(f)|^2\frac{d\varepsilon}{\varepsilon}\right)^{1/2} = S(f).$$

Thus by Theorem 4, $f\to M(f)$ is bounded on L^p, $1<p\leqslant 2$. On the other hand $A_\varepsilon(f)\geqslant cA_\delta(f)$ if $\delta\leqslant\varepsilon\leqslant 2\delta$, so that

$$\frac{1}{2\delta}\int_0^{2\delta} A_\varepsilon(f)d\varepsilon \geqslant \frac{1}{2\delta}\int_\delta^{2\delta} cA_\delta(f)d\varepsilon = c'A_\delta(f).$$

Thus $\sup_{0<\delta<1/2}A_\delta(f)$ is bounded on L^p. Finally, $\sup_{1/2<\delta<1}A_\delta(f)\leqslant cA_1(f)$, and this is easily seen to be bounded on L^p. Therefore the corollary is proved.

COROLLARY 2. *Suppose* $f\in L^p$, $1<p\leqslant\infty$. *Then*

$$\lim_{\varepsilon\to 0} A_\varepsilon(f) = f.$$

This follows from the previous corollary in the usual way.

Proof of Theorem B. With Theorems 3 and 4, the proof of Theorem B is immediate. In fact, the same arguments used to handle R_2 in the proof of Theorem A shows that the operator $M_2 f$ defined by

$$(M_2 f)(P) = \sup_{\alpha\leqslant\delta\leqslant 1}\frac{1}{|B(p,\delta)|}\int_{B(P,\delta)} |f(Q)|\,d\sigma_P(Q)$$

will send $L^p(\Omega)$ to $L^p(\Omega_2)$ for fixed $\alpha>0$. Taking α small enough and using a partition of unity the problem reduces to the study of averages over balls which are contained in coordinate patches of the type introduced in the Corollary of Section 1. We are brought back then to the setting of Theorems 3 and 4 (with a C^∞ extra factor $\gamma(x,t;y,s)$ under each integral which is due to changes of variables; such factors obviously do not affect the arguments there). The desired estimates follow. Q.E.D.

Appendix: Pseudo-differential operators of class $S_{1/2,1/2}$

In this appendix we describe a proposition giving the L^2 boundedness for pseudo-differential operators with operator-valued symbols of the class $S_{1/2,1/2}$, which is used in §§ 3 and 4 above. Part of the complication of the formulation and proof of the result below is due to the fact that we need to be conservative in the degree of smoothness required of our symbols.

First some preliminary definitions. $\mathcal{H}_1$ and $\mathcal{H}_2$ are two separable Hilbert spaces; $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the bounded linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$, with norm $\|\cdot\|$. The space $L^2(\mathbf{R}^n, \mathcal{H})$ consists of the usual square-integrable $\mathcal{H}$-valued functions on $\mathbf{R}^n$. When $0<m_1<1$, and $0<m_2<1$ we define Λ^{m_1, m_2} to be the Banach space of functions a: $\mathbf{R}^n\times\mathbf{R}^n\to\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ so that

$$\|a(x,\xi)\| \leq A, \quad \left\|\Delta_x^{h_1} a(x,\xi)\right\| \leq A|h_1|^{m_1},$$
$$\left\|\Delta_\xi^{h_2} a(x,\xi)\right\| \leq A|h_2|^{m_2} \quad \text{and} \quad \left\|\Delta_x^{h_1}\Delta_\xi^{h_2} a(x,\xi)\right\| \leq A|h_1|^{m_1}|h_2|^{m_2}. \tag{A.1}$$

Here $\Delta_x^{h_1}(b(x,\xi))=b(x+h_1,\xi)-b(x,\xi)$, etc. We take the norm of Δ^{m_1, m_2} to be the smallest A for which the inequalities (A.1) hold. When m_1, m_2 are non-integral positive numbers we extend the definition of Δ^{m_1, m_2} by requiring that

$$\partial_x^\beta \partial_\xi^\alpha a(x,\xi) \in \Delta^{m_l-[m_1], m_2-[m_2]}, \tag{A.2}$$

for all α, β, so that $|\alpha|\leq[m_2]$, $|\beta|\leq[m_1]$ with $[m]$ denoting the largest integer in m. We take the norm of a to be the sum of the norms (in $\Delta^{m_1-[m_1], m_2-[m_2]}$) of $\partial_x^\beta\partial_\xi^\alpha a(x,\xi)$, with $|\beta|\leq[m_2]$, $|\alpha|\leq[m_1]$.

Next with m_1 and m_2 fixed we define the symbol class $S_{0,0}$ to consist of those a, for which $a\in\Delta^{m_1, m_2}$. In addition we define the symbol class $S_{1/2,1/2}$ as follows. Fix scalar-valued functions $\varphi, \psi\in C_0^\infty(\mathbf{R}^n)$, so that $\varphi(\xi)=1$ for $|\xi|$ near 1, $\varphi=0$, and $\psi=1$ near the origin. Then $a\in S_{1/2,1/2}$ if for each $0<u<1$, $a(ux, u^{-1}\xi)\varphi(u\xi)\in S_{0,0}$ uniformly (i.e. uniformly with respect to the Δ^{m_1, m_2} norm) in u, $0<u\leq 1$, and $a(x,\xi)\psi(\xi)\in S_{0,0}$. It is not difficult to see that the definition given for $S_{1/2,1/2}$ is in fact independent of the particular φ and ψ used. We are interested in the operator.

$$(Tf)(x) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n} e^{i\xi\cdot x} a(x,\xi)\hat{f}(\xi)\,d\xi, \tag{A.3}$$

defined for appropriate functions f which take their values in $\mathcal{H}_1$. Then, of course, Tf will take its values in $\mathcal{H}_2$.

PROPOSITION 1. *Suppose* $a\in S_{0,0}$, *with* $m_1>n/2$, $m_2>n$. *Then* T *defined by* (A.3) *extends to a bounded operator from* $L^2(\mathbf{R}^n, \mathcal{H}_1)$ *to* $L^2(\mathbf{R}^n, \mathcal{H}_2)$.

PROPOSITION 2. *Suppose* $a \in S_{1/2,\,1/2}$, *with* $m_1 > n/2$, $m_2 > n$. *Then T defined by* (A.3) *extends to a bounded operator from* $L^2(\mathbf{R}^n, \mathcal{H}_1)$ *to* $L^2(\mathbf{R}^n, \mathcal{H}_2)$.

In the case a is scalar-valued (i.e. $\mathcal{H}_1$ and $\mathcal{H}_2$ are one-dimensional) then a sharper result, requiring only $m_1 > n/2$, $m_2 > n/2$ holds, and this is essentially contained in Coifman and Meyer (see [7], Théorème 7, p. 30). However there is a difficulty in passing from the scalar case to the case of operator-valued symbols. It is this: Plancherel's theorem holds for vector-valued (more precisely, Hilbert-space valued) functions; but it fails for operator-valued functions, (unless one would substitute the Hilbert-Schmidt norm for the operator norm, which is inappropriate here). However given this *caveat* one can follow the broad lines of the argument in [7] to prove the above propositions. This we shall now outline.

For simplicity of presentation we shall restrict ourselves to the one-dimensional case, $n=1$, since anyway this is the case we apply above. The general case requires only slight changes. We can then take $m_1 > 1/2$, and $m_2 > 1$. For any non-negative m we define $L^2_m(\mathbf{R}, \mathcal{H})$ to consist of those $f \in L^2(\mathbf{R}, \mathcal{H})$ for which

$$\left(\int_{\mathbf{R}} (1+|\xi|^2)^m \, |\hat{f}(\xi)|^2 \, d\xi\right)^{1/2} = \|f\|_{L^2_m} < \infty.$$

When $0 < m < 1$ an equivalent norm is given by

$$\left\{\int_{\mathbf{R}} |f(x)|^2 \, dx + \int_{\mathbf{R}}\int_{\mathbf{R}} |f(x-y)-f(x)|^2 \frac{dx\,dy}{|y|^{1+2m}}\right\}^{1/2}. \tag{A.4}$$

(See [47], p. 139–140 in the scalar case; the proof is the same if f takes its values in a Hilbert space.)

LEMMA 1. *Suppose* $x \to a(x)$ *takes its values in* $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ *and* $\|a(x)\| \le A$, $\|a(x+h)-a(x)\| \le A|h|^{m_1}$, $x, h \in \mathbf{R}$. *If* $f \in L^2_m(\mathcal{H}_1)$ *and* $g(x) = a(x)\cdot f(x)$, *the* $g \in L^2_m(\mathcal{H}_2)$, *whenever* $0 < m < 1$ *and* $m < m_1$. *Moreover*, $\|g\|_{L^2_m} \le A' \|f\|_{L^2_m}$.

Proof. According to (A.4) it suffices to estimate

$$\int_{\mathbf{R}} |g(x)|^2 \, dx \quad \text{and} \quad \int_{\mathbf{R}}\int_{\mathbf{R}} |g(x-y)-g(x)|^2 \frac{dx\,dy}{|y|^{1+2m}}.$$

We have $g(x-y)-g(x) = a(x-y)\cdot[f(x-y)-f(x)] + (a(x-y)-a(x))\cdot f(x)$. Thus

$$|g(x-u)-g(x)| \le A|f(x-y)-f(x)| + A \min\left[1, |y|^{m_1}\right] |f(x)|.$$

Also $|g(x)|\leq A|f(x)|$, and so the required estimates follow from the corresponding ones for f, in view of the fact that $m_1>m$.

LEMMA 2. *Suppose $f_k\in L^2_m(\mathscr{H})$, $m>1/2$ and $\Sigma\,\|f_k\|^2_{L^2_m}<\infty$. Then f defined by*

$$f(x)=\sum_k f_k(x)\,e^{ikx} \tag{A.5}$$

belongs to $L^2(\mathscr{H})$, and

$$\|f\|_{L^2(\mathscr{H})}\leq A\left(\sum_k \|f_k\|^2_{L^2_m(\mathscr{H})}\right)^{1/2}.$$

This is the "almost-orthogonality" lemma of [7, p. 13]. The proof in the Hilbert space case is the same as in the scalar-valued case.

LEMMA 3. *Suppose $a\in S_{0,0}$ (i.e. $a\in\Lambda^{m_1,m_2}$) with $m_1>1/2$, $m_2>1$. Also assume that $a(x,\xi)$ is supported in $|\xi|<1$. Then T defined by* (A.3) *maps $L^2(\mathbf{R},\mathscr{H}_1)$* to *$L^2_m(\mathbf{R},\mathscr{H}_2)$, if $m<m_1$; the bound depends only on m and the norm of a.*

We follow in part the proof in [7, p. 17]. Write

$$\tilde{a}(x,\xi)=\sum_k a(x,\xi+2k\pi)=\sum_k a_k(x)\,e^{ik\xi}$$

with $a_k(x)=(1/2\pi)\int_0^{2\pi}\tilde{a}(x,\xi)\,e^{-ik\xi}\,d\xi$. Also $a(x,\xi)=\tilde{a}(x,\xi)\,\varphi(\xi)$ where φ is a suitable C_0^∞ function. Then $T(f)=\tilde{T}(g)$, with $\hat{g}(\xi)=\varphi(\xi)\,\hat{f}(\xi)$, and $\tilde{T}(g)=\Sigma_k\,a_k(x)\cdot g(x+k)$. Clearly g belongs to L^2_m, for every m, and so do the $g(x+k)$, $k\in\mathbf{Z}$ all with the same norm. However

$$ika_k(x)=\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial}{\partial\xi}\tilde{a}(x,\xi)\,e^{-ik\xi}\,d\xi,$$

and

$$ik(e^{ikh}-1)\,a_k(x)=\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\partial}{\partial\xi}\tilde{a}(x,\xi+h)-\frac{\partial}{\partial\xi}\tilde{a}(x,\xi)\right)e^{-ik\xi}\,d\xi.$$

Therefore $|k(e^{ikh}-1)|\,\|a_k(x)\|\leq A|h|^{m_2}$, for all $k\in\mathbf{Z}$, $h\in\mathbf{R}$. It follows that $\|a_k(x)\|\leq A(|k|+1)^{-m_2}$, if we take $h=1/k$. Similarly

$$ik(e^{ikh_2}-1)\,\Delta_x^{h_1} a_k(x) = \frac{1}{2\pi}\int_0^{2\pi} \Delta_x^{h_1}\,\Delta_\xi^{h_2}\left(\frac{\partial}{\partial\xi}a(x,\xi)\right)e^{-ik\xi}\,d\xi$$

from which we see that $\|\Delta_x^{h_1} a_k(x)\| \leq A|h_1|^{m_1}(|k|+1)^{m_2}$, because $a\in\Lambda^{m_1,m_2}$. Since $\Sigma\, 1/(|k|+1)^{m_2}$ converges ($m_2>1$), our conclusion follows from Lemma 1.

Proof of Proposition 1. (See [7], p. 15.) Fix a real C_0^∞ function φ, supported in $|\xi|\leq 1$, so that $\Sigma_k(\varphi(\xi-k))^2=1$. Write $b_k(x,\xi)=a(x,\xi+k)\,\varphi(\xi)$, and T_k the operator with symbol b_k. Define f_k by $\hat f_k(\xi)=\varphi(\xi)\hat f(\xi+k)$. Then

$$(Tf)(x) = \sum T_k(f_k)(x)\,e^{ikx},$$

and the proposition follows from Lemmas 2 and 3, if we choose m so that $1/2<m<m_1$, and observe that

$$\sum \|f_k\|_{L^2}^2 = \sum_k \int |\hat f_k(\xi)|^2\,d\xi = \sum\int (\varphi(\xi-k))^2\,|\hat f(\xi)|^2\,d\xi$$
$$= \|f\|_{L^2}^2.$$

Proposition 2 follows from Proposition 1 exactly as in the argument given in [7, pp. 35–36].

References

[1] Agmon, S., Douglis, A. & Nirenberg, L., Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I. *Comm. Pure Appl. Math.*, 12 (1959), 623–727; II. ibid., 17 (1964), 35–92.

[2] Beals, M., Fefferman, C. & Grossman, R., Strictly pseudo-convex domains. *Bull. Amer. Math. Soc.*, 8 (1983), 125–322.

[3] Boutet de Monvel, L., Comportement d'un opérateur pseudo-differentiel sur une variété à bord I. *J. Analyse Math.*, 17 (1966), 241–253; II. ibid., 16 (1966), 255–304.

[4] Boutet de Monvel, L. & Sjöstrand, J., Sur la singularité des noyaux de Bergman et de Szegö. *Astérisque*, 34–35 (1976), 123–164.

[5] Carleson, L., On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116 (1966), 135–157.

[6] Chiang, C., To appear.

[7] Coifman, R. & Meyer, Y., *Au delà des opérateurs pseudo-différentiels,* 57 (1978).

[8] Coifman, R. & Weiss, G., *Analyse harmonique non-commutative sur certains espaces homogènes*. Lecture Notes in Math., 242 (1971). Springer-Verlag.

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

[10] ESKIN, G., *Boundary value problems for elliptic pseudo-differential equations*. American Mathematical Society, Providence, R.I., 1981.

[11] FEFFERMAN, C., The Bergman kernel and biholomorphic mappings of pseudoconvex domains. Invent. Math., 26 (1974), 1–66.

[12] FEFFERMAN, C. & STEIN, E. M., H^p spaces in several variables. *Acta Math.*, 129 (1972), 137–193.

[13] FOLLAND, G. & KOHN, J. J., *The Neumann problem for the Cauchy-Riemann complex*. Ann. Math. Studies 75 (1972). Princeton University Press.

[14] FOLLAND, G. & STEIN, E. M., Estimates for the $\bar\partial_b$-complex and analysis on the Heisenberg group. *Comm. Pure Appl. Math.*, 27 (1974), 429–522.

[15] — *Hardy spaces on homogeneous groups*. Math. Notes 28 (1982). Princeton University Press.

[16] GELLER, D. & STEIN, E. M., Singular convolution operators on the Heisenberg group. *Bull. Amer. Math. Soc.*, 6 (1982), 99–103.

[17] — Estimates for singular convolution operators on the Heisenberg group. *Math. Ann.*, 267 (1984), 1–15.

[18] GREENLEAF, A., Pointwise convergence of singular convolution operators. To appear.

[19] GREINER, P. & STEIN, E. M., *A parametrix for the $\bar\partial$-Neumann problem*. Math. Notes 19 (1977). Princeton University Press.

[20] GROSSMAN, A., LOUPIAS, G. & STEIN, E. M., An algebra of pseudo-differential operators and quantum mechanics in phase space. *Ann. Institut Fourier*, 18 (1969), 343–368.

[21] GUILLEMIN, V. & STERNBERG, S., *Geometric asymptotics*. Amer. Math. Society, 1977.

[22] GUILLEMIN, V. & UHLMANN, G., Oscillatory integrals and singular symbols. *Duke Math. J.*, 48 (1981), 251–267.

[23] HARVEY, R. & POLKING, J., The $\bar\partial$ Neumann kernel in the ball in $\mathbf{C}^n$. *Trans. Amer. Math. Soc.*

[24] HÖRMANDER, L., *Linear partial differential operators*. Springer-Verlag, 1963.

[25] — Pseudo-differential operators and non-elliptic boundary problems. *Ann. of Math.*, 83 (1966), 129–209.

[26] — Fourier integral operators I. *Acta Math.*, 127 (1971), 79–183.

[27] — Oscillatory integrals and multipliers on FL^p. *Ark. Mat.*, 11 (1973), 1–11.

[28] HOWE, R., Quantum mechanics and partial differential equations. *J. Funct. Anal.*, 38 (1980), 188–254.

[29] LIEB, I. & RANGE, M., On integral representations and a priori estimates. *Math. Ann.*, 265 (1983), 221–251.

[30] MAUCERI, G., PICARDELLO, M. A. & RICCI, F., Twisted convolution, Hardy spaces, and Hörmander multipliers. *Supp. Rend. Circ. Mat. Palermo*, 1 (1981), 191–202.

[31] MELROSE, R. & UHLMANN, G., Lagrangian intersection and the Cauchy problem. *Comm. Pure Appl. Math.*, 32 (1979), 483–519.

[32] — Microlocal structure in involutive conical refraction. *Duke Math. J.*, 46 (1979), 571–582.

[33] MÜLLER, D., Calderón-Zygmund kernels carried by linear subspaces of homogeneous nilpotent Lie algebras, *Invent. Math.*, 73 (1983), 467–489.

[34] — Singular kernels supported by homogeneous submanifolds. *J. Reine Angew. Math.*, 356 (1985), 90–118.

[35] NAGEL, A., STEIN, E. M. & WAINGER, S., Hilbert transforms and maximal functions related to variable curves. *Proc. Symp. Pure Math.*, 35 (1979), 95–98.

[36] PHONG, D. H., On integral representations for the Neumann operator. *Proc. Nat. Acad. Sci. U.S.A.*, 76 (1979), 1554–1558.

[37] PHONG, D. H. & STEIN, E. M., Estimates for the Bergman and Szego projections on strongly pseudoconvex domains. *Duke Math. J.*, 44 (1977), 695–704.

[38] — Some further classes of pseudo-differential and singular integral operators arising in boundary value problems I. *Amer. J. Math.*, 104 (1982), 141–172.
[39] — Singular integrals with kernels of mixed homogeneities. *Conference on harmonic analysis in honor of A. Zygmund*. Eds. Beckner, W., Calderón, A., Fefferman, R., Jones, P., Wadsworth. 327–339 (1982).
[40] — Singular integrals related to the Radon transform and boundary value problems. *Proc. Nat. Acad. Sci. U.S.A.*, 80 (1983), 7697–7701.
[41] REMPEL, S. & SCHULZE, B. W., *Index theory of boundary value problems*. Akademie-Verlag, Berlin, 1983.
[42] RICCI, F., Calderón-Zygmund kernels on nilpotent Lie groups, in *Lecture Notes in Math.*, 908 (1982), 217–227. Springer-Verlag.
[43] SAMPSON, G., Oscillating kernels that map H^1 into L^1. *Ark. Math.*, 18 (1981), 125–144.
[44] SJÖLIN, P., Convergence almost everywhere of certain singular integrals and multiple Fourier series. *Ark. Mat.*, 9 (1971), 65–90.
[45] — Convolution with oscillating kernels. *Indiana Univ. Math. J.*, 30 (1981), 47–56.
[46] STANTON, N. K., The solution of the $\bar{\partial}$-Neumann problem in a strictly pseudoconvex Siegel domain. *Invent. Math.*, 65 (1981), 137–174.
[47] STEIN, E. M., *Singular integrals and differentiability of functions*. Princeton University Press, 1970.
[48] — Singular integrals and estimates for the Cauchy-Riemann equations. *Bull. Amer. Math. Soc.*, 79 (1973), 440–445.
[49] STEIN, E. M. & WAINGER, S., Problems in harmonic analysis related to curvature. *Bull. Amer. Math. Soc.*, 84 (1978), 1239–1295.

Received December 13, 1984
Received in revised form April 16, 1985