A T(1) Theorem for Singular Radon Transforms

Michael Greenblatt

1. Introduction

The purpose of this paper is to extend and clarify the methods of the papers [G1] and [G2] by using the methods of [G3], incorporating ideas from Carnot-Caratheodory geometry such as those of the fundamental paper [NSW]. We will thereby prove an analogue for singular Radon transforms to the $T(1)$ theorem of David and Journe [DJ]. As in [G1] and [G2], we will associate a singular integral operator to a singular Radon transform. The crux of this paper consists of showing that the difference of the singular integral operator and the singular Radon transform is bounded on L^2, analogous to [G1] [G2] and their predecessors. Consequently, in some sense this $T(1)$ theorem for singular Radon transforms follows from "lifting" to the traditional $T(1)$ theorem for the associated singular integral operator.

The operators under consideration in this paper are of the form

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

Here $\gamma(x,t)$ is a smooth function defined in a neighborhood of the origin in $\mathbf{R}^n \times \mathbf{R}^k$ for some $n > k \geq 1$. We assume that $\gamma(x,t)$ is "centered" at x, in other words that

$$\gamma(x,0) = x \tag{1.2a}$$

We further assume that for all x we have

$$t \to \gamma(x,t) \text{ is a one to one immersion} \tag{1.2b}$$

Equation (1.1) is a way of saying $Tf(x)$ is an average of $f(x)$ along the surface $t \to \gamma(x,t)$ with respect to the kernel $K(x,t)$. Although (1.2b) is not required in some works on this subject, such as [CNSW] and [G2], for us to be able to use Carnot-Caratheodory geometry it is necessary to assume it. We also assume that γ satisfies the curvature condition of [CNSW], which we will have more to say about shortly.

The function $K(x,t)$ is assumed to be C^1 in the support of γ outside $t \neq 0$ and to satisfy the estimates

$$|K(x,t)| < C|t|^{-k}, \;\; |\nabla_t K(x,t)| < C|t|^{-k-1} \tag{1.3}$$

So that there is no issue as to whether the operator T is well-defined, we assume that for each x the function $K(x,t)$ is a distribution in the t variables, and that (1.1) holds for each Schwarz function f. We will also assume some regularity on $K(x,t)$ in the x-variable, which we will be able to describe once we have some facts about the metric associated to γ; see (1.20) below.

The distance associated to $\gamma(x,t)$

Given two points x and y, we define the distance $d(x,y)$ between x and y by

$$d(x,y) = \inf\{r : \text{ for some } m \text{ there exists } t^0, ..., t^{2m+1} \in \mathbf{R}^k \text{ and } x_0, ..., x_{m+1} \in \mathbf{R}^n \text{ such}$$

$$\text{that } x = x_0,\ y = x_{m+1},\ \sum_{i=0}^{2m+1} |t^i| \leq r, \text{ and } \gamma(x_i, t^{2i}) = \gamma(x_{i+1}, t^{2i+1})\} \text{ for all } i. \tag{1.4}$$

If no such path between x and y exists, we say that $d(x,y) = \infty$. Denote the balls with respect to this distance by $B(x,r)$. The curvature condition of [CNSW] at a point $x \in \mathbf{R}^n$ can be restated as follows.

Curvature Condition: The curvature condition holds at x_0 if and only if there is a (Euclidean) neighborhood E containing x_0 and an $r_0 > 0$ such that

1) If $x \in E$, then $d(x, x_0) < \infty$.

2) There exists a constant C such that if $r < r_0$ and $B(x, 2r) \subset E$, then $|B(x, 2r)| < C|B(x,r)|$.

We will not prove here the equivalence of the above with the curvature condition of [CNSW]; it follows fairly directly from the similar restatement of the curvature condition given in [G2]. Note also that when the curvature condition holds at x_0, then the distance $d(x,y)$ satisfies the conditions of the generalized Calderon-Zygmund theorem of Coifman and Weiss [CW] near $x = x_0$. Throughout this paper, we assume the curvature condition holds. Specifically, we assume that the curvature condition holds at $x_0 = 0$, and that the support of $K(x,t)$ is sufficiently small that the x-projection of this support is contained in the neighborhood E for which **1)** and **2)** hold.

As might be expected from its definition, the distance d is closely related to a certain Carnot-Caratheodory metric. To see this, we move to a viewpoint related to that of [H]. Let E be as in the statement of the curvature condition for the operator T at $x = 0$. For a sufficiently small $r_2 > 0$ that things are well-defined, we define the manifold M by

$$M = \{(x, \gamma(x,t)) \in E \times E : |t| < r_2\} \tag{1.5}$$

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

$$T^* f(y) = \int_{\mathbf{R}} f(\gamma^*(y,t))\eta(y,t)K(\gamma^*(y,t),t)\,dt \tag{1.6}$$

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

$$M = \{(\gamma^*(y,t),y) : (\gamma^*(y,t),y) \in E \times E,\ |t| < r_2\} \tag{1.7}$$

There are natural vector fields $X_1, ..., X_k$ and $Y_1, ..., Y_k$ on the manifold M. The vector field X_i at $(x, \gamma(x,t))$ is the image of the vector field dt_i under the map $t \to (x, \gamma(x,t))$, while the vector field Y_i at $(\gamma^*(y,t), y)$ is the image of the vector field dt_i under the map $t \to (\gamma^*(y,t), y)$. So we can write

$$X_i = (0, \sum_j \frac{d\gamma_j(x,t)}{dt_i} dy_j),\ \ Y_i = (\sum_j \frac{d\gamma_j^*(y,t)}{dt_i} dx_j, 0) \tag{1.8}$$

The curvature condition can be expressed in terms of the vector fields $\{X_i\}, \{Y_i\}$. Namely, the curvature condition above is equivalent to these vector fields satisfying Hormander's condition:

Curvature Condition The operator T satisfies the *curvature condition* at a point x if the vector fields $\{X_i\}$, $\{Y_i\}$ and their iterated commutators 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 equivalent to the curvature condition given in section 1; however we will not prove the equivalence here. We refer to [CNSW] for more information on these issues. It is natural to examine the metric one obtains by applying the techniques of [NSW] to the vector fields $\{X_i\}$ and $\{Y_i\}$ on M. Let $D((x,y),r)$ denote the ball of radius r in this metric centered at a point $(x,y) \in M$. As is well-known (see Chapter 1 of [Gr] for example), the balls of the metric can be taken to be the following:

$$D((x,y),r) = \{e^{s_{2m+1}X_{i_{2m+1}}} e^{s_{2m}Y_{i_{2m}}} ... e^{s_2 Y_{i_2}} e^{s_1 X_{i_1}}(x,y) : m > 0,\ \sum_{j=1}^{2m+1} |s_j| < r\} \tag{1.9}$$

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

$$D((x,y),r) = \{e^{s_{2m+1}X_{i_{2m+1}}} e^{s'_{2m}Y_{i_{2m}}} ... e^{s_2 Y_{i_2}} e^{s_1 X_{i_1}}(x,y) : |s_j| < r \text{ for all } j\} \tag{1.10a}$$

$$D((x,y),r) = \{e^{s_{2m}Y_{i_{2m}}} e^{s_{2m-1}X_{i_{2m-1}}} ... e^{s_2 Y_{i_2}} e^{s_1 X_{i_1}}(x,y) : |s_j| < r \text{ for all } j\} \tag{1.10b}$$

Let π_1 and π_2 be the projections of M onto the first n and last n coordinates respectively. In view of (1.9) we have

$$\pi_1(D(x,x),r) = B(x,r) \tag{1.11}$$

Lemma 1.1: There are constants c_1 and c_2 such that

$$D((x,x),c_1r) \subset \{(y,\gamma(y,t)) : y \in B(x,r),\ |t| < r\} \subset D((x,x),c_2r) \tag{1.12}$$

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

$$C|D((x,x),r)| < r^k|B(x,r)| < C'|D((x,x),r)| \tag{1.13}$$

Proof: Equation (1.12) follows directly from (1.9). Using the nondegeneracy condition (1.2b), (1.12) implies that

$$C_1|D((x,x),c_1r)| < r^k|B(x,r)| < C_2|D((x,x),c_2r)| \tag{1.14}$$

By the [NSW] theory, we have

$$C_3|D((x,x),r)| < |D((x,x),c_1r)| \tag{1.15a}$$

$$|D((x,x),c_2r)| < C_4|D((x,x),r)| \tag{1.15b}$$

Combining (1.14) and (1.15a) – (1.15b) gives (1.13), and the lemma follows.

We can use this Caratheodory metric to define privileged coordinates on the balls $B(x,r)$, in analogy to the privileged coordinates of [NSW] on the Carnot-Caratheodory balls $D((x,x),r)$ on M. Let T_1,...,T_p be an enumeration of commutators of the vector fields X_i and Y_i on M, such that the list T_1,...,T_p contains commutators of sufficiently high degree to include those used in privileged coordinates for the Carnot-Caratheodory balls on M. Let d_i denote the degree of T_i.

Fix some x and r. By the theory of [NSW], we can find vector fields $V_1, ..., V_{n+k}$ among $r^{d_1-1}T_1, ..., r^{d_p-1}T_p$ such that for some large constant C_0 dictated by our subsequent arguments the following map, which we denote by $A_{x,r}(s_1, ..., s_{n+k})$, gives privileged coordinates on $D((x,x),C_0r)$ for $|s_i| < r$.

$$A_{x,r}(s_1, ..., s_{n+k}) = exp(s_1V_1 + ... + s_{n+k}V_{n+k})(x,x) \tag{1.16}$$

For notational convenience, we assume that the degrees of the vector fields V_i are *decreasing* with i. If we can find k independent nonvanishing vector fields amongst the X_i and Y_i, we may let $V_{n+1}, ..., V_{n+k}$ be these vector fields. By (1.2b), we can do this in our current setting by letting $V_{n+i} = X_i$ for $1 \leq i \leq k$.

For $1 \leq i \leq n$, define $W_i = d\pi_1 V_i$, where π_1 is left projection.

Lemma 1.2: Define the map $\beta_{x,r}$ by

$$\beta_{x,r}(s_1, ..., s_n) = exp(s_1W_1 + ... + s_nW_n)(x,x) \tag{1.17}$$

Then $\beta_{x,r}$ provides privileged coordinates on $B(x,r)$ in the following sense. First, the image of $[-r,r]^n$ under the map $\beta_{x,r}$ contains $B(x,C_0 r)$ and is contained in $B(x,Cr)$ for some C. Secondly, the determinant of $\beta_{x,r}$ is within a constant factor of a fixed value $D_{x,r}$ on $[-r,r]^n$, such that $|B(x,r)| \sim |D_{x,r}| r^n$.

Proof: We view the n vector fields $W_1, ..., W_n$ as vector fields in $n+k$ dimensions, and extend them to a set of $n+k$ vector fields $W_1, ..., W_{n+k}$ by defining $W_i = V_i$ for $i > n$. Observe that for $i \leq n$, $W_i = V_i + \sum_{l=n+1}^{n+k} O(r^{d_i - 1}) V_l$ with $d_i > 1$. Consequently, if we define $B_{x,r}$ by

$$B_{x,r}(s_1, ..., s_{n+k}) = exp(s_1 W_1 + ... + s_{n+k} W_{n+k})(x,x) \tag{1.18}$$

then the Jacobian determinant of the map $A_{x,r} B_{x,r}^{-1}$ is of the form $I + rF$, where I is the identity matrix and F has bounded coefficients. As a result, if a point has coordinates $s = (s_1, ..., s_{n+k})$ in the coordinates of (1.16) and coordinates $s' = (s_1', ..., s_{n+k}')$ in the coordinates of (1.18), then we have

$$|s - s'| < Cr|s| \tag{1.19}$$

In particular, the function $B_{x,r}$ on $[-r,r]^{n+k}$ gives privileged coordinates on $D((x,x),r)$. Observe that by the way the W_i are defined, the points (y_1, y_2) in the image of $B_{x,r}$ with $y_2 = x$ are exactly the points $B_{x,r}(s_1, ..., s_n, 0, ..., 0) = (\beta_{x,r}(s_1, ..., s_n), x)$. Furthermore, the determinant of $B_{x,r}$ is comparable to the determinant of $\beta_{x,r}$ since the vectors $W_1, ..., W_n$ are orthogonal to the vectors $W_{n+1}, ..., W_{n+k}$ and the latter vectors have degree 0 in the map $B_{x,r}$. The conclusions of Lemma 1.2 thus follow from the corresponding facts about the coordinate map $B_{x,r}$.

Next, we use the vector fields $W_1, ..., W_n$ above to give the x-derivative bounds on $K(x,t)$ that we assume in this paper. Namely, we assume that there exists a $\delta > 0$ such for each i and for each t with $|t| < r$ we have

$$|W_i K(x,t)| < C|t|^{-k-1-\delta} \tag{1.20}$$

To be perfectly clear, the vector field W_i acts in the x variable. The extra δ is necessary for the arguments of this paper, as will become clearer in sections 3 and 4. The author believes it is likely that Theorem 1.1 does not hold in general for $\delta = 0$.

One also has a notion of bump functions for the balls $B(x,r)$. We fix some integer $M > n$, and we say that a smooth function ψ is a normalized bump function for a ball $B(x,r)$ if ψ is supported in $B(x,r)$ and for any nonnegative integers α_1,...,α_n with $\sum_i \alpha_i \leq M$ the following inequalities hold:

$$\sup |W_1^{\alpha_1} ... W_n^{\alpha_n} \psi| < |r|^{-\sum_i \alpha_i} \tag{1.21}$$

The $T(1)$ theorem (see [C] or [S] Ch. 7 for example) says that a singular integral operator S satisfying appropriate differentiability conditions is bounded on L^2 if and only if S and

S^* are L^2-bounded on all bump functions; that is, if for all x and r and all ψ satisfying (1.21) we have

$$||S\psi||_{L^2} < C|B(x,r)|^{\frac{1}{2}} \tag{1.22a}$$

$$||S^*\psi||_{L^2} < C|B(x,r)|^{\frac{1}{2}} \tag{1.22b}$$

Our main theorem is analogous:

Theorem 1.1: Suppose $(1.2a)-(1.2b)$ holds and $\gamma(x,t)$ satisifies the curvature condition at $x=0$. Suppose $K(x,t)$ satisfies (1.3) and (1.20). Then there exists a neighborhood F of the origin in $\mathbf{R}^n \times \mathbf{R}^k$ such that if $K(x,t)$ is supported in F, then T is bounded on L^2 if and only if for each x and r and all smooth ψ satisfying (1.21) we have

$$||T\psi||_{L^2} < C|B(x,r)|^{\frac{1}{2}} \tag{1.23a}$$

$$||T^*\psi||_{L^2} < C|B(x,r)|^{\frac{1}{2}} \tag{1.23b}$$

Some history

Metrics have been used in the study of singular Radon transforms for some time. 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. The curvature condition was developed in [CNSW], where L^p boundedness for a singular Radon transform is shown when the curvature condition holds, under some restrictions on the kernel $K(x,t)$. The paper [CNSW] uses, among other things, 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], using the metric here the author found another proof of L^p boundedness of singular Radon transforms under the curvature condition, again assuming some restrictions on $K(x,t)$. The methods of [G2] are also related to those of [Fa] [PS] [SW] and [NSW], with [PS] being the most influential.

The issue of proving $T(1)$ theorems for singular Radon transforms over hypersurfaces when $\gamma(x,t)$ satisfies a strong nondegeneracy condition called rotational curvature was taken up in the thesis of Pottinton [Po], where several interesting theorems are proven. In [Po], one also proceeds by showing L^2 boundedness of the difference between a singular Radon transform and an associated singular integral operator. The rotational curvature assumption allows one to directly define the singular integral operator without the use of Carnot-Caratheodory metrics or the associated machinery that was developed here and in [G3].

2. Carnot-Caratheodory geometry and the distance function $d(x, y)$

We now define some notation which will be used at various places in this paper.

The functions $\beta^l(x, t^1, ..., t^l)$ for small $|x|$, $|t^i|$ are defined inductively for $l \geq 2$ as follows. For $l = 2$, we define

$$z = \beta^2(x, t^1, t^2) \iff \gamma(x, t^1) = \gamma(z, t^2) \tag{2.1}$$

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:

$$\beta^l(x, t^1, t^2, ..., t^l) = \beta^2(\beta^{l-2}(x, t^1, t^2, ..., t^{l-2}), t^{l-1}, t^l) \tag{2.3}$$

In this notation, by (1.10) and (1.12), for sufficiently large l the metric $d(x, y)$ is equivalent to the one with balls $B^l(x, r)$ defined by

$$B^l(x, r) = \{\beta^l(x, t^1, ..., t^l) : |t^i| < r \text{ for all } i\} \tag{2.4}$$

As in [G3], we also define a parameterized family of functions $\beta^l_{y^1,...,y^l}(x, t^1, ..., t^l)$. We fix x and r, and let T_1,...,T_p be the commutators of the X_i's and Y_j's as before. Define $U_i = r^{d_i - 1} d\pi_1 T_i$, where d_i is the degree of T_i. For $|y| < r$ we define

$$\gamma_y(x, t) = exp(y_1 U_1 + ... + y_p U_p)\gamma(x, t) \tag{2.5}$$

In analogy to (2.1) – (2.3), we define

$$z = \beta^2_{y^1,y^2}(x, t^1, t^2) \iff \gamma_{y^1}(x, t^1) = \gamma_{y^2}(z, t^2) \tag{2.6}$$

For $l > 2$ odd:

$$\beta^l_{y^1,...,y^l}(x, t^1, t^2, ..., t^l) = \gamma(\beta^{l-1}_{y^1,...,y^{l-1}}(x, t^1, ..., t^{l-1}), t^l) \tag{2.7}$$

For $l > 2$ even:

$$\beta^l_{y^1,...,y^l}(x, t^1, t^2, ..., t^l) = \beta^2_{y^1,y^2}(\beta^{l-2}_{y^1,...,y_{l-2}}(x, t^1, t^2, ..., t^{l-2}), t^{l-1}, t^l) \tag{2.8}$$

Observe that $\beta^l_{0,...0}(x, t^1, t^2, ..., t^l)$ is just $\beta^l(x, t^1, t^2, ..., t^l)$.

For the remainder of the section, we will develop some technical ideas that are necessary for the proof of Theorem 1.1. They are quite similar to the concepts developed in the second half of section 2 of [G3], and will be used for an analogous purpose; in fact all we will use from the remainder of this section is Lemma 2.3, which will be used as a substitute for Corollary 2.6 of [G3] when applying the arguments of [G3]. Suppose $l \geq n$, and I is a subset of $\{1,...,l\} \times \{1,...,k\}$ of cardinality n. (Recall k is the t-dimension). Write $I = \{(p_1,q_1),...,(p_n,q_n)\}$ where $p_1 \leq p_2 < ... \leq p_n$. We define the function $det_I\, \beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ to be the determinant of the matrix whose jth column is $\partial_{t^{p_j}_{q_j}} \beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$. In other words, $det_I \beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ is the Jacobian determinant of $\beta^l_{y^1,...,y^l}$ in the $t^{p_1}_{q_1},...,t^{p_n}_{q_n}$ variables.

Next, we define $M_{y^1,...,y^l}(x,r)$ by

$$M_{y^1,...,y^l}(x,r) = \sup_I \sup_{|t^p_q|<r \ for\ all\ p,q} |det_I \beta^l_{y^1,...,y^l}(x,t^1,...,t^l)| \tag{2.9}$$

By (1.10), each $\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ is of finite type in the t variables for l sufficiently large. As a result, for a given l the measure of the image of $\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ in the t variables for $|t^p_q| < r$ satisfies

$$|Image(\beta^l_{y^1,...,y^l}(x,t^1,...,t^l))| > CM_{y^1,...,y^l}(x,r)r^n \tag{2.10}$$

On the other hand, by the definition of $\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ there exists a C' such that

$$Image(\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)) \subset B(x,C'r) \tag{2.11}$$

Hence we have

$$M_{y^1,...,y^l}(x,r) < C'' \frac{|B(x,r)|}{r^n} \tag{2.12}$$

It is also true that for some C''' we have

$$M_{0,...,0}(x,r) > C''' \frac{|B(x,r)|}{r^n} \tag{2.13}$$

Equation (2.13) follows from the constructions of [NSW]. Namely, by [NSW] the measure of the Carnot-Caratheodory ball $D((x,x),r)$ is comparable to the maximum determinant of (1.10a) or (1.10b) in n of the s-variables. Equation (2.13) then follows from left projection onto $B(x,r)$, using Lemma 1.1 to translate into the current situation. Combining (2.12) and (2.13), we conclude

$$C''' \frac{|B(x,r)|}{r^n} < M_{0,...,0}(x,r) < C'' \frac{|B(x,r)|}{r^n} \tag{2.14}$$

We now show that it is also true that $M_{y^1,...,y^l}(x,r) \sim \frac{|B(x,r)|}{r^n}$ for $|y^i| < cr$ for a small constant c:

Lemma 2.1: For a given l, there is a constant c such that if $|y^i| < cr$ for each i then

$$C_1 \frac{|B(x,r)|}{r^n} < M_{y^1,...,y^l}(x,r) < C_2 \frac{|B(x,r)|}{r^n}$$

Proof: The right inequality follows from (2.12) so it suffices to prove the left-hand inequality. Let I and $(u^1, ..., u^l)$ be such that

$$det_I \beta^l(x, u^1, ..., u^l) = M_{0,...,0}(x, r) \tag{2.15}$$

By the Bernstein inequalities (see Lemma 3.2 of [G3] for example) if all $|t^i|, |y^i| < r$ we have

$$|\nabla_{t,y} det_I \beta^l_{y^1,...,y^l}(x, t^1, ..., t^l)| < C \sup_y \frac{M_{y^1,...,y^l}(x, r)}{r} < C' \frac{M_{0,...,0}(x, r)}{r} \tag{2.16}$$

The right-hand inequality follows from (2.14). By the mean value theorem, if $(y^1, ..., y^l)$ satisfies $|y^i| < \frac{C'}{2l} r$ for all i, where C' is as in (2.16), then

$$|det_I \beta^l_{y^1,...,y^l}(x, u^1, ..., u^l) - det_I \beta^l(x, u^1, ..., u^l)| < \frac{1}{2} M_{0,...,0}(x, r) \tag{2.17}$$

Combining (2.15) and (2.17), we have

$$M_{y^1,...,y^l}(x, r) > |det_I \beta^l_{y^1,...,y^l}(x, u^1, ..., u^l)| \geq \frac{1}{2} M_{0,...,0}(x, r) \tag{2.18}$$

Thus letting $c = \frac{C'}{2l}$ gives the lemma and we are done.

Lemma 2.2: Let x and r be fixed, and let $W_1, ..., W_n$ denote the privileged coordinates for $B(x, r)$ from Lemma 1.2. Let $W_i(z)$ denote the value of the vector field at a point z.

a) For $|s_i|, |t| < r$ consider the map $\Phi_1(s, t)$ defined by

$$\Phi_1(s_1, ..., s_p, t) = exp(\sum_{i=1}^{p} s_i U_i)\gamma(x, t) \tag{2.19}$$

Then for each i and each s and t we can find $a_1, ..., a_n$ with $|a_l| < C'$ for all l such that

$$\frac{\partial \Phi_1}{\partial s_i}(s_1, ..., s_p, t) = \sum_{l=1}^{n} a_l W_l(\Phi_1(s_1, ..., s_p, t)) \tag{2.20}$$

b) For $|s_i|, |t| < r$ define the map $\Phi_2(s, t)$ by

$$\Phi_2(s_1, ..., s_p, t) = \gamma^*(exp(\sum_{i=1}^{p} -s_i U_i)x, t) \tag{2.21}$$

For each i and each s and t we can find $a_1, ..., a_n$ with $|a_l| < C'$ for all l such that

$$\frac{\partial \Phi_2}{\partial s_i}(s_1, ..., s_p, t) = \sum_{l=1}^{n} a_l W_l(\Phi_2(s_1, ..., s_p, t)) \tag{2.22}$$

c) Let N be a positive integer. For $|y^1|,...|y^N|,|t^1|,...,|t^N| < r$ define the map $\Phi_3(y,t)$ by

$$\Phi_3(y^1,...,y^N,t^1,...,t^N) = \beta^N_{y^1,...,y^N}(x,t^1,...,t^N) \tag{2.23}$$

Then for any i,i',j and j', and any $|y^1|,...|y^N|,|t^1|,...,|t^N| < r$ there exist $a_1,...,a_n$ and $a'_1,...,a'_n$ with $|a_l|,\ |a'_l| < C'$ for all l such that

$$\frac{\partial\Phi_3}{\partial y_i^j}(y^1,...,y^N,t^1,...,t^N) = \sum_{l=1}^n a_l W_l(\Phi_3(y^1,...,y^N,t^1,...,t^N)) \tag{2.24a}$$

$$\frac{\partial\Phi_3}{\partial t_{i'}^{j'}}(t^1,...,t^N,t^1,...,t^N) = \sum_{l=1}^n a'_l W_l(\Phi_3(y^1,...,y^N,t^1,...,t^N)) \tag{2.24b}$$

Proof: Define $\Phi(q,s,t,y,u)$ by $\Phi(q_1,..,q_n,s_1,...,s_p,t^1,..,t^N,y^1,..,y^N,u_1,...,u_p) =$

$$exp(\sum_{i=1}^n q_iW_i)exp(\sum_{i=1}^p s_iU_i)\beta^l_{y^1,...,y^N}(exp(\sum_{i=1}^p -u_iU_i)x,t^1,..,t^N) \tag{2.25}$$

Define $P(x,r)$ by

$$P(x,r) = \sup_{J,q,s,t,y,u} |det_J\Phi(q,s,t,y,u)| \tag{2.26}$$

Here J is a subset of cardinality n of the various q,s,t,y, and u variables, and det_J denotes the determinant in those variables. The J,q,s,t,y,u supremum is taken over all possible n-tuples J and all q,s,t,y,u with $|q_i|,|s_i|,|t^j|,|y^j|,|u_i| < r$. The determinant of Φ in the q variables is at least $C\frac{|B(x,r)|}{r^n}$ since the W_i are a privileged coordinate system for the ball $B(x,r)$. Consequently we must have $P(x,r) \geq C\frac{|B(x,r)|}{r^n}$. Furthermore, since the image of Φ is contained in the ball $B(x,C'r)$ for some C', exactly as in (2.10) – (2.12) we also have $P(x,r) < C''\frac{|B(x,r)|}{r^n}$. Thus we have

$$C\frac{|B(x,r)|}{r^n} < P(x,r)r^n < C''\frac{|B(x,r)|}{r^n} \tag{2.27}$$

Let v be any of the variables s_i, t_i^j, y_i^j, u_i. Then I claim that for any (s,t,y,u), equation (2.27) implies the existence of constants $b_1,...,b_n$ with $|b_l| < C$ for all l such that

$$\partial_v\Phi(0,s,t,y,u) = \sum_l b_lW_l(\Phi(0,s,t,y,u)) \tag{2.28a}$$

Or equivalently,

$$\partial_v\Phi(0,s,t,y,u) = \sum_l b_l\partial_{q_l}\Phi(0,s,t,y,u) \tag{2.28b}$$

For if there were not b_l such that (2.28a)–(2.28b) holds, then the vector $\partial_v\Phi(0,s,t,y,u)$ and the vectors $\partial_{q_l}\Phi(0,s,t,y,u) = W_l(\Phi(0,s,t,y,u))$ would generate too large a parallelopiped for the right-hand side of (2.27) to be satisfied, given that the vectors $W_l(\Phi(0,s,t,y,u))$ already generate a parallelopiped of measure at least $C\frac{|B(x,r)|}{r^n}$ by virtue of their being privileged coordinates. Letting v be an s variable and setting $t^2 = ... = t^N = 0$, $y^1 = ... = y^N = 0$ and $u = 0$ in (2.28b) gives part a) of this lemma. Letting v be a u variable and setting $s = 0$, $y^1 = ... = y^N = 0$, $t^1 = 0$, and $t^3 = ... = y^N = 0$ gives part b). Setting Letting v be a t or y variable and setting $s = 0$, $u = 0$ gives part c). This completes the proof of Lemma 2.2.

Lemma 2.3: Let $l > n$. Suppose $(t^1, ..., t^l)$ satisfies $|t^j| < r$ for all j, and $(y^1, ..., y^l)$ is such that $|y^j| < cr$ for each j, where c is as in Lemma 2.1. Suppose I is a set of t_i^j variables of cardinality n. For a given i and j let $det_I^{y_j^i,p}\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ denote the determinant of the matrix obtained by replacing the pth column of the Jacobian matrix of $\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$ in the I variables by the vector $\frac{\partial}{\partial y_j^i}\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)$. Assuming l is sufficiently large, if $\epsilon > 0$ is such that $|det_I\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)| = \epsilon M_{y^1,...,y^l}(x,r)$ then we have

$$\frac{|det_I^{y_j^i,p}\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)|}{|det_I\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)|} < \frac{C}{\epsilon} \tag{2.29}$$

Proof: Let A denote the matrix whose determinant is being taken in the numerator of (2.29). By Lemma 2.2 c), each column of A is of the form $\sum_i a_i W_i(\beta^l_{y^1,...,y^l}(x,t^1,...,t^l))$ for $|a_i|$ bounded by a constant. Since the W_i are a coordinate system for $B(x,r)$, we have that the determinant of the matrix whose columns are the W_i's is comparable to $M_{0,...,0}(x,r)$, which by Lemma 2.1 is comparable to $M_{y^1,...,y^l}(x,r)$. As a result, we have

$$|det_I^{y_j^i,p}\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)| < CM_{y^1,...,y^l}(x,r) \tag{2.30}$$

On the other hand,

$$|det_I\beta^l_{y^1,...,y^l}(x,t^1,...,t^l)| = \epsilon M_{y^1,...,y^l}(x,r) \tag{2.31}$$

Combining (2.30) and (2.31) gives the lemma.

3. Associating a singular integral to a singular Radon transform; kernel estimates

As in [G1] and [G2], in our proofs we will make use of a singular integral associated to the singular Radon transform T, even though this singular integral does not explicitly appear in the statement of Theorem 1.1. In order to define this singular integral operator, we will effectively have to define bump functions for the balls $B(x,r)$. One can apply the results of [NS] to define bump functions on $B(x,r)$; however, we need some rather delicate cancellation conditions particular to our singular integrals that require the bump functions to be defined somewhat differently.

Let $\alpha_0(t)$ be a nonnegative bump function on $\mathbf{R}^k$ supported for $|t| > 2$ with $\alpha_0(t) = 1$ for $|t| \leq 1$. Let $\alpha(t) = \alpha_0(t) - \alpha_0(2t)$, and let $K_j(x,t) = \alpha(2^j t)K(x,t)$. We define the operator T_j by

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

Note that we have

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

As before, Let T_1,...,T_p be an enumeration of the commutators of the vector fields X_i and Y_i on M, such that T_1,...,T_p contains commutators of sufficiently high degree as to include all those used in privileged coordinates for the Carnot-Caratheodory balls on M. Next, assuming j is fixed we let $U_i = 2^{-j(deg(T_i)-1)}d\pi_1 T_i$, where π_1 is left projection as before. Let $W_1, ..., W_n$ be vector fields amongst the U_i that give privileged coordinates on $B(x, 2^{-j})$ in the sense of Lemma 1.2. Let $\phi(x)$ be a nonnegative bump function on $\mathbf{R}^k$ with integral 1 that is supported on $|x| < 1$. Define the operator S_j by

$$S_j f(x) = 2^{jp}\int f(exp(s_1U_1 + ... + s_pU_p)\gamma(x,t))K_j(x,t)\phi(2^j s_1)\ldots\phi(2^j s_p)\,ds_1\ldots ds_p\,dt \tag{3.3}$$

The **singular integral associated to** T is defined to be the operator S given by

$$S = \sum_j S_j \tag{3.4}$$

The exact definition of the S_j is not really important here; what is necessary is that the S_j satisfy the conditions of the following lemma:

Lemma 3.1: Let $L_j(x,y)$ be the kernel of S_j. Let W_i^x denote the vector field W_i of the privileged coordinates on $B(x,2^{-j})$, and let $W_i^x L_j(x,y)$ denote the action of W_i^x on $L_j(x,y)$ in the y variables. Similarly, let W_i^y denote the vector field W_i of the privileged coordinates on $B(y,2^{-j})$, and let $W_i^y L_j(x,y)$ denote the action of W_i^y on $L_j(x,y)$ in the x variables. We have

$$|L_j(x,y)| < C|B(x,2^{-j})|^{-1} \tag{3.5}$$

$$|W_i^x L_j(x,y)|\,,\ |W_i^y L_j(x,y)| < C2^j|B(x,2^{-j})|^{-1} \tag{3.6}$$

Furthermore, if δ is as in (1.20), for any x and y we have the cancellation conditions

$$|S_j(1)(x) - T_j(1)(x)|\,,\ |S_j^*(1)(y) - T_j^*(1)(y)| < C2^{-\delta j} \tag{3.7}$$

(The idea behind (3.7) is that if the kernel of $T_j(x,y)$ could have been written as a function $M_j(x,y)$, then the cancellation condition (3.7) could have been cast in the more recognizable form

$$|\int M_j(x,y) - L_j(x,y)\,dy|\,,\ |\int M_j(x,y) - L_j(x,y)\,dx| < C2^{-\delta j} \tag{3.8}$$

Proof: We write $L_j(x,y)$ in the form

$$2^{jp}\int \delta(y - exp(\sum_{l=1}^{p} s_l U_l)\gamma(x,t))K_j(x,t)\phi(2^j s_1)\ldots\phi(2^j s_p)\, ds_1 \ldots ds_p\, dt \tag{3.9}$$

This can be rewritten as

$$2^{jp}\int[\int \delta(y - exp(\sum_{l=1}^{n} z_l W_l + \sum_{l=1}^{p-n} s_{m_l} U_{m_l})\gamma(x,t))\phi(2^j z_1)\ldots\phi(2^j z_n)dz_1\ldots dz_n]$$

$$\times K_j(x,t)\phi(2^j s_{m_1})\ldots\phi(2^j s_{m_{p-n}})ds_{m_1}\ldots ds_{m_{p-n}}\, dt \tag{3.10}$$

We now do the dz integrations. Define $A(z)$ by

$$A(z_1,\ldots,z_n) = exp(\sum_{l=1}^{n} z_l W_l + \sum_{l=1}^{p-n} s_{m_l} U_{m_l})\gamma(x,t) \tag{3.11}$$

Of course, the function $A(z)$ also depends on j, x, t, and the s_{m_l}, but for notational convenience those variables are not mentioned explicitly. We change variables in the dz integral of (3.10) from z to $y' = A(z)$. Writing $\Phi(2^j z) = \phi(2^j z_1)\ldots\phi(2^j z_n)$, we get

$$L_j(x,y) = 2^{jp}\int[\int \delta(y-y')|detA^{-1}(y')|\Phi(2^j A^{-1}(y'))dy']K_j(x,t)\phi(2^j s_{m_1})\ldots\phi(2^j s_{m_{p-n}})$$

$$ds_{m_1}\ldots ds_{m_{p-n}}\, dt \tag{3.12}$$

In Lemma 1.2, the vector fields W_i were defined to be canonical coordinates on $B(x, C_0 2^{-j})$ for some large C_0. As a result, assuming C_0 was selected large enough, the vectors $s_{m_l}U_{m_l}$ appearing in (3.11) are small enough such that the map $A(z)$ is a perturbation of the privileged coordinate map in the sense that $|detA^{-1}(y')| \sim \frac{2^{-jn}}{|B(x,2^{-j})|}$. Therefore

$$|L_j(x,y)| < C\frac{2^{j(p-n)}}{|B(x,2^{-j})|}|\int |K_j(x,t)||\phi(2^j s_{m_1})|\ldots|\phi(2^j s_{m_{p-n}})|ds_{m_1}\ldots ds_{m_{p-n}}\, dt$$

$$< \frac{C}{|B(x,2^{-j})|} \tag{3.13}$$

This gives (3.5). Moving on to proving (3.6), we write $W_i^x L_j(x,y)$ as

$$W_i^x L_j(x,y) = 2^{jp}\int[\int W_i^x \delta(y - A(z))\phi(2^j z_1)\ldots\phi(2^j z_n)dz_1\ldots dz_n]$$

$$\times K_j(x,t)\phi(2^j s_{m_1})\ldots\phi(2^j s_{m_{p-n}})ds_{m_1}\ldots ds_{m_{p-n}}\, dt \tag{3.14}$$

The strategy will be to write $W_i^x\delta(y-A(z))$ in the form $\sum_q b_q(y,z)\partial_{z_q}\delta(y-A(z))$, do an integration by parts in (3.14), and then bound terms. To this end, observe that by the chain rule

$$\nabla_z[\delta(y-A(z))] = -\nabla_y\delta(y-A(z))dA(z)$$

This can be rewritten as

$$\nabla_y\delta(y-A(z)) = -\nabla_z[\delta(y-A(z))]dA(z)^{-1} \tag{3.15}$$

Taking the dot product of (3.15) with $W_i^x(y)$, we get

$$W_i^x\delta(y-A(z)) = -\nabla_z[\delta(y-A(z))]dA(z)^{-1}W_i^x(y) \tag{3.16}$$

The right-hand side of (3.16) is in the form $\sum_q b_q(y,z)\partial_{z_q}\delta(y-A(z))$. The functions $b_q(y,z)$ are readily determined via Cramer's rule. Namely, let $dA_i^q(y,z)$ denote the matrix obtained by replacing the qth column of $dA(z)$ by $W_i^x(y)$. Then by applying Cramer's rule to (3.16) we get

$$W_i^x\delta(y-A(z)) = -\sum_q \frac{det\, dA_i^q(y,z)}{det\, dA(z)}\partial_{z_q}\delta(y-A(z)) \tag{3.17}$$

Thus to prove (3.6), we must bound each term

$$-2^{jp}\int\Big[\int \frac{det\, dA_i^q(y,z)}{det\, dA(z)}\partial_{z_q}\delta(y-A(z))\phi(2^jz_1)\ldots\phi(2^jz_n)dz_1\ldots dz_n\Big]$$

$$\times K_j(x,t)\phi(2^js_{m_1})\ldots\phi(2^js_{m_{p-n}})ds_{m_1}\ldots ds_{m_{p-n}}\,dt \tag{3.18}$$

We do the integration by parts in (3.18) in the z_q variable. We get several terms, depending on where the derivative lands. First suppose it lands on the $\phi(2^jz_q)$ factor. In this case, the term of (3.18) is bounded by

$$2^{jp}\int\Big[\int \Big|\frac{det\, dA_i^q(y,z)}{det\, dA(z)}\Big|\delta(y-A(z))|\phi(2^jz_1)|\ldots|2^j\phi'(2^jz_q)|\ldots|\phi(2^jz_n)|\,dz_1\ldots dz_n\Big]$$

$$\times|K_j(x,t)||\phi(2^js_{m_1})|\ldots|\phi(2^js_{m_{p-n}})|\,ds_{m_1}\ldots ds_{m_{p-n}}\,dt \tag{3.19}$$

By Lemma 2.2 part a), each column of $det\, dA(z)$ can be written as a linear combination of $W_i^x(A(z))$ with bounded coefficients. Due to the delta function in (3.19), we can set $A(z)=y$ and say that each column of $det\, dA(z)$ is a linear combination of $W_i^x(y)$ with bounded coefficients. As a result, each column of $det\, dA_i^q(y,z)$ is also a linear combination of $W_i^x(y)$ with bounded coefficients. Since the W_i^x are privileged coordinates, it follows that

$$|det\, dA_i^q(y,z)| < C\frac{|B(x,2^{-j})|}{2^{-jn}} \tag{3.20}$$

Furthermore, as in (3.13), the fact that $A(z)$ is a perturbation of privileged coordinates implies that

$$|det\, dA(z)| > C' \frac{|B(x, 2^{-j})|}{2^{-jn}} \tag{3.21}$$

Hence (3.19) is bounded by

$$2^{jp} \int [\int \delta(y - A(z))|\phi(2^j z_1)|...|2^j \phi'(2^j z_q)|...|\phi(2^j z_n)|dz_1 \dots dz_n]$$

$$\times |K_j(x,t)||\phi(2^j s_{m_1})|...|\phi(2^j s_{m_{p-n}})|ds_{m_1}...ds_{m_{p-n}}\, dt \tag{3.22}$$

If we change variables from z to $A(z)$ like we did in (3.5), and then argue like before, the dz integration gives a factor of $C\frac{2^{-jn+j}}{|B(x,2^{-j})|}$. The extra 2^j factor comes from the $2^j \phi'(2^j z_q)$ factor. (3.22) is thus at most

$$C\frac{2^{j(p-n+1)}}{|B(x, 2^{-j})|} \int |K_j(x,t)||\phi(2^j s_{m_1})|...|\phi(2^j s_{m_{p-n}})|ds_{m_1} \dots ds_{m_{p-n}}\, dt \tag{3.23}$$

$$< C\frac{2^j}{|B(x, 2^{-j})|} \tag{3.24}$$

This is the estimate we seek for this term. We next consider the term where the derivative lands on the $det\, dA_i^q(y, z)$ factor. There are two differences between this term and the previous term. First, instead of having a factor of $2^j \phi'(2^j z_q)$ appearing, we have a factor of $\phi(2^j z_q)$ showing up. This changes the estimate by a factor of $C2^{-j}$. Secondly, instead of having the $det\, dA_i^q(y, z)$ in the equation, we have $\partial_{z_q} det\, dA_i^q(y, z)$. This changes the equation by a factor of $\frac{\partial_{z_q} det\, dA_i^q(y,z)}{det\, dA_i^q(y,z)}$, whose absolute value by the Bernstein inequalities is bounded by the reciprocal of the radius of the z-domain. Since this radius is $C2^{-j}$, we get a net gain of $C2^j$ for this factor. This cancels out the $C2^{-j}$ from the $\phi(2^j z_q)$, and we once again get the estimate (3.24).

The observant reader might ask how we know we can use the Bernstein inequalities here, since they assume that $det\, dA_i^q(y, z)$ is of finite-type uniformly in the s, x, y, and z variables. By continuity, one needs just that $det\, dA_i^q(y, z)$ is of finite type at $s = x = y = z = 0$. Since the W_i^x give privileged coordinates, we know that $det\, dA(z)$ is uniformly of finite type. If it so happens that $W_i^x(y)$ is orthogonal, to infinite order, to the $n-1$ columns of $det\, dA(z)$ other than the qth column at $x = y = s = z = 0$, then we just write

$$det\, dA_i^q(y, z) = det\, dB_i^q(y, z) + det\, dA_{i'}^q(y, z) \tag{3.25}$$

Here $dB_i^q(y, z)$ is the matrix obtained by replacing the qth column of $dA(z)$ by $W_i^x(y) - U_{i'}^x(y)$ for an appropriate i'. The functions $det\, dB_i^q(y, z)$ and $det\, dA_{i'}^q(y, z)$ will then be of finite-type uniformly, so if we plug (3.25) into (3.18) and consider each of the two terms obtained separately, the argument of the last paragraph will hold in both cases.

Lastly, we consider the term of (3.18) where the z_q derivative lands on $\frac{1}{det\, dA(z)}$. In this case, there are again two differences between this term and the first one. First, again we have $\phi(2^j z_q)$ in place of $2^j\phi'(2^j z_q)$, resulting in a factor of $C2^{-j}$ over the first term. Secondly, after applying the derivative, the function $\frac{1}{det\, dA(z)}$ factor becomes $\frac{-\partial_{z_q} det\, dA(z)}{(det\, dA(z))^2}$. As a result, we get an additional factor of $\frac{-\partial_{z_q} det\, dA(z)}{det\, dA(z)}$, which by the Bernstein inequalities is bounded in absolute value by $C2^j$. Hence we again get the estimate (3.24). This completes the proof of the left inequality of (3.6). Notice that this time there is no issue about whether we can use the Bernstein inequalities; since the W_i's give privileged coordinates the function $det\, dA(z)$ will necessarily be of finite type.

Proving the right-hand inequality of (3.6) is much the same as proving the left-hand inequality, looking at the adjoint of S_j instead of S_j. Let us write out what this adjoint operator is. For functions $f(x)$ and $g(x)$, we have

$$\int S_j f(x) g(x)\, dx =$$

$$2^{jp}\int f(exp(s_1U_1+...+s_pU_p)\gamma(x,t))g(x)K_j(x,t)\phi(2^js_1)\,...\,\phi(2^js_p)\,ds_1\,...\,ds_p\,dt\,dx \quad (3.26)$$

Changing variables from x to $\gamma(x,t)$, we get

$$2^{jp}\int f(exp(s_1U_1+...+s_pU_p)x)g(\gamma^*(x,t))K_j^0(\gamma^*(x,t),t)\phi(2^js_1)\,...\,\phi(2^js_p)\,ds_1\,...\,ds_p\,dt\,dx \quad (3.27)$$

Here K_j^0 is K_j multiplied by the smooth Jacobian of the coordinate change. Next, we change variables from x to $exp(s_1U_1+...+s_pU_p)x$ in (3.27). Once this is done, (3.27) becomes

$$2^{jp}\int f(x)g(\gamma^*(exp(-s_1U_1-...-s_pU_p)x,t))K_j^1(\gamma^*(exp(-s_1U_1-...-s_pU_p)x,t),s,t)$$
$$\times\phi(2^js_1)\,...\,\phi(2^js_p)\,ds_1\,...\,ds_p\,dt\,dx \quad (3.28)$$

Here K_j^1 is K_j^0 times the smooth Jacobian of the coordinate change. As a result,

$$S_j^*g(x) = 2^{jp}\int g(\gamma^*(exp(-s_1U_1-...-s_pU_p)x,t))K_j^1(\gamma^*(exp(-s_1U_1-...-s_pU_p)x,t),s,t)$$
$$\times\phi(2^js_1)\,...\,\phi(2^js_p)\,ds_1\,...\,ds_p\,dt \quad (3.29)$$

This may look different from (3.3), but the same arguments that gave the left inequality of (3.6) still work here. The two (minor) differences worth mentioning are as follows: First, the kernel $K(x,t)$ is replaced by $K_j^1(\gamma^*(exp(-s_1U_1-...-s_pU_p)x,t),s,t)$. This is relevant since when an z_q integration by parts occurs, the derivative might land on this factor. However, the differentiability condition (1.20) coupled with part b) of Lemma 2.2 ensures that taking this derivative incurs a factor of at most $2^{(1-\delta)j}$ and won't affect the estimates obtained. The other difference is that instead of having $f(exp(s_1U_1 + ... + s_pU_p)\gamma(x,t))$ in the integral we have $g(\gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t))$. One can still use the coordinates W_i and do the integration by parts. The definition of $A(z)$ changes but not the estimates; this time one uses part b) of Lemma 2.2 in place of part a) to justify them. This completes the proof of (3.6).

Proceeding now to the cancellation conditions (3.7), observe that by (3.3)

$$S_j(1)(x) = 2^{jp}\int K_j(x,t)\phi(2^js_1)\dots\phi(2^js_p)\,ds_1\dots ds_p\,dt \tag{3.30}$$

Since the function ϕ was defined to have integral 1, we have

$$S_j(1)(x) = \int K_j(x,t)dt = T_j(1)(x)$$

This gives half of (3.7). As for the other half, we first determine an expression for $T_j^*g(x)$. To do this, observe that

$$\int T_jf(x)g(x)\,dx = \int f(\gamma(x,t))K_j(x,t)g(x)\,dx\,dt \tag{3.31}$$

Changing variables from x to $\gamma(x,t)$, this becomes

$$\int f(x)K_j^0(\gamma^*(x,t),t)g(\gamma^*(x,t))\,dx\,dt \tag{3.32}$$

Consequently,

$$T_j^*g(x) = \int g(\gamma^*(x,t))K_j^0(\gamma^*(x,t),t)\,dt \tag{3.33}$$

Next, observe that by (3.29)

$$S_j^*(1)(x) = 2^{jp}\int K_j^1(\gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t),s,t)\phi(2^js_1)...\phi(2^js_p)\,ds_1\dots ds_p\,dt \tag{3.34}$$

Part b) of Lemma 2.2 and the differentiability conditions (1.20) imply that for each i

$$|\partial_{s_i}K_j^1(\gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t),s,t)| < C2^{(k+1-\delta)j} \tag{3.35}$$

(Recall that k denotes the t-dimension). As a result, since each $|s_i| < 2^{-j}$ we have

$$|K_j^1(\gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t),s,t) - K_j^1(\gamma^*(x,t),0,t)| < C2^{(k-\delta)j} \tag{3.36}$$

Consequently,

$$S_j^*(1) = 2^{jp}\int K_j^1(\gamma^*(x,t),0,t)\phi(2^js_1)\dots\phi(2^js_p)\,ds_1\dots ds_p\,dt + O(2^{-\delta j}) \tag{3.37}$$

Recall that K_j^1 is equal to K_j^0 times the Jacobian of a coordinate change, and that this Jacobian is equal to the identity when $s = 0$. As a result, for the values of s and t in question, the Jacobian determinant is $1 + O(2^{-j})$ and we get

$$S_j^*(1) = 2^{jp} \int K_j^0(\gamma^*(x,t),t)\phi(2^j s_1) \ldots \phi(2^j s_p)\, ds_1 \ldots ds_p\, dt + O(2^{-j}) + O(2^{-\delta j})$$

$$= \int K_j^0(\gamma^*(x,t),t)dt + O(2^{-j}) + O(2^{-\delta j}) \tag{3.38}$$

The left term of (3.38) is exactly $T_j^*(1)$, and so the right-hand inequality of (3.7) follows. This completes the proof of Lemma 3.1.

Since $S_j(x,y)$ is supported for $d(x,y) < C2^{-j}$, (3.6) implies the following corollary.

Corollary 3.2:

$$\int |L_j(x,y)|\, dx,\ \int |L_j(x,y)|\, dy < C \tag{3.39}$$

4. Main Estimates

To prove Theorem 1.1, it suffices to show that $S - T$ is bounded on L^2:

Lemma 4.1: Suppose $S - T$ is bounded on L^2. Then Theorem 1.1 holds.

Proof: Assume that $S - T$ is bounded on L^2. We must show that if T satisfies $(1.23a)$ – $(1.23b)$, then T is bounded on L^2, as the converse direction is trivial. So assume T satisfies $(1.23a)$ – $(1.23b)$. By the boundedness of $S - T$ on L^2, the operator S satisfies $(1.22a)$ – $(1.22b)$.

Note that the kernel $L_j(x,y)$ of each S_j in the sum $S = \sum_j S_j$ is supported on $d(x,y) < C2^{-j}$ and satisfies the estimates (3.5) and (3.6). As a result, the usual $T(1)$ theorem (see [C]) implies that S is bounded on L^2. By the assumption that $S - T$ is bounded on L^2, we conclude that T is also bounded on L^2. This completes the proof of Lemma 4.1.

We now proceed to the main argument, proving the boundedness of $S - T$ on L^2. We use almost-orthogonality on the sum $S - T = \sum_j S_j - T_j$. In particular, we will prove the estimates

$$||S_j - T_j||_{L^2 \to L^2} < C \tag{4.1}$$

$$||(S_i - T_i)(S_j^* - T_j^*)||_{L^2 \to L^2},\ ||(S_i^* - T_i^*)(S_j - T_j)||_{L^2 \to L^2} < C2^{-\mu|i-j|} \tag{4.2}$$

Here μ is a small positive constant. We begin with the straightforward proof of (4.1). We will show that for all j we have $||T_j||$, $||S_j|| < C$. Observe that for each t, the map

$x \to \gamma(x,t)$ is a perturbation of the identity map. As a result, the map $f(x) \to f(\gamma(x,t))$ is bounded on L^2, uniformly for small $|t|$. Consequently, by (3.1) we have

$$||T_j|| < C \int |K(x,t)|\, dt < C' \tag{4.3}$$

Similarly, for each s and t, the map $f(x) \to f(exp(s_1U_1 + ... + s_pU_p)\gamma(x,t))$ is bounded on L^2, uniformly for small $|s|$ and $|t|$. As a result, (3.3) implies that

$$||S_j|| < C2^{jp} \int |K_j(x,t)\phi(2^j s_1) ... \phi(2^j s_p)|\, ds_1 ... ds_p\, dt < C'' \tag{4.4}$$

Equations (4.3) and (4.4) imply (4.1).

We proceed now to the orthogonality estimates (4.2). Since the two estimates are proven the same way, we will restrict our attention to the estimates on $(S_i^* - T_i^*)(S_j - T_j)$. Replacing this operator by its adjoint and switching the letters i and j if necessary, we may assume that $i > j$. For a small constant ϵ dictated by our subsequent arguments, define the operator S_j' by

$$S_j' f(x) = 2^{p(j+\epsilon|i-j|)} \int f(exp(s_1U_1 + ... + s_pU_p)\gamma(x,t))K_j(x,t)$$

$$\times\phi(2^{j+\epsilon|i-j|}s_1) ... \phi(2^{j+\epsilon|i-j|}s_p)\, ds_1 ... ds_p\, dt \tag{4.5}$$

In other words, the definition of S_j' is like that of S_j in (3.3), except the powers 2^j are replaced by $2^{j+\epsilon|i-j|}$. Thus the operator S_j' can be viewed as being slightly closer to the operator T_j than S_j is. Let $L_j'(x,y)$ denote the kernel of $S_j'(x,y)$. The exact same arguments that produced (3.5) and (3.6) now give

$$|L_j'(x,y)| < C2^{n\epsilon|i-j|}|B(x,2^{-j})|^{-1} \tag{4.6}$$

$$|W_i^x L_j'(x,y)|,\ |W_i^y L_j'(x,y)| < C2^{(n+1)\epsilon|i-j|+j}|B(x,2^{-j})|^{-1} \tag{4.7}$$

We write the operator $(S_i^* - T_i^*)(S_j - T_j)$ as $O_1 + O_2$, where

$$O_1 = (S_i^* - T_i^*)(S_j - S_j') \tag{4.8a}$$

$$O_2 = (S_i^* - T_i^*)(S_j' - T_j) \tag{4.8b}$$

The operators O_1 and O_2 will be treated separately, starting with O_1.

Analysis of the operator O_1

Let $\{\delta_n\}$ be a sequence of functions converging to the delta function, and let $T_i^{*(n)}$ be the operator with kernel $M_i^{*(n)}(x,y) =$

$$2^{ip} \int \delta_n(y - \gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t))K_i^1(\gamma^*(exp(-s_1U_1 - ... - s_pU_p)x,t),s,t)$$

$$\times\phi(2^i s_1)\ldots\phi(2^i s_p)\,ds_1\ldots ds_p\,dt \tag{4.9}$$

We will prove that for some $\mu > 0$, uniformly in n we have

$$||(S_i^* - T_i^{*(n)})(S_j - S_j')|| < C2^{-\mu|i-j|} \tag{4.10}$$

Taking the limit of this as n goes to infinity gives the desired estimate for O_1:

$$||(S_i^* - T_i^*)(S_j - S_j')|| < C2^{-\mu|i-j|} \tag{4.11}$$

Observe that $\int M_i^{*(n)}(x,y)\,dy = T_i^{*(n)}(1)(x) = T_i^*(1)(x)$, so that if L_i^* denotes the kernel of S_i^*, by (3.7) we have

$$|\int (L_i^*(x,y) - M_i^{*(n)}(x,y))\,dy| = |S_i^*(1)(x) - T_i^*(1)(x)| < C2^{-\delta i}$$

$$< C'2^{-\delta|i-j|} \tag{4.12}$$

Similar to the proof of (4.1), it is a straightforward consequence of the definition (4.9) of $M_i^{*(n)}$ that

$$\int |M_i^{*(n)}(x,y)|\,dx,\ \int |M_i^{*(n)}(x,y)|\,dy < C \tag{4.13}$$

As a result, by Corollary 3.2 we also have

$$\int |L_i^*(x,y) - M_i^{*(n)}(x,y))|\,dx,\ \int |L_i^*(x,y) - M_i^{*(n)}(x,y))|\,dy < C \tag{4.14}$$

The estimate (4.10) will now be proved basically the same way that orthogonality estimates are proved for singular integral operators with a good cancellation condition, and equations (4.12) and (4.14) will be key in accomplishing this. We write $S_j'' = S_j - S_j'$, and let $L_j''(x,y)$ be the kernel of S_j''. Then by (3.5) – (3.6) and (4.6) – (4.7) we have

$$|L_j''(x,y)| < C2^{n\epsilon|i-j|}|B(x,2^{-j})|^{-1} \tag{4.15}$$

$$|W_i^x L_j''(x,y)|,\ |W_i^y L_j''(x,y)| < C2^{(n+1)\epsilon|i-j|+j}|B(x,2^{-j})|^{-1} \tag{4.16}$$

The kernel of $O_1 = (S_i^* - T_i^{*(n)})S_j''$ is given by

$$N(x,y) = \int (L_i^*(x,z) - M_i^{*(n)}(x,z))L_j''(z,y)\,dz \tag{4.17}$$

We will use Schur's test to bound O_1, by showing that for some $\mu > 0$ we have

$$\int |N(x,y)|\,dy < C2^{-\mu|i-j|} \tag{4.18}$$

This suffices to prove (4.10) by Schur's test since the estimate $\int |N(x,y)|\,dx < C$ directly from (4.13) and (3.39). Rewrite $N(x,y)$ as

$$\int (L_i^*(x,z) - M_i^{*(n)}(x,z))L_j''(x,y)\,dz + \int (L_i^*(x,z) - M_i^{*(n)}(x,z))(L_j''(z,y) - L_j''(x,y))\,dz \tag{4.19}$$

Then

$$\int |N(x,y)|\,dy \leq \int \left| \int (L_i^*(x,z) - M_i^{*(n)}(x,z))L_j''(x,y)\,dz \right| dy$$
$$+ \int\int |(L_i^*(x,z) - M_i^{*(n)}(x,z))(L_j''(z,y) - L_j''(x,y))|\,dz\,dy \tag{4.20}$$

By (4.12), the first term in (4.20) is bounded by

$$C2^{-\delta|i-j|}\int |L_j''(x,y)|\,dy \tag{4.21}$$

By (4.15) and the fact that $|L_j''(x,y)|$ is supported on $d(x,y) < C2^{-j}$, (4.21) is at most

$$C2^{(n\epsilon-\delta)|i-j|} \tag{4.22}$$

As long as $\epsilon < \frac{\delta}{n}$, this gives a term bounded by $C2^{-\mu|i-j|}$ for some positive μ, which is the desired estimate. Moving on to the second term of (4.20), observe that in order for the factor $L_i^*(x,z) - M_i^{*(n)}(x,z)$ to be nonzero, we must have $d(x,z) < C2^{-i}$. Consequently for such x and z, by the differentiability conditions (4.16), we have

$$|L_j''(z,y) - L_j''(x,y)| < C\frac{2^{(n+1)\epsilon|i-j|-|i-j|}}{|B(x,2^{-j})|} \tag{4.23}$$

As a result, as long as ϵ was chosen such that $(n+1)\epsilon < \frac{1}{2}$, we have

$$|L_j''(z,y) - L_j''(x,y)| < C\frac{2^{-|i-j|/2}}{|B(x,2^{-j})|} \tag{4.24}$$

Substituting this back into the second term of (4.20), we get a term bounded by

$$C\frac{2^{-|i-j|/2}}{|B(x,2^{-j})|}\int |L_i^*(x,z) - M_i^{*(n)}(x,z)|\,dy\,dz \tag{4.25}$$

Here the dy integral is only over y for which the factors $L_i^*(x,z) - M_i^{*(n)}(x,z)$ and $L_j''(z,y) - L_j''(x,y)$ are nonzero, which is a subset of the y for which $d(x,y) < C2^{-j}$. Hence (4.25) is bounded by

$$C2^{\frac{-|i-j|}{2}}\int |L_i^*(x,z) - M_i^{*(n)}(x,z)|\,dz \tag{4.26}$$

By (4.14) this is at most

$$C2^{\frac{-|i-j|}{2}} \tag{4.27}$$

Consequently, the second term of (4.20) contributes at most $C2^{-\frac{|i-j|}{2}}$ to $\int |N(x,y)|\,dy$. We have now shown that both terms of (4.20) contribute at most $C2^{-\mu|i-j|}$ to $\int |N(x,y)|\,dy$ for some $\mu > 0$. As a result, (4.18) holds and by Schur's lemma we have

$$||(S_i^* - T_i^{*(n)})(S_j - S_j')|| < C2^{-\frac{\mu|i-j|}{2}} \tag{4.28}$$

This gives (4.10), and taking limits as n goes to infinity gives (4.11) (replacing $\frac{\mu}{2}$ by μ. This is the desired estimate for the operator O_1.

Analysis of the operator O_2. We now analyze $O_2 = (S_i^* - T_i^*)(S_j' - T_j)$, with the goal of proving $||O_2|| < C2^{-\mu|i-j|}$ for some $\mu > 0$; this will complete the proof of Theorem 1.1. By (4.1), we have $||S_i^* - T_i^*|| = ||S_i - T_i|| < C$, so it suffices to show that

$$||S_j' - T_j|| < C2^{-\mu|i-j|} \tag{4.29}$$

For a p-tuple $s = (s_1, ..., s_p)$, as in (2.5) let $\gamma_s(x,t) = exp(s_1U_1 + ... + s_pU_p)\gamma(x,t)$. We define the operator T_j^s by

$$T_j^s f(x) = \int f(\gamma_s(x,t))K_j(x,t)\,dt \tag{4.30}$$

Lemma 4.2: Suppose we can prove that for some $\alpha > 0$, for $|s| < Cr$ we have

$$||T_j^s - T_j|| < C|s2^j|^\alpha \tag{4.31}$$

Then (4.29) (and therefore Theorem 1.1) holds.

Proof: $(S_j' - T_j)f(x)$ can be rewritten as

$$2^{p(j+\epsilon|i-j|)}\int f(\gamma_s(x,t))K_j(x,t)\phi(2^{j+\epsilon|i-j|}s_1)\,...\,\phi(2^{j+\epsilon|i-j|}s_p)\,ds_1\,...\,ds_p\,dt$$

$$-2^{p(j+\epsilon|i-j|)}\int[\int f(\gamma(x,t))K_j(x,t)\,dt]\phi(2^{j+\epsilon|i-j|}s_1)\,...\,\phi(2^{j+\epsilon|i-j|}s_p)\,ds_1\,...\,ds_p \tag{4.32}$$

$$= 2^{p(j+\epsilon|i-j|)}\int[\int f(\gamma_s(x,t))K_j(x,t)\,dt - \int f(\gamma(x,t))K_j(x,t))\,dt]$$

$$\times\phi(2^{j+\epsilon|i-j|}s_1)\,...\,\phi(2^{j+\epsilon|i-j|}s_p)\,ds_1\,...\,ds_p$$

$$= 2^{p(j+\epsilon|i-j|)}\int(T_j^s f(x) - T_jf(x))\phi(2^{j+\epsilon|i-j|}s_1)...\phi(2^{j+\epsilon|i-j|}s_p)\,ds_1\,...\,ds_p \tag{4.33}$$

In other words, we have

$$S'_j - T_j = 2^{p(j+\epsilon|i-j|)} \int (T_j^s - T_j)\phi(2^{j+\epsilon|i-j|}s_1)...\phi(2^{j+\epsilon|i-j|}s_p)\, ds_1 \ldots ds_p \tag{4.34}$$

Consequently, if (4.31) holds, then

$$||S'_j - T_j|| \leq 2^{p(j+\epsilon|i-j|)} \int |s|^\alpha \phi(2^{j+\epsilon|i-j|}s_1)...\phi(2^{j+\epsilon|i-j|}s_p)\, ds_1 \ldots ds_p \tag{4.35}$$

In the support of $\phi(2^{j+\epsilon|i-j|}s_1) \ldots \phi(2^{j+\epsilon|i-j|}s_p)$, we have $|s2^j|^\alpha < C2^{-\epsilon\alpha|i-j|}$, so

$$||S'_j - T_j|| \leq C2^{-\epsilon\alpha|i-j|} \int 2^{p(j+\epsilon|i-j|)}\phi(2^{j+\epsilon|i-j|}s_1) \ldots \phi(2^{j+\epsilon|i-j|}s_p)\, ds_1 \ldots ds_p$$

$$= C2^{-\epsilon\alpha|i-j|} \tag{4.36}$$

This implies (4.29) for $\mu = \epsilon\alpha$ and we are done.

Hence our goal is now to prove that $||T_j^s - T_j|| < C|s2^j|^\alpha$ for some $\alpha > 0$. The argument from here on is all but identical to the argument in [G3] proving $||T_j^y - T_j|| < C(2^{jk}|y|)^\epsilon 2^{-j\delta}$ (in the notation of section 5 of [G3]). Namely, the N-fold iterate $(T_j^s - T_j)^*(T_j^s - T_j)...(T_j^s - T_j)^*(T_j^s - T_j)$ can be written in the form

$$\sum_{l=1}^{2^N} \pm T_j^{s_{N,l}}(T_j^{s_{N-1,l}})^*...T_j^{s_{2,l}}(T_j^{s_{1,l}})^* \tag{4.37}$$

Here each $s_{i,l}$ is either s or 0. Observe that in the sum (4.37), 2^{N-1} terms have a plus sign and 2^{N-1} terms have a minus sign. As a result, (4.37) is exactly

$$\sum_{l=1}^{2^N} \pm T_j^{s_{N,l}}(T_j^{s_{N-1,l}})^*...T_j^{s_{2,l}}(T_j^{s_{1,l}})^* - T_jT_j^*...T_jT_j^* \tag{4.38}$$

As a result, in analogy to Lemma 5.2 of [G3] we have

Lemma 4.3: Suppose we can show that for some N, for each i and for any $(s_1, ...s_N)$ with $|s| < C2^{-j}$ we have

$$||\partial_{s_i}(T_j^{s_N}T_j^{s_{N-1}*}...T_j^{s_2}T_j^{s_1*})|| < C2^j \tag{4.39}$$

Then each term of (4.38) has operator norm at most $C|s2^j|$. Consequently, $||T_j^s - T_j|| < C|s2^j|^{\frac{1}{N}}$ and Theorem 1.1 follows.

In the notation of $(2.5)-(2.8)$, for a given function $f(x)$, $\partial_{s_i}(T_j^{s_N}T_j^{s_N*}...T_j^{s_2}T_j^{s_1*})f(x)$ can be written as

$$\int \partial_{s_i}[f(\beta^N_{s_1,...,s_N}(x,t^1,...,t^N))\tilde{K}_j(x,s_1,...s_N,t^1,...,t^N)]\,dt^1\,...\,dt^N \tag{4.40}$$

Here $\tilde{K}_j(x,s_1,...,s_N,t^1,...,t^N)$ is a function that satisfies

$$|\tilde{K}_j(x,s_1,...,s_N,t^1,...,t^N)|<C2^{kjN},\ |\nabla_{s,t}\tilde{K}_j(x,s_1,...,s_N,t^1,...,t^N)|<C2^{kjN+j} \tag{4.41}$$

The analogue in [G3] to (4.40) is equation (5.13) of that paper, whose derivative is (using the notation of that paper)

$$\int \partial_{y_i^j}[f(\beta^N_{y^1,...,y^N}(x,t_1,...,t_N))k_j(t_1)...k_j(t_N)\xi(x,y^1,...,y^N,t_1,...,t_N)]\,dt_1\,...\,dt_N \tag{4.42}$$

The analysis of (4.40) is nearly identical to (5.13). Rather than rehash a long technical argument essentially verbatim, we will highlight three (minor) differences. First, the function $\tilde{K}_j(x,s_1,...s_N,t^1,...,t^N)$ and its first derivatives are a factor of $C2^{j\delta N}$ larger than the function $k_j(t_1)...k_j(t_N)\xi(x,y^1,...,y^N,t_1,...,t_N)$ and its corresponding derivatives (here δ is as in [G3]). Thus in the present situation we incur a factor of $C2^{j\delta N}$ over the estimates of [G3]. Secondly, we must insert Lemma 2.3 of this paper instead of the analogous Corollary 2.6 of [G3]. Lastly, the range of y in [G3] was $|y|<2^{-jk}$, while the scaling in this paper is such that the range of s is $|s|<2^{-j}$. So the correct estimate for the operator (4.40) will be what one obtains by taking the estimate for (4.42) in [G3], replacing the 2^{jk} by 2^j, and multiplying the result by $2^{j\delta N}$. The result is exactly (4.39) and we are done.

5. Final Comments

For $1<p<\infty$, L^p boundedness of singular Radon transforms under the assumptions $(1.23a)-(1.23b)$ can be proven from the results of this paper using standard methods (see [G1]-[G2] [PS] for example), as can L^p boundedness of the associated maximal function. Namely, in section 4 we showed that under the assumptions $(1.23a)-(1.23b)$, $S-T$ is bounded on L^2. Consequently S and S^* are also bounded on bump functions, and the $T(1)$ theorem then implies that S is bounded on L^2. So by the Calderon-Zygmund theory, S is bounded on each L^p for $1<p<\infty$ as well. If we can prove that $S-T$ is bounded on L^p, it will then follow that T is bounded on L^p. To accomplish this, one defines operators S_{ij} by $S_{ij}f(x)=$

$$2^{(i+j)p}\int f(exp(s_1U_1+...+s_pU_p)\gamma(x,t))K_j(x,t)\phi(2^{i+j}s_1)\,...\,\phi(2^{i+j}s_p)\,ds_1\,...\,ds_p\,dt$$

$$-2^{(i+j+1)p}\int f(exp(s_1U_1+...+s_pU_p)\gamma(x,t))K_j(x,t)\phi(2^{i+j+1}s_1)\,...\,\phi(2^{i+j+1}s_p)\,ds_1\,...\,ds_p\,dt \tag{5.1}$$

Define S^i by

$$S^i = \sum_j S_{ij} \tag{5.2}$$

Observe that we have

$$S - T = \sum_{i \geq 0} S^i \tag{5.3}$$

Straightforward adaptations of the arguments giving (4.2) show that for some $\eta > 0$ we have

$$||S^i||_{L^2 \to L^2} < C2^{-\eta i} \tag{5.4}$$

Using the nonisotropic Calderon-Zygmund theory (see for example Ch. 1 of [S]), one obtains that for $1 < p < \infty$ and a fixed $0 < \eta' < \eta$ we have

$$||S^i||_{L^p \to L^p} < C_p 2^{-\eta' i} \tag{5.5}$$

Summing this in i gives

$$||S - T||_{L^p \to L^p} < C_p' \tag{5.6}$$

We conclude that under $(1.23a) - (1.23b)$ we have L^p boundedness of T for $1 < p < \infty$.

The boundedness of the maximal function associated to T follows from the results of this paper by applying standard square-function methods to the difference between the maximal function associated with T and the maximal function associated to the singular integral S, whose L^p boundedness follows from the Calderon-Zygmund theory.

6. Acknowledgement

This paper can be viewed as the culmination of on-again off-again work stretching back nearly a decade to the beginning of the author's graduate school days at Princeton. As a result, the author owes a great debt of gratitude to Eli Stein for his generous assistance during the first few years of this effort, which resulted in [G1].

7. References

[C] M. Christ, *A T(b) theorem with remarks on analytic capacity*, Colloq. Math. **50/51** (1990), 601-628.

[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 Caldern-Zygmund operators.* Ann. of Math. (2) **120** (1984), no. 2, 371–397.

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

[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, *An analogue to a theorem of Fefferman and Phong for averaging operators along curves with singular fractional integral kernel*, submitted.

[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.

[NS] A. Nagel, E. M. Stein, *Differentiable control metrics and scaled bump functions*, J. Diff. Geom. **(57)** (2001) 457-491.

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

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

[Po] D. Pottinton, *A T(1) theorem for singular Radon transforms*, Princeton Univ. PhD thesis, 1998.

[S] E. Stein, *Harmonic analysis; real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematics Series Vol. 43, Princeton University Press, Princeton, NJ, 1993.

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

[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.