

MULTI-PARAMETER SINGULAR RADON TRANSFORMS

ELIAS M. STEIN AND BRIAN STREET

ABSTRACT. The purpose of this announcement is to describe a development given in a series of forthcoming papers by the authors that concern operators of the form

$$f \mapsto \psi(x) \int f(\gamma_t(x)) K(t) dt,$$

where $\gamma_t(x) = \gamma(t, x)$ is a C^∞ function defined on a neighborhood of the origin in $(t, x) \in \mathbb{R}^N \times \mathbb{R}^n$ satisfying $\gamma_0(x) \equiv x$, $K(t)$ is a “multi-parameter singular kernel” supported near $t = 0$, and ψ is a cutoff function supported near $x = 0$. This note concerns the case when K is a “product kernel.” The goal is to give conditions on γ such that the above operator is bounded on L^p for $1 < p < \infty$. Associated maximal functions are also discussed. The “single-parameter” case when K is a Calderón-Zygmund kernel was studied by Christ, Nagel, Stein, and Wainger. The theory here extends these results to the multi-parameter context and also deals effectively with the case when γ is real-analytic.

1. Introduction

The purpose of this note is to announce the results from a three part series of papers by the authors [18, 16, 17], and to give an overview of the main ideas in a somewhat simpler context. The object is to study the L^p ($1 < p < \infty$) boundedness of operators of the form

$$(1.1) \quad Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt,$$

where $\gamma_t(x) = \gamma(t, x) : \mathbb{R}_0^N \times \mathbb{R}_0^n \rightarrow \mathbb{R}^n$ is a C^∞ function with $\gamma_0(x) \equiv x$, $\psi \in C_0^\infty(\mathbb{R}^n)$ is supported on a small neighborhood of $0 \in \mathbb{R}^n$, and K is an appropriate “multi-parameter” singular kernel supported on a small neighborhood of $0 \in \mathbb{R}^N$. Here we have written \mathbb{R}_0^N for a neighborhood of the origin in \mathbb{R}^N . The case when K is a Calderón-Zygmund kernel (that is, when the number of parameters, ν , equals 1) was studied in the paper of Christ, Nagel, Stein, and Wainger [7]. Given the complexity of the formulation and proof of our general results, we will here restrict ourselves to the case when K is a “product kernel” and give only an outline of the main points of the argument contained in [18, 16, 17]. Some of these results, valid in a more general setting, are indicated in Section 9.

In the definition of a product kernel given in Section 3, we fix a decomposition of \mathbb{R}^N into ν factors, $\mathbb{R}^N = \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_\nu}$, and write $t = (t_1, \dots, t_\nu)$. The main goal of the papers [18, 16] is to give conditions on γ for which the operator T given by (1.1) is bounded on L^p ($1 < p < \infty$) for *every* such product kernel (supported on a

Received by the editors December 10, 2010.

sufficiently small neighborhood of $0 \in \mathbb{R}^N$). Under these same conditions, one also obtains the L^p boundedness of the corresponding maximal function:

$$(1.2) \quad \mathcal{M}f(x) = \psi(x) \sup_{0 < \delta_1, \dots, \delta_\nu < 1} \int_{|t| \leq 1} |f(\gamma_{\delta_1 t_1, \dots, \delta_\nu t_\nu}(x))| dt.$$

To simplify the presentation we limit ourselves to the special case when γ is given by

$$\gamma_t(x) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha \right) x,$$

where the X_α are C^∞ vector fields. I.e., γ is an exponential of a *finite* sum of vector fields. In our cited work, more general γ are considered. See Section 9 for some comments on this.

In [17], it is shown that when γ is assumed to be real analytic, many of the assumptions necessary to carry out to proofs in our work hold automatically. This is discussed in Section 2.

1.1. Statement of main results. In this section, we state the main results from this paper. Consider a C^∞ function

$$(1.3) \quad \gamma_t(x) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha \right) x,$$

where the X_α are C^∞ vector fields on \mathbb{R}^n . Here, and in what follows, we will be restricting attention to (t, x) in some small neighborhood of the origin $\mathbb{R}^N \times \mathbb{R}^n$.

Decompose $\mathbb{R}^N = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_\nu}$ as in the introduction. We will be considering product kernels $K(t_1, \dots, t_\nu)$ relative to this decomposition of \mathbb{R}^N (see Definition 3.2). For a multi-index $\alpha \in \mathbb{N}^N$, we obtain a corresponding decomposition $\alpha = (\alpha_1, \dots, \alpha_\nu)$ where $\alpha_\mu \in \mathbb{N}^{N_\mu}$ and $t^\alpha = \prod_\mu t_\mu^{\alpha_\mu}$.

Our goal is to give conditions on the vector fields X_α such that the operator T given by (1.1) and the operator \mathcal{M} given by (1.2) are bounded on L^p ($1 < p < \infty$).

Corresponding to a multi-index $\alpha \in \mathbb{N}^\nu$, we assign a formal “degree” $\deg(\alpha) = (|\alpha_1|, \dots, |\alpha_\nu|) \in \mathbb{N}^\nu$. This assigns to each vector field X_α the formal degree $\deg(\alpha)$. We say α is a “pure power” if $\deg(\alpha)$ is non-zero in precisely one component; otherwise we say α is a “non-pure power.” This yields two finite sets of vector fields, paired with formal degrees:

$$\begin{aligned} \mathcal{P} &:= \{(X_\alpha, \deg(\alpha)) : \alpha \text{ is a pure power}\}, \\ \mathcal{N} &:= \{(X_\alpha, \deg(\alpha)) : \alpha \text{ is a non-pure power}\}. \end{aligned}$$

Let \mathcal{S} be the smallest set of vector fields paired with formal degrees such that

- $\mathcal{P} \subseteq \mathcal{S}$ and
- if $(X, d), (Y, e) \in \mathcal{S}$, then $([X, Y], d + e) \in \mathcal{S}$.

We are now prepared to state the two assumptions we make on the vector fields X_α .

- A “finite-type” condition: there exists a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that for every $(X, d) \in \mathcal{S}$, we have

$$(1.4) \quad X = \sum_{\substack{(Y,e) \in \mathcal{F} \\ e \leq d}} c_X^Y Y,$$

where $c_X^Y \in C^\infty$, and $e \leq d$ means that the inequality holds coordinatewise (that is, $e_\mu \leq d_\mu$ for each μ).

- An “algebraic” condition: for every $(X, d) \in \mathcal{N}$, there exists a finite subset $\mathcal{F} \subseteq \mathcal{S}$ such that

$$(1.5) \quad X = \sum_{\substack{(Y,e) \in \mathcal{F} \\ e \leq d}} c_X^Y Y,$$

where $c_X^Y \in C^\infty$. Note that, if the finite type condition holds, one may take \mathcal{F} as in that condition.

Theorem 1.1. *Under the above two conditions, the operator T given by (1.1) (for any product kernel $K(t_1, \dots, t_\nu)$ with small support) and the operator \mathcal{M} given by (1.2) are bounded on L^p ($1 < p < \infty$).*

Remark 1.2. A more general version of our result can be found in our cited work, where the following issues are addressed:

- γ is not restricted to be of the form (1.3) (see Section 9).
- A more general class of kernels than product kernels is studied.
- The sums (1.4) and (1.5) are replaced with more general (and more complicated) conditions.

Remark 1.3. Our general results for the case $\nu = 1$ in fact have wider scope than those in [7]. See also the comments in Section 9. Note that when $\nu = 1$ the algebraic condition (1.5) is vacuous, and the finite type condition follows from Hörmander’s condition: that the Lie algebra generated by the $\{X_\alpha\}$ spans the tangent space at 0. We remark that when $\nu > 1$, Hörmander’s condition implies neither the finite type condition, nor the algebraic condition. For example, decompose $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ and write $(t_1, t_2) \in \mathbb{R} \times \mathbb{R}$. Then for $x \in \mathbb{R}$ ($n = 1$)

$$(1.6) \quad \gamma_{t_1, t_2}(x) = \exp\left(t_1^3 \frac{\partial}{\partial x} + t_1 t_2 \frac{\partial}{\partial x} + t_2^3 \frac{\partial}{\partial x}\right) x = x + t_1^3 + t_1 t_2 + t_2^3$$

does not satisfy the algebraic condition. This would have been true, as well, if we had taken t_1^2 and t_2^2 instead of t_1^3 and t_2^3 . However, the more general algebraic condition in our cited work allows for the case of t_1^2 and t_2^2 . Furthermore, there exist product kernels $K(t_1, t_2)$ such that if γ is given by (1.6), then operator T given by (1.1) is not bounded on L^2 (which is not true if we use t_1^2 and t_2^2)—see [18]. For $(x, y) \in \mathbb{R}^2$, if we set $X_1 = \frac{\partial}{\partial x}$, $X_2 = e^{-\frac{1}{x^2}} \frac{\partial}{\partial y}$, and $X_3 = \frac{\partial}{\partial y}$, then

$$\exp(t_1 X_1 + t_1^2 X_2 + t_2 X_3)(x, y)$$

does not satisfy the finite type condition, while X_1 , X_2 , and X_3 clearly satisfy Hörmander’s condition.

Remark 1.4. If the finite type condition holds, then the involutive distribution¹ generated by the vector fields in \mathcal{P} is finitely generated as a C^∞ module (the vector fields in \mathcal{F} generate it). In this context, the Frobenius theorem applies to foliate the manifold into leaves. Further, if the algebraic condition also holds, then the vector fields corresponding to the non-pure powers are tangent to these leaves. More is true: the finite-type and algebraic conditions can be thought of as “scale invariant” versions of the above. This is discussed in Section 4. In the single parameter case ($\nu = 1$) this scale invariance comes “for free.” However, in the multi-parameter case ($\nu > 1$) this scale invariance is an essential point.

Remark 1.5. The algebraic condition excludes examples like

$$f \mapsto \psi(x) \int f(x-st) K(s,t) ds dt,$$

where $K(s,t)$ is a product kernel corresponding to $(s,t) \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ (here, $x \in \mathbb{R}$). It is well known that there exist product kernels such that the above operator is not bounded on L^2 . This dates back to [13]. See [18] for a further discussion.

Remark 1.6. Some other work that overlaps with our results is in the recent work “Singular integrals with flag kernels on graded groups: I” by Nagel, Ricci, Stein, and Wainger. Also, earlier work by Carbery, Wright, and Wainger [2, 3, 4]. Some other related works include [5, 14].

2. When γ is real analytic

The third paper in the series [17] deals with the special case when γ is assumed to be real analytic (in *both* variables). Once again, we assume in this note that γ is given by (1.3). In this case assuming that γ is real analytic is the same as assuming that each X_α is real analytic. The essential point in this situation is the following proposition.

Proposition 2.1. *When γ is real analytic (equivalently when each X_α is real analytic), the finite type condition holds automatically.*

Proposition 2.1 is a consequence of the Weierstrass preparation theorem.² We refer the reader to [17] for details.

Corollary 2.2. *If each X_α is real analytic, and the algebraic condition holds, then T and \mathcal{M} are bounded on L^p ($1 < p < \infty$).*

Proof. This follows directly from combining Theorem 1.1 and Proposition 2.1. \square

Corollary 2.3. *If $\nu = 1$ (i.e., when K is a Calderón-Zygmund kernel) and if each X_α is real analytic, then T and \mathcal{M} are bounded on L^p ($1 < p < \infty$).*

Proof. This follows from Corollary 2.2 since the algebraic condition is vacuous when $\nu = 1$. \square

¹Here, we are using distribution to mean a C^∞ module of vector fields, and an involutive distribution is one which is closed under Lie brackets.

²This is closely related to the fact that the involutive distribution generated by a finite collection of real analytic vector fields is automatically locally finitely generated as a C^∞ module and the Frobenius theorem applies. This dates back to [9, 8].

Actually, more is true for the maximal function. Indeed, we have the following result.

Theorem 2.4. *For any $\nu > 0$, if each X_α is real analytic, then \mathcal{M} is bounded on L^p ($1 < p \leq \infty$).*

Proof idea. Consider,

$$\mathcal{M}f(x) \lesssim \psi(x) \sup_{j_1, \dots, j_\nu \in \mathbb{N}} \int_{|t| \leq a} \left| f \left(\exp \left\{ \sum_{0 < |\alpha| \leq L} 2^{-j_1 |\alpha_1|} \dots 2^{-j_\nu |\alpha_\nu|} t^\alpha X_\alpha \right\} x \right) \right| dt,$$

where $a > 0$ is some small number (depending on the X_α). Following an idea of Christ [6], we study the stronger maximal operator

$$\widetilde{\mathcal{M}}f(x) = \psi(x) \sup_{k_\alpha \in \mathbb{N}} \int_{|t| \leq a} \left| f \left(\exp \left\{ \sum_{0 < |\alpha| \leq L} 2^{-k_\alpha} t^\alpha X_\alpha \right\} x \right) \right| dt.$$

It is easy to see that $\mathcal{M}f(x) \lesssim \widetilde{\mathcal{M}}f(x)$, since one may take $k_\alpha = j_1 |\alpha_1| + \dots + j_\nu |\alpha_\nu|$. Furthermore, because of the form of $\widetilde{\mathcal{M}}$, each X_α behaves like a vector field corresponding to a pure power. Thus, the algebraic condition holds automatically. Since the finite type condition holds by Proposition 2.1, the proof of Theorem 1.1 (with minor modifications) goes through to prove that $\widetilde{\mathcal{M}}$ is bounded on L^p . \square

The above results generalize to the case when γ is merely assumed to be a germ of a real analytic function satisfying $\gamma_0(x) \equiv x$. I.e., γ not necessarily of the form (1.3). We take a moment, here, to discuss the generalizations of Corollary 2.3 and Theorem 2.4. Proposition 2.1 and Corollary 2.2 can also be generalized, but we defer a discussion of this to [17].

Theorem 2.5. *Suppose $\nu = 1$ (i.e., K is a Calderón-Zygmund kernel) and γ is a real analytic function with $\gamma_0(x) \equiv x$. Then the operator T given by (1.1) is bounded on L^p ($1 < p < \infty$).*

Theorem 2.6. *Suppose γ is a real analytic function satisfying $\gamma_0(x) \equiv x$. Then the operator \mathcal{M} given by (1.2) (for any ν) is bounded on L^p ($1 < p \leq \infty$).*

Theorem 2.6 generalizes some results from the literature. A well known result of Bourgain [1] deals with the special case of Theorem 2.6 when $\nu = 1$, $p = 2$, and $\gamma_t(x) = x + tv(x)$, where $x \in \mathbb{R}^2$, $t \in \mathbb{R}$, and v is a real analytic vector field on \mathbb{R}^2 —Theorem 2.6 extends this to $1 < p \leq \infty$ and allows $x \in \mathbb{R}^n$ for any n (instead of only $n = 2$). Also, a result of Christ [6] deals with a special case of Theorem 2.4 where each X_α is a left invariant vector field on a nilpotent Lie group (and therefore real analytic). In fact, the proof of Theorem 2.6 incorporates ideas both from [6] and [1]. We refer the reader to these references and [17] for further details on these issues.

3. Product Kernels

In this section, we discuss the notion of a product kernel. Our main reference for the following is [10] and we refer the reader there for further information.

We fix a decomposition $\mathbb{R}^N = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_\nu}$ as in the introduction. Our goal is to define the notion of a product kernel $K(t_1, \dots, t_\nu)$, relative to this decomposition.

Definition 3.1. A k -normalized bump function on \mathbb{R}^N is a C^k function supported on the unit ball with C^k norm bounded by 1.

The definitions that follow turn out to be independent of the choice of $k \geq 1$, and we therefore refer to normalized bump functions (which can be taken to mean 1-normalized bump functions, for instance).

Definition 3.2. A product kernel relative to the decomposition $\mathbb{R}^N = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_\nu}$ is a distribution $K(t_1, \dots, t_\nu)$, which corresponds with a C^∞ function away from the coordinate axes $t_\mu = 0$ and which satisfies:

- For each multi-index $\alpha = (\alpha_1, \dots, \alpha_\nu)$, there is a constant C_α such that

$$|\partial_{t_1}^{\alpha_1} \dots \partial_{t_\nu}^{\alpha_\nu} K(t)| \leq C_\alpha |t_1|^{-N_1 - |\alpha_1|} \dots |t_\nu|^{-N_\nu - |\alpha_\nu|}.$$

- We proceed recursively on ν :
 - For $\nu = 1$, given any normalized bump function ϕ and and $R > 0$,

$$\int K(t) \phi(Rt) dt$$

is bounded independently of ϕ and R .

- For $\nu > 1$, given any $1 \leq \mu \leq \nu$, any normalized bump function ϕ on \mathbb{R}^{N_μ} , and any $R > 0$, the distribution

$$K_{\phi,R}(t_1, \dots, t_{\mu-1}, t_{\mu+1}, \dots, t_\nu) = \int K(t) \phi(Rt_\mu) dt_\mu$$

is a product kernel on the lower dimensional space $\mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_{\mu-1}} \times \mathbb{R}^{N_{\mu+1}} \times \dots \times \mathbb{R}^{N_\nu}$, uniformly in R and ϕ .

Remark 3.3. The setting in [10] was slightly more general: product kernels were defined relative to a decomposition of \mathbb{R}^N into subspaces which were homogeneous under certain dilations. Here, we have chosen the standard dilations on each \mathbb{R}^{N_μ} . The framework in our cited work takes into account these more general product kernels, and even kernels which are not product kernels. We do not discuss this here, though.

Given $\delta = (\delta_1, \dots, \delta_\nu) \in [0, \infty)^\nu$, define $\delta(t_1, \dots, t_\nu) = (\delta_1 t_1, \dots, \delta_\nu t_\nu)$. Further, for $j = (j_1, \dots, j_\nu) \in \mathbb{Z}^\nu$, define $2^j = (2^{j_1}, \dots, 2^{j_\nu}) \in (0, \infty)^\nu$. Finally, for a function $f(t)$ define

$$f^{(2^j)}(t) = 2^{j_1 N_1 + \dots + j_\nu N_\nu} f(2^j t).$$

Note that $\int f^{(2^j)}(t) dt = \int f(t) dt$.

Let $B^N(a)$ denote the ball in \mathbb{R}^N of radius $a > 0$. We use the following characterization of product kernels, which is a slight modification of a result of [10].

Proposition 3.4. Suppose $\{\eta_j\}_{j \in \mathbb{N}^\nu}$ is a bounded subset of $C_0^\infty(B^N(a))$ satisfying

$$(3.1) \quad \int \eta_j(t) dt_\mu = 0, \text{ if } j_\mu \neq 0.$$

Then the sum

$$\sum_{j \in \mathbb{N}^\nu} \eta_j^{(2^j)}$$

converges in distribution to a product kernel which is supported in $B^N(a)$. Conversely, every product kernel supported in $B^N(a)$ can be decomposed in this way.

Recall our object of study. We are interested in operators of the form

$$Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt,$$

where $K(t)$ is a product kernel supported in $B^N(a)$ for some small $a > 0$. Proposition 3.4 allows us to decompose

$$K(t) = \sum_{j \in \mathbb{N}^\nu} \eta_j^{(2^j)}(t),$$

where $\{\eta_j\}_{j \in \mathbb{N}^\nu} \subset C_0^\infty(B^N(a))$ is a bounded set, satisfying (3.1). This yields a corresponding decomposition of T . Indeed, define

$$(3.2) \quad T_j f(x) = \psi(x) \int f(\gamma_t(x)) \eta_j^{(2^j)}(t) dt.$$

Then, $T = \sum_{j \in \mathbb{N}^\nu} T_j$.

4. Scaling and the Frobenius theorem

The heart of the proof that T is bounded on L^p ($1 < p < \infty$) is the idea that for $j, k \in \mathbb{N}^\nu$, $j \neq k$, T_j and T_k are essentially of the same form, but at different “scales.” Because they reside at different scales, we will obtain, for instance, the almost orthogonality estimate,

$$(4.1) \quad \|T_j^* T_k\|_{L^2 \rightarrow L^2}, \|T_j T_k^*\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon|j-k|},$$

for some $\epsilon > 0$. The L^2 boundedness of T then follows immediately from the Cotlar-Stein lemma.

The key to proving this almost orthogonality is to use an appropriate scaling map which is adapted to the vector fields. This scaling map was developed in [19], based on ideas from [12, 20]. As was pointed out in [19], and as will be discussed below, these scaling maps can be viewed as the coordinate charts defining the leaves in a quantitative version of the Frobenius theorem on involutive distributions.

Let \mathcal{F} be the finite set of vector fields with formal degrees given in the finite type condition. Enumerate the set \mathcal{F} to obtain a finite list of vector fields with formal degrees $(X_1, d_1), \dots, (X_q, d_q)$. Notice, by the finite type assumption, we have

$$(4.2) \quad [X_j, X_k] = \sum_{d_l \leq d_j + d_k} c_{j,k}^l X_l,$$

where $d_l \leq d_j + d_k$ denotes that the inequality holds coordinatewise, and $c_{j,k}^l \in C^\infty$.

Let \mathcal{D} be the distribution³ generated by X_1, \dots, X_q . Note that \mathcal{D} is involutive: if $X, Y \in \mathcal{D}$, then $[X, Y] \in \mathcal{D}$. This follows directly from (4.2). Because of this, the classical Frobenius theorem applies to foliate the ambient space into leaves; the tangent bundle to each leaf given by \mathcal{D} .

However, more is true. Let $\delta \in [0, 1]^\nu$ and multiply both sides of (4.2) by $\delta^{d_j + d_k}$ to obtain,

$$[\delta^{d_j} X_j, \delta^{d_k} X_k] = \sum_{d_l \leq d_j + d_k} (\delta^{d_j + d_k - d_l} c_{j,k}^l) \delta^{d_l} X_l.$$

³Here, we are using distribution to mean C^∞ module of vector fields.

Defining

$$c_{j,k}^{l,\delta} = \begin{cases} \delta^{d_j+d_k-d_l} c_{j,k}^l & \text{if } d_l \leq d_j + d_k \\ 0 & \text{otherwise,} \end{cases}$$

we have that $c_{j,k}^{l,\delta} \in C^\infty$ uniformly in δ , and

$$[\delta^{d_j} X_j, \delta^{d_k} X_k] = \sum_l c_{j,k}^{l,\delta} \delta^{d_l} X_l.$$

Because of this one might hope that the Frobenius theorem holds when applied to the vector fields $\delta^{d_j} X_j$ “uniformly” in δ in the sense that appropriate coordinate charts defining the leaves may be controlled “uniformly” in δ . Indeed, this turns out to be the case and is the main theorem of [19].

We turn to describing this theorem now. Let $Z_j = \delta^{d_j} X_j$ so that

$$[Z_j, Z_k] = \sum_l \tilde{c}_{j,k}^l Z_l,$$

with $\tilde{c}_{j,k}^l \in C^\infty$ uniformly in δ . Fix $x_0 \in \mathbb{R}^n$. Our goal is to create a “good” coordinate chart near x_0 on the leaf passing through x_0 generated by Z_1, \dots, Z_q (note that this leaf is independent of $\delta \in (0, 1)^\nu$, but we wish to have good estimates on this chart in terms of Z_1, \dots, Z_q ; in particular, the theorem will be meaningful even if Z_1, \dots, Z_q span the tangent space and the leaf is the entire ambient space).

For $\xi > 0$, $x_0 \in \mathbb{R}^n$, define

$$B_Z(x_0, \xi) = \left\{ y \mid \exists \gamma : [0, 1] \rightarrow \mathbb{R}^n, \gamma(0) = x_0, \gamma(1) = y, \right. \\ \left. \gamma'(t) = \sum_{j=1}^q a_j(t) Z_j(\gamma(t)), |a_j(t)| < \xi, \forall t \right\};$$

$B_Z(x_0, \xi)$ is called a Carnot-Carathéodory ball. Note that $B_Z(x_0, \xi)$ is an open neighborhood of the point x_0 on the leaf passing through x_0 generated by Z_1, \dots, Z_q . Let $n_0 = \dim \text{span} \{Z_1(x_0), \dots, Z_q(x_0)\}$ (the dimension of this leaf). The quantitative Frobenius theorem we will use is

Theorem 4.1 ([19]). *There exist $\xi, \eta \approx 1$ and a smooth map $\Phi : B^{n_0}(\eta) \rightarrow B_Z(x_0, 1)$ such that:*

- Φ is one-to-one.
- $B_Z(x_0, \xi) \subseteq \Phi(B^{n_0}(\eta))$.
- If we let Y_j be the pull back of Z_j via the map Φ to $B^{n_0}(\eta)$, then Y_1, \dots, Y_q span the tangent space. More precisely,

$$1 \lesssim \inf_{u \in B^{n_0}(\eta)} \sup_{1 \leq j_1, \dots, j_{n_0} \leq q} |\det [Y_{j_1}(u) : \dots : Y_{j_{n_0}}(u)]|.$$

- Y_1, \dots, Y_q are smooth. That is,

$$\|Y_j\|_{C^m(B^{n_0}(\eta))} \lesssim 1.$$

In the above, all implicit constants can be chosen to depend only on n , upper bounds for a finite number of the C^m norms of the Z_j and $\tilde{c}_{j,k}^l$, and q . In particular, when $Z_j = \delta^{d_j} X_j$ as in our primary example, the implicit constants are independent of δ .

Notice that the implicit constants in Theorem 4.1 do *not* depend on lower bounds for quantities like

$$\left| \det_{n_0 \times n_0} [Z_1(x_0) : \cdots : Z_q(x_0)] \right|,$$

where $\det_{n_0 \times n_0} A$ denotes the vector whose coordinates are the determinates of the $n_0 \times n_0$ submatrices of A . This is what sets Theorem 4.1 apart from classical versions of the Frobenius theorem.

In particular, recall the sets \mathcal{P} and \mathcal{N} of vector fields with formal degrees corresponding to the pure powers and non-pure powers, respectively. From the finite type condition we obtain the vector fields $\mathcal{F} = \{(X_1, d_1), \dots, (X_q, d_q)\}$. For each $j_0 \in \mathbb{N}^\nu$, and $x_0 \in \mathbb{R}^n$ (x_0 near 0), we consider the vector fields

$$Z_1 = 2^{-j_0 \cdot d_1} X_1, \dots, Z_q = 2^{-j_0 \cdot d_q} X_q,$$

which satisfy the hypotheses of Theorem 4.1 uniformly in j_0, x_0 . We therefore obtain the map $\Phi = \Phi_{x_0, j_0}$ as in Theorem 4.1. Let Y_1, \dots, Y_q be the pull backs of Z_1, \dots, Z_q so that Y_1, \dots, Y_q span the tangent space uniformly in j_0, x_0 .

Further, we let Y_α be the pull back of $2^{-j_0 \cdot \alpha} X_\alpha$, where

$$\gamma_t(x) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha \right) x.$$

By (1.4) and (1.5) Y_α is a finite linear combination, with coefficients in C^∞ of the Y_l . Thus $Y_\alpha \in C^\infty$ uniformly in j_0 . Furthermore, since each Y_l is in the Lie algebra generated by Y_α where α is a pure power (this comes from the corresponding statement about the X_l), and since the Y_l span the tangent space uniformly in j_0 , we have that $\{Y_\alpha : \alpha \text{ is a pure power}\}$ satisfies Hörmander's condition, uniformly in j_0 .

Let $j, k \in \mathbb{N}^\nu$ and set $j_0 = j \wedge k \in \mathbb{N}^\nu$. I.e., the μ th coordinate of j_0 is the minimum of the μ th coordinates of j and k . Consider,

$$\gamma_{2^{-k}t} \circ \gamma_{2^{-j}s}^{-1}(x) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha 2^{-k \cdot \deg(\alpha)} X_\alpha \right) \exp \left(- \sum_{0 < |\alpha| \leq L} s^\alpha 2^{-j \cdot \deg(\alpha)} X_\alpha \right) x.$$

Hence, if we define,

$$\theta_{t,s}(u) = \Phi_{x_0, j_0}^{-1} \circ \gamma_{2^{-k}t} \circ \gamma_{2^{-j}s}^{-1} \circ \Phi_{x_0, j_0}(u),$$

then,

(4.3)

$$\theta_{t,s}(u) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha 2^{-(k-j_0) \cdot \deg(\alpha)} Y_\alpha \right) \exp \left(- \sum_{0 < |\alpha| \leq L} s^\alpha 2^{-(j-j_0) \cdot \deg(\alpha)} Y_\alpha \right) u.$$

We have suppressed the dependence of θ on j, k, x_0 .

Recall, α is a pure power if $\deg(\alpha)$ is non-zero in only one component. Thus, if α is a pure power, either $(k-j_0) \cdot \deg(\alpha)$ or $(j-j_0) \cdot \deg(\alpha)$ will be 0. It follows that each vector field Y_α corresponding to a pure power appears *unscaled* in at least one of the two exponentials in (4.3). Hence, the vector fields in the exponentials of (4.3) (namely, $\{2^{-(k-j_0) \cdot \deg(\alpha)} Y_\alpha, 2^{-(j-j_0) \cdot \deg(\alpha)} Y_\alpha\}$) satisfy Hörmander's condition

uniformly in j, k, x_0 . This argument is the main reason we had to distinguish between pure powers and non-pure powers.

The way θ comes up in our argument is as follows,

$$T_j^* T_k f(x) = \psi(x) \int f(\gamma_{2^{-k}t} \circ \gamma_{2^{-j}s}^{-1}(x)) \kappa(s, t, x) \overline{\eta_j(s)} \eta_k(t) ds dt,$$

where $\kappa \in C^\infty$. Thus, if we let $\Phi_{x_0, j_0}^\#$ denote the map $\Phi_{x_0, j_0}^\# f = f \circ \Phi_{x_0, j_0}$, then we see,

$$(4.4) \quad \Phi_{x_0, j_0}^\# T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} g(u) = \psi(\Phi_{x_0, j_0}(u)) \int g(\theta_{t,s}(u)) \tilde{\kappa}(s, t, u) \overline{\eta_j(s)} \eta_k(t) ds dt.$$

(4.4) alone is not quite enough to directly show (4.1), however it is close. We outline the proof of the L^2 theorem in Section 5.

5. The L^2 theorem

In this section, we outline the proof of the case $p = 2$ of Theorem 1.1 (for the operator T). This describes work from [18].

As discussed in Section 4, our goal is to show

$$(5.1) \quad \|T_j^* T_k\|_{L^2 \rightarrow L^2}, \|T_j T_k^*\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon|j-k|},$$

where T_j and T_k are defined in (3.2). We focus on the estimate for $T_j^* T_k$, the other estimate being similar. In what follows $\epsilon > 0$ is a constant that may change from line to line, but is always independent of j, k . To prove (5.1) it suffices to show

$$(5.2) \quad \left\| (T_k^* T_j T_j^* T_k)^{n+1} \right\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon|j-k|},$$

which proves (5.1) with ϵ replaced by $\frac{\epsilon}{2(n+1)}$. Recall, n is the dimension of the ambient space. It is easy to see that

$$\|T_k^* T_j T_j^* T_k\|_{L^1 \rightarrow L^1} \lesssim 1,$$

and therefore to prove (5.2), interpolation shows that it suffices to show

$$(5.3) \quad \left\| (T_k^* T_j T_j^* T_k)^{n+1} \right\|_{L^\infty \rightarrow L^\infty} \lesssim 2^{-\epsilon|j-k|}.$$

Fix a point x_0 and a bounded function f . To show (5.3) we show that

$$(5.4) \quad \left| (T_k^* T_j T_j^* T_k)^{n+1} f(x_0) \right| \lesssim 2^{-\epsilon|j-k|} \|f\|_{L^\infty}.$$

Recall

$$(5.5) \quad T_j f(x) = \psi(x) \int f(\gamma_t(x)) \eta_j^{(2^j)}(t) dt,$$

where $\eta_j \in C_0^\infty(B^N(a))$ (and satisfies certain cancellation conditions); with a similar formula for T_k . Here, $a > 0$ is small if the support of K is small, and therefore, we may take $a > 0$ as small as we like (since the Theorem 1.1 is about kernels K with small support).

Since T_k^* and T_j^* are essentially of the same form as T_k and T_j (with γ_t replaced by γ_t^{-1}), by taking $a > 0$ small enough, (5.5) shows that $(T_k^* T_j T_j^* T_k)^{n+1} f(x_0)$ depends

only on the values of f in $B_Z(x_0, \xi)$ where $j_0 = j \wedge k$, $Z_1 = 2^{-j_0 \cdot d_1} X_1, \dots, Z_q = 2^{-j_0 \cdot d_q} X_q$, and ξ is as in Theorem 4.1. Thus to prove (5.4) it suffices to show

$$\left| \Phi_{x_0, j_0}^\# (T_k^* T_j T_j^* T_k)^{n+1} \left(\Phi_{x_0, j_0}^\# \right)^{-1} g(0) \right| \lesssim 2^{-\epsilon|j-k|} \|g\|_{L^\infty},$$

where g is supported on $B^{n_0}(\eta)$ (and $n_0, \eta > 0$ are as in Theorem 4.1).

Putting all this together, to prove (5.1) it suffices to show

$$(5.6) \quad \left\| \left(\Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} \right)^{n+1} \right\|_{L^\infty \rightarrow L^\infty} \lesssim 2^{-\epsilon|j-k|},$$

where $\epsilon > 0$ and the implicit constant are independent of j, k , and x_0 . Fortunately, (5.6) follows from methods developed in the single parameter case by Christ, Nagel, Stein, and Wainger [7].

We close this section by briefly outlining the proof of (5.6); we refer the reader to [18] for more precise details. By a computation like the one leading up to (4.4) and an application of the Campbell-Hausdorff formula it is easy to see that

$$\Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} g(u) = \tilde{\psi}(u) \int f(\Theta_t(u)) \tilde{\kappa}(t, u) \tilde{\eta}(t) dt,$$

where $\tilde{\eta} \in C_0^\infty$ and is supported near $0 \in \mathbb{R}^{4N}$, $\tilde{\kappa} \in C^\infty$, and⁴

$$(5.7) \quad \Theta_t(u) \sim \exp \left(\sum_{0 < |\alpha|} t^\alpha W_\alpha \right) u,$$

and $\{W_\alpha\}$ satisfy Hörmander's condition, uniformly in any relevant parameters. Using results from [7] it is then easy to show that

$$(5.8) \quad \left[\Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} \right]^n g(u) = \tilde{\psi}(u) \int g(v) h(u, v) dv,$$

where $h(u, v)$ has a certain level of "smoothness" in v . Indeed, one can show that (uniformly in u), $h(u, \cdot) \in L_\delta^1$ for some $\delta > 0$, where L_δ^1 is the space of those $f \in L^1$ such that,

$$\int |f(y-z) - f(y)| dy \lesssim |z|^\delta.$$

There is also "cancellation" in the operator $\Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1}$ induced by the cancellation condition in η_j, η_k . One can use this cancellation together with the smoothness from (5.8) to show that

$$\left\| \left[\Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} \right]^n \Phi_{x_0, j_0}^\# T_k^* T_j T_j^* T_k \left(\Phi_{x_0, j_0}^\# \right)^{-1} \right\|_{L^\infty \rightarrow L^\infty} \lesssim 2^{-\epsilon|j-k|},$$

which completes the proof of (5.6) and therefore of (5.1).

Remark 5.1. A significant issue dealt with is related to the fact that T_j and T_k^* do not necessarily commute. The proof of (5.6) exploits this possibility, allowing us to take vectors in (4.3) from *both* exponentials to obtain a collection that satisfies Hörmander's condition. This allows one to exploit, for instance, the commutator of

⁴(5.7) means that $\Theta_t(u) = \exp \left(\sum_{0 < |\alpha| \leq L} t^\alpha W_\alpha \right) u + O(|t|^{L+1})$ as $t \rightarrow 0$.

two vector fields, one from each exponential. This is in contrast to the case when $\nu = 1$, where one needs only look at one of the exponentials. This accounts for one of the differences between our cited work and [7]. Of course, there are cases where T_j and T_k^* (approximately) commute (for instance, certain cases when all the X_α commute), and our theorem applies. In these cases, one does not need to exploit the commutator of two vector fields, one from each of the exponentials in (4.3).

6. The Littlewood-Paley theory

In order to extend the results from L^2 to L^p ($1 < p < \infty$) we use a Littlewood-Paley theory adapted to the operator T . Recall the vector fields with formal degrees $(X_1, d_1), \dots, (X_q, d_q)$ from Section 4 (these vector fields are nothing more than an enumeration of the set \mathcal{F} from the finite-type condition).

For each μ , $1 \leq \mu \leq \nu$, consider those vector fields with formal degrees (X_j, d_j) such that d_j is non-zero in only the μ th component. Think of these vector fields as vector fields with *one*-parameter formal degrees, where the one parameter is the μ th component of d_j . Enumerate these vector fields and one-parameter formal degrees $(X_1^\mu, d_1^\mu), \dots, (X_{q_\mu}^\mu, d_{q_\mu}^\mu)$ where $d_{q_\mu}^\mu \in (0, \infty)$.

For $(t_1, \dots, t_{q_\mu}) \in \mathbb{R}^{q_\mu}$ and $j \in \mathbb{Z}$, define $2^j(t_1, \dots, t_{q_\mu}) = (2^{jd_1^\mu} t_1, \dots, 2^{jd_{q_\mu}^\mu} t_{q_\mu})$. For a function $f(t)$ (with $t \in \mathbb{R}^{q_\mu}$), define

$$f(2^j) = 2^{j(d_1^\mu + \dots + d_{q_\mu}^\mu)} f(2^j t).$$

Decompose

$$\delta_0 = \sum_{j=0}^{\infty} \eta_j^{(2^j)}(t),$$

where $\{\eta_j\} \subset C_0^\infty(B^{q_\mu}(a))$ is a bounded set and $\int \eta_j = 0$ if $j \neq 0$ (here, $a > 0$ is a small number).

Fix $\psi_0 \in C_0^\infty(\mathbb{R}^n)$ which is 1 on a neighborhood of 0 and supported on a small neighborhood of 0. Define,

$$D_j^\mu f(x) = \psi_0(x) \int_{\mathbb{R}^{q_\mu}} f(e^{t \cdot X^\mu} x) \psi_0(e^{t \cdot X^\mu} x) \eta_j^{(2^j)}(t) dt,$$

where $X^\mu = (X_1^\mu, \dots, X_{q_\mu}^\mu)$. Note that $\sum_{j=0}^{\infty} D_j^\mu = \psi_0^2$.

For $j = (j_1, \dots, j_\nu) \in \mathbb{N}^\nu$, define,

$$D_j = D_{j_1}^1 \cdots D_{j_\nu}^\nu.$$

Note that $\sum_{j \in \mathbb{N}^\nu} D_j = \psi_0^{2\nu}$.

Theorem 6.1 (Littlewood-Paley square function). *For f supported sufficiently close to $0 \in \mathbb{R}^n$ and $1 < p < \infty$,*

$$\left\| \left(\sum_{j \in \mathbb{N}^\nu} |D_j f|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \approx \|f\|_{L^p}.$$

For $M \in \mathbb{N}$, define

$$U_M f(x) = \sum_{j \in \mathbb{N}^\nu} \sum_{\substack{k \in \mathbb{N}^\nu \\ |k-j| \leq M}} D_j D_k.$$

Theorem 6.2 (Calderón reproducing formula). *Fix $\psi_1 \prec \psi_0$, $\psi_1 \in C_0^\infty$. For every p ($1 < p < \infty$) there exists an M , and an operator $V_M : L^p \rightarrow L^p$ such that $\psi_1 U_M V_M = \psi_1 = V_M U_M \psi_1$.*

Proof sketch. Note that $\psi_0^{4\nu} = U_M + R_M$, where

$$R_M = \sum_{j \in \mathbb{N}^\nu} \sum_{\substack{k \in \mathbb{N}^\nu \\ |k-j| > M}} D_j D_k.$$

For each fixed p , $1 < p < \infty$, one has,

$$(6.1) \quad \lim_{M \rightarrow \infty} \|R_M\|_{L^p \rightarrow L^p} = 0.$$

Taking $M = M(p)$ such that $\|R_M\|_{L^p \rightarrow L^p}$ is sufficiently small, V_M can be constructed using a Neumann series. For $p = 2$, (6.1) follows from the ideas in Section 5. For other p , an interpolation argument is used. \square

The use of Theorem 6.2 is that instead of proving the L^p boundedness of T , it suffices to prove the L^p boundedness of TU_M for some $M = M(p)$ (here we have used $T\psi_1 = T$ if the support of T is chosen to be on a sufficiently small neighborhood of 0—where we have taken $\psi_1 \equiv 1$ on a neighborhood of 0).

We close this section by outlining the proof of the \lesssim part of Theorem 6.1. We defer further discussion of all other results in this section to [16].

By a well-known reduction, to prove the \lesssim part of Theorem 6.1 it suffices to show that for any ν sequences $\epsilon_j^\mu \in \{-1, 1\}$, $j \in \mathbb{N}$, $1 \leq \mu \leq \nu$, we have

$$\sum_{j_1, \dots, j_\nu=0}^{\infty} \epsilon_{j_1}^1 D_{j_1}^1 \cdots \epsilon_{j_\nu}^\nu D_{j_\nu}^\nu,$$

is bounded on L^p (with bound independent of the choice of the sequences ϵ^μ). As a consequence, it suffices to show for each μ ,

$$S_\mu := \sum_{j=0}^{\infty} \epsilon_j^\mu D_j^\mu$$

is bounded on L^p (with bound independent of the sequence).

If the vector fields $X_1^\mu, \dots, X_{q_\mu}^\mu$ span the tangent space at each point, then it is easy to see that S_μ is a Calderón-Zygmund singular integral operator associated to single parameter Carnot-Carathéodory balls generated by $(X_1^\mu, d_1^\mu), \dots, (X_{q_\mu}^\mu, d_{q_\mu}^\mu)$ (which, in turn, endow the ambient space with the structure of a space of homogeneous type—see [12]). The L^p boundedness of S_μ then follows from classical results (see, e.g., [15]).

It does not follow from our assumptions that $X_1^\mu, \dots, X_{q_\mu}^\mu$ span the tangent space. However, it *does* follow from our assumptions that $[X_j^\mu, X_k^\mu]$ can be written as a C^∞ linear combination of $X_1^\mu, \dots, X_{q_\mu}^\mu$ and therefore the classical theorem of Frobenius

applies to foliate the ambient space into leaves; $X_1^\mu, \dots, X_{q_\mu}^\mu$ spanning the tangent space to each leaf. One would then like to apply the above argument (seeing S_μ as a Calderón-Zygmund singular integral operator) to each leaf (uniformly in the leaf), and use this to prove the L^p boundedness of S_μ on the entire space. To do this, one needs appropriately uniform control of the coordinate charts defining the leaves. Theorem 4.1 gives just such control, and this argument then yields the L^p boundedness of S_μ .

Along a similar line, let $\sigma \in C_0^\infty(\mathbb{R}^{q_\mu})$ be a non-negative function which is 1 on a neighborhood of 0, and supported on a small neighborhood of 0. Define,

$$\begin{aligned} A_j^\mu f(x) &= \psi_0(x) \int f(e^{t \cdot X^\mu} x) \psi_0(e^{t \cdot X^\mu} x) \sigma(2^j)(t) dt \\ &= \psi_0(x) \int f(e^{(2^{-j}t) \cdot X^\mu} x) \psi_0(e^{(2^{-j}t) \cdot X^\mu} x) \sigma(t) dt, \end{aligned}$$

and define

$$\mathcal{M}_\mu f(x) = \sup_{j \in \mathbb{N}} A_j^\mu |f|(x).$$

By a similar argument using the Frobenius theorem to reduce the question to the classical theory of spaces of homogeneous type, we have $\|\mathcal{M}_\mu f\|_{L^p} \lesssim \|f\|_{L^p}$ for $1 < p \leq \infty$.

7. Auxiliary operators

To prove Theorem 1.1, we need to use several auxiliary operators. Recall, we have decomposed $T = \sum_{j \in \mathbb{N}^\nu} T_j$, we have defined a Littlewood-Paley square function in terms of operators D_j in Section 6, and we have introduced some “averaging” operators A_j^μ at the end of Section 6. In what follows, we will have many operators with a subscript with values in \mathbb{N}^ν (e.g., T_j , $j \in \mathbb{N}^\nu$). For such operators we define T_j for $j \in \mathbb{Z}^\nu \setminus \mathbb{N}^\nu$ to be 0.

For $k_1, k_2 \in \mathbb{Z}^\nu$, we define the operator \mathcal{T}_{k_1, k_2} on sequences of measurable functions $\{f_j\}_{j \in \mathbb{N}^\nu}$ by

$$\mathcal{T}_{k_1, k_2} \{f_j\}_{j \in \mathbb{N}^\nu} = \{D_j T_{j+k_1} D_{j+k_2} f_j\}_{j \in \mathbb{N}^\nu}.$$

The use of \mathcal{T}_{k_1, k_2} comes from the following proposition:

Proposition 7.1. *Fix p , $1 < p < \infty$. If there exists $\epsilon > 0$ such that*

$$\|\mathcal{T}_{k_1, k_2}\|_{L^p(\ell^2(\mathbb{N}^\nu)) \rightarrow L^p(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon(|k_1| + |k_2|)},$$

then $T : L^p \rightarrow L^p$. Here, $L^p(\ell^q(\mathbb{N}^\nu))$ denotes the L^p space of functions on \mathbb{R}^n taking values in the Banach space $\ell^q(\mathbb{N}^\nu)$.

Proof. Take $M = M(p)$ as in Theorem 6.2, so that it suffices to prove $TU_M : L^p \rightarrow L^p$. Consider, using Theorem 6.1 and the triangle inequality, we have

$$\begin{aligned}
\|TU_M f\|_{L^p} &\approx \left\| \left(\sum_{j \in \mathbb{N}^\nu} |D_j TU_M f|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \\
&= \left\| \left(\sum_{j \in \mathbb{N}^\nu} \left| \sum_{\substack{j_1, j_2 \in \mathbb{N}^\nu \\ |l| \leq M}} D_j T_{j_1} D_{j_2} D_{j_2+l} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \\
&= \left\| \left(\sum_{j \in \mathbb{N}^\nu} \left| \sum_{\substack{k_1, k_2 \in \mathbb{Z}^\nu \\ |l| \leq M}} D_j T_{j+k_1} D_{j+k_2} D_{j+k_2+l} f \right|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \\
&\leq \sum_{\substack{k_1, k_2 \in \mathbb{Z}^\nu \\ |l| \leq M}} \left\| \left(\sum_{j \in \mathbb{N}^\nu} |D_j T_{j+k_1} D_{j+k_2} D_{j+k_2+l} f|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \\
&= \sum_{\substack{k_1, k_2 \in \mathbb{Z}^\nu \\ |l| \leq M}} \left\| \mathcal{T}_{k_1, k_2} \{D_{j+k_2+l} f\}_{j \in \mathbb{N}^\nu} \right\|_{L^p(\ell^2(\mathbb{N}^\nu))} \\
&\lesssim \sum_{\substack{k_1, k_2 \in \mathbb{Z}^\nu \\ |l| \leq M}} 2^{-\epsilon(|k_1|+|k_2|)} \left\| \{D_{j+k_2+l} f\}_{j \in \mathbb{N}^\nu} \right\|_{L^p(\ell^2(\mathbb{N}^\nu))} \\
&\lesssim \left\| \left(\sum_{j \in \mathbb{N}^\nu} |D_j f|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \\
&\lesssim \|f\|_{L^p},
\end{aligned}$$

which completes the proof. \square

Our proof of the L^p boundedness of T will proceed by applying Proposition 7.1. A similar idea is used for \mathcal{M} , however \mathcal{M} does not have the inherent cancellation that T does. Because of this, we introduce this cancellation in an *ad hoc* way. To do this we proceed by induction on ν (the base case $\nu = 0$ will be trivial—this will be explained more in what follows). We will construct an operator B_j ($j \in \mathbb{N}^\nu$) which has cancellation similar to T_j and which (under our inductive hypothesis) satisfies $\|\sup_j |B_j f|\|_{L^p} \lesssim \|f\|_{L^p}$ ($1 < p < \infty$) if and only if $\|\mathcal{M}f\|_{L^p} \lesssim \|f\|_{L^p}$.

Let $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$. For a subset $E \subseteq \{1, \dots, \nu\}$ and $j = (j_1, \dots, j_\nu) \in \mathbb{N}^\nu$, define $j_E \in \mathbb{N}_\infty^\nu$ to be equal to j_μ for those components $\mu \in E$ and equal to ∞ in the rest of the components. For $t \in \mathbb{R}^N$, we define $2^{-j_E t}$ in the usual way, with the convention that $2^{-\infty} = 0$; thus $2^{-j_E t}$ is zero in every coordinate corresponding to t_μ where $\mu \in E^c$. We may think of these dilations as $|E|$ -parameter dilations acting on

the lower dimensional space consisting of those coordinates t_μ with $\mu \in E$. Notice that $j_\emptyset = (\infty, \dots, \infty)$ and $j_{\{1, \dots, \nu\}} = j$.

Let $\sigma_0 \in C_0^\infty(\mathbb{R})$ be a non-negative function supported on a small neighborhood of 0, with $\sigma_0 \geq 1$ on a neighborhood of 0. For $t = (t_1, \dots, t_m) \in \mathbb{R}^m$ (for any m), we define $\sigma(t) = \sigma_0(t_1) \cdots \sigma_0(t_m)$. We use this σ to define the operators A_j^μ as at the end of Section 6. Let $\psi_1, \psi_2 \in C_0^\infty(\mathbb{R}^n)$, with $\psi_1, \psi_2 \geq 0$, $\psi_2 \prec \psi_1 \prec \psi_0$, and $\psi_1, \psi_2 \geq 1$ on a neighborhood of 0. We define, for $j \in \mathbb{N}_\infty^\nu$,

$$M_j f(x) = \psi_2(x) \int f(\gamma_{2^{-j}t}(x)) \psi_2(\gamma_{2^{-j}t}(x)) \sigma(t) dt.$$

Notice,

$$(7.1) \quad M_{(\infty, \dots, \infty)} f(x) = M_{j_\emptyset} f(x) = \psi_2^2(x) \left[\int \sigma(t) dt \right] f(x).$$

It is immediate to see

$$\mathcal{M}f(x) \lesssim \sup_{j \in \mathbb{N}^\nu} M_j |f|(x).$$

Thus, to prove \mathcal{M} is bounded on L^p ($1 < p \leq \infty$), it suffices to show

$$(7.2) \quad \left\| \sup_{j \in \mathbb{N}^\nu} |M_j f| \right\|_{L^p} \lesssim \|f\|_{L^p},$$

for $1 < p \leq \infty$.

Recall the operators A_j^μ from Section 6 (replace ψ_0 with ψ_1 in the definition of A_j^μ). For $j = (j_1, \dots, j_\nu) \in \mathbb{N}_\infty^\nu$, define

$$A_j = A_{j_1}^1 \cdots A_{j_\nu}^\nu.$$

Note that

$$A_{(\infty, \dots, \infty)} = \left[\int \sigma(t) dt \right]^\nu \psi_1^{2\nu}.$$

Thus, to prove (7.2) it suffices to show

$$(7.3) \quad \left\| \sup_{j \in \mathbb{N}^\nu} |A_{j_\emptyset} M_{j_{\{1, \dots, \nu\}}} f| \right\|_{L^p} \lesssim \|f\|_{L^p}.$$

It is easy to see that M_{j_E} is of the same form as M_j but with ν replaced by $|E|$. Thus, we may assume (for induction) that $\left\| \sup_{j \in \mathbb{N}^\nu} |M_{j_E} f| \right\|_{L^p} \lesssim \|f\|_{L^p}$ for $1 < p \leq \infty$ and $E \subsetneq \{1, \dots, \nu\}$. In light of (7.1), the base case $E = \emptyset$ is trivial.

Fix p ($1 < p \leq 2$).⁵ As discussed at the end of Section 6,

$$\left\| \sup_{j \in \mathbb{N}_\infty^\nu} |A_j f| \right\|_{L^p} \lesssim \|f\|_{L^p}.$$

It follows that, for $E \subsetneq \{1, \dots, \nu\}$,

$$\left\| \sup_{j \in \mathbb{N}^\nu} |A_{j_{E^c}} M_{j_E} f| \right\|_{L^p} \lesssim \|f\|_{L^p}.$$

⁵The maximal result for $p > 2$ follows from the result for $p = 2$ and interpolation with the trivial result $p = \infty$.

Define the operator,

$$B_j = \sum_{E \subseteq \{1, \dots, \nu\}} (-1)^{|E|} A_{j_{E^c}} M_{j_E}.$$

By the above discussion, showing (7.2) is equivalent to showing

$$\left\| \sup_{j \in \mathbb{N}^\nu} |B_j f| \right\|_{L^p} \lesssim \|f\|_{L^p}.$$

In fact, we show the stronger result:

$$(7.4) \quad \left\| \left(\sum_{j \in \mathbb{N}^\nu} |B_j f|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \lesssim \|f\|_{L^p}.$$

For $k \in \mathbb{Z}^\nu$, define the vector valued operator

$$\mathcal{B}_k \{f_j\}_{j \in \mathbb{N}^\nu} = \{B_j D_{j+k} f_j\}_{j \in \mathbb{N}^\nu}.$$

From the above discussion, and a proof along the lines of Proposition 7.1, we have the following.

Proposition 7.2. *Fix p , $1 < p < \infty$. If there exists $\epsilon > 0$ such that*

$$\|\mathcal{B}_k\|_{L^p(\ell^2(\mathbb{N}^\nu)) \rightarrow L^p(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon|k|},$$

then \mathcal{M} is bounded on L^p .

8. Completion of the proof

In this section, we outline the remaining points in the proof of Theorem 1.1. Since the L^2 adjoint of T is essentially of the same form as T (with $\gamma_t(x)$ replaced by $\gamma_t^{-1}(x)$) and since \mathcal{M} is trivially bounded on L^∞ , it suffices to verify Theorem 1.1 for $1 < p \leq 2$. In light of Propositions 7.1 and 7.2 we wish to show that for every p , $1 < p \leq 2$, there exists $\epsilon = \epsilon(p) > 0$ such that

$$(8.1) \quad \|\mathcal{T}_{k_1, k_2}\|_{L^p(\ell^2(\mathbb{N}^\nu)) \rightarrow L^p(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon(|k_1| + |k_2|)}, \quad \|\mathcal{B}_k\|_{L^p(\ell^2(\mathbb{N}^\nu)) \rightarrow L^p(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon|k|}.$$

Step 1. Using the methods of Section 5, we show

$$\|D_j T_{j+k_1} D_{j+k_2}\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon_2(|k_1| + |k_2|)}, \quad \|B_j D_{j+k}\|_{L^2 \rightarrow L^2} \lesssim 2^{-\epsilon_2|k|},$$

for some $\epsilon_2 > 0$. As a consequence, we obtain

$$(8.2) \quad \|\mathcal{T}_{k_1, k_2}\|_{L^2(\ell^2(\mathbb{N}^\nu)) \rightarrow L^2(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon_2(|k_1| + |k_2|)}, \quad \|\mathcal{B}_k\|_{L^2(\ell^2(\mathbb{N}^\nu)) \rightarrow L^2(\ell^2(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon_2|k|}.$$

It is easy to see that

$$(8.3) \quad \|\mathcal{T}_{k_1, k_2}\|_{L^1(\ell^1(\mathbb{N}^\nu)) \rightarrow L^1(\ell^1(\mathbb{N}^\nu))} \lesssim 1, \quad \|\mathcal{B}_k\|_{L^1(\ell^1(\mathbb{N}^\nu)) \rightarrow L^1(\ell^1(\mathbb{N}^\nu))} \lesssim 1.$$

Interpolating (8.2) and (8.3) we obtain for $1 < p \leq 2$,

$$(8.4) \quad \|\mathcal{T}_{k_1, k_2}\|_{L^p(\ell^p(\mathbb{N}^\nu)) \rightarrow L^p(\ell^p(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon_p(|k_1| + |k_2|)}, \quad \|\mathcal{B}_k\|_{L^p(\ell^p(\mathbb{N}^\nu)) \rightarrow L^p(\ell^p(\mathbb{N}^\nu))} \lesssim 2^{-\epsilon_p|k|},$$

where $\epsilon_p > 0$.

Step 2. (8.2) shows that \mathcal{M} is bounded on L^2 . A bootstrapping argument like the one from [11] then can be used (in conjunction with (8.4)) to show that \mathcal{M} is bounded on L^p ($1 < p \leq 2$), thereby completing the proof for \mathcal{M} .

Step 3. The L^p boundedness of \mathcal{M} can be used to show that

$$\|\mathcal{T}_{k_1, k_2}\|_{L^p(\ell^\infty(\mathbb{N}^\nu)) \rightarrow L^p(\ell^\infty(\mathbb{N}^\nu))} \lesssim 1.$$

Interpolating this with (8.4) yields (8.1) and completes the proof.

9. More general results

The discussion in this note has been restricted to γ of the form

$$\gamma_t(x) = \exp\left(\sum_{0 < |\alpha| \leq L} t^\alpha X_\alpha\right)x,$$

an exponential of a *finite* sum of vector fields. However, our cited work is not limited to such γ .

If γ were of the form

$$(9.1) \quad \gamma_t(x) = \exp(A(t))x,$$

where $A(t)$ is a vector field depending smoothly on t with $A(0) \equiv 0$, then the discussion in this note would be easy to generalize. Unfortunately, not every γ is of this form. Fortunately, there is a simple alternative. Given γ define a vector field depending on t , $W(t)$, by

$$(9.2) \quad W(t) = W(t, x) = \frac{d}{d\epsilon} \bigg|_{\epsilon=1} \gamma_{\epsilon t} \circ \gamma_t^{-1}(x) \in T_x \mathbb{R}^n.$$

Note that $W(0) \equiv 0$. Moreover, the map $\gamma \mapsto W$ is a bijection (in an appropriate sense). Everywhere in this note where one might wish to use A , it suffices to use W instead.

We now informally state a special case of the main result of [18, 16], but refer the reader to those papers for a more rigorous statement in full generality. The setting is the same as Theorem 1.1, and so we are given a decomposition $\mathbb{R}^N = \mathbb{R}^{N_1} \times \dots \times \mathbb{R}^{N_\nu}$. We are interested in the L^p boundedness of operators of the form

$$Tf(x) = \psi(x) \int f(\gamma_t(x)) K(t) dt,$$

where K is a product kernel relative to this decomposition (with small support). γ is a C^∞ function satisfying $\gamma_0(x) \equiv x$, but we do not assume that γ is given by (9.1).

For a set of vector fields \mathcal{V} , let $\mathcal{D}(\mathcal{V})$ denote the involutive distribution generated by \mathcal{V} ; that is, let $\mathcal{D}(\mathcal{V})$ be the smallest C^∞ module containing \mathcal{V} and closed under Lie brackets.

Given γ define W by (9.2). Decompose W as a Taylor series in t ,

$$W(t) \sim \sum_{0 < |\alpha|} t^\alpha X_\alpha,$$

where each X_α is a C^∞ vector field. We define pure powers and non-pure powers just as in Section 1.1. Our assumptions are (informally stated) as follows:

(i) For every $\delta \in (0, 1]^\nu$,

$$\mathcal{D}_\delta = \mathcal{D} \left(\left\{ \delta_1^{|\alpha_1|} \cdots \delta_\nu^{|\alpha_\nu|} X_\alpha : \alpha \text{ is a pure power} \right\} \right)$$

is finitely generated as a C^∞ module “uniformly” in δ .

(ii) $W(\delta t) = W(\delta_1 t_1, \dots, \delta_\nu t_\nu) \in \mathcal{D}_\delta$ “uniformly” in $\delta \in (0, 1]^\nu$.

Remark 9.1. Note that, if it were not for the “uniform” part of the above assumptions, they would be independent of δ . Thus it is the uniform part which is the heart of the above assumptions. We refer the reader to [18] for a more detailed discussion of these assumptions.

Theorem 9.2. *Under the above assumptions (made precise in [18]) the operator $T : L^p \rightarrow L^p$ ($1 < p < \infty$).*

Remark 9.3. There is also a corresponding maximal result. See [16].

Remark 9.4. If $\gamma_t(x) = \exp(A(t))x$, where $A(0) \equiv 0$, then the above assumptions can be (equivalently) stated with W replaced by A . The point of using W is that that every γ has a corresponding W , while not every γ has a corresponding A .

Remark 9.5. In the single parameter case ($\nu = 1$), the assumptions of [7] can be equivalently stated as: the $\{X_\alpha\}$ satisfy Hörmander’s condition. This is a special case of the above assumptions. If the X_α do not satisfy Hörmander’s condition, the above assumptions still hold (in the case $\nu = 1$) if:

- (i) The involutive distribution generated by the $\{X_\alpha\}$ is finitely generated as a C^∞ -module.
- (ii) W is tangent to the leaves of the foliation corresponding to this distribution in an appropriate “scale invariant” way.

When the $\{X_\alpha\}$ satisfy Hörmander’s condition, this involutive distribution is the entire space of smooth sections of the tangent bundle (and is therefore clearly finitely generated) and there is only one leaf (the entire space). Thus, W is trivially tangent to this leaf (the scale invariance turns out to follow as well).

Remark 9.6. We have not separated the above conditions into analogs of the finite type and algebraic conditions discussed in Section 1.1, however this can be done. See [17].

There is another way in which our cited work is more general than what is discussed here: we deal with a larger class of singular kernels K than product kernels. We defer a general discussion to [18], and instead discuss a special case that illustrates the ideas. The setting is the three dimensional Heisenberg group \mathbb{H}^1 . As a manifold \mathbb{H}^1 is \mathbb{R}^3 and we write (x, y, t) for the coordinates of \mathbb{H}^1 . For $j = (j_1, j_2) \in \mathbb{Z}$ and $(x, y, t) \in \mathbb{H}^1$, write $2^j(x, y, t) = (2^{j_1}x, 2^{j_2}y, 2^{j_1+j_2}t)$; and for a function $f(x, y, t)$ write $f^{(2^j)}(x, y, t) = 2^{2j_1+2j_2}f(2^{j_1}x, 2^{j_2}y, 2^{j_1+j_2}t)$.

Let $\{\eta_j\}_{j \in \mathbb{Z}^2} \subseteq C_0^\infty(B^3(1))$ be a bounded set such that

$$\int \eta_j(x, y, t) dx dt = 0 = \int \eta_j(x, y, t) dy dt, \quad \forall j \in \mathbb{Z}^2.$$

Let K be a distribution defined by

$$K(x, y, t) = \sum_{j \in \mathbb{Z}^2} \eta_j^{(2^j)}(x, y, t).$$

Proposition 9.7. *The operator defined by $f \mapsto f * K$ is bounded on L^p ($1 < p < \infty$) where the convolution is taken in the sense of \mathbb{H}^1 .*

The idea of Proposition 9.7 is the following. We may write $f * K$ as

$$f * K(\xi) = \int f(e^{xX+yY+tT}\xi) K(x, y, t) dt,$$

where X, Y, T are the usual left invariant vector fields on \mathbb{H}^1 satisfying $[X, Y] = T$ and T is in the center of the Lie algebra. Because of our dilations we assign the formal degrees $(X, (1, 0))$, $(Y, (0, 1))$, and $(T, (1, 1))$; so that T is a vector field corresponding to a non-pure power. Note that $(T, (1, 1)) = ([X, Y], (1, 0) + (0, 1))$ and so the algebraic condition is satisfied. The finite type condition trivially holds.

Remark 9.8. If the convolution in Proposition 9.7 is replaced with the convolution using the usual group structure of \mathbb{R}^3 , then there are kernels K (of the above type) such that the above operator is not bounded on L^p for any p . Moreover, the algebraic condition is not satisfied. Indeed, if the usual Euclidean group structure is used, then X, Y , and T would be replaced with $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, and $\frac{\partial}{\partial t}$, respectively. Since $\left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right] = 0$, the algebraic condition fails.

Acknowledgements

The first author was partially supported by NSF DMS-0901040 and the second author was partially supported by NSF DMS-0802587.

References

- [1] J. Bourgain, *A remark on the maximal function associated to an analytic vector field*, in Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987), Vol. 137 of *London Math. Soc. Lecture Note Ser.*, 111–132, Cambridge Univ. Press, Cambridge (1989).
- [2] A. Carbery, S. Wainger, and J. Wright, *Double Hilbert transforms along polynomial surfaces in \mathbb{R}^3* , *Duke Math. J.* **101** (2000), no. 3, 499–513.
- [3] ———, *Singular integrals and the Newton diagram*, *Collect. Math.* (2006), no. Vol. Extra, 171–194.
- [4] ———, *Triple Hilbert transforms along polynomial surfaces in \mathbb{R}^4* , *Rev. Mat. Iberoam.* **25** (2009), no. 2, 471–519.
- [5] Y.-K. Cho, S. Hong, J. Kim, and C. W. Yang, *Triple Hilbert transforms along polynomial surfaces*, *Integral Equations Operator Theory* **65** (2009), no. 4, 485–528.
- [6] M. Christ, *The strong maximal function on a nilpotent group*, *Trans. Amer. Math. Soc.* **331** (1992), no. 1, 1–13.
- [7] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, *Singular and maximal Radon transforms: analysis and geometry*, *Ann. of Math. (2)* **150** (1999), no. 2, 489–577.
- [8] C. Lobry, *Contrôlabilité des systèmes non linéaires*, *SIAM J. Control* **8** (1970) 573–605.
- [9] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, *J. Math. Soc. Japan* **18** (1966) 398–404.
- [10] A. Nagel, F. Ricci, and E. M. Stein, *Singular integrals with flag kernels and analysis on quadratic CR manifolds*, *J. Funct. Anal.* **181** (2001), no. 1, 29–118.
- [11] A. Nagel, E. M. Stein, and S. Wainger, *Differentiation in lacunary directions*, *Proc. Nat. Acad. Sci. U.S.A.* **75** (1978), no. 3, 1060–1062.

- [12] ———, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. **155** (1985), no. 1-2, 103–147.
- [13] A. Nagel and S. Wainger, *L^2 boundedness of Hilbert transforms along surfaces and convolution operators homogeneous with respect to a multiple parameter group*, Amer. J. Math. **99** (1977), no. 4, 761–785.
- [14] F. Ricci and E. M. Stein, *Multiparameter singular integrals and maximal functions*, Ann. Inst. Fourier (Grenoble) **42** (1992), no. 3, 637–670.
- [15] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, Vol. 43 of *Princeton Mathematical Series*, Princeton University Press, Princeton, NJ (1993), ISBN 0-691-03216-5. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [16] E. M. Stein and B. Street, *Multi-parameter singular Radon transforms II: the L^p theory* (2010). In preparation.
- [17] ———, *Multi-parameter singular Radon transforms III: real analytic surfaces* (2010). In preparation.
- [18] B. Street, *Multi-parameter singular Radon transforms I: the L^2 theory* (2010). Preprint.
- [19] ———, *Multi-parameter Carnot-Carathéodory balls and the theorem of Frobenius* (2011). To appear in Rev. Mat. Iberoam.
- [20] T. Tao and J. Wright, *L^p improving bounds for averages along curves*, J. Amer. Math. Soc. **16** (2003), no. 3, 605–638 (electronic).

PRINCETON UNIVERSITY, DEPARTMENT OF MATHEMATICS, FINE HALL, WASHINGTON ROAD,
PRINCETON NJ 08544-1000 USA

E-mail address: `stein@math.princeton.edu`

UNIVERSITY OF WISCONSIN-MADISON, DEPARTMENT OF MATHEMATICS, 480 LINCOLN DRIVE,
MADISON WI 53706-1388 USA

E-mail address: `street@math.wisc.edu`