An Analogue to a Theorem of Fefferman and Phong for Averaging Operators Along Curves with Singular Fractional Integral Kernel

Michael Greenblatt

1. Introduction

In this paper, we consider an operator T acting on L^2 functions defined near the origin of $\mathbf{R}^n$, $n > 1$, of the form

$$Tf(x) = \int_{\mathbf{R}} f(\gamma(x,t))k(t)\,dt \tag{1.1}$$

Here $\gamma(x,t)$ is a smooth function defined in a neighborhood of the origin in $\mathbf{R}^n \times \mathbf{R}$ satisfying

$$\gamma(x,0) = x,\ \frac{\partial\gamma}{\partial t}(x,t) \neq 0 \tag{1.2}$$

$\gamma(x,t)$ is also assumed to satisfy the curvature condition of [CNSW]. We will have more to say about the curvature condition in section 2. The kernel $k(t)$ is assumed to be a compactly supported function satisfying the following conditions for some sufficiently small δ; how small δ has to be will be determined later:

$$|k| < C|t|^{-1+\delta},\ \ |\frac{dk}{dt}| < C|t|^{-2+\delta} \tag{1.3}$$

Thus $Tf(x)$ can be viewed as averaging the function f over the curve $t \rightarrow \gamma(x,t)$ with respect to the fractional integral kernel $k(t)$. The condition $\gamma(x,0) = x$ is a way of saying the curve at x is "centered at x", and the condition that $\frac{\partial\gamma}{\partial t}(x,t) \neq 0$ ensures that the averaging is smooth. In fact, the arguments of this paper will go through with $k(t)$ replaced by $k(x,t)$ satisfying appropriate x-derivative conditions, as will be described at the end of this paper. However, to simplify the exposition we asssume (1.3). It should be pointed out that recent work of Seeger and Wainger [SW] has also dealt with Radon transforms with fractional integral kernel, proving L^p to L^q estimates under rather different hypotheses.

Our goal will be to prove sharp L^2 regularity estimates for T for sufficiently small $\delta > 0$ that are analogous to a well-known theorem of Fefferman and Phong [FP] in subelliptic PDE's, and will hopefully be useful in that subject. To this end, we suppose L is a linear partial differential operator on a compact manifold M of dimension greater than two which is of the following form in local coordinates:

$$L = -\sum_{i,j=1}^{n} a_{ij}(x)\partial_{x_i}\partial_{x_j} + \sum_{k=1}^{n} b_k(x)\partial_{x_k} + c(x) \tag{1.4}$$

Here the coefficients are real, and the matrix (a_{ij}) is positive semidefinite. L is said to be *subelliptic* if the following is satisfied for some $\epsilon > 0$.

$$||u||_{L^2_{2\epsilon}} < C(||u||_{L^2} + ||Lu||_{L^2}) \tag{1.5}$$

In the paper [FP], Fefferman and Phong define a metric that characterizes the ϵ for which (1.5) holds. A vector $v = (v_1, ..., v_n)$ is said to be *subunit for* L *at* x if we have

$$|v \cdot \xi|^2 \leq \sum_{i,\,j=1}^{n} a_{ij}(x)\,\xi_i\xi_j \text{ for all } \xi \in \mathbf{R}^n \tag{1.6}$$

A path $\gamma : [0, b] \rightarrow M$ is *subunit for* L if $\gamma'(t)$ is subunit for each $t \in [0, b]$. The metric of Fefferman and Phong is defined by

$$d_L(x, y) = \inf\{b : \gamma \text{ is a subunit path for } L,\ \gamma(0) = x,\ \gamma(b) = y\} \tag{1.7}$$

Let $\{B_L(x, r)\}$ denote the balls generated by the distance d_L and let $\{E(x, r)\}$ denote the usual Euclidean balls on M. It is shown in [FP] that (1.5) being true is equivalent to the existence of a constant C and numbers r_0, $\epsilon > 0$ such that for each $x \in M$, each $r < r_0$ we have

$$E(x, r) \subset B_L(x, Cr^\epsilon) \tag{1.8}$$

Hence the supremum of the ϵ that (1.5) holds is equal to the supremum of the ϵ for which (1.8) holds. For a subelliptic operator L, if $Lu = f$ we have the representation

$$u(x) = \int G(x, y) f(y)\, dy \tag{1.9}$$

Here $G(x, y)$ is a singular kernel satisfying the estimates

$$|G(x, y)| < C \frac{d_L(x, y)^2}{|B_L(x, r)|} \tag{1.10}$$

with corresponding estimates on the derivatives of $G(x, y)$ (See [FSa]). As the work of Fabes [Fa] and [P] illustrates, frequently the kernel $G(x, y)$ and its lower-order derivatives can be expressed as an average of operators of the form (1.1) or of singular Radon transforms As a result, sharp L^2 regularity estimates for operators (1.1) can be used to prove regularity results for integral operators (1.9) or their derivatives; this is done in [Fa] and [P].

In the author's paper [G2] on singular Radon transforms, a metric was associated to $\gamma(x, t)$. Suppose $B(x, r)$ denotes a ball of this metric. Then in analogy to (1.8), by [G2] the curvature condition of [CNSW] is equivalent to the existence of a constant C and numbers r_0, $\epsilon > 0$, such that for that for each sufficiently small $|x|$, each $r < r_0$ we have

$$E(x, r) \subset B(x, Cr^\epsilon) \tag{1.11}$$

Let $\frac{1}{k}$ be the supremum of the ϵ for which (1.11) holds in a neighborhood of the origin for some r_0. Our main theorem is as follows.

Theorem 1.1: There is a neighborhood V of the origin in $\mathbf{R}^n$, a neighborhood W of the origin in $\mathbf{R}$, and a constant C such that if $\delta > 0$ is sufficiently small, then for $f(x)$ supported in V and $k(t)$ supported in W the following holds.

$$||Tf||_{L^2_{\frac{\delta}{k}}} < C||f||_{L^2} \tag{1.12}$$

Furthermore, if for sufficiently small $|t|$ we have $|k(t)| > c|t|^{-1+\delta}$ for some constant c then (1.12) is sharp; there is no neighborhood V' of the origin such that there is a constant C' for which for some $\epsilon > \frac{\delta}{k}$ the following holds for all f supported in V':

$$||Tf||_{L^2_\epsilon} < C||f||_{L^2} \tag{1.13}$$

Since δ is required to be quite small, Theorem 1.1 does not directly imply subelliptic estimates using methods such as those in [F] or [P]. For example, if $G(x,y)$ is to be written as the average of operators (1.1), δ is equal to 2, which corresponds to the exponent 2ϵ appearing in (1.5). However, if the operator L or the selfadjoint LL^* can be well-approximated by U^N for some large N, where U can also be written as an average of Radon transforms (1.1) then for U, δ will be $\frac{2}{N}$ and the theorem can be applied. At any rate, Theorem 1.1 can be viewed as analogous to (1.5) for the operators in question. Also, as will be explained at the end of the paper, we also have L^p to $L^p_{\frac{\delta}{k}-\epsilon}$ estimates for p near 2, which might be of interest to people interested in subelliptic equations.

2. The Metric Associated to $\gamma(x,t)$ and its properties

We will now describe the above-mentioned metric that one can associate to $\gamma(x,t)$. Metrics have been used in the study of singular Radon transforms for some time. A singular Radon transform is an operator of the form (1.1) except the δ in (1.3) is taken to be zero and some cancellation conditions on $k(x,t)$ are assumed in the t variable. An early example of associating a metric to a singular Radon transform occurs in [Fa] where the parabolic singular Radon transform $\gamma(x,t) = (x_1 - t, x_2 - t^2)$ is considered and L^2 boundedness is proven in this case. General L^p boundedness was later shown by Nagel, Riviere, and Wainger [NRW1], and L^p boundedness of the associated maximal operator was proved by Stein and Wainger [NRW2]. In fact, [NRW2] shows L^p boundedness of both the singular Radon transform and the associated maximal operator for any finite-type, translation-invariant $\gamma(x,t)$. We refer to [G2] for more details on these situations. A nontranslation-invariant analogue of finite-type, suitable for analyzing Radon transforms, was developed in [CNSW], where this finite-type condition is called "the curvature condition". L^p boundedness for a singular Radon transform is shown in [CNSW] under some conditions on the kernel $k(x,t)$, using the technique of lifting to nilpotent Lie groups. We refer the reader to [CNSW] for a detailed history of the ideas that led up to it, as well as an extensive discussion of the meaning of the curvature condition. In [G2], the author showed how a natural metric could be associated to a $\gamma(x,t)$ satisfying this curvature condition (the curvature condition can actually be defined in terms of the metric), and then used this metric in an analogous (but substantially more intricate) fashion to [Fa], [SW], and other papers to prove L^p boundedness of a singular Radon transform under the curvature condition, again assuming some restrictions on $k(x,t)$.

We now come to the definition of the metric of [G2] (and of this paper). First we define functions $\beta^l(x, t_1, ..., t_l)$ for small $|x|$, $|t_i|$ inductively on $l \geq 2$ as follows. For $l = 2$, we define

$$y = \beta^2(x, t_1, t_2) \iff \gamma(x, t_1) = \gamma(y, t_2) \tag{2.1}$$

For $l > 2$ odd:

$$\beta^l(x, t_1, t_2, ..., t_l) = \gamma(\beta^{l-1}(x, t_1, ..., t_{l-1}), t_l) \tag{2.2}$$

For $l > 2$ even:

$$\beta^l(x, t_1, t_2, ..., t_l) = \beta^2(\beta^{l-2}(x, t_1, t_2, ..., t_{l-2}), t_{l-1}, t_l) \tag{2.3}$$

For x in an appropriately small neighborhood U of the origin, we define a ball $B(x,r)$ of our metric by

$$B(x,r) = \{\beta^l(x, t_1, ..., t_l) : \beta^l(x, t_1, ..., t_l) \in U,\ l > 0,\ |t_1| + ... + |t_l| < r\} \tag{2.4}$$

The balls $B(x,r)$ clearly define a metric, assuming the distance between any two points is finite; this will follow from the curvature condition that we will be assuming. It is also true that under the curvature condition the balls satisfy the axioms of the generalized Calderon-Zygmund theorem of Coifman and Weiss [CW]. In particular we have

Lemma 2.1: There exists an $r_1 > 0$ such that for a constant $M > 0$ the following hold for $|x|$, $r < r_1$.

$$|B(x, 2r)| < M|B(x,r)| \tag{$CZ1$}$$

$$B(x_1, r) \cap B(x_2, r) \neq \emptyset \to B(x_2, r) \subset B(x_1, 2r) \tag{$CZ2$}$$

The equation $(CZ2)$ automatically holds by the definition of the balls $B(x,r)$. We defer the proof of $(CZ1)$ for now, as we develop a different viewpoint. For a sufficiently small $r_2 > 0$ that things are well-defined, we define the manifold M by

$$M = \{(x, \gamma(x,t)) \in U \times U : |t| < r_2\} \tag{2.5}$$

Let $\gamma^*(y,t)$ be defined by

$$\gamma^*(\gamma(x,t), t) = x$$

It is not hard to verify that the adjoint of T is given by

$$T^*f(y) = \int_{\mathbf{R}} f(\gamma^*(y,t))\eta(y,t)k(t)\,dt \tag{2.6}$$

Here $\eta(y,t)$ is the Jacobian of the coordinate change. M can also be expressed as

$$M = \{(\gamma^*(y,t), y) : (\gamma^*(y,t), y) \in U_0\} \tag{2.7}$$

Here U_0 is a subset of $U \times U$ containing the diagonal. There are natural vector fields X and Y on the manifold M. The vector field X at $(x, \gamma(x,t))$ is the image of the vector field dt under the map $t \to (x, \gamma(x,t))$, while the vector field Y at $(\gamma^*(y,t), y)$ is the image of the vector field dt under the map $t \to (\gamma^*(y,t), y)$. So we can write

$$X = (0, \sum_i \frac{d\gamma_i(x,t)}{dt} dy_i),\ Y = (\sum_i \frac{d\gamma_i^*(y,t)}{dt} dx_i, 0) \tag{2.8}$$

The curvature condition used in [CNSW], [G2], and [Se] can be expressed in terms of the vector fields X, Y. Namely, the curvature condition is equivalent to the vector fields satisfying Hormander's condition:

Definition: The operator T satisfies the *curvature condition* at a point x if the vector fields X, Y and their iterated commuators span the tangent space to M at (x, x).

This formulation of the curvature condition is perhaps most similar to that of Seeger [Se]. It is natural to examine the metric one obtains by applying the techniques of [NSW] to M with respect to the vector fields X and Y. Let $D((x,y), r)$ denote the ball of radius r in this metric centered at a point $(x,y) \in M$. By the general theory of such metrics (see Chapter 1 of [Gr] for example), the balls of the metric can be taken to be the following:

$$D((x,y), r) = \{e^{s_{m+1}X} e^{s'_m Y} ... e^{s'_1 Y} e^{s_1 X}(x,y) : m > 0, \sum_i^{m+1} |s_i| + \sum_i^{m} |s'_i| < r\} \tag{2.9}$$

If l is sufficiently large the balls may also be taken to be either of the following:

$$D((x,y), r) = \{e^{s_{l+1}X} e^{s'_l Y} ... e^{s'_1 Y} e^{s_1 X}(x,y) : |s_i|, |s'_i| < r \text{ for all } i\} \tag{2.10a}$$

$$D((x,y), r) = \{e^{s'_l Y} e^{s_l X} ... e^{s'_1 Y} e^{s_1 X}(x,y) : |s_i|, |s'_i| < r \text{ for all } i\} \tag{2.10b}$$

Let π_1 and π_2 be the projections of M onto the first n and last n coordinates respectively (The reader might recognize the similarity here with the setup Helgason uses in work on Radon transforms; see [H] for example). By (2.9) we have that

$$\pi_1(D(x,x), r) = B(x, r) \tag{2.11}$$

We have the following lemma.

Lemma 2.2: There are constants c_1 and c_2 such that

$$D((x,x), c_1 r) \subset \{(y, \gamma(y,t)) : y \in B(x,r),\ |t| < r\} \subset D((x,x), c_2 r) \tag{2.12}$$

In addition, there exist constants C and C' such that

$$C|D((x,x),r)| < r|B(x,r)| < C'|D((x,x),r)| \tag{2.13}$$

Proof: Equation (2.12) follows directly from (2.9). Using the nondegeneracy condition $c' > |\nabla_t \gamma(x,t)| > c$, (2.12) implies that

$$C_1|D((x,x),c_1 r)| < r|B(x,r)| < C_2|D((x,x),c_2 r)| \tag{2.14}$$

By the [NSW] theory applied to the balls $D((x,y),r)$, we have

$$C_3|D((x,x),r)| < |D((x,x),c_1 r)|,\ |D((x,x),c_2 r)| < C_4|D((x,x),r)| \tag{2.15}$$

Combining (2.14) and (2.15) gives (2.13), and the lemma follows.

Corollary: (CZ1) of Lemma 2.1 holds.

Proof: This follows directly from (2.13) and the fact that (CZ1) holds for the balls $D((x,x),r)$.

Note that by (2.10) and (2.12), for sufficiently large l the metric whose balls are $B(x,r)$ is equivalent to the one with balls $B^l(x,r)$ defined by

$$B^l(x,r) = \{\beta^l(x,t_1,...,t_l) : |t_i| < r \text{ for all } i\} \tag{2.16}$$

Lemma 2.3: The supremum of the ϵ for which (1.11) holds in a neighborhood of the origin for some r_0 is of the form $\frac{1}{k}$, where k is an integer.

Proof: By (2.12), Lemma 2.3 will follow if we can prove the analogous result for the balls $D((x,x),r)$ in place of $B(x,r)$. But this follows readily from the general theory of Carnot-Caratheodory metrics; see [NSW] for example. Namely, at the origin there are privileged coordinates $(t_1,...,t_{n+1}) \rightarrow exp(\sum_i t_i r^{k_i} Z_i)$ for $|t_i| < 1$ whose image contains $D((0,0),c_1 r)$ and is contained in $D((0,0),c_2 r)$ for some c_1 and c_2. Each Z_i is an iterated commutator of X and Y, and k_i is the order of this commutator. There is an integer k such that k is the maximum of the k_i for r is sufficiently small. The number k is also the number of commutators it takes to span the whole tangent space to X at the origin. Lemma 2.3 then holds for this value of k and we are done.

Define $\gamma_y(x,t) = \gamma(x+y,t)$. We introduce some notation analogous to (2.1) $-$ (2.3):

$$z = \beta^2_{y_1,y_2}(x,t_1,t_2) \iff \gamma_{y_1}(x,t_1) = \gamma_{y_2}(z,t_2) \tag{2.18}$$

For $l > 2$ odd:

$$\beta^l_{y_1,...,y_l}(x,t_1,t_2,...,t_l) = \gamma_{y_l}(\beta^{l-1}_{y_1,...,y_{l-1}}(x,t_1,...,t_{l-1}),t_l) \tag{2.19}$$

For $l > 2$ even:

$$\beta^l_{y_1,...,y_l}(x,t_1,t_2,...,t_l) = \beta^2_{y_{l-1},y_l}(\beta^{l-2}_{y_1,...,y_{l-2}}(x,t_1,t_2,...,t_{l-2}),t_{l-1},t_l) \tag{2.20}$$

Suppose r is fixed and y is such that $|y| < r^k$. Then so long as x is sufficiently close to the origin, by Lemma 2.3 we have that the distance from x to $x+y$ in the metric (2.4) is at most Cr. As a result, if $|t| < r$, the distance from x to $\gamma(x+y,t)$ in this metric is at most $2r$. Inductively, if each $|t_i| < r$ and each $|y_i| < r^k$, one has that the distance from x to $\beta^n_{y_1,...y_n}(x,t_1,...,t_n)$ is at most $C'r$. Hence we have

$$|\{\beta^n_{y_1,...y_n}(x,t_1,...,t_n) : |t_i| < r\}| < C|B(x,r)| \tag{2.21}$$

Suppose $N \geq n$ is an integer, and I is a subset of $\{1,...,N\}$ of cardinality n which we write as $I = \{m_1,...,m_n\}$ where $m_1 < m_2 < ... < m_n$. We define $det_I\ \beta^N_{y_1,...y_N}(x,t_1,...,t_N)$ to be the determinant of the matrix whose jth column is $\partial_{t_{m_j}}\beta^N_{y_1,...y_N}(x,t_1,...,t_N)$. In other words, $det_I\ \beta^N_{y_1,...y_N}(x,t_1,...,t_N)$ is the Jacobian determinant of $\beta^N_{y_1,...y_N}$ in the $t_{m_1},...,t_{m_n}$ variables.

Next, we define $M_{y_1,...,y_N}(x,r)$ by

$$M_{y_1,...,y_N}(x,r) = \sup_I \sup_{|t_i|<r} |det_I\ \beta^N_{y_1,...y_N}(x,t_1,...,t_N)| \tag{2.22}$$

Since each $det_I\ \beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ is a finite-type function in the t variables for a given $(y_1,...,y_n)$ in a neighborhood of the origin (this was shown in [G2] for example), the image of $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ in the t variables for $|t| < Cr$ has volume comparable to the measure of its domain times its maximum determinant; we write this as

$$|\text{Image } \beta^N_{y_1,...y_N}| \sim M_{y_1,...,y_N}(x,r)r^n \tag{2.23}$$

Since this image is contained in the ball $B(x,Cr)$ for some C we have

$$M_{y_1,...,y_N}(x,r)r^n < C'|B(x,Cr)| < C''|B(x,r)| \tag{2.24}$$

In addition, by [G2] we have

$$M_{0,...,0}(x,r)r^n > C'''|B(x,r)| \tag{2.25}$$

One may also show this by lifting to the ball $B((x,x),r)$ and examining the geometry of the map

$$(t_1,...,t_{N+1}) \rightarrow e^{t_{N+1}Y}....e^{t_2Y}e^{t_1X}(x,x)$$

if N is odd, or the map

$$(t_1,...,t_{N+1}) \rightarrow e^{t_{N+1}X}....e^{t_2Y}e^{t_1X}(x,x)$$

if N is even.

We can go further and say the following:

Lemma 2.4: There exist $c_N, C_1, C_2 > 0$ such that whenever each y_i satisfies $|y_i| < c_N r^k$, we have

$$C_1|B(x,r)| < M_{y_1,...,y_N}(x,r)r^n < C_2|B(x,r)| \tag{2.26}$$

Proof: The right-hand inequality follows from (2.24), so we must prove the left-hand inequality. Each $det_I\, \beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ is a finite-type function. As a result, by the Bernstein inequalities of Lemma 3.2 applied to the set $\{(y_1, ..., y_N, t_1, ..., t_N) : |y_i| < r^k, |t_i| < r$ for all $i\}$, for each I we have the following.

$$\sup_{|y_i|<r^k, |t_i|<r} |\nabla_y det_I\, \beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)| < \frac{C}{r^k} \sup_{|y_i|<r^k, |t_i|<r} |det_I\, \beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)| \tag{2.27}$$

$$< \frac{C}{r^{k+n}}|B(x,r)| \tag{2.28}$$

The last inequality follows from the right-hand side of (2.26). By (2.25), there exists an I and a $(t_1', ..., t_N')$ with $|t_i'| < r$ for all i such that

$$|det_I\ \beta^N_{0,...,0}(x, t_1', ..., t_N')| > \frac{C}{r^n}|B(x,r)| \tag{2.29}$$

By (2.28) and (2.29), for some c_N, whenever y is such that $|y_i| < c_N r^k$ for each i, we have

$$|det_I\ \beta^N_{y_1,...,y_N}(x, t_1', ..., t_N')| > \frac{C'}{r^n}|B(x,r)| \tag{2.30}$$

Since $M_{y_1,...,y_N}(x,r) \geq |det_I\ \beta^N_{y_1,...,y_N}(x, t_1', ..., t_N')|$, the left-hand side of (2.26) follows and we are done.

We now fix (x,r) and consider privileged coordinates on the ball $D((x,x),r)$ of our Carnot-Caratheodory space. Namely, we take higher order commutators of X and Y which we denote by $Z_1, ..., Z_{n+1}$, where Z_i is of order m_i, such that the image of $\{(t_1, ..., t_{n+1}) : |t_i| < r\}$ under the following map contains a ball $D((x,x), c_1 r)$ and is contained in a ball $D((x,x), c_2 r)$:

$$(t_1, ..., t_{n+1}) \rightarrow exp(t_1 r^{m_1 - 1} Z_1 + ... + t_{n+1} r^{m_{n+1}-1} Z_{n+1}) \tag{2.31}$$

The map (2.31) can also be used to give canonical coordinates for the balls $B(x,r)$. If n is even, we may take $Z_{n+1} = X$, and project $r^{m_1-1}Z_1, ..., r^{m_n-1}Z_n$ at (x,x) onto the left n coordinates to obtain vector fields $W_1, ..., W_n$ such that following map gives canonical coordinates on $B(x,r)$ for $|t_i| < r$:

$$(t_1, ..., t_n) \rightarrow exp(t_1 W_1 + ... + t_n W_n) \tag{2.32}$$

If n is odd, we can take $Z_{n+1} = Y$ and project $r^{m_1-1}Z_1$,..., $r^{m_n-1}Z_n$ at (x,x) onto the right n coordinates to obtain vector fields W_1,...,W_n such that (2.32) gives a canonical coordinate system for $B(x,r)$.

Lemma 2.5: Suppose y_1,...,y_N are such that $|y_i| < c_N r^k$ for all i, where c_N is as in Lemma 2.4, and suppose I, $(t_1',...,t_N')$ and $\epsilon < 1$ are such that $|det_I\, \beta^N_{y_1,...,y_N}(x,t_1',...,t_N')| = \epsilon M_{y_1,...,y_N}(x,r)$.

Then the set $\{\sum_{i\in I} c_i \partial_{t_i}\beta^N_{y_1,...,y_N}(x,t_1',...,t_N') : |c_i| < 1\}$ contains a ball centered at the origin of radius $C\epsilon r^{k-1}$, where C is a uniform constant and k is as in Lemma 2.3.

Proof: Consider the map $\Phi(s_1,...,s_n,t_1,...,t_N)$ defined by

$$\Phi(s_1,...,s_n,t_1,...,t_N) = exp(s_1W_1 + ... + s_nW_n)\beta^N_{y_1,...,y_N}(x,t_1,...,t_N) \tag{2.33}$$

So viewed as a function of $s_1,...,s_n$, the map $\Phi(s_1,...,s_n,t_1,...,t_N)$ is the exponential flow along $s_1W_1 + ... + s_nW_n$ starting at $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$. Replacing the coordinates on $B(x,r)$ with those on $B(x,Cr)$ for some large C if necessary, we may assume that $\Phi(s_1,...,s_n,t_1,...,t_N)$ is defined for all $|s_i|,|t_i| < r$. Define $P_{y_1,...,y_N}(x,r)$ by

$$P_{y_1,...,y_N}(x,r) = \sup_{J,s,t} |det_J \Phi(s_1,...,s_n,t_1,...,t_N)| \tag{2.34}$$

Here J is a subset of $\{\{1,...,n\},\{1,...,N\}\}$ of cardinality n, and det_J denotes the determinant in the corresponding subset of the $s_1,...,s_n,t_1,...,t_N$ variables. The s,t supremum is taken over all $|s_i| < r$, $|t_i| < r$. Then since the image of $\Phi(s_1,...,s_n,t_1,...,t_N)$ is contained in the ball $B(x,Cr)$ for some C, analogous to (2.24) we have

$$P_{y_1,...,y_N}(x,r)r^n < C'B(x,r) \tag{2.35}$$

Next, since $W_1,...,W_n$ are canonical coordinates we also have

$$|det_{s_1,...,s_n}\Phi(s_1,...,s_n,t_1,...,t_N)|r^n > C''B(x,r) \tag{2.36}$$

Hence we may conclude that for some constants C_1 and C_2 we have

$$C_1|B(x,r)| < P_{y_1,...,y_N}(x,r)r^n < C_2|B(x,r)| \tag{2.37}$$

Equations (2.36) – (2.37) also imply that for each i and each s_1,...,s_n, t_1,...,t_N we have that $\partial_{t_i}\Phi(s_1,...,s_n,\, t_1,...,t_N)$ can be written in the following form, where $|c_1|$,...,$|c_n|$ are bounded by a uniform constant:

$$\partial_{t_i}\Phi(s_1,...,s_n,t_1,...,t_N) = \sum_j c_j\partial_{s_j}\Phi(s_1,...,s_n,t_1,...,t_N) \tag{2.38}$$

For if (2.38) did not hold, thenin view of (2.36) the vectors $\partial_{t_i}\Phi(s_1,...,s_n,t_1,...,t_N)$ and $\{\partial_{s_j}\Phi(s_1,...,s_n,t_1,...,t_N)\}$ would generate too large a parallelopiped for the right-hand side of (2.37) to be satisfied. In particular, (2.38) holds when s_1,...s_n are all zero. In this case (2.38) translates to

$$\partial_{t_i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N) = \sum_j c_j W_j \tag{2.39}$$

Here the vector fields W_j are evaluated at the point $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$. Since (2.39) holds for each i, including i in I, we have for some constant C:

$$\{\sum_{i\in I} c_i\partial_{t_i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N) : |c_i| < 1\} \subset \{\sum_j c_j W_j : |c_j| < C\} \tag{2.40}$$

By Lemma 2.4 and the fact that the W_i are coordinates for $B(x,r)$, the volume of the right-hand side of (2.40) is comparable to $M_{y_1,...,y_N}(x,r)$. Hence if $(t_1',...,t_N')$ satisfies $|det\,\beta^N_{y_1,...,y_N}(x,t_1',...,t_N')| = \epsilon M_{y_1,...,y_N}(x,r)$, then the measure of the left-hand side of (2.40) is at least $C\epsilon$ times that of the right- hand side. By Lemma 2.3 and the fact that the W_i are canonical coordinates, this right-hand side contains a ball of radius $C'r^{k-1}$. We conclude that the left hand side of (2.40) contains a ball of radius $C''\epsilon r^{k-1}$ and we are done.

Corollary 2.6: Suppose $(t_1,...,t_N)$ satisfies $|t_i| < r$ for all i, and $(y_1,...,y_N)$ is such that $|y_i| < c_N r^k$ for each i, where c_N is as before. Suppose $I \subset \{1,...,N\}$ is of cardinality n such that

$$|det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)| = \epsilon M_{y_1,...,y_N}(x,r) \tag{2.41}$$

If $v = (v_1,...,v_n)$ satisfies $|v| = 1$ and $det_I^{v,j}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ denotes the determinant of the matrix obtained by replacing the jth column of the Jacobian matrix of $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ in the I variables by the vector v, then we have

$$\frac{|det_I^{v,j}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)|}{|det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)|} < \frac{C}{\epsilon r^{k-1}} \tag{2.42}$$

Proof: Since $|v| = 1$, by Lemma 2.5 we can find numbers a_1,...,a_n with $|a_i| < \frac{C}{\epsilon r^{k-1}}$ for all i such that

$$v = \sum_{i\in I} a_i\partial_{t_i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N) \tag{2.43}$$

The determinant of the matrix obtained by replacing the jth column of the Jacobian in I of $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ by v is exactly the determinant of the matrix obtained by replacing the jth column of the Jacobian in I of $\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ by $a_j\partial_{t_j}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$, since the other terms lead to determinants of matrices with two equal columns. Hence the determinant in question is exactly a_j times $det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$. As $|a_j| < \frac{C}{\epsilon r^{k-1}}$, the corollary follows.

3. Beginning of the Proof; Analytic Lemmas

The general outline of the proof is remniscent of that of [G2], whose origins trace back to [PS]. The biggest difference is that since we are proving L^2 regularity here, one does not need to associate a singular integral operator to T as in [G2]. Instead one needs to consider certain convolutions of bump functions with the operators called T_j below; the arguments become somewhat different, in particular in the almost-orthogonality arguments. Let $\eta(t)$ be a compactly supported function on $\mathbf{R}$, equal to 0 on a neighborhood of the origin that satisfies $\sum_{j=-\infty}^{\infty} \eta(2^j t) = 1$. Denote $\eta(2^j t)k(t)$ by $k_j(t)$, and define the operator T_j by

$$T_j f(x) = \int_{\mathbf{R}} f(\gamma(x,t))k_j(t)\,dt \tag{3.1}$$

Thus we have

$$\sum_j T_j = T \tag{3.2}$$

Let $\zeta(x)$ be a Schwarz function $\mathbf{R}^n$ whose Fourier transform is compactly supported and equal to 1 on a neighborhood of the origin, and let $2^{mn}\zeta(2^m x)$ be denoted by $\zeta_m(x)$. Let k be the integer of Lemma 2.3. We define operators $\bar{T}_j$ and T_{ij} by

$$\bar{T}_j f = \zeta_{jk} * T_j f,\ T_{ij} f = \zeta_{jk+i+1} * T_j f - \zeta_{jk+i} * T_j f \tag{3.3}$$

We let $T^i = \sum_j T_{ij}$ and $\bar{T} = \sum_j \bar{T}_j$. We have

$$\bar{T} + \sum_{i \geq 0} T^i = T \tag{3.4}$$

In order to prove Theorem 1.1, it suffices to show that the operator $U = \triangle^{\frac{\delta}{2k}} T$ is bounded on L^2. To this end, we decompose U analogously to (3.4). Namely, we let $\xi = \triangle^{\frac{\delta}{2k}} \zeta$ and $\xi_m = \triangle^{\frac{\delta}{2k}} \zeta_m$, and we define the following operators.

$$\bar{U}_j f = \xi_{jk} * T_j f,\ U_{ij} f = (\xi_{jk+i+1} - \xi_{jk+i}) * T_j f$$

$$U^i = \sum_j U_{ij},\ \bar{U} = \sum_j \bar{U}_j$$

Hence we have

$$\bar{U} + \sum_{i \geq 0} U^i = U \tag{3.5}$$

Observe that because $\triangle^{\frac{\delta}{2k}}$ has Fourier multiplier $|y|^{\frac{\delta}{k}}$ and ζ's Fourier transform is smooth and compactly supported, we have that $\xi = \triangle^{\frac{\delta}{2k}} \zeta$ is smooth with

$$\xi(x) \sim |x|^{-n-\frac{\delta}{k}} \text{ as } |x| \to \infty \tag{3.6}$$

Furthermore, since $\zeta_m(x)$ has Fourier transform $\hat{\zeta}(2^{-m}x)$, we have

$$\xi_m(x) = 2^{m(n+\frac{\delta}{k})}\xi(2^m x)$$

As a result,

$$|\xi_m(x)| < C\frac{2^{m(n+\frac{\delta}{k})}}{1+(2^m|x|)^{n+\frac{\delta}{k}}},\ |\nabla\xi_m(x)| < C\frac{2^{m(n+1+\frac{\delta}{k})}}{1+(2^m|x|)^{n+1+\frac{\delta}{k}}} \tag{3.7}$$

In addition, since the Fourier transform of $\xi_{m+1}(x) - \xi_m(x)$ is equal to $|x|^{\frac{\delta}{k}}[\zeta(2^{-m-1}x) - \zeta(2^{-m}x)]$, a smooth function, we also have the following estimates for any N:

$$|\xi_{m+1}(x) - \xi_m(x)| < C_N\frac{2^{m(n+\frac{\delta}{k})}}{1+(2^m|x|)^N},\ |\nabla\xi_{m+1}(x) - \nabla\xi_m(x)| < C_N\frac{2^{m(n+1+\frac{\delta}{k})}}{1+(2^m|x|)^N} \tag{3.8}$$

To ensure our operators will all be well-defined, it is appropriate to truncate the $\xi_m(x)$ to functions $\chi(x)\xi_m(x)$ for a function $\chi(x)$ with appropriately small support. This will not affect the validity of our estimates; by (3.7)$-$(3.8) for any C we have $\sum_m \int_{|x|>C} |\xi_{m+1}(x) - \xi_m(x)|\,dx < C'$, so by Schur's test $\sum_{i,j} ||U_{ij} - \text{truncation of } U_{ij}||_{L^2\to L^2} < C''$. Hence in our future arguments we may assume $\xi_m(x)$ has been replaced by $\chi(x)\xi_m(x)$. Note that (3.7) and (3.8) still hold.

Observe that $\xi_m(x)$ satsifies the following cancellation condition:

$$\int \xi_m(x) < C \tag{3.9}$$

In other words, the cancellation gives us a gain of $2^{-\frac{m\delta}{k}}$.

In order to prove Theorem 1.1, we will prove the following estimates, where ν_1, ν_2, and ν_3 are positive.

$$||U_{ij_1}U^*_{ij_2}||_{L^2\to L^2},\ ||U^*_{ij_1}U_{ij_2}||_{L^2\to L^2} < C2^{\nu_1 i}2^{-\nu_2|j_1-j_2|} \tag{3.11}$$

$$||\bar{U}_{j_1}\bar{U}^*_{j_2}||_{L^2\to L^2},\ ||\bar{U}^*_{j_1}\bar{U}_{j_2}||_{L^2\to L^2} < C2^{-\nu_2|j_1-j_2|} \tag{3.12}$$

$$||U_{ij}||_{L^2\to L^2} < C2^{-\nu_3 i} \tag{3.13}$$

$$||\bar{U}_j||_{L^2\to L^2} < C \tag{3.14}$$

Equations (3.11) – (3.14) imply Theorem 1.1 in the following way. By (3.11) and (3.13), for any small $\beta > 0$ we have

$$||U_{ij_1}U^*_{ij_2}||_{L^2\to L^2},\ ||U^*_{ij_1}U_{ij_2}||_{L^2\to L^2} < C(2^{\nu_1 i}2^{-\nu_2|j_1-j_2|})^\beta(2^{-\nu_3 i})^{1-\beta} \tag{3.15}$$

Taking β sufficiently small in (3.15) gives the following for some $\nu_4, \nu_5 > 0$:

$$||U_{ij_1} U^*_{ij_2}||_{L^2 \rightarrow L^2},\ ||U^*_{ij_1} U_{ij_2}||_{L^2 \rightarrow L^2} < C 2^{-\nu_4 i} 2^{-\nu_5 |j_1 - j_2|} \tag{3.16}$$

Combining (3.13) and (3.16), Cotlar-Stein almost-orthogonality gives

$$||U^i||_{L^2 \rightarrow L^2} < C 2^{-\nu_6 i} \tag{3.17}$$

Similarly by almost-orthogonality, (3.12) implies

$$||\bar{U}||_{L^2 \rightarrow L^2} < C \tag{3.18}$$

In view of (3.10), equations (3.17) and (3.18) imply

$$||U||_{L^2 \rightarrow L^2} < C \tag{3.19}$$

This gives Theorem 1.1.

We now give a couple of analytic lemmas that will be important in proving Theorem 1.1. The first is a standard lemma concerning the growth of finite-type functions. It for example appears in [G2] (Lemma 3.2a).

Lemma 3.1: Suppose f is a C^∞ function on the closure of a domain U such that there is an $\epsilon > 0$, a direction v, and a positive integer a such that on U we have

$$\left|\frac{\partial^a f}{\partial v^a}\right| > \epsilon$$

Let B be an n-dimensional cube in U with side lengths equal to r. Let M denote the maximum of f on B. If r is sufficiently small, depending on ϵ, a, n, and the C^{a+1} norm of f on U, there is a constant C_0 such that for each $1 > \delta > 0$ we have

$$|\{x \in B : |f(x)| < \delta\, M\}| < C_0 \delta^{\frac{1}{a}} r^n$$

The constant C_0 also depends on ϵ, a, n, and the C^{a+1} norm of f on U.

The next lemma contains the Bernstein-type inequalities needed for this paper. This is the m-dimensional generalization of the two-dimensional Lemma 3.5 in [G3]. The proof is essentially identical, so it is not included here. It is conceivable that it has also appeared somewhere else before.

Lemma 3.2: Let $f(x)$ be a smooth real-valued function defined in a neighborhood of the origin in R^m for some $m \geq 2$, and let α be a multiindex such that

$$\frac{\partial^\alpha f}{\partial x^\alpha}(0, ..., 0) = \mu \neq 0 \tag{3.20}$$

Suppose that $\epsilon > 0$. There is a neighborhood U of the origin such that if $R \subset U$ is an $r_1 \times ... \times r_m$ rectangle with $r_i \leq r_j^\epsilon$ for each i and j, then for any multiindex β we have

$$\sup_{x \in R} |\frac{\partial^\beta f}{\partial x^\beta}(x)| < C r_1^{-\beta_1} ... r_m^{-\beta_m} \sup_{x \in R} |f(x)| \tag{3.21}$$

Furthermore, if S is a subrectangle of R of dimensions $\frac{r_1}{2} \times ... \times \frac{r_m}{2}$, then we also have

$$\sup_{x \in R} |f(x)| < C \sup_{x \in S} |f(x)| \tag{3.22}$$

The constants in (3.21) and (3.22) and the neighborhood U depend on α, β, ϵ, μ and the suprema of the absolute value of finitely many derivatives of $f(x)$ on U.

The following lemma provides a local Whitney-type decomposition of the complement of the zero set of a finite-type function into disjoint squares of approximately maximal size on each of which the function is within a factor of two of a fixed value. This will be necessary for the detailed analysis of section 5; this will be discussed further after (5.35).

Lemma 3.3: Let $f(x)$ be a smooth real-valued function defined in a neighborhood of the origin in R^m for some $m \geq 1$, and let α be a multiindex such that

$$\frac{\partial^\alpha f}{\partial x^\alpha}(0, ..., 0) = \mu \neq 0 \tag{3.23}$$

There is a cube Q centered at the origin with axes parallel to the coordinate axes, with side length 2^{-l_Q} for some integer l_Q such that one of the following two possibilities holds:

1) There is a number A_Q such that on Q we have

$$A_Q \leq |f| \leq 3A_Q \tag{3.23}$$

2) There is a collection B of dyadic subcubes of Q with disjoint interiors such that the union of the cubes in B is $\{x \in Q : f(x) \neq 0\}$. B has the property that there exists a $\delta > 0$ such that for each cube R in B, if 2^{-l_R} denotes the side length of R and R' denotes the concentric dilate of R with $1 + \delta$ times the side length, then the following hold:

i) There exist constants C_1 and C_2 such that

$$\frac{C_1}{2^{-l_R}} \sup_R |f| < \sup_R |\nabla f| < \frac{C_2}{2^{-l_R}} \sup_R |f| \tag{3.24}$$

ii): There exists a number A_R such that on R' we have:

$$A_R \leq |f| \leq 6A_R \tag{3.25}$$

iii) Suppose R_1 and R_2 are cubes in B with $R_1' \cap R_2' \neq \emptyset$. There exists a constant C_3 such that

$$|l_{R_1} - l_{R_2}| < C_3 \tag{3.26}$$

iv) There exists a partition of unity $\{\phi_R : R \in B\}$ such that each ϕ_R is nonnegative and supported in R', such that on Q we have

$$\sum_R \phi_R = 1, \tag{3.27}$$

Furthermore, there is a constant C_4 such that

$$|\nabla \phi_R| < \frac{C_4}{2^{-l_R}} \tag{3.28}$$

The constants δ, C_1, C_2, C_3, and C_4, as well as the cube Q, depend on α, μ, m and the suprema of the absolute value of finitely many derivatives of $f(x)$ on U.

Proof: Let Q be a cube centered at the origin with axes parallel to the coordinate axes, with side length 2^{-l_Q} for some integer l_Q, such that Q is small enough that $|\frac{\partial^\alpha f}{\partial x^\alpha}| > \frac{\mu}{2}$ on Q and the conclusions of Lemma 3.2 hold for $\epsilon = 1$ on Q. Let C_2 be small enough so that if $\sup_Q |\nabla f| < \frac{C_2}{2^{-l_Q}} \sup_Q |f|$, then (3.23) holds. In this case, **1)** of the lemma automatically holds and there is nothing to prove. So we assume that $\sup_Q |\nabla f| \geq \frac{C_2}{2^{-l_Q}} \sup_Q |f|$.

We proceed inductively as follows. We first divide Q into 2^m dyadic cubes of side length $2^{-(l_Q+1)}$. For each cube R amongst these, if $\sup_R |\nabla f| < \frac{C_2}{2^{-(l_Q+1)}} \sup_R |f|$ then we add R to the collection B. If R satisfies $\sup_R |\nabla f| \geq \frac{C_2}{2^{-(l_Q+1)}} \sup_R |f|$, then we do not add R to the collection B. Instead, we divide R into 2^m subcubes of diameter $2^{-(l_Q+2)}$. For each cube S amongst these, we check if $\sup_S |\nabla f| < \frac{C_2}{2^{-(l_Q+2)}} \sup_S |f|$ holds. If it does, we add S to B. Otherwise, we divide S into 2^m more subcubes, and proceed analogously, ad infinitum.

By construction, the cubes in B are disjoint and the right-hand side of (3.24) holds. Furthermore, if R is a cube in B and y and z are any points in R, if C_2 were chosen sufficiently small the mean-value theorem and the right-hand side of (3.24) imply that

$$|f(y) - f(z)| < 2^{-l_R}\sqrt{m} \sup_R |\nabla f| < \frac{1}{2} \sup_R |f| \tag{3.29}$$

In particular, we may take y to be the point for which $|f(y)| = \sup_R |f|$, in which case (3.29) says that for all z in R we have

$$\frac{1}{2}|f(y)| < |f(z)| \leq |f(y)| \tag{3.30}$$

Next, since we are not in case **1**) of this lemma, each cube R of B was derived from a cube S of side length 2^{-l_R+1} from the previous stage which was not chosen to be a member of B. Since S was not chosen we must have

$$\sup_S |\nabla f| \geq \frac{C_2}{2^{-l_R+1}} \sup_S |f| \tag{3.31}$$

Therefore, by (3.22) of Lemma 3.2 there exists a C such that

$$\begin{aligned}\sup_R |\nabla f| &\geq \frac{C}{2^{-l_R+1}} \sup_S |f| \\ &\geq \frac{C}{2^{-l_R+1}} \sup_R |f|\end{aligned}$$

Setting $C_1 = \frac{C}{2}$ gives the left-hand side of (3.24). This gives conclusion i) of the lemma.

Next, we move to proving ii). Suppose $R \in B$. By (3.21) and (3.22) of Lemma 3.2, there is a constant C such that

$$\sup_{R'} |\nabla f| < \frac{C}{2^{-l_{R'}}} \sup_{R'} |f| < \frac{C'}{2^{-l_R}} \sup_R |f| \tag{3.32}$$

Hence if y is any point in R', and z is the point in R nearest to y, then $|y-z| \leq \sqrt{m}\delta 2^{-l_R}$ and we have

$$\begin{aligned}|f(y) - f(z)| &\leq |y-z| \sup_{R'} |\nabla f| \leq \sqrt{m}\delta 2^{-l_R} \times \frac{C}{2^{-l_R}} \sup_R |f| \\ &= C'\delta \sup_R |f|\end{aligned}$$

As a result, as long as δ is small enough so that $C'\delta < \frac{1}{6}$, we have

$$|f(y) - f(z)| < \frac{1}{6} \sup_R |f| \tag{3.33}$$

By (3.30), on R the function f is always between $\frac{1}{2}\sup_R |f|$ and $\sup_R |f|$. In view of this fact and equation (3.33), we conclude there exists an A_R such that on R' we have

$$A_R \leq |f| \leq 6A_R \tag{3.34}$$

This gives ii) of this lemma.

Moving now to proving iii), suppose R_1 and R_2 are cubes in B such that $R_1' \cap R_2' \neq \emptyset$. Since the cubes R_1' and R_2' intersect, by (3.34) we have

$$\frac{1}{36} \leq \frac{A_{R_1}}{A_{R_2}} \leq 36 \tag{3.35}$$

Now suppose that $l_{R_1} \leq l_{R_2}$, the case where $l_{R_1} \geq l_{R_2}$ is done in the same way. Since R_1' and R_2' have nonempty intersection, the smaller cube R_2' is contained in the dilate of R_1' by a factor of 4. Consequently, by two applications of equation (3.22) of Lemma 3.2 to the first partial derivatives of f, we have

$$\sup_{R_2'} |\nabla f| < C \sup_{R_1'} |\nabla f| \tag{3.36}$$

(In the case that one or more of the first partials of f has a zero of infinite order at the origin, we have to do a little more work. If they all have a zero of infinite order, so that $\alpha = 0$ in (3.23), then case **1)** of this lemma automatically holds. If only some of these first partials have a zero of infinite order at the origin, instead of applying (3.22) to the first partials of f, we apply it to combinations $\sum c_i \partial_i f$ which do not have zeroes of infinite order at the origin.)

Next, by (3.21) of Lemma 3.2 and (3.34) we have

$$\sup_{R_1'} |\nabla f| < \frac{C}{2^{-l_{R_1'}}} \sup_{R'} |f| \leq \frac{6C}{2^{-l_{R_1}}} A_{R_1} \tag{3.37}$$

On the other hand by the left hand side of (3.24), we have

$$\sup_{R_2'} |\nabla f| \geq \sup_{R_2} |\nabla f| > \frac{C_1}{2^{-l_{R_2}}} A_{R_2} \tag{3.38}$$

Consequently, (3.36) – (3.38) imply that

$$\frac{1}{2^{-l_{R_1}}} A_{R_1} \geq \frac{C'}{2^{-l_{R_2}}} A_{R_2} \tag{3.39}$$

Since A_{R_1} and A_{R_2} are within a factor of 36 of one another by (3.35), we conclude that

$$2^{-l_{R_2}} > \frac{C'}{36} 2^{-l_{R_1}} \tag{3.40}$$

Consequently,

$$l_{R_2} < l_{R_1} + C'' \tag{3.41}$$

Since we are assuming $l_{R_1} \leq l_{R_2}$, we have (3.26). This completes the proof of iii).

Moving to iv), let $\rho(x)$ be a nonnegative smooth function which is 1 on the cube centered at the origin of side length 1, and supported in the cube centered at the origin of side length $1 + \delta$. For each R in B, let c_R denote the center of R and define $\psi_R(x) = \rho(2^{l_R}(x - c_R))$. We define $\phi_R(x)$ by

$$\phi_R(x) = \frac{\psi_R(x)}{\sum_{S \in B} \psi_S(x)} \tag{3.42}$$

Then $\phi_R(x)$ is nonnegative and supported on R', and $\sum_R \phi_R = 1$. Suppose x_0 is such that $\phi_R(x_0) \neq 0$. By iii), if S is in the sum of (3.42) and $\psi_S(x_0) \neq 0$, then the side length of S is within a factor of C of the side length of R and there can only be boundedly many such S. Equation (3.28) is now seen to hold and we are done.

4. Almost-Orthogonality

In this section, we will prove (3.11) and (3.12). Let V_{ij} be the operator defined by

$$V_{ij}f = \xi_{jk+i} * T_j f \tag{4.1}$$

By the definitions (3.5) of U_{ij} and $\bar{U}_{ij}$, equations (3.11) and (3.12) will follow if we can show that for $|i_1 - i_2| \leq 1$ the following holds, where $i = \max(i_1, i_2)$:

$$||V_{i_1j_1}V^*_{i_2j_2}||_{L^2\to L^2},\ ||V^*_{i_1j_1}V_{i_2j_2}||_{L^2\to L^2} < C2^{\nu_1 i}2^{-\nu_2|j_1-j_2|} \tag{4.2}$$

Observe that $V_{ij} = \int_t V^t_{ij}k_j(t)\,dt$, where

$$V^t_{ij}f(x) = \int f(\gamma(y,t))\xi_{jk+i}(x-y)\,dy \tag{4.3}$$

Hence, by the bounds (1.3), in order to prove (4.2) it suffices to show that for all t_1 and t_2 that

$$||V^{t_1}_{i_1j_1}{V^{t_2}_{i_2j_2}}^*||_{L^2\to L^2},\ ||{V^{t_1}_{i_1j_1}}^*V^{t_2}_{i_2j_2}||_{L^2\to L^2} < C2^{\delta(j_1+j_2)}2^{\nu_1 i}2^{-\nu_2|j_1-j_2|} \tag{4.4}$$

We examine the kernels of $V^{t_1}_{ij}$ and ${V^{t_2}_{ij}}^*$. Changing coordinates in (4.3) via $y_{old} = \gamma^*(y_{new}, t)$, we get

$$V^t_{ij}f(x) = \int f(y)\xi_{jk+i}(x-\gamma^*(y,t))\psi(y,t)dy \tag{4.5}$$

Here $\psi(y,t)$ is the Jacobian of the coordinate change. As a result, the kernel of V^t_{ij} is given by

$$K^t_{ij}(x,y) = \xi_{jk+i}(x-\gamma^*(y,t))\psi(y,t) \tag{4.6}$$

By (3.7), we have the bounds

$$|K^t_{ij}(x,y)| < \frac{C2^{(jk+i)(n+\frac{\delta}{k})}}{1+(2^{jk+i}|x-\gamma^*(y,t)|)^{n+\frac{\delta}{k}}} \tag{4.7a}$$

$$|\nabla_x K_{ij}(x,y)|,\ |\nabla_y K_{ij}(x,y)| < \frac{C2^{(jk+i)(n+1+\frac{\delta}{k})}}{1+(2^{jk+i}|x-\gamma^*(y,t)|)^{n+\frac{\delta}{k}}} \tag{4.7b}$$

Since the function $y \to x - \gamma^{-1}(y,t)$ has Jacobian nearly the identity matrix and is equal to zero for $y = \gamma(x,t)$, we analogously have

$$|K^t_{ij}(x,y)| < \frac{C2^{(jk+i)(n+\frac{\delta}{k})}}{1+(2^{jk+i}|y-\gamma(x,t)|)^{n+\frac{\delta}{k}}} \tag{4.8a}$$

$$|\nabla_x K_{ij}(x,y)|,\ |\nabla_y K_{ij}(x,y)| < \frac{C2^{(jk+i)(n+1+\frac{\delta}{k})}}{1+(2^{jk+i}|y-\gamma(x,t)|)^{n+\frac{\delta}{k}}} \tag{4.8b}$$

We also need some cancellation conditions on K_{ij}^t. By (4.6) and (3.9) we clearly have

$$|\int K_{ij}^t(x,y)\,dx| < C \tag{4.9}$$

In addition, changing coordinates back in (4.6) we have

$$|\int K_{ij}^t(x,y)\,dy| = |\int \xi_{jk+i}(x-y)\,dy| < C \tag{4.10}$$

We now are in a position to prove the estimates (4.4). Since they are totally symmetric, we only prove the estimates on $V_{i_1j_1}^{t_1}{V_{i_2j_2}^{t_2}}^*$. Replacing this operator by its adjoint if necessary, we may assume that $j_1 \geq j_2$. The operator $V_{i_1j_1}^{t_1}{V_{i_2j_2}^{t_2}}^*$ has kernel $L(x,y)$ given by

$$L(x,y) = \int K_{i_1j_1}^{t_1}(x,z)K_{i_2j_2}^{t_2}(z,y)\,dz \tag{4.11}$$

We will prove (4.4) using Schur's test on $L(x,y)$. For the x integral, observe that by (4.7) we have

$$\int |L(x,y)|\,dx < C2^{\frac{\delta}{k}(j_1k+j_2k+i_1+i_2)} < C'2^{\delta(j_1+j_2+2\frac{i}{k})} \tag{4.12}$$

As earlier, $i = \max(i_1,i_2)$. To do the y integral, we rewrite $L(x,y)$ as

$$\int K_{i_1j_1}^{t_1}(x,z)K_{i_2j_2}^{t_2}(\gamma(x,t_1),y)\,dz + \int K_{i_1j_1}^{t_1}(x,z)[K_{i_2j_2}^{t_2}(z,y) - K_{i_2j_2}^{t_2}(\gamma(x,t_1),y)]\,dz \tag{4.13}$$

Let $L_1(x,y)$ and $L_2(x,y)$ be the two terms in (4.13), so that $L(x,y) = L_1(x,y)+L_2(x,y)$. By (4.10), $|L_1(x,y)| < C|K_{i_2j_2}^{t_2}(\gamma(x,t_1),y)|$, so by (4.7) we have

$$\int |L_1(x,y)|\,dy < C2^{\frac{i\delta}{k}+j_2\delta} \tag{4.14}$$

Next, we have

$$\int |L_2(x,y)|\,dy < \int |K_{i_1j_1}^{t_1}(x,z)||K_{i_2j_2}^{t_2}(z,y) - K_{i_2j_2}^{t_2}(\gamma(x,t_1),y)|\,dy\,dz \tag{4.15}$$

We break the integral into two parts depending on whether or not $|y-\gamma(x,t_1)|$ is greater than or less than twice $|z-\gamma(x,t_1)|$. Denote these two integrals by I_1 and I_2 respectively. For I_1, the mean-value theorem and (4.8) give

$$|K_{i_2j_2}^{t_2}(z,y) - K_{i_2j_2}^{t_2}(\gamma(x,t_1),y)| < C|z-\gamma(x,t_1)|\frac{2^{(j_2k+i)(n+1+\frac{\delta}{k})}}{1+(2^{j_2k+i}|y-z^*|)^{n+\frac{\delta}{k}}}$$

Here z^* is a point on the line segment joining z to $\gamma(x,t_1)$. Since $|z-\gamma(x,t_1)| < \frac{1}{2}|y-\gamma(x,t_1)|$, we must have $|y-z^*| \sim |y-\gamma(x,t_1)|$, and the previous equation implies that

$$|K^{t_2}_{i_2j_2}(z,y) - K^{t_2}_{i_2j_2}(\gamma(x,t_1),y)| < C|z-\gamma(x,t_1)|\frac{2^{(j_2k+i)(n+1+\frac{\delta}{k})}}{1+(2^{j_2k+i}|y-\gamma(x,t_1)|)^{n+\frac{\delta}{k}}}$$

As a result,

$$I_1 < C\int |K^{t_1}_{i_1j_1}(x,z)||z-\gamma(x,t_1)|\,dz \int \frac{2^{(j_2k+i)(n+1+\frac{\delta}{k})}}{1+(2^{j_2k+i}|y-\gamma(x,t_1)|)^{n+\frac{\delta}{k}}}\,dy \tag{4.16}$$

$$< C2^{(j_2k+i)(1+\frac{\delta}{k})}\int |K^{t_1}_{i_1j_1}(x,z)||z-\gamma(x,t_1)|\,dz \tag{4.17}$$

Using (4.8), this is at most

$$C2^{(j_2k+i)(1+\frac{\delta}{k})}\int \frac{2^{(j_1k+i)(n+\frac{\delta}{k})}|z-\gamma(x,t_1)|}{1+(2^{j_1k+i}|z-\gamma(x,t_1)|)^{n+\frac{\delta}{k}}}\,dz \tag{4.18}$$

$$< C2^{(j_2k+i)(1+\frac{\delta}{k})+\frac{\delta}{k}(j_1k+i)-j_1k-i} < C2^{(j_2-j_1)k+\delta(j_1+j_2)+2i\frac{\delta}{k}} \tag{4.19}$$

This is the needed estimate for I_1. For I_2, we use (4.7) to directly bound

$$|K^{t_2}_{i_2j_2}(z,y) - K^{t_2}_{i_2j_2}(\gamma(x,t_1),y)| < C2^{(j_2k+i)(n+\frac{\delta}{k})} \tag{4.20}$$

As a result, doing the y integration first in (4.15), keeping in mind that $|y-\gamma(x,t_1)| \leq 2|z-\gamma(x,t_1)|$, we obtain

$$I_2 < C2^{(j_2k+i)(n+\frac{\delta}{k})}\int |K^{t_1}_{i_1j_1}(x,z)||z-\gamma(x,t_1)|^n\,dz \tag{4.21}$$

Using (4.8a), this is at most

$$C2^{(j_2k+i)(n+\frac{\delta}{k})}\int \frac{2^{(j_1k+i)(n+\frac{\delta}{k})}|z-\gamma(x,t_1)|^n}{1+(2^{j_1k+i}|z-\gamma(x,t)|)^{n+\frac{\delta}{k}}}\,dz \tag{4.22}$$

$$< C2^{(j_2k+i)(n+\frac{\delta}{k})-(j_1k+i)n+(j_1k+i)\frac{\delta}{k}} < C2^{(j_2-j_1)kn+(j_1+j_2)\delta+2i\frac{\delta}{k}} \tag{4.23}$$

Observe this is the type of estimate we had in (4.19), with a coefficient of kn in front of j_2-j_1 instead of k. So we conclude that

$$\int |L_2(x,y)|\,dy < C2^{(j_2-j_1)k+(j_1+j_2)\delta+2i\frac{\delta}{k}}$$

Since the bounds (4.14) for $\int |L_1(x,y)|\,dy$ were smaller than this, we conclude that

$$\int |L(x,y)|\,dy < C2^{(j_2-j_1)k+(j_1+j_2)\delta+2i\frac{\delta}{k}} \tag{4.24}$$

By Schur's test, the norm on L^2 of the operator with kernel $L(x,y)$, namely $V_{i_1j_1}^{t_1}{V_{i_2j_2}^{t_2}}^*$, is bounded by the square root of the product of the x and y-integrals, given by (4.12) and (4.24) respectively. Hence we conclude that

$$||V_{i_1j_1}^{t_1}{V_{i_2j_2}^{t_2}}^*|| < C2^{\frac{j_2-j_1}{2}k+(j_1+j_2)\delta+2i\frac{\delta}{k}} \tag{4.25}$$

Since we are assuming $j_1 \geq j_2$, this is exactly the bound (4.4), and as we saw this implies the orthogonality estimates (3.11) and (3.12).

5. Estimates on an individual U_{ij}; the main argument

In this section we prove (3.13) and (3.14); by the results of section 4 and the discussion above Lemma 3.1, this is what is needed to prove Theorem 1.1. Since (3.14) follows directly from applying Schur's test to the kernel of $\bar{U}_j$, we restrict our attention to proving (3.13). Recall that $U_{ij}f = (\xi_{jk+i+1} - \xi_{jk+i}) * T_jf$. Define

$$\xi_0(x) = 2^{n+\frac{\delta}{k}}\xi(2x) - \xi(x) = \triangle^{\frac{\delta}{2k}}(2^n\zeta(2x) - \zeta(x))$$

Since the Fourier transform of $\zeta(x)$ is assumed to be compactly supported and equal to 1 in a neighborhood of the origin, $\xi_0(x)$ is a Schwarz function. By (3.9) we have that $\xi_{m+1}(x) - \xi_m(x) = 2^{m(n+\frac{\delta}{k})}\xi_0(2^mx)$, and hence denoting $\xi_{m+1}(x) - \xi_m(x)$ by $\theta_m(x)$ we have

$$U_{ij}f = \theta_{jk+i} * T_jf \tag{5.1}$$

Writing this out explicitly,

$$U_{ij}f(x) = \int(\int f(\gamma(x+y,t))k_j(t)\,dt)2^{(jk+i)(n+\frac{\delta}{k})}\xi_0(2^{jk+i}y)\,dy \tag{5.2}$$

Since ξ_0 has integral zero, we also have

$$U_{ij}f(x) = \int[\int f(\gamma(x+y,t))k_j(t)\,dt - \int f(\gamma(x,t))k_j(t)\,dt]2^{(jk+i)(n+\frac{\delta}{k})}\xi_0(2^{jk+i}y)\,dy \tag{5.3}$$

We define W_j^y by

$$W_j^yf(x) = \int f(\gamma(x+y,t))k_j(t)\,dt - \int f(\gamma(x,t))k_j(t)\,dt \tag{5.4}$$

We will show that for some $\epsilon > 0$ independent of δ, that for all $|y| < C2^{-jk}$ we have

$$||W_j^yf(x)|| < C(2^{jk}|y|)^\epsilon 2^{-j\delta} \tag{5.5}$$

This will imply (3.13) and therefore Theorem 1.1 by the following lemma.

Lemma 5.1: If δ is sufficiently small, then equation (5.5) implies that $||U_{ij}|| < C2^{-\mu i}$ for some $\mu > 0$.

Proof: By (5.3), we have

$$||U_{ij}|| < \int 2^{(jk+i)(n+\frac{\delta}{k})}|\xi_0(2^{jk+i}y)|\,||W_j^y||\,dy$$

$$= \int_{|y|<2^{-jk}} 2^{(jk+i)(n+\frac{\delta}{k})}|\xi_0(2^{jk+i}y)|\,||W_j^y||\,dy + \int_{|y|\geq 2^{-jk}} 2^{(jk+i)(n+\frac{\delta}{k})}|\xi_0(2^{jk+i}y)|\,||W_{ij}^y||\,dy \tag{5.6}$$

Denote the two integrals in (5.6) by I_1 and I_2. For I_1, we substitute in (5.5) and we obtain

$$I_1 < C2^{(jk+i)(n+\frac{\delta}{k})+jk\epsilon-j\delta}\int |y|^\epsilon|\xi_0(2^{jk+i}y)|\,dy \tag{5.7}$$

Scaling $y \to 2^{-(jk+i)}y$, (5.7) implies that

$$I_1 < C2^{(jk+i)(n+\frac{\delta}{k})+jk\epsilon-j\delta-(jk+i)(n+\epsilon)} = C2^{i\frac{\delta}{k}-i\epsilon} \tag{5.8}$$

As a result, so long as $\delta < k\epsilon$, I_1 will contribute no more than $C2^{-\mu i}$ to $||U_{ij}||$. As for I_2, first note that by direct integration a.k.a Schur's test, we have $||W_j^y|| < C2^{-j\delta}$. So since ξ_0 is a Schwarz function, for any N we have that

$$I_2 < C_N\int_{|y|\geq 2^{-jk}} 2^{(jk+i)(n+\frac{\delta}{k}-N)-j\delta}|y|^{-N}\,dy = C_N 2^{(jk+i)(n+\frac{\delta}{k}-N)-j\delta+jk(N-n)}$$

$$= 2^{i(n+\frac{\delta}{k}-N)} \tag{5.9}$$

Hence by taking $N > n+\frac{\delta}{k}$, I_2 also is seen to contribute no more than $C2^{-\mu i}$ to $||U_{ij}||$. As a result, under the assumptions of this lemma, we have $||U_{ij}|| < C2^{-\mu i}$. This completes the proof of Lemma 5.1.

As in section 2, we let $\gamma_y(x,t) = \gamma(x+y,t)$. Define the operator T_j^y by

$$T_j^y f(x) = \int \gamma_y(x,t)k_j(t)\,dt = \int \gamma(x+y,t)k_j(t)\,dt \tag{5.10}$$

Thus $W_j^y = T_j^y - T_j$. Suppose $N \geq n$ is even. Then the N-fold iterate $W_j^{y*}W_j^y...W_j^{y*}W_j^y$ can be written in the form

$$\sum_{l=1}^{2^N} \pm T_j^{y_N}T_j^{y_{N-1}*}...T_j^{y_2}T_j^{y_1*} \tag{5.11}$$

Here each y_k is either y or 0. Observe that in the sum (5.11), exactly 2^{N-1} terms have a plus sign and 2^{N-1} terms have a minus sign. As a result, (5.11) is exactly

$$\sum_{l=1}^{2^N} \pm(T_j^{y_N}T_j^{y_{N-1}*}...T_j^{y_2}T_j^{y_1*} - T_jT_j^*...T_jT_j^*)$$

Consequently, we have

Lemma 5.2: Suppose that for N sufficiently large, for each $Y = (y_1, ..., y_N)$ with $|Y| < 2^{-jk}$ the following holds.

$$||T_j^{y_N} T_j^{y_{N-1}*} ... T_j^{y_2} T_j^{y_1*} - T_j T_j^* ... T_j T_j^*|| < C2^{-jN\delta + jk}|Y| \tag{5.12}$$

Then equation (3.13) holds, thereby proving Theorem 1.1.

Incidentally, it is in proving (5.12) that the curvature condition is used. (It has not been used so far). For as we will see, the curvature condition implies that the kernels of $T_j T_j^* ... T_j T_j^*$ and $T_j^{y_N} T_j^{y_{N-1}*} ... T_j^{y_2} T_j^{y_1*}$ are spread out near one another well enough that (5.12) holds. To be more precise, we will examine the partial derivatives of the kernel of $T_j^{y_N} T_j^{y_{N-1}*} ... T_j^{y_2} T_j^{y_1*}$ with respect to the y variables, and show that the associated operator has norm bounded by $C2^{-jN\delta + jk}$ on L^2.

To this end, using the formulas (1.1) and (2.6) for T and T^*, in the notation of (2.18) – (2.20) we have the following expressions for some smooth ξ.

$$T_j^{y_N} T_j^{y_{N-1}*} ... T_j^{y_2} T_j^{y_1*} f(x) =$$

$$\int f(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) k_j(t_1)...k_j(t_N) \xi(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 ... dt_N \tag{5.13}$$

$$T_j T_j^* ... T_j T_j^* f(x) =$$

$$\int f(\beta^N(x, t_1, ..., t_N)) k_j(t_1)...k_j(t_N) \xi(x, 0, ..., 0, t_1, ..., t_N)\, dt_1 ... dt_N \tag{5.14}$$

Taking the y_l derivative of (5.13), we have

$$\frac{\partial}{\partial y_l}(T_j^{y_N} T_j^{y_{N-1}*} ... T_j^{y_2} T_j^{y_1*} f(x)) =$$

$$\int \nabla f(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) \cdot \frac{\partial \beta^N_{y_1,...,y_N}}{\partial y_l}(x, t_1, ..., t_N) k_j(t_1)...k_j(t_N)$$

$$\times \xi(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 ... dt_N \tag{5.15a}$$

$$+ \int f(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) k_j(t_1)...k_j(t_N) \frac{\partial \xi}{\partial y_l}(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 ... dt_N \tag{5.15b}$$

The expression (5.15b) is easily bounded; viewed as a superposition of delta functions it has norm bounded by $(\int |k_j|)^N \times \sup \frac{\partial \xi}{\partial y_l}$, which is at most $C2^{-jN\delta}$, better than the estimate $C2^{-jN\delta + jk}$ we need. Thus we need only bound (5.15a). To do this it suffices for each l and m to bound the operator U_{lm} defined by

$$U_{lm} f(x) = \int (\partial_m f)(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) k_j(t_1)...k_j(t_N)$$

$$\times \Xi_{lm}(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 \ldots dt_N \tag{5.16}$$

Here $\Xi_{lm} = \xi \times \partial_{y_l}(\beta^N_{y_1,\ldots,y_N})_m$, where $(\beta^N_{y_1,\ldots,y_N})_m$ denotes the mth component of $\beta^N_{y_1,\ldots,y_N}$. The exact form of Ξ_{lm} will not be relevant, just the fact that it is a smooth function. The necessary bounds for the operator Ξ_{lm} are given by the following theorem.

Theorem 5.3: For sufficiently small δ, for each l, m we have $||U_{lm}|| < C2^{-jN\delta+jk}$. Consequently, Theorem 1.1 holds.

Before being able to prove Theorem 5.3, we will have to learn more about the function $\beta^N_{y_1,\ldots,y_N}(x, t_1, ..., t_N)$ for large N. We first define the function $\alpha^N_y(x, t_1, ..., t_N)$ by

$$\alpha^N_y(x, t_1, ..., t_N) = \sum_I (det_I \beta^N_{y_1,\ldots,y_N}(x, t_1, ..., t_N))^2 \tag{5.17}$$

The sum is over all $I \subset \{1, ..., N\}$ of cardinality n. The function α^N_y is of finite-type in the t variables. In fact, each function $det_I \beta^N_{y_1,\ldots,y_N}(x, t_1, ..., t_N)$ is finite-type in the variables $\{t_{q_i} : i \in I\}$, as was shown in [G2]. Let $M_{y_1,\ldots,y_N}(x, r)$ be as in (2.22), so that

$$C_1' M_{y_1,\ldots,y_N}(x, r) < \sup_{|t_l|<r} \alpha^N_y(x, t_1, ..., t_N) < C_2' M_{y_1,\ldots,y_N}(x, r) \tag{5.18}$$

Fix some even $N_0 \geq n$. By Lemma 3.1 and (5.18), we may let $a > 0$ be such that for each x close enough to the origin and r sufficiently small we have

$$|\{(t_1, ..., t_{N_0}) : |t_l| < r \text{ for all l, } \alpha^{N_0}_y(x, t_1, ..., t_{N_0}) < \delta M_{y_1,\ldots,y_N}(x, r)\}| < C\delta^a r^{N_0} \tag{5.19}$$

The following lemma says that if we replace N_0 by a large enough multiple of N_0, then we can make the power of δ appearing in (5.19) arbitrarily large; this will give regularity to the kernel of U_{lm} that we will exploit in our subsequent arguments.

Lemma 5.4: Fix an even N_0, and let a be as in (5.19). Then for any integer $p > 0$ we have

$$|\{(t_1, ..., t_{pN_0}) : |t_l| < r \text{ for all l, } \alpha^{pN_0}_y(x, t_1, ..., t_{pN_0}) < \delta M_{y_1,\ldots,y_N}(x, r)\}| < C\delta^{ap} r^{pN_0} \tag{5.20}$$

Proof: Let $\chi^{pN_0}_y(x, t_1, ..., t_{pN_0})$ denote the characteristic function of set on the left-hand side of (5.20). We integrate $\chi^{pN_0}_y(x, t_1, ..., t_{pN_0})$ in the final N_0 variables, using (5.19) in these variables on the function $(t_{(p-1)N_0+1}, ..., t_{pN_0}) \rightarrow \beta^n_{y_{(p-1)N_0+1},\ldots,y_N}(w, t_1, ..., t_{N_0})$, where $w = \beta^{(p-1)N_0}_{y_1,\ldots,y_{(p-1)N_0}}(t_1, ..., t_{(p-1)N_0})$. We obtain

$$\int \chi^{pN_0}_y(x, t_1, ..., t_{pN_0}) dt_{(p-1)N_0+1} \ldots dt_{pN_0} < C\delta^a r^{N_0} \chi^{(p-1)N_0}_y(x, t_1, ..., t_{(p-1)N_0}) \tag{5.21}$$

Similarly, we can then integrate in the $t_{(p-2)N_0+1}, ..., t_{(p-1)N_0}$ variables to obtain

$$\int \chi_y^{pN_0}(x, t_1, ..., t_{pN_0}) dt_{(p-2)N_0+1} ... dt_{pN_0} < C\delta^{2a} r^{2N_0} \chi_y^{(p-2)N_0}(x, t_1, ..., t_{(p-2)N_0}) \quad (5.22)$$

Proceeding thus inductively, after we exhaust all pN_0 of the t variables, we obtain (5.20). This completes the proof of the lemma.

In our future arguments, we assume $N = pN_0$ such that Lemma 5.4 holds for $ap \geq 3$. For a given x and y, we apply Lemma 3.3 to the function $\alpha_y^N(x, t_1, ..., t_N)$ in the t variables. Let $\{\phi_R\}$ be the corresponding partition of unity of (3.27). By (3.25), $\alpha_y^N(x, t_1, ..., t_N)$ is within a factor of 6 of a fixed value on the support of a given ϕ_R. However, in our subsequent arguments, we will need that the largest term $(det_I \beta_{y_1,...,y_N}$ $(x, t_1, ..., t_N))^2$ of the sum (5.17) also stays within a constant factor of a fixed value on the support of a member of our partititon of unity. To accomplish this, we will refine the partition of unity by writing each ϕ_R as the sum of boundedly many terms; the resulting partition of unity will have the desired property. We do this as follows. Let $\zeta(x)$ be a function on $\mathbf{R}^N$ that is equal to 1 on $[\frac{1}{2}, \frac{1}{2}]^N$, and is supported in $[-1, 1]^N$. Let $\theta(x) = \frac{\zeta(x)}{\sum_v \zeta(x-v)}$, where the sum is taken over all v with integral coordinates. Note that $\{\theta(x - v)\}$ are a partition of unity, as are $\{\theta(B2^{l_R}x - v)\}$, where 2^{-l_R} denotes the side length of R like before and B is a large constant to be determined in Lemma 5.5. Observe that $O(B^{-N})$ of the functions $\{\theta(B2^{l_R}x - v)\phi_R(x)\}$ are nonzero and the diameter of the support of each of them is $O(B^{-1}2^{-l_R})$ The properties of $\{\theta(B2^{l_R}x - v)\}$ that we need are given by the next lemma. Equation (5.23) below is needed for the reasons outlined in the above discussion, while (5.24) ensures that the support cubes are large enough that the functions $\theta(B2^{l_R}x - v)\phi_R(x)$ don't induce spurious large terms when a derivative lands on them.

Lemma 5.5: Let $\{\sigma_{v,R}^B(x)\}$ be the set of all functions of the form $\theta(B2^{l_R}x - v)\phi_R(x)$, where R is a cube of the partition of unity $\{\phi_R\}$ and v has integer coordinates.

If the constant B is sufficiently large, then there is a constant $c > 0$ such that for each v and R, there is an I such that on the support of $\sigma_{v,R}^B$ the following holds, where A_R is as in (3.25):

$$cA_R \leq |det_I \beta_{y_1,...,y_N}^N(x, t_1, ..., t_N)| \leq 6A_R \quad (5.23)$$

Furthermore, if ϵ is such that $A_R = \epsilon M_{y_1,...,y_N}(x, r)$, there is a constant C such that the side length 2^{-l_R} of R satisfies

$$2^{-l_R} > C\epsilon r \quad (5.24)$$

Proof: Fix v and R. Let $w = (w_1, ..., w_N)$ denote the point $\frac{v}{B2^{l_R}}$. We may select I such that $(det_I \beta_{y_1,...,y_N}^N(x, w_1, ..., w_N))^2$ is the largest term in the sum (5.17) for $(t_1, ..., t_N) = (w_1, ..., w_N)$. In this case, by (3.25) there exists a constant $c > 0$ such that

$$|det_I \beta_{y_1,...,y_N}^N(x, w_1, ..., w_N)| > cA_R \quad (5.25)$$

If $z = (z_1, ..., z_N)$ is any other point in the support of $\sigma^B_{v,R}$, then by the mean value theorem we have

$$|det_I \beta^N_{y_1,...,y_N}(x, z_1, ..., z_N) - det_I \beta^N_{y_1,...,y_N}(x, w_1, .., w_N)| \leq |z - w| \sup_{R'} |\nabla_t det_I \beta^N_{y_1,..,y_N}(x, \cdot)| \tag{5.26}$$

$$< C \frac{2^{-l_R}}{B} \sup_{R'} |\nabla_t det_I \beta^N_{y_1,...,y_N}(x, \cdot)| \tag{5.27}$$

By the Bernstein inequality (3.21), (5.27) is at most

$$\frac{C'}{B} \sup_{R'} |det_I \beta^N_{y_1,...,y_N}(x, \cdot)|$$

$$\leq \frac{6C'' A_R}{B} \tag{5.28}$$

The last equation follows from (3.25). Hence if B is sufficiently small that $\frac{6C''}{B} < \frac{c}{2}$, where c is as in (5.25), then by (5.25) and (5.28), for z in the support of $\sigma^B_{v,R}$ we have

$$|det_I \beta^N_{y_1,...,y_N}(x, z_1, ..., z_N)| > \frac{c}{2} A_R \tag{5.29}$$

This gives the left inequality of (5.23). The right-hand inequality of (5.23) follows from the fact that $det_I \beta^N_{y_1,...,y_N}(x, z_1, ..., z_N) \leq \alpha^N_y(x, t_1, ..., t_N)$ and the fact that $\alpha^N_y(x, t_1, ..., t_N) \leq 6A_N$ by (3.25). Thus we have proven (5.23).

As for (5.24), let ϵ be such that $A_R = \epsilon M_{y_1,...,y_N}(x, r)$. By the left side of (3.24) and (5.23) we have

$$2^{-l_R} > C_1 \frac{\sup_R |\alpha^N_y(x, t_1, ..., t_N)|}{\sup_R |\nabla_t \alpha^N_y(x, t_1, ..., t_N)|} \geq \frac{C_1 \epsilon M_{y_1,...,y_N}(x, r)}{\sup_R |\nabla_t \alpha^N_y(x, t_1, ..., t_N)|} \tag{5.30}$$

By the Bernstein inequality (3.21), we have

$$\sup_R |\nabla_t \alpha^N_y(x, t_1, ..., t_N)| < C' \frac{A_R}{r} \leq C' \frac{M_{y_1,...,y_N}(x, r)}{r}$$

Consequently, (5.30) leads to

$$2^{-l_R} > \frac{C_1 c \epsilon r}{C'}$$

This gives (5.24) and we are done.

We are now in a position to prove Lemma 5.3. For a fixed x and y, we use the partition of unity $\{\sigma^B_{v,R}\}$ as follows. We write the operator U_{lm} as $U_{lm} = \sum_p U^p_{lm}$, where each $U^p_{lm} f(x)$ is of the form

$$\int (\partial_m f)(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) \sigma^B_{v,R}(t_1, ..., t_N) k_j(t_1)...k_j(t_N)$$

$$\times \Xi_{lm}(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 \ldots dt_N \tag{5.31}$$

The idea now is as follows. By Lemma 5.5, on the support of a given $\sigma^B_{v,R}(t_1, ..., t_N)$ there exists an $I = \{q_1, ..., q_N\}$ such that (5.23) holds. On this support, the determinant of $\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ in the t_{q_i} variables is nonzero. As a result, we can find functions a_1,...,a_n such that

$$(\partial_m f)(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)) = \sum_i a_i \frac{\partial}{\partial t_{q_i}}[f(\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N))] \tag{5.32}$$

We will find explicit formulas for the a_i, and substitute them into (5.32). We will then subsitute (5.32) into the formula (5.31) for $U^p_{lm}f(x)$, and integrate by parts with respect to the t_{q_i} variables. We will then bound our terms and use Schur's lemma to find the needed L^2 bounds for U_{lm}, thereby proving Theorem 5.3 and thus Theorem 1.1.

The explicit formulas for the a_i are determined as follows. Define $d\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ to be the derivative matrix of $\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ in the t_{q_i} variables. Let $\mathbf{e_m}$ denote the mth unit coordinate vector, and $\mathbf{a}$ the column vector $(a_1, ..., a_n)$. Then for (5.32) to hold we need that

$$d\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)\, \mathbf{a} = \mathbf{e_m} \tag{5.33}$$

By Cramer's rule, using the notation of (2.42) we have

$$a_i = \frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)}{det_I\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)} \tag{5.34}$$

Denote the operator of the ith term of (5.32) by U^p_{ilm}, so that $\sum_i U^p_{ilm} = U^p_{lm}$. We will integrate by parts in t_{q_i} in the following expression for $U^p_{ilm}f(x)$:

$$U^p_{ilm}f(x) = \int \frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)}{det_I\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)} \frac{\partial}{\partial t_{q_i}}[f(\beta_{y_1,...,y_N}(x, t_1, ..., t_N))]$$

$$\times \sigma^B_{v,R}(t_1, ..., t_N) k_j(t_1)...k_j(t_N) \Xi_{lm}(x, y_1, ..., y_N, t_1, ..., t_N)\, dt_1 \ldots dt_N \tag{5.35}$$

In a sense (5.35) is the crux of this paper. In converting x_m derivative of f into t-derivatives, one gains an additional factor of $\frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}$. By Corollary 2.6, this factor is at least $C2^{j(k-1)}$, and the integration by parts will lead to an additional factor of $C2^j$, so overall we will get an operator with norm $C2^{jk}$ beyond what we would get if we were just looking for L^2 to L^2 estimates. By taking N sufficiently large and invoking Lemma 5.4 on the appropriate $det_I\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$, we can ensure that the points near the zero set of $det_I\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ do not have enough measure for (5.35) to satisfy worse estimates than the portion where $det_I\beta^N_{y_1,...,y_N}(x, t_1, ..., t_N)$ is large. The Whitney type decompositon of Lemma 3.3 is used to ensure two things. First, that one can use a single I in the determinants of (5.35); otherwise we'd have to use variable-coefficient determinants which would be difficult to handle. The second reason for the use of Lemma 3.3 is to ensure that the cutoff functions $\sigma^B_{v,R}$ have large enough support that the t_{q_i} derivatives landing on them don't give large factors that would worsen our estimates.

Denote the kernel of U^p_{ilm} by $K^p_{ilm}(x,x')$. We will bound $\int |K^p_{ilm}(x,x')|\,dx'$, with the goal of eventually using it in applying Schur's test to $U_{lm} = \sum_{i,p} U^p_{ilm}$. To this end, we now do the integration by parts in t_{q_i} in (5.35). We get several terms, depending on where t_{q_i} derivative lands. We examine each separately.

First, we examine the term where the derivative lands on the Ξ_{lm} factor. Denote the kernel of the operator for this term by $K^{p1}_{ilm}(x,x')$. We can write $K^{p1}_{ilm}(x,x')$ as the average of delta functions as

$$\int \delta(x' - \beta^N_{y_1,...,y_N}(x,t_1,...,t_N))\frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}\sigma^B_{v,R}(t_1,...,t_N)$$

$$\times k_j(t_1)...k_j(t_N)\frac{\partial \Xi_{lm}}{\partial t_{q_i}}(x,y_1,...,y_N,t_1,...,t_N)\,dt_1\,...\,dt_N \tag{5.36}$$

Thus we have

$$\int |K^{p1}_{ilm}(x,x')|\,dx' <$$

$$\int |\frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}\sigma^B_{v,R}(t_1,...,t_N)$$

$$\times k_j(t_1)...k_j(t_N)\frac{\partial \Xi_{lm}}{\partial t_{q_i}}(x,y_1,...,y_N,t_1,...,t_N)|\,dt_1\,...\,dt_N \tag{5.37}$$

By Lemma 5.5, there is an A_R such that on the support of $\sigma^B_{v,R}$ the function $|det_I\beta^N_{y_1,...,y_N}|$ is between cA_R and $6A_R$. Define ϵ to be the number such that $A_R = \epsilon M_{y_1,...,y_N}(x,2^{-j})$. By Corollary 2.6, the factor $\frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}$ appearing in (5.37) satisfies

$$|\frac{det_I^{\mathbf{e_m},i}\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}| < \frac{C}{\epsilon 2^{-j(k-1)}} \tag{5.38}$$

By (1.3), the factor $k_j(t_1)...k_j(t_N)$ satisfies

$$|k_j(t_1)...k_j(t_N)| < C2^{jN(1-\delta)} \tag{5.39}$$

The $\frac{\partial \Xi_{lm}}{\partial t_{q_i}}$ factor we just bound by a constant. As a result, we have

$$\int |K^{p1}_{ilm}(x,x')|\,dx' <$$

$$C\frac{2^{j(k-1)+jN(1-\delta)}}{\epsilon}\int\sigma^B_{v,R}(t_1,...,t_N)\,dt_1\dots dt_N$$

$$=C\frac{2^{j(k-1)+jN(1-\delta)}}{\epsilon}|R| \tag{5.40}$$

Next, we consider the term where the derivative lands on the $k_j(t_1)...k_j(t_N)$ factor. The only difference here is that by (1.3) one gains an additional 2^j factor. As a result, if the kernel for this term is denoted by $K^{p2}_{ilm}(x,x')$, we get

$$\int|K^{p2}_{ilm}(x,x')|\;dx'<C\frac{2^{jk+jN(1-\delta)}}{\epsilon}|R| \tag{5.41}$$

Next, we consider the term where the derivative lands on the $\sigma^B_{v,R}(t_1,...,t_N)$ factor. Recall that this factor is a partition of unity function for a box of width $\frac{2^{-l_R}}{B}$. As a result taking the derivative leads to a an additional factor of $C2^{l_R}$, which is at most $\frac{C}{\epsilon 2^{-j}}$ by Lemma 5.5. Thus if the kernel of this term is denoted by $K^{p3}_{ilm}(x,x')$, we have

$$\int|K^{p3}_{ilm}(x,x')|\;dx'<C\frac{2^{jk+jN(1-\delta)}}{\epsilon^2}|R| \tag{5.42}$$

Next, we consider the term where the derivative lands on the $det^{\mathbf{em},i}_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ factor. One gets an additional factor of $\frac{\frac{\partial}{\partial t_{q_i}}det^{\mathbf{em},i}_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det^{\mathbf{em},i}_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}$, which by the Bernstein inequality (3.21) of Lemma 3.2 is bounded by $\frac{C}{2^{-l_R}}<\frac{C'}{\epsilon 2^{-j}}$ on the support of $\sigma^B_{v,R}(t_1,...,t_N)$. Hence as in the previous term, if we denote the kernel for this term by $K^{p4}_{ilm}(x,x')$, we get

$$\int|K^{p4}_{ilm}(x,x')|\;dx'<C\frac{2^{jk+jN(1-\delta)}}{\epsilon^2}|R| \tag{5.43}$$

Finally, we examine the term where the derivative lands on the $det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)$ in the denominator. In this case, we get an additional factor of $\frac{-\frac{\partial}{\partial t_{q_i}}det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}{det_I\beta^N_{y_1,...,y_N}(x,t_1,...,t_N)}$. This is bounded by $\frac{C'}{\epsilon 2^{-j}}$ by the Bernstein inequalities exactly as in the last term, so denoting the kernel of this term by $K^{p5}_{ilm}(x,x')$, once again we get

$$\int|K^{p5}_{ilm}(x,x')|\;dx'<C\frac{2^{jk+jN(1-\delta)}}{\epsilon^2}|R| \tag{5.44}$$

Adding (5.40) $-$ (5.44), we conclude that

$$\int|K^{p}_{ilm}(x,x')|\;dx'<C\frac{2^{jk+jN(1-\delta)}}{\epsilon^2}|R| \tag{5.45}$$

Recalling that $U^p_{lm} = \sum_i U^p_{ilm}$ (see above (5.35)), if we denote the kernel of U^p_{lm} by $K^p_{lm}(x,x')$, then (5.45) implies

$$\int |K^p_{lm}(x,x')|\, dx' < C\frac{2^{jk+jN(1-\delta)}}{\epsilon^2}|R| \tag{5.46}$$

Next, recall that $U_{lm} = \sum_p U^p_{lm}$ (see above (5.31)). Denote the kernel of U_{lm} by $K_{lm}(x,x')$. We will determine bounds on $\int |K_{lm}(x,x')|\, dx'$ by adding up the corresponding bounds (5.46). We do this as follows. For each integer a, we together group the terms (5.46) for $2^{-a} < \epsilon \leq 2^{-a+1}$. The sum of these terms is bounded by

$$C2^{jk+jN(1-\delta)+2a} \times \sum_{\{R: 2^{-a}M_{y_1,\ldots,y_N}(x,2^{-j}) \leq A_R \leq 2^{-a+1}M_{y_1,\ldots,y_N}(x,2^{-j})\}} |R| \tag{5.47}$$

Recall we are assuming that N is large enough so that Lemma 5.4 holds with $ap \geq 3$. As a result, by Lemma 5.4, (5.47) is bounded by

$$C2^{jk-jN\delta-a}$$

Adding this up over all a, we get that

$$\int |K_{lm}(x,x')|\, dx' < C2^{jk-jN\delta} \tag{5.48}$$

These are the x'-integral estimates we seek in proving Theorem 5.3 by Schur's test. Fortunately, we have already done all the work needed for the x-integral estimates. For so long as N is even, which we are assuming, the adjoint U^*_{lm} can be written as the sum of finitely many terms of the form (5.16), with the functions Ξ_{lm} replaced by other smooth functions. Thus one can deal with U^*_{lm} exactly as we dealt with U_{lm}, and we have

$$\int |K_{lm}(x,x')|\, dx < C2^{jk-jN\delta} \tag{5.49}$$

By Schur's test, (5.48) and (5.49) taken together give Theorem 5.3, and therefore Theorem 1.1.

6. Final Comments

As mentioned before, the arguments of this paper still hold when $k(t)$ is replaced by a function $k(x,t)$ under certain differentiability conditions on $k(x,t)$ in the x-variable. For example, our arguments work if we assume that for some $\rho > 0$ we have

$$|\nabla_x k(x,t)| < C|t|^{-2+\delta+\rho} \tag{6.1}$$

The author does not know exactly how small δ must be in Theorem 1.1. If one examines the proof, one gets upper bounds on δ in terms of the growth of the distribution function of

$\sum_I (det_I \beta^N(x, t_1, ..., t_N))^2$ for a sufficiently large N, for x near the origin. However there is no compelling reason to think this is the best possible.

Using interpolations with easy L^1 estimates, one can argue as in [G2] and earlier papers to get almost-sharp L^p to $L^p_{\frac{\delta}{k}-\epsilon}$ boundedness of T if p is sufficiently close to 2. This may be useful in the study of the L^p theory of subelliptic PDE's, which is still poorly understood.

7. Acknowledgements

The author would like to thank David Jerison for useful discussions concerning subelliptic PDE's.

8. References

[CNSW] M. Christ, A. Nagel, E. Stein, S. Wainger, *Singular and maximal Radon transforms: analysis and geometry*, Annals of Mathematics (2) **150** (1999), no. 2, 489-577.

[CW] R. R. Coifman, G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogenes*, Lecture Notes in Math. **242**, Springer-Verlag.

[DJ] G. David, J.-L. Journe, *A boundedness criterion for generalized Calderon-Zygmund operators*, Ann. of Math. **120** (1984), 371-397.

[Fa] E. Fabes, *Singular integrals and partial differential equations of parabolic type*, Studia Mathematica **28** (1966), 81-131.

[FP] C. Fefferman, D.H. Phong, *Subelliptic eigenvalue problems*, Conference on harmonic analysis in honor of Antoni Zygmund, 590-606, Wadsworth Math. Series, Wadsworth, Belmont, California, 1983.

[FSa] C. Fefferman, A. Sanchez-Calle, *Fundamental solutions for second-order subelliptic operators*, Ann. of Math. **124** (1986), 247-272.

[G1] M. Greenblatt, *A method for proving L^p boundedness of singular Radon transforms in codimension one for* $1 < p < \infty$, Duke Math J. **108** (2001) 363-393.

[G2] M. Greenblatt, *Boundedness of singular Radon transforms under a finite-type condition*, Amer. Journal of Math **123** (2001), 1009-1053.

[G3] M. Greenblatt, *Sharp estimates for one-dimensional oscillatory integral operators with C^∞ phase*, to appear, Amer. J Math.

[Gr] M. Gromov, *Carnot-Caratheodory spaces seen from within*, A. Bellaiche ed, Prog.

Math **144** (1996), 79-323, Birkhauser, Basel.

[H] S. Helgason, *A duality in integral geometry on symmetric spaces*, in Proc. U.S.-Japan Seminar in Differential Geometry, Kyoto 1965 (1966), 37-56.

[NRW1] A. Nagel, N. Riviere, S. Wainger, *On Hilbert transforms along curves II.*, Amer. J. Math, **98** (1976), no. 2, 395-403.

[NRW2] A. Nagel, N. Riviere, S. Wainger, *A maximal function associated to the curve* (t, t^2). Proc. Nat. Acad. Sci. U.S.A. **73** (1976), no. 5, 1416–1417.

[NSW] A. Nagel, E. Stein, S. Wainger, *Balls and metrics defined by vector fields I. Basic properties*, Acta. Math. 155 (1985), 103-147.

[P] D. H. Phong, *On integral representations for the Neumann operator*, Proceedings of the National Academy of Sciences USA **76** (1979), no.4, 1554-1558.

[PS] D. H. Phong, E. M. Stein, *Hilbert integrals, singular integrals and Radon transforms I*, Acta Math. 157 (1986), 99-157.

[Se] A. Seeger, *Radon transforms and finite type conditions*, Journal of the American Mathematical Society **11** (1998) no.4, 869-897.

[SeW] A. Seeger, S. Wainger, *Bounds for singular fractional integrals and related Fourier integral operators*, Journal of Functional Analysis **199** (2003), 48-91.

[SW] E. Stein, S. Wainger, *Maximal functions associated to smooth curves*, Proceedings of the National Academy of Sciences, USA **73** (1976), no. 12, 4295 - 4296.

1001 N Vermont St #811
Arlington, VA 22201
mgreenbla@yahoo.com