SPHERICAL MEANS ON THE HEISENBERG GROUP: STABILITY OF A MAXIMAL FUNCTION ESTIMATE

THERESA C. ANDERSON   LAURA CLADEK
MALABIKA PRAMANIK   ANDREAS SEEGER

Abstract. Consider the surface measure μ on a sphere in a nonvertical hyperplane on the Heisenberg group $\mathbb{H}^n$, $n \geq 2$, and the convolution $f * \mu$. Form the associated maximal function $Mf = \sup_{t>0} |f * \mu_t|$ generated by the automorphic dilations. We use decoupling inequalities due to Wolff and Bourgain-Demeter to prove L^p-boundedness of M in an optimal range.

1. Introduction

Let $\mathbb{H}^n$ be the Heisenberg group of Euclidean dimension $2n+1$, with the group law

$$(x, u) \cdot (y, v) = (x + y, u + v + x^{\intercal} J y)$$

with J is a nondegenerate skew symmetric $2n \times 2n$ matrix. Consider $\mathbb{H}^n$ as the vector space $\mathbb{R}^{2n+1}$ and let $\mathbb{V}$ be a linear subspace of dimension $2n$ which does not contain the center of $\mathbb{H}^n$, i.e.

$$\mathbb{V} \equiv \mathbb{V}_\lambda = \{(x, \lambda(x))\}$$

is the graph of a linear functional $\lambda : \mathbb{R}^{2n} \to \mathbb{R}$. Let Σ be a convex hypersurface in $\mathbb{V}$, with *nonvanishing curvature*, which contains the origin in its interior. Note that Σ is a surface of codimension two in $\mathbb{H}^n$. Let μ be a smooth density on Σ, that is $\mu = \chi d\sigma$ where $d\sigma$ is surface measure on Σ and $\chi \in C_c^\infty$.

The natural dilation group $\{\mathrm{Dil}_t\}_{t>0}$ of automorphisms on $\mathbb{H}^n$ is given by

$$(x, u) \mapsto \mathrm{Dil}_t(x, u) = (tx, t^2 u)$$

where $x \in \mathbb{R}^{2n}$, $u \in \mathbb{R}$. Define the dilated measure $\mu_t \equiv \mathrm{Dil}_t \mu$ by its action on Schwartz functions f,

$$\langle \mu_t, f \rangle = \int f(\mathrm{Dil}_t(x, u)) d\mu(x, u) \tag{1.1}$$

and consider the maximal function

$$\mathcal{M} f = \sup_{t>0} |f * \mu_t| \tag{1.2}$$

2010 *Mathematics Subject Classification.* 42B25, 22E25, 43A80, 35S30.

Preliminary version. Research supported in part by the National Science Foundation.

where the convolution refers to the noncommutative convolution on the Heisenberg group (see (2.1)).

The purpose of this paper is to prove a sharp result on L^p boundedness for $n \geq 2$; a corresponding estimate on $\mathbb{H}^1$ will remain open. The problem for $n \geq 2$ was first taken up in a paper by Nevo and Thangavelu [11] who considered the spherical measure on $\mathbb{R}^{2n} \times \{0\}$ (i.e. the case $\lambda = 0$) and proved L^p boundedness for the maximal operator in the non-optimal range $p > \frac{2n-1}{2n-2}$. An optimal result for $\lambda = 0$ was proved by D. Müller and one of the authors in [9]. There it is shown that L^p boundedness holds for $p > \frac{2n}{2n-1}$ when $n \geq 2$ and $\lambda = 0$. In the case $\lambda \neq 0$ the paper [9] only has a non-optimal result, proving L^p boundedness of the maximal operator for $p > \frac{2n-1/3}{2n-4/3}$. It was also conjectured that boundedness for $\lambda \neq 0$ remains true in the larger range $p > \frac{2n}{2n-1}$. The results in [9] actually cover the larger class of *Métivier groups* which strictly contains the class of groups of Heisenberg type (with possibly higher dimensional center). We note that for the case $\mathbb{H}^n$, $n \geq 2$, $\lambda = 0$ an alternative proof of the result in [9] was given by Narayanan and Thangavelu [10], who used spectral theoretic arguments and the representation theory of the Heisenberg group.

The crucial difference between the two cases $\lambda = 0$ and $\lambda \neq 0$ is that the automorphic dilations $\{\mathrm{Dil}_t\}$ act on V_0 but not on V_λ for $\lambda \neq 0$. We refer to [9] for an explanation of this phenomenon in terms of the geometry of the underlying Fourier integral operators with folding canonical relations. For the case $\lambda \neq 0$ the L^2 methods of both [9] and [10] are no longer applicable to obtain the optimal range of L^p-boundedness. Here we use different L^p methods based on Wolff's decoupling inequality [18], [8] and its recent improvements by Bourgain and Demeter [1] to prove the conjecture in [9] for the Heisenberg groups $\mathbb{H}^n$, $n \geq 2$, for all subspaces V_λ. The approach is motivated by previous results on L^q-Sobolev estimates for averaging operators associated to families of curves in [12], [13]. In an early version for generalized Radon transforms associated with families of curves in three dimensions ([14]) the relevant decoupling inequalities are proved by an induction procedure where a scaled version of the constant coefficient decoupling inequality is combined with a nonlinear change of variables, at every stage in the iteration. We use this idea here as well. The resulting theorem can be interpreted as a stability result for the maximal function estimate in [9].

Theorem 1.1. *Let $n \geq 2$, $\Sigma \subset V_\lambda$ as above and μ be a smooth density on Σ. Let M as in* (1.2) *and $p > \frac{2n}{2n-1}$. Then M extends to a bounded operator on $L^p(\mathbb{H}^n)$.*

Remark: It is instructive to note that an analogous stability result may fail for other more regular measures. For example, if we let ν_λ be the measure on V_λ given by $\mathbb{1}_B dS$ where dS is the $2n$-dimensional Lebesgue measure on $\mathbb{V}_\lambda$ and $\mathbb{1}_B$ is the characteristic function of the unit ball centered at the origin. Then results on maximal and singular Radon transforms [15] show that for $\lambda = 0$ the maximal operator $f \mapsto \sup_t |f * \mathrm{Dil}_t \nu_0|$ is bounded on $L^p(\mathbb{H}^n)$ for

$1 < p < \infty$. However for $\lambda \neq 0$ the maximal operator $f \mapsto \sup_t |f * \mathrm{Dil}_t \nu_\lambda|$ is bounded on $L^p(\mathbb{H}^n)$ only for $\frac{2n+1}{2n} < p < \infty$. The local analogue of the latter maximal operator (with dilations parameters in $[1, 2]$) shares some properties with the spherical maximal operator (*cf.* [16]), due to the rotation effect of the nonisotropic dilation structure. One has an example that shows unboundedness that is similar to the example that shows unboundedness of the spherical maximal operator $L^p(\mathbb{R}^d)$, when $p \leq \frac{d}{d-1}$ (cf. [17], [16]).

This paper. In §2 consider regularizations of the measure defined by dyadic frequency decompositions and prove a crucial L^p-Sobolev inequality for the convolution $f * \mu$ when acting on L^p functions with compact support. As a consequence we obtain an estimate for a restricted version of the maximal operator where the dilation parameter is taken in a compact subinterval of $\mathbb{R}^+$. In §4 we describe the basic decoupling step. In §5 an iteration and combination with known L^2 estimates leads to the proof of the main Proposition 2.1. In §6 we use Calderón-Zygmund type arguments to extend this result to obtain Theorem 1.1. The appendix contains a basic integration by parts lemma which is useful in checking the details of the decoupling step in §4.

2. Main estimates

The convolution $f_1 * f_2(x, u) = \int f_1(y, v) f_2((y, v)^{-1}(x, u)) dy dv$ on the Heisenberg group can be written as

$$\begin{aligned} f_1 * f_2(x, u) &= \int f_1(y, v) f_2(x - y, u - v + x^\intercal J y) dy dv \\ &= \int f_2(y, v) f_1(x - y, u - v - x^\intercal J y) dy dv \end{aligned} \tag{2.1}$$

Here J is a nondegenerate skew symmetric matrix on $\mathbb{R}^{2n}$.

Split $x = (\underline{x}, x_{2n})$ where $\underline{x} \in \mathbb{R}^{2n-1}$. We consider a localization of the measure to a graph $x_{2n} = g(\underline{x})$ on V_λ, where the Hessian of g is nondegenerate. We will use permutation of variables to reduce to this situation (*cf.* the remark in §5).

The localized measure μ can be represented as an oscillatory integral distribution by

$$\eta(x, u) \iint e^{i(\sigma(x_{2n} - g(\underline{x})) + \tau(u - \lambda(x)))} d\sigma d\tau \tag{2.2}$$

where η is a smooth compactly supported function.

Let $\varsigma_0 \in C_0^\infty(\mathbb{R})$ be an even smooth function such that $\varsigma_0(s) = 1$ if $|s| \leq 1$ and such that the support of ς_0 is contained in $(-2, 2)$. Let $\varsigma_1(s) = \varsigma_0(s/2) - \varsigma_0(s)$ and let, for $k \geq 1$, $\zeta_k(s) = \varsigma_1(2^{1-k}s)$. Also let $\zeta_0 = \varsigma_0$. Then for $k \geq 1$, ζ_k is supported in $\{s : 2^{k-1} \leq |s| \leq 2^{k+1}\}$, $k \geq 1$, and we have $\sum_{k=0} \zeta_k = 1$.

Let

$$\mu_k(x, u) = \eta(x, u) \iint e^{i(\sigma(x_{2n} - g(\underline{x})) + \tau(u - \lambda(x)))} \zeta_k((\sigma^2 + \tau^2)^{1/2}) \, d\sigma \, d\tau \tag{2.3}$$

and we decompose (2.2) as $\sum_{k=0}^{\infty} \mu_k$ in the sense of distributions.

The maximal function $\sup_{t>0} |f * \mathrm{Dil}_t \mu_k|$ is dominated by $C(k)$ times the analogue of the Hardy-Littlewood maximal function on the Heisenberg group. Therefore it suffices to consider the case $k \gg 1$.

Our main proposition will be

Proposition 2.1. *(i) For* $q > \frac{4n}{2n-1}$,

$$\|f * \mu_k\|_{L^q(\mathbb{H}^n)} \lesssim 2^{-k\frac{2n-1}{q}} \|f\|_{L^q(\mathbb{H}^n)} .$$

(ii) For $p < \frac{4n}{2n+1}$

$$\|f * \mu_k\|_{L^p(\mathbb{H}^n)} \lesssim 2^{-k(2n-1)(1-\frac{1}{p})} \|f\|_{L^p(\mathbb{H}^n)} .$$

Moreover, for $1 \le s \le 2$

$$\Big\| \frac{d}{ds} f * \mathrm{Dil}_s \mu_k \Big\|_{L^p(\mathbb{H}^n)} \lesssim 2^k 2^{-k(2n-1)(1-\frac{1}{p})} \|f\|_{L^p(\mathbb{H}^n)} .$$

The implicit constants are uniform if λ is taken from a compact subset of $(\mathbb{R}^{2n})^*$.

A well known Sobolev imbedding argument gives a sharp bound for the restricted maximal function:

Corollary 2.2. *For* $p < \frac{4n}{2n+1}$,

$$\Big\| \sup_{1/2<s<2} |f * \mathrm{Dil}_s \mu_k| \Big\|_{L^p(\mathbb{H}^n)} \lesssim 2^{k(\frac{2n}{p}-2n+1)} \|f\|_{L^p(\mathbb{H}^n)}.$$

We use a further decomposition in the σ-variables, as in [9]. Let

$$\begin{aligned}
\zeta_{k,0}(\sigma,\tau) &= \varsigma_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})(1-\varsigma_0(2^{-k}\sigma)) \\
\zeta_{k,l}(\sigma,\tau) &= \varsigma_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})\varsigma_1(2^{l-k}\sigma) \\
\widetilde{\zeta}_k(\sigma,\tau) &= \varsigma_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})\varsigma_0(2^{[k/3]-k-1}\sigma).
\end{aligned}$$

so that

$$\zeta_k((\sigma^2+\tau^2)^{1/2}) = \widetilde{\zeta}_k + \sum_{0\le l<k/3} \zeta_{k,l}.$$

Let $\mu_{k,l}$ be defined as in (2.3) but with $\zeta_k((\sigma^2+\tau^2)^{1/2})$ replaced by $\zeta_{k,l}$ when $l < k/3$ and by $\widetilde{\zeta}_k$ when $l = [k/3]$. We shall prove the following refined version of Proposition 2.1.

Proposition 2.3. *Let* $\varepsilon > 0$ *and* $2 \le q < \frac{4n+2}{2n-1}$. *Then there is* $C_\varepsilon > 0$ *such that for* $0 \le l \le [k/3]$,

$$\|f * \mu_{k,l}\|_q \le C_\varepsilon 2^{-k\frac{2n-1}{q}} 2^{l(\varepsilon+\frac{2n}{q}-\frac{2n-1}{2})} \|f\|_q.$$

If $\frac{4n}{2n-1} < q < \frac{4n+2}{2n-1}$ and $\varepsilon > 0$ is sufficiently small then we can sum in l and obtain part (i) of Proposition 2.1. The L^p inequality for $p < \frac{4n}{2n+1}$ follows by duality and the estimate for $\frac{d}{ds} f * \mathrm{Dil}_s \mu_k$ is proved similarly.

Remark. We also have, by interpolation with an easy L^∞ estimate,

$$\|f * \mu_{k,l}\|_q \le C_\varepsilon 2^{-k\frac{2n-1}{q}} 2^{-l\frac{1-\varepsilon}{q}} \|f\|_q, \quad \frac{4n+2}{2n-1} \le q \le \infty. \tag{2.4}$$

3. *Background and idea of the proof*

The idea in the proof of Proposition 2.1 is to consider the fibers of the fold surface which curved varying cones; this goes back to the paper [5] by Greenleaf and one of the authors which dealt with $L^2 \to L^p$ inequalities for classes of generalized Radon transforms. One then would like to apply decoupling for localizations to plates adapted to neighborhoods of these cones. The cones vary with the base points and some approximation and preparations via changes of variables have to be used, *cf.* [14].

Concretely if $\chi_1(x,u)$ and $\chi_2(y,v)$ are compactly supported C_c^∞ functions we want to examine the functions $(f\chi_2)*\mu_{k,l}(x,u)\chi_1(x,u)$ which are written in the form

$$\int \mathcal{K}_{k,l}(x,u,y,v) f(y,v)\,dy\,dv$$

where the Schwartz kernel $\mathcal{K}_{k,l}$ is given by

$$\chi_1(x,u)\chi_2(y,v) \iint \zeta(2^{l-k}\sigma)\chi_1(2^{-k}\sqrt{\sigma^2+\tau^2}) e^{i\varphi(\sigma,\tau,x,u,y,v)} d\sigma d\tau$$

and the phase function is defined by

$$\varphi(\sigma,\tau,x,u,y,v) = \sigma(x_{2n} - y_{2n} - g(\underline{x} - \underline{y})) + \tau(u - v + x^\intercal J y) \tag{3.1}$$

where J is a skew symmetric nondegenerate $2n \times 2n$ matrix (for example a skew symmetric perturbation of the standard symplectic matrix). Note we do not assume that J is orthogonal.

With $\varphi(\sigma,\tau,x,u,y,v)$ as in (3.1) the cones in question are given, for each (x,u), by

$$\begin{aligned} &\{(\varphi_x, \varphi_u) :\ \sigma = 0,\, x_{2n} - y_{2n} - g(\underline{x} - \underline{y}) = 0,\, v = u + x^\intercal J y\} \\ =&\{(\tau J y, \tau) :\ y_{2n} = x_{2n} - g(\underline{x} - \underline{y})\} \end{aligned}$$

which is actually independent of u. Denote this conic surface by Σ_x and let

$$\Gamma^x(\underline{y}) = (\underline{y}, x_{2n} - g(\underline{x} - \underline{y})). \tag{3.2}$$

Then

$$\Sigma_x = \{(\tau J \Gamma^x(\underline{y}), \tau)\}. \tag{3.3}$$

We wish to use the decoupling inequalities in [1] (or the previous paper [8] if n is sufficiently large) for thin neighborhoods of the cones Σ_x, for suitable frozen x. Note that by our assumptions on g the cones are maximally curved (i.e. $d-2=2n-1$ principal curvatures are nonzero). The basic decoupling step will be described in the next section.

4. The decoupling step

Let $\delta_0 > 2^{-l}$ and let $\delta_1 < \delta_0$ be such that

$$\delta_1 \ge (2^{-l}\delta_0)^{1/2} \tag{4.1}$$

Fix $a \in \mathbb{R}^{2n}$, $\underline{b} \in \mathbb{R}^{2n-1}$. Suppose we are given a family of disjoint cubes $\{Q_\nu\}$ in $\mathbb{R}^{2n-1}$ of side length δ_1 contained in the reference cube

$$Q := \{\underline{y} \in \mathbb{R}^{2n-1} : |y_i - b_i| \le \delta_0, i = 1, \dots 2n-1\}. \tag{4.2}$$

Suppose in what follows that for each ν the function $(y,v) \mapsto f_\nu(y,v)$ is supported in $Q_\nu \times \mathbb{R} \times \mathbb{R}$.

We fix $\varepsilon > 0$ and let

$$\delta_1 \ge 2^{-l(1-\varepsilon)}. \tag{4.3}$$

Let φ be as in (3.1)

Let χ_{l,a,u° be a smooth function supported in a ball of sidelength 2^{-l} centered at (a,u°), satisfying $|\partial^\alpha \chi_{l,a,u^\circ}| \le C_\alpha 2^{l|\alpha|}$ for all multiindices α. Let ζ be a smooth function supported in $(-2,2)$. Let $K = K_{k,l,a,u^\circ}$ be given by

$$\begin{aligned} &K(x,u,y,v) \\ &\quad = \chi_{l,a,u^\circ}(x,u) \iint \zeta(2^{l-k}\sigma)\chi_1(2^{-k}\sqrt{\sigma^2+\tau^2}) e^{i\varphi(\sigma,\tau,x,u,y,v)} d\sigma d\tau \end{aligned} \tag{4.4a}$$

which after a change of variable (replacing $2^{-k}(\sigma,\tau)$ by (σ,τ)) we can write

$$K(x,u,y,v) = 2^{2k} \iint \gamma(\sigma,\tau,x,u,y,v) e^{i2^k\varphi(\sigma,\tau,x,u,y,v)} d\sigma d\tau \tag{4.4b}$$

with

$$\gamma(\sigma,\tau,x,u,y,v) = 0 \text{ if } |x-a| + |u-u^\circ| + |\sigma| \gtrsim 2^{-l} \tag{4.5a}$$

$$|\partial^M_{\sigma,\tau,x,u,y,v}\gamma| \le C_M 2^{lM}. \tag{4.5b}$$

Here $\partial^M_{\dots}$ stands for any differentiation of order M in the variables indicated.

We let T denote any such operator with kernel K and γ as above.

Proposition 4.1. *Let $2 \le q \le \frac{4n+2}{2n-1}$. Let $0 < \varepsilon \le 1$, $k \gg 1$, $l \le k/3$, $\delta_1 \ge 2^{-l(1-\varepsilon)}$. With the above specifications on Q and $\{Q_\nu\}$ we have, for any $\varepsilon_1 \in (0,\varepsilon)$, and $N \in \mathbb{N}$,*

$$\begin{aligned} \Big\| T[\sum_\nu f_\nu] \Big\|_q \le C(\varepsilon_1)(\delta_0/\delta_1)^{(2n-1)(\frac12-\frac1q)+\varepsilon_1} \Big(\sum_\nu \|Tf_\nu\|_q^q\Big)^{1/q} \\ + \tilde C(\varepsilon,N) 2^{-kN} \sup_\nu \|f_\nu\|_q. \end{aligned} \tag{4.6}$$

4.1. *A model case.* We first consider the model situation δ_0, δ_1 as in (4.1), Q as in (4.2), φ as (3.1), such that

$$a = 0, \ \underline{b} = 0, \ \nabla g(0) = 0. \tag{4.7}$$

As pointed out above the crucial tool is the decoupling estimate from [1]. The relevant cones Σ_0 are given by

$$(\underline{y}, \tau) \mapsto \tau(J\Gamma(\underline{y}), 1), \tag{4.8}$$

with $\Gamma \equiv \Gamma^0$ as in (3.2), i.e.

$$\Gamma(\underline{y}) = (\underline{y}, -g(-\underline{y})). \tag{4.9}$$

Let

$$N(\underline{y}) = e_{2n} - \sum_{i=1}^{2n-1} \partial_i g(-\underline{y}) e_i \tag{4.10}$$

which is normal to $\underline{y} \mapsto \Gamma(\underline{y})$. Let $\underline{y}^\nu \in Q_\nu$ and let $N_\nu = N(\underline{y}^\nu)/|N(\underline{y}^\nu)|$. Let $J^\#$ be the contragredient matrix, i.e. $J^\# = (J^{-1})^\intercal$. Since J is skew symmetric we have $J^\# = -J^{-1}$. $J^\# N(\underline{y})$ is normal to $J\Gamma$ and

$$\langle J \frac{\partial^2 \Gamma(y)}{\partial y_j \partial y_k}, \frac{J^\# N(\underline{y})}{|J^\# N(\underline{y})|} \rangle = \frac{|N(\underline{y})|}{|J^\# N(\underline{y})|} \langle \frac{\partial^2 \Gamma(y)}{\partial y_j \partial y_k}, \frac{N(\underline{y})}{|N(\underline{y})|} \rangle.$$

This relates the curvature form for Γ to the curvature form for $J\Gamma$, and the decoupling estimates from [1], in the version for general curved cones ([13]), are applicable.

It turns out that, in order to perform the decoupling step via the Bourgain-Demeter inequality we will have to make a change of variable in the (x, u) variables, using a quadratic shear transformation. Thus we consider instead the operator $\mathcal{T}$ defined by

$$\mathcal{T} f(x, u) = T f(x, u + \tfrac{1}{2} \langle Sx, x \rangle)$$

where S is a suitable symmetric linear transformation. Obviously, by changing variables, (4.6) holds with T if and only if it holds with $\mathcal{T}$.

It will be important in the proof to choose S such that the following crucial assumption

$$S J^\# e_{2n} = -e_{2n} \tag{4.11}$$

is satisfied. To see that S can be chosen in smooth dependence on J we notice that $u_1 := e_{2n}$ and $u_2 = J^\# e_{2n}/|J^\# e_{2n}|$ form an orthonormal basis on $\mathbb{V} = \operatorname{span}(J^\# e_{2n}, e_{2n})$ which can be extended to an orthonormal basis $\{u_1, \dots, u_{2n}\}$ of $\mathbb{R}^{2n}$. Let $c = |J^\# e_{2n}|$ and we let $Su_2 = -c^{-1} u_1$, $Su_1 = -c^{-1} u_2$ and $Su_i = u_i$ for $i = 3, \dots, 2n$. Then S is symmetric, invertible and

$$\min\{1, |J^\# e_{2n}|^{-1}\} \le \|S\| \le \max\{1, |J^\# e_{2n}|^{-1}\}$$

where $\|S\|$ denotes the $\ell^2 \to \ell^2$ operator norm on $2n \times 2n$ matrices.

The Schwartz kernel of $\mathcal{T}$ is given by

$$2^{2k} \iint \gamma_1(\sigma,\tau,x,u,y,v) e^{i2^k \Phi(\sigma,\tau,x,u,y,v)} d\sigma d\tau \tag{4.12}$$

with

$$\gamma_1(\sigma,\tau,x,u,y,v) = \gamma(\sigma,\tau,x,u+\tfrac{1}{2}\langle Sx,x\rangle,y,v)$$
$$\Phi(\sigma,\tau,x,u,y,v) = \varphi(\sigma,\tau,x,u,y,v) + \frac{\tau}{2}\langle Sx,x\rangle$$

We now define nonisotropic cylinders (or "plates") associated to the cone (4.8). We use the notation

$$\xi = (\overline{\xi}, \xi_{2n+1}) = (\underline{\xi}, \xi_{2n}, \xi_{2n+1}).$$

The tangent space to the cone at $\tau(J\Gamma(\underline{y}^\nu),1)$ is spanned by

$$\{J\Gamma(\underline{y}^\nu),1\} \cup \{(J\partial_i\Gamma(\underline{y}^\nu),0) : i=1,\dots,2n-1\},$$

and a normal vector is given by

$$\begin{pmatrix} J^{\#}N_\nu \\ -\langle J\Gamma(\underline{y}^\nu), J^{\#}N_\nu\rangle \end{pmatrix} = \begin{pmatrix} J^{\#}N_\nu \\ -\langle \Gamma(\underline{y}^\nu), N_\nu\rangle \end{pmatrix}. \tag{4.13}$$

The relevant plates are $2^k\Pi_\nu(\delta_1)$ where the normalized plates $\Pi_\nu(\delta_1)$ are defined by the inequalities

$$C^{-1} \le \sqrt{|\overline{\xi}|^2 + \xi_{2n+1}^2} \le C \tag{4.14a}$$

$$|\overline{\xi} - \xi_{2n+1}J\Gamma(\underline{y}^\nu)| \le C\delta_1 \tag{4.14b}$$

$$\big|\langle \overline{\xi} - \xi_{2n+1}J\Gamma(\underline{y}^\nu), J^{\#}N_\nu\rangle\big| \le C\delta_1^2. \tag{4.14c}$$

The Bourgain-Demeter decoupling theorem gives that

$$\Big\|\sum_\nu F_\nu\Big\|_q \le C(\varepsilon)(\delta_1/\delta_0)^{-\varepsilon}\Big(\sum_\nu \|F_\nu\|_q^2\Big)^{1/2}, \quad 2 \le q \le \frac{4n+2}{2n-1},$$

provided that the Fourier transforms $\widehat{F}_\nu$ are supported in $2^k\Pi_\nu(\delta_1)$. We have some freedom in the choice of the constant C ranging over a compact subset of $(0,\infty)$. Let η_ν be a bump function which is equal to 1 on $\Pi_\nu(\delta_1)$ and is supported on its double, and η_ν satisfies the natural differential inequalities. Specifically consider the radial tangential, nonradial tangential, and normal differentiation operators:

$$\mathcal{V}_{\nu,0} = \langle J\Gamma(\overline{y}^\nu), \nabla_{\overline{\xi}}\rangle + \frac{\partial}{\partial \xi_{2n+1}},$$
$$\mathcal{V}_{\nu,i} = \frac{\partial}{\partial \xi_i} - \langle J\Gamma(\overline{y}^\nu), e_i\rangle \frac{\partial}{\partial \xi_{2n+1}}, \quad i=1,\dots,2n-1,$$
$$\mathcal{V}_\nu = \langle J^{\#}N_\nu, \nabla_{\overline{\xi}}\rangle - \langle \Gamma(\overline{y}^\nu), N_\nu\rangle \frac{\partial}{\partial \xi_{2n+1}}.$$

Then

$$\big|\mathcal{V}_{\nu,0}^{\alpha_0}\mathcal{V}_{\nu,1}^{\alpha_1}\dots\mathcal{V}_{\nu,2n-1}^{\alpha_{2n-1}}\mathcal{V}_\nu^{\alpha_{2n}}\eta_\nu(\xi)\big| \le C_{\alpha_0,\alpha_1,\dots,\alpha_{2n}}(1/\delta_1)^{2\alpha_{2n}+\sum_{i=1}^{2n-1}\alpha_i}.$$

Define the Euclidean convolution operator $P_{k,\nu}$ in the multiplier formulation by

$$\widehat{P_{k,\nu}f}(\xi) = \eta_\nu(2^{-k}\xi)\widehat{f}(\xi).$$

Then by the decoupling inequality

$$\Big\| \sum_\nu P_{k,\nu}\mathcal{T}f_\nu \Big\|_q \le C(\varepsilon)(\delta_1/\delta_0)^{-\varepsilon}\Big(\sum_\nu \|\mathcal{T}f_\nu\|_q^2\Big)^{1/2}, \qquad 2 \le q \le \frac{4n+2}{2n-1}. \tag{4.15}$$

We need to analyze the Schwartz kernel of $f \mapsto (I - P_{k,\nu})\mathcal{T}$ when acting on f_ν. Thus we consider for $y \in Q_\nu$,

$$2^{(2n+3)k}\int\iint e^{i2^k\langle x-\tilde{x},\bar{\xi}\rangle + i2^k(u-\tilde{u})\xi_{2n+1}}(1-\eta_\nu(2^{-k}\xi)) \times \iint e^{i2^k\Phi(\sigma,\tau,\tilde{x},\tilde{u},y,v)}\gamma_1(\sigma,\tau,\tilde{x},\tilde{u},y,v)d\sigma d\tau\, d\tilde{x}d\tilde{u}\, d\xi\,.$$

We can replace $(I-P_{k,\nu})\mathcal{T}f_\nu$ with $(I-P_{k,\nu})L_k\mathcal{T}f_\nu$ with L_k a Littlewood-Paley cutoff operator localizing to frequencies $C_1^{-1}2^k \le |\xi| \le C_1 2^k$ (as the remaining error terms are handled by standard integration by part arguments, see e.g.[7, §4.2]). Also if $|x|+|u| > C$ we see that by an additional integration by parts arguments in the (ξ,σ,τ) variables we get a bound $\lesssim C2^{-(k-l)N}(|x|+|u|)^{-N}$. Hence we need to show that the L^∞ norm of the integral is $O(2^{-kN})$. Since the $(\tilde{x},\tilde{u})$ integral is over a compact set it suffices to show that the Fourier transform of $\mathcal{T}f_\nu$ in the complement of the double plate has norm $O(2^{-kN})$. Moreover, if $|\tau - \xi_{2n+1}| \ge 2^{l-k}$ then we can integrate by parts with respect to $\tilde{u}$ and show that the resulting integral is $O(2^{-kN})$.

It remains to analyze, for $|\tau| \approx |\xi_{2n+1}|$ and $y \in Q_\nu$, the behavior of

$$\iiint e^{i2^k(\Phi(\sigma,\tau,x,u,y,v)-\langle x,\bar{\xi}\rangle - u\xi_{2n+1})}\gamma_1(\sigma,\tau,x,u,y,v)d\sigma\, dx\, du \tag{4.16}$$

assuming that $\xi \notin \Pi_\nu(\delta)$ with C in (4.14) large. For better readability we have changed the notation from $(\tilde{x},\tilde{u})$ to (x,u) in (4.16). In what follows we will set

$$\Psi(\sigma,\tau,x,u,y,v,\xi) = \Phi(\sigma,\tau,x,u,y,v) - \langle x,\bar{\xi}\rangle - u\xi_{2n+1}\,. \tag{4.17}$$

In order to estimate Fourier transforms in the complement of plates we need bounds for certain directional derivatives of the phase function Φ (which will then turned into lower bounds for the directional derivatives of Ψ when ξ is away from the plate).

Lemma 4.2. *There is a constant $A \ge 1$ so that the following statements hold for all $y \in Q_\nu$, $|\sigma| \approx 2^{-l}$, $|\tau| \approx 1$.*

(i) Let $\vec{V}_{\nu,i} = e_i - \langle J\Gamma(\underline{y}^\nu), e_i\rangle e_{2n+1}$. Then

$$\big|\langle \vec{V}_{\nu,i}, \nabla_{x,u}\Phi\rangle\big| \le A\big(\delta_1 + |x_{2n} - y_{2n} - g(\underline{x}-\underline{y})|\big). \tag{4.18}$$

(ii) Let $\vec{V}_\nu$ be the normal vector in (4.13)*. Then*

$$\big|\langle \vec{V}_\nu, \nabla_{x,u}\Phi\rangle\big| \le A\big(\delta_1^2 + |x_{2n} - y_{2n} - g(\underline{x}-\underline{y})|\big). \tag{4.19}$$

4.1.1. *Proof of Lemma 4.2.* To see (i) we have for $i = 1, \dots, 2n$,

$$\begin{aligned}\langle \vec{V}_{\nu,i}, \nabla_{x,u}\Phi\rangle &= \frac{\partial \Phi}{\partial x_i} - \langle J\Gamma(\underline{y}^\nu), e_i\rangle \frac{\partial \Phi}{\partial u}\\ &= \sigma\langle N(\underline{y}), e_i\rangle + \tau(\langle Jy, e_i\rangle - \langle J\Gamma(\underline{y}^\nu), e_i\rangle + \langle Sx, e_i\rangle).\end{aligned}$$

We have $\sigma = O(2^{-l})$, $Sx = O(2^{-l})$ and

$$\begin{aligned}Jy - J\Gamma(\underline{y}^\nu) &= (y_{2n} + g(-\underline{y}))Je_{2n} + (g(-\underline{y}^\nu) - g(-\underline{y}))Je_{2n}\\ &= (y_{2n} + g(-\underline{y}))Je_{2n} + O(\delta_1)\end{aligned}$$

Split

$$\begin{aligned}|y_{2n} + g(-\underline{y}^\nu)| \le\ & |y_{2n} - x_{2n} + g(\underline{x}-\underline{y})| + |x_{2n}|\\ &+ |g(-\underline{y}) - g(\underline{x}-\underline{y})| + |g(-\underline{y}^\nu) - g(-\underline{y})|.\end{aligned}$$

The second and the third terms are $O(2^{-l})$, by our localization in x. The fourth term is $O(\delta_0\delta_1)$ since $\underline{y} \in Q_\nu$ and $\nabla g = O(\delta_0)$ in Q_ν. Consequently

$$|y_{2n} + g(-\underline{y}^\nu)| \lesssim (2^{-l} + \delta_0\delta_1) + |x_{2n} - y_{2n} - g(\underline{x}-\underline{y})| \tag{4.20}$$

and thus (4.18) follows easily.

Next we verify (ii). We have

$$\begin{aligned}\langle \vec{V}_\nu, \nabla_{x,u}\Phi\rangle &= \langle J^\# N_\nu, \nabla_x \Phi\rangle - \langle J^\# N_\nu, J\Gamma(\underline{y}^\nu)\rangle \frac{\partial\Phi}{\partial u}\\ &= I(x,y,\sigma) + II(y,\sigma) + III(x,y,\tau) + IV(x,\tau)\end{aligned} \tag{4.21}$$

where

$$\begin{aligned}I(x,y,\sigma) &= -\sigma \sum_{i=1}^{2n-1} \langle J^\# N_\nu, \partial_i g(\underline{x}-\underline{y}) - \partial_i g(-\underline{y})\rangle,\\ II(y,\sigma) &= -\sigma \sum_{i=1}^{2n-1} \langle J^\# N_\nu, \partial_i g(-\underline{y})\rangle + \sigma\langle J^\# N_\nu, e_{2n}\rangle,\\ III(x,y,\tau) &= \tau\langle J^\# N_\nu, Jy\rangle - \tau\langle N_\nu, \Gamma(\underline{y}^\nu)\rangle,\\ IV(x,\tau) &= \tau\langle J^\# N_\nu, Sx\rangle.\end{aligned}$$

Since $\sigma = O(2^{-l})$ and $|x| = O(2^{-l})$ we have $|I(x,y,\sigma)| \lesssim 2^{-2l}$ which is of course $O(\delta_1^2)$. By (4.10) we have

$$\begin{aligned}II(y,\sigma) &= \sigma\langle J^\# N_\nu, N(\underline{y})\rangle\\ &= \sigma\langle J^\# N_\nu, N(\underline{y}) - N(\underline{y}^\nu)\rangle + \sigma|N(\underline{y}^\nu)|\langle J^\# N_\nu, N_\nu\rangle\end{aligned}$$

and by the skew symmetry of $J^{\#}$ the last summand drops out and we get $\sigma\langle J^{\#}N_\nu, N(y)\rangle = O(2^{-l}\delta_1)$ which is $O(\delta_1^2)$. For the third term, we write

$$\begin{aligned}\tau^{-1}III(x,y,\tau) &= \langle N_\nu, y-\Gamma(\underline{y}^\nu)\rangle\\ &= \langle N_\nu, \Gamma(\underline{y})-\Gamma(\underline{y}^\nu)\rangle + \langle N_\nu, y-\Gamma(\underline{y})\rangle.\end{aligned}$$

By definition of N_ν and Taylor expansion,

$$\langle N_\nu, \Gamma(\underline{y})-\Gamma(\underline{y}^\nu)\rangle = O(|\underline{y}-\underline{y}^\nu|^2) = O(\delta_1^2).$$

Next observe $N_\nu = e_{2n} + O(\delta_0)$ and

$$\begin{aligned}x_{2n} - g(x-\underline{y}) + g(-\underline{y}) &= x_{2n} - \langle \underline{x}, \nabla g(-\underline{y})\rangle + O(2^{-2l})\\ &= \langle x, N(y)\rangle \ + O(2^{-2l})\end{aligned}$$

and thus

$$\begin{aligned}&\langle N_\nu, y-\Gamma(\underline{y})\rangle = \langle N_\nu, e_{2n}\rangle(y_{2n} + g(-\underline{y}))\\ =& -\langle N_\nu, e_{2n}\rangle(x_{2n}-y_{2n}-g(\underline{x}-\underline{y})) + \langle N_\nu, e_{2n}\rangle(x_{2n}-g(x-\underline{y})+g(-\underline{y}))\\ =& \langle N_\nu, e_{2n}\rangle\langle x, N(y)\rangle + O\big(2^{-2l} + |x_{2n}-y_{2n}-g(\underline{x}-\underline{y})|\big),\end{aligned}$$

and furthermore

$$\begin{aligned}\langle N_\nu, e_{2n}\rangle\langle x, N(y)\rangle &= \langle N(y), x\rangle(1+O(\delta_0))\\ &= \langle N_\nu, x\rangle + O(2^{-l}\delta_0)).\end{aligned}$$

Adding $\tau^{-1}IV(x,\tau)$ we get

$$\begin{aligned}\langle N_\nu, x\rangle + \tau^{-1}IV(x,\tau) &= \langle (I+SJ^{\#})N_\nu, x\rangle\\ = \langle (I+SJ^{\#})e_{2n}, x\rangle &+ O(2^{-l}\delta_0) \ = \ O(2^{-l}\delta_0)\end{aligned}$$

where in the last equation we have used the crucial assumption (4.11) on the choice of S, i.e. that e_{2n} is in the nullspace of $I+SJ^{\#}$. Collecting these estimates we obtain

$$|\tau|^{-1}|III(x,y,\tau)+IV(x,\tau)| \lesssim (\delta_1^2 + 2^{-l}\delta_0 + |x_{2n}-y_{2n}-g(\underline{x}-\underline{y})|).$$

By our assumption (4.1) we have $2^{-l}\delta_0 \lesssim \delta_1^2$ and the asserted estimate in (ii) follows. The proof is complete. □

4.1.2. *Estimation of Fourier transforms in the complement of the plates.* We apply Lemma A.2 in a two-dimensional setting where the w_1-derivative will be replaced with the directional derivative for a vector $\vec{V}$ in $\mathbb{R}^{2m+1}$ and where $w_2 = \sigma$.

We first assume that (4.14b) does not hold, i.e.

$$|\xi_i - \langle J\Gamma(\underline{y}^\nu), e_i\rangle\xi_{2n+1}| \geq C_1\delta_1 \tag{4.22}$$

for some $i \in \{1,\dots,2n\}$ and $C_1 > 2A$, with $A \geq 1$ as in Lemma 4.2.

Note that $\frac{\partial\Phi}{\partial\sigma} = x_{2n} - y_{2n} - g(\underline{x} - \underline{y})$. Hence if (4.22) holds with $C_1 \geq 2A \geq 2$ then, we get from part (i) of Lemma 4.2

$$\big|\langle \vec{V}_{\nu,i}, \nabla_{x,u}\Psi\rangle\big| + \big|\frac{\partial\Psi}{\partial\sigma}\big| \geq \delta_1/2. \tag{4.23}$$

Indeed the left hand side is equal to

$$\Big|\frac{\partial\Phi}{\partial x_i} - \langle J\Gamma(\underline{y}^\nu), e_i\rangle\frac{\partial\Phi}{\partial u} - \xi_i + \langle J\Gamma(\underline{y}^\nu), e_i\rangle\xi_{2n+1}\Big| + \Big|\frac{\partial\Phi}{\partial\sigma}\Big|$$
$$\geq \max\{0, (C_1 - A)\delta_1 - A|x_{2n} - y_{2n} - g(\underline{x} - \underline{y})|\} + |x_{2n} - y_{2n} - g(\underline{x} - \underline{y})|$$
$$\geq \frac{C_1 - A}{2A}\delta_1 \geq \frac{\delta_1}{2}.$$

We now use integration by parts. We assume (4.22) and define a differential operator L by

$$Lh = \langle \vec{V}_{\nu,i}, \nabla\rangle\Big(\frac{\langle \vec{V}_{\nu,i}, \nabla\Psi\rangle\langle \vec{V}_{\nu,i}, \nabla h\rangle}{|\langle \vec{V}_{\nu,i}, \nabla\Psi\rangle|^2 + |\frac{\partial\Psi}{\partial\sigma}|^2}\Big) + \frac{\partial}{\partial\sigma}\Big(\frac{\frac{\partial\Psi}{\partial\sigma}\frac{\partial h}{\partial\sigma}}{|\langle \vec{V}_{\nu,i}, \nabla\Psi\rangle|^2 + |\frac{\partial\Psi}{\partial\sigma}|^2}\Big). \tag{4.24}$$

The integral (4.16) becomes

$$i^N 2^{-kN} \iiint e^{i2^k(\Psi(\sigma,\tau,x,u,y,v,\xi))} L^N\gamma_1(\sigma,\tau,x,u,y,v)d\sigma dx du \tag{4.25}$$

and $|L^N\gamma| \lesssim_N (2^l\delta_1^{-1})^N$ by a straightforward analysis using Lemma A.2. Since $2^{k-l}\delta_1 \gtrsim 2^{k/3}$ we gain a factor $2^{-kN_1/3}$ with N_1 integrations by parts.

Next consider the more subtle case where (4.14c) does not hold, i.e. we have

$$\big|\langle \bar{\xi} - \xi_{2n+1}J\Gamma(\underline{y}^\nu), J^\# N_\nu\rangle\big| \geq C_1\delta_1^2 \tag{4.26}$$

provided that $C_1 \geq 2A$. We now have by part (ii) of Lemma 4.2

$$\big|\langle \vec{V}_\nu, \nabla\Psi\rangle\big| + \big|\frac{\partial\Psi}{\partial\sigma}\big| \geq \delta_1^2/2. \tag{4.27}$$

To see this observe that the left hand side is equal to

$$\Big|\langle J^\# N_\nu, \nabla_x\Phi\rangle - \langle\Gamma(\underline{y}^\nu), N_\nu\rangle\frac{\partial\Phi}{\partial u} - \langle J^\# N_\nu, \bar{\xi}\rangle + \langle\Gamma(\underline{y}^\nu), N_\nu\rangle\xi_{2n+1}\Big| + \Big|\frac{\partial\Phi}{\partial\sigma}\Big|$$
$$\geq \max\{0, (C_1 - A)\delta_1^2 - A|x_{2n} - y_{2n} - g(\underline{x} - \underline{y})|\} + |x_{2n} - y_{2n} - g(\underline{x} - \underline{y})|$$
$$\geq \frac{C_1 - A}{2A}\delta_1^2 \geq \frac{\delta_1^2}{2}.$$

We use for our integration by parts the operator $\tilde{L}$ defined by

$$\tilde{L}h = \langle \vec{V}_\nu, \nabla\rangle\Big(\frac{\langle \vec{V}_\nu, \nabla\Psi\rangle\langle \vec{V}_\nu, \nabla h\rangle}{|\langle \vec{V}_\nu, \nabla\Psi\rangle|^2 + |\frac{\partial\Psi}{\partial\sigma}|^2}\Big) + \frac{\partial}{\partial\sigma}\Big(\frac{\frac{\partial\Psi}{\partial\sigma}\frac{\partial h}{\partial\sigma}}{|\langle \vec{V}_\nu, \nabla\Psi\rangle|^2 + |\frac{\partial\Psi}{\partial\sigma}|^2}\Big). \tag{4.28}$$

Again we get the formula (4.25) with L replaced by $\tilde{L}$ and we need to examine the symbol $(\tilde{L})^N\gamma$ using the crucial lower bound (4.27). We use the terminology in the appendix (*cf.* Definition A.1). Analyzing the terms of type (A,j) we thus get a bound $O((2^l\delta_1^{-2})^j)$. For the terms $(B,1)$ (second derivative of Ψ divided by the square of a gradient) we notice that pure (σ,u) derivatives of second order are zero and pure x derivatives of second order carry the factor $\sigma = O(2^{-l})$. We need to get an upper bound for mixed derivatives of second order and notice that

$$\frac{\partial}{\partial\sigma}\langle\vec{V}_\nu,\nabla\rangle\Psi = -\langle J^\# N_\nu, N(y)\rangle = O(\delta_1) \tag{4.29}$$

for $y\in Q_\nu$. This shows that the type $(B,1)$ terms are $\lesssim \delta_1^{-3}+2^{-l}\delta_1^{-4}$ and hence $\lesssim 2^l\delta_1^{-2}$. The type (B,j) terms for $j\ge 2$ are $O(\delta^{-2(j+1)})$. Hence if β is a product of a bounded term, a term of type (A,j_A), M_1 terms of type $(B,1)$ and M_2 terms of type (B,κ_i) with $\kappa_i\ge 2$, and if $j_A+M_1+\sum_{i=1}^{M_2}\kappa_i = N_1$, we get a bound

$$|\beta| \lesssim (2^l\delta_1^{-2})^{j_A+M_1}\prod_{i=1}^{M_2}\delta_1^{-2(\kappa_i+1)}$$

and hence

$$\begin{aligned}2^{-kN_1}|(\tilde{L})^{N_1}\gamma| &\lesssim_{N_1} (2^{k-l}\delta_1^2)^{-j_A-M_1}\prod_{i=1}^{M_2}\big(2^{-k}\delta_1^{-2\frac{\kappa_i+1}{\kappa_i}}\big)^{\kappa_i}\\ &\lesssim_{N_1} (2^{k-l}\delta_1^2)^{-N_1} \lesssim 2^{-kN_1}2^{l(3-2\varepsilon)N_1} \lesssim 2^{-2kN_1\varepsilon/3}\end{aligned}$$

since we are assuming $\delta_1 > 2^{-l(1-\varepsilon)}$ and $l\le k/3$. We choose N_1 large, say $N_1 > (2N+10n)/\varepsilon$, and from (4.15) and the above error analysis we get the bound

$$\begin{aligned}&\Big\|\sum_\nu P_{k,\nu}Tf_\nu\Big\|_q\\ &\qquad\le C(\varepsilon)(\delta_1/\delta_0)^{-\varepsilon}\Big(\sum_\nu\|Tf_\nu\|_q^2\Big)^{1/2} + C_3(\varepsilon,N)2^{-kN}\sup_\nu\|f_\nu\|_q.\end{aligned} \tag{4.30}$$

Now apply Hölder's inequality in the ν sum (which has $\lesssim (\delta_1/\delta_0)^{-(2n-1)}$ terms) to also get (4.6). This finishes the proof of Proposition 4.1 under the additional assumption (4.7).

4.2. *Changes of variables.* We now complete the proof of Proposition 4.1 by reducing the general case to the model case (4.7).

Let again δ_1,δ_0 be as in (4.1). We are now given a family of disjoint cubes $\{Q_\nu\}$ in $\mathbb{R}^{2n-1}$ of sidelength δ_1 contained in a reference cube

$$Q := \{\underline{y}\in\mathbb{R}^{2n-1} : |y_i-b_i|\le\delta_0, i=1,\dots 2n-1\}$$

and suppose that the function f_ν is supported in $Q_\nu\times\mathbb{R}\times\mathbb{R}$. We consider the operator T with Schwartz kernel as in (4.4b) but do not assume that ∇g vanishes at the reference point $(\underline{a},\underline{b})$. Decomposing the cutoff function in (x,u) into a finite number of pieces (with the number depending on upper bounds for g' we may assume that

$$\text{supp}\,(\chi_{l,a,u^\circ}) \subset \{|x-a|\le c_0 2^{-l},\ |u-u^\circ|\le c_0 2^{-l}\} \tag{4.31}$$

for some small $c_0 > 0$. We also set $a = (\underline{a}, a_{2n})$, and $b = (\underline{b}, 0)$.

Define $G \equiv G_{a,\underline{b}}$ by

$$G(\underline{w}) = g(\underline{a} - \underline{b} + \underline{w}) - a_{2n} - \underline{w}^\intercal \nabla_x g(\underline{a} - \underline{b})$$

so that

$$G(0) = -a_{2n} + g(\underline{a} - \underline{b}), \quad G'(0) = 0\,.$$

We now introduce the change of variables $(X, U) = (\underline{X}, X_{2n}, U)$, $(Y, V) = (\underline{Y}, Y_{2n}, V)$, defined by

$$\begin{aligned}
\underline{X} &= \underline{x} - \underline{a}, \\
X_{2n} &= x_{2n} - a_{2n} - (\underline{x} - \underline{a})^\intercal \nabla g(\underline{a} - \underline{b}) \\
U &= u + \lambda(x) + (x - a)^\intercal J b
\end{aligned}$$

and

$$\begin{aligned}
\underline{Y} &= \underline{y} - \underline{b}, \\
Y_{2n} &= y_{2n} - (\underline{y} - \underline{b})^\intercal \nabla g(\underline{a} - \underline{b}) \\
V &= v + \lambda(y) - a^\intercal J y
\end{aligned}$$

Then we have

$$x_{2n} - y_{2n} - g(\underline{x} - \underline{y}) = X_{2n} - Y_{2n} - G(\underline{X} - \underline{Y}). \tag{4.32a}$$

Setting $B = \begin{pmatrix} I_{2n-1} & 0 \\ g'(\underline{a} - \underline{b}) & 1 \end{pmatrix}$ we also have

$$\begin{aligned}
u - v + x^\intercal J y + \lambda(x - y) &= U - V + (x - a)^\intercal J (y - b) \\
&= U - V + X^\intercal B^\intercal J B Y. \qquad (4.32\text{b})
\end{aligned}$$

Here $B^\intercal J B$ belongs to a compact family of invertible skew symmetric $2n \times 2n$ matrices. Now, after a decomposition into a finite number of pieces with (x, u)-support of diameter $< 2^{-l}$ we can reduce to the situation where we can apply the estimates in §4.1.

5. Proof of Proposition 2.1

We now iterate the estimates in Proposition 4.1. We give the argument for $f \mapsto f * \mu_{k,l}$, $l \le [k/3]$ (see the definitions ahead of the statement of Proposition 2.3). We can use the Heisenberg translations to reduce to the case that f is supported in $\{(y, v) : |y| \le 1, |v| \le 1\}$. Then $f * \mu_{k,l}$ is supported in $\{(x, u) : |x| + |u| \le C\}$ for some fixed constant.

We shall work with a partition of unity in (x, u) space

$$\sum_{(x^\circ, u^\circ) \in \mathcal{Z}_l} \chi_{x^\circ, u^\circ} = 1$$

where $\mathcal{Z}_l$ is a grid of $c2^{-l}$ separated points and the bump functions χ_{x°, u° are associated in a natural way with cubes of diameter $O(2^{-l})$ centered at the points in the grid.

Then for f with support in a fixed ball near the origin we get

$$\|f * \mu_{k,l}\|_q \lesssim \Big(\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l} \big\|\chi_{x^\circ,u^\circ}(f*\mu_{k,l})\big\|_q^q\Big)^{1/q}$$

We define numbers δ_j of the form 2^{-m_j} with $m_j \in \mathbb{N}$ as follows. Let $m_0 = [l\varepsilon/100n]$ and $\delta_0 = 2^{-m_0}$. Define for $j \geq 1$

$$m_j = \Big\lfloor \frac{l+m_{j-1}}{2}\Big\rfloor,$$

note that $m_j \to l$. We will stop the process when $m_j > l(1-\frac{\varepsilon}{10n})$. Let j_* be the smallest integer greater than $l(1-\frac{\varepsilon}{10n})$.

Decompose $\mathbb{R}^{2n-1}$ into disjoint dyadic cubes of side length 2^{-m_0}, call these cubes Q^0_ν. Let $f_{0,\nu}(y,v) = f(y,v)\mathbb{1}_{Q^0_\nu}(y,v)$. Then by Minkowski's and Hölder's inequality

$$\begin{aligned}\|f*\mu_{k,l}\|_q &\lesssim \sum_{\nu_0}\Big(\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l}\big\|\chi_{x^\circ,u^\circ}(f_{0,\nu_0}*\mu_{k,l})\big\|_q^q\Big)^{1/q} \\ &\lesssim 2^{m_0(2n+1)/q'}\Big(\sum_{\nu_0}\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l}\big\|\chi_{x^\circ,u^\circ}(f_{0,\nu_0}*\mu_{k,l})\big\|_q^q\Big)^{1/q}.\end{aligned} \tag{5.1}$$

Fix $j \leq j_*$ and let $\{Q^j_\nu\}$ be the collection of dyadic cubes of sidelength 2^{-m_j}. Set $f_{j,\nu} = f\mathbb{1}_{Q^j_\nu}$. We claim, for some $C(\varepsilon_1)$ and any $N_2 \in \mathbb{N}$ the bound

$$\begin{aligned}\|f*\mu_{k,l}\|_q \leq{}& C_0C_1(\varepsilon_1)^j 2^{m_0\frac{2n+1}{q'}}2^{(m_j-m_0)((2n-1)(\frac12-\frac1q)+\varepsilon_1)} \\ &\times\Big(\sum_{\nu_j}\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l}\big\|\chi_{x^\circ,u^\circ}(f_{j,\nu_j}*\mu_{k,l})\big\|_q^q\Big)^{1/q} \\ &+ jC_2(\varepsilon,N_1)C_1(\varepsilon_1)^{j-1}2^{(2n+1)l}2^{-kN_1}\|f\|_q.\end{aligned} \tag{5.2}$$

For $j=0$ this holds by (5.1). Suppose (5.2) holds for $j = J < j_*$. We apply Proposition 4.1 to bound

$$\begin{aligned}\|f*\mu_{k,l}\|_q \leq{}& C(\varepsilon_1)^{J+1}2^{m_0(2n+1)/q'}2^{(m_{J+1}-m_J+m_J-m_0)((2n-1)(\frac12-\frac1q)+\varepsilon_1)} \\ &\times\Big(\sum_{\nu_J}\sum_{\nu_{J+1}:Q^{J+1}_{\nu_{J+1}}\subset Q^J_{\nu_J}}\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l}\big\|\chi_{x^\circ,u^\circ}(f_{J+1,\nu_{J+1}}*\mu_{k,l})\big\|_q^q\Big)^{1/q} \\ &+(J+1)C(\varepsilon,N_1)C(\varepsilon_1)^J2^{(2n+1)l}2^{-kN_1}\|f\|_q.\end{aligned}$$

Choosing $N_1 > N_2 + 2n + 1$ gives

$$\begin{aligned}\|f*\mu_{k,l}\|_q \leq{}& C(\varepsilon_1)^{J+1}2^{m_0\frac{2n+1}{q'}}2^{(m_{J+1}-m_0)((2n-1)(\frac12-\frac1q)+\varepsilon_1)} \\ &\times\Big(\sum_{\nu}\sum_{(x^\circ,u^\circ)\in\mathcal{Z}_l}\big\|\chi_{x^\circ,u^\circ}(f_{J+1,\nu}*\mu_{k,l})\big\|_q^q\Big)^{1/q} \\ &+(J+1)C(\varepsilon_1)^J\tilde{C}(\varepsilon,N_2)2^{-kN_2}\|f\|_q.\end{aligned}$$

We apply this for $J = j_* - 1$ and observe that $j_* \lesssim \varepsilon^{-1} + \log l$.

$$\begin{aligned}\|f * \mu_{k,l}\|_q \lesssim C(\varepsilon_1)^{C_2(\log l + \varepsilon^{-1})} 2^{m_0 \frac{2n+1}{q'}} 2^{(m_{j_*} - m_0)((2n-1)(\frac12 - \frac1q) + \varepsilon_1)} \\ \times \Big(\sum_\nu \sum_{(x^\circ, u^\circ) \in \mathcal{Z}_l} \big\| \chi_{x^\circ, u^\circ} (f_{j_*,\nu} * \mu_{k,l}) \big\|_q^q \Big)^{1/q} \\ + C(\varepsilon_1)^{C_2(\log l + \varepsilon^{-1})} C(\varepsilon, N_2)(\varepsilon^{-1} + \log l) 2^{-kN_2} \|f\|_q\end{aligned}$$

We use the definition of m_0 and m_{j_*}, and that $A^{\log j} \le C(\delta, A) 2^{\delta j}$ for any $\delta > 0$. Thus, for $2 \le q \le \frac{4n+2}{2n-1}$,

$$\|f * \mu_{k,l}\|_q \lesssim_{\varepsilon, \varepsilon_1} 2^{l\varepsilon/10} 2^{l(1-\varepsilon)(2n-1)(\frac12 - \frac1q + \varepsilon_1)} \Big(\sum_\nu \|f_{j_*,\nu} * \mu_{k,l}\|_q^q \Big)^{1/q} + C_{N,\varepsilon} 2^{-kN} \|f\|_q \tag{5.3}$$

We use the L^2 estimates for Fourier integral operators associated with folding canonical relations, as in [9] (relying on the version in [5]). We get for a bounded set $\mathcal{U}$

$$\Big(\sum_\nu \|f_{j_*,\nu} * \mu_{k,l}\|^2_{L^2(\mathcal{U})} \Big)^{1/2} \lesssim 2^{-k(2n-1)/2} 2^{l/2} \|f\|_2 \tag{5.4}$$

for $l \le [k/3]$. We also have a trivial L^∞ bound, using that the projection of the support of $f_{j_*,\nu}$ to $\mathbb{R}^{2n-1}$ is contained in a ball of radius $c2^{-m_{j_*}} \approx 2^{-l(1-\frac{\varepsilon}{10n})}$. We get

$$\sup_\nu \|f_{j_*,\nu} * \mu_{k,l}\|_\infty \lesssim 2^{-l(1-\frac{\varepsilon}{10n})(2n-1)} \|f\|_\infty. \tag{5.5}$$

By interpolation

$$\Big(\sum_\nu \|f_\nu * \mu_{k,l}\|^q_{L^q(\mathcal{U})} \Big)^{1/q} \lesssim_\varepsilon 2^{-k(2n-1)/q} 2^{-l(2n-1) + l(4n-1)/q} 2^{l\varepsilon/2} \|f\|_q \tag{5.6}$$

We combine (5.3) and (5.6) and obtain

$$\|f * \mu_{k,l}\|_{L^q(\mathcal{U})} \lesssim_\varepsilon 2^{l\varepsilon} 2^{l(\frac{2n}{q} - \frac{2n-1}{2})} 2^{-k\frac{2n-1}{q}} \|f\|_q \tag{5.7}$$

for $l \le [k/3]$ and (choosing ε small) we can sum in l if $q > \frac{4n}{2n-1}$.

Equivalently we obtain

$$\|f * \mu_k\|_q \le C(q) 2^{-k(2n-1)/q} \|f\|_q \text{ for } \frac{4n}{2n-1} < q \le \frac{4n+2}{2n-1}, \tag{5.8}$$

provided that f is supported in a fixed ball centered at the origin, say in $Q = [0,1)^{2n+1}$.

We now remove this assumption on the support of f in (5.9). For $m \in \mathbb{Z}^{2n+1}$ let

$$Q_m = m \cdot Q = \{(\overline{m} + y, m_{2n+1} + v + \overline{m}^\intercal J y : y \in Q\}.$$

Let μ_k be supported in $R_A = \{(x,u) : |x| \le A, |u| \le A\}$ for some $A \ge 1$. Then $(f\mathbb{1}_Q) * \mu_k$ is supported on the set

$$\{(x,u) : |x| \le A + \sqrt{2n},\ |u| \le A + 1 + (A + \sqrt{2n}\|J\|\sqrt{2n}\};$$

to see this write $x^\intercal Jy = (x-y)^\intercal Jy$. Thus $(f\mathbb{1}_Q) * \mu_k$ is supported in R_B with $B = B(A)$ (and $B(A)$ is the maximum of the bounds for $|x|$ and $|u|$ in the displayed formula). By left translation $(f\mathbb{1}_{Q_m}) * \mu_k$ is supported in $m \cdot R_B$ and we have

$$\|(f\mathbb{1}_{Q_m}) * \mu_k\|_q \le C(q)2^{-k(2n-1)/q}\|f\mathbb{1}_{Q_m}\|_q \tag{5.9}$$

for $\frac{4n}{2n-1} < q \le \frac{4n+2}{2n-1}$, uniformly in m. Now for each $m \in \mathbb{Z}^{2n+1}$ the cardinality of

$$\{\widetilde{m} \in \mathbb{Z}^{2n+1} : m \cdot R_B \cap \widetilde{m} \cdot R_B \ne \emptyset\}$$

is bounded above by $C(B)^{2n+1}$. Indeed, let $m \cdot R_B \cap \widetilde{m} \cdot R_B \ne \emptyset$ which is equivalent with $R_B \cap m^{-1}\widetilde{m}R_B \ne \emptyset$. Let $(w,t) = m \cdot \widetilde{m}^{-1}$ and $(x,u) \in R_B$. Then $(w,t)\cdot(x,u) = (w+x, t+u+w^\intercal Jx)$. if $|w| \ge 2B$ then $(w,t)\cdot(x,u) \notin R_B$. If $|w| \le 2B$ and $t > 2B + 2B^2\|J\|$ then $|t + u + w^\intercal Jx| > B$ and again $(w,t)\cdot(x,u) \notin R_B$. Apply this with

$$(w,t) = m^{-1} \cdot \tilde{m} = (\overline{\widetilde{m}} - \overline{m}, \tilde{m}_{2n+1} - m_{2n+1} - \overline{m}^\intercal J(\overline{\widetilde{m}} - \overline{m})),$$

and clearly for fixed m there are only $C(B)^{2n+1}$ integer vectors $\tilde{m}$ with $|m^{-1} \cdot \tilde{m}| \le 2B + 2B^2\|J\|$.

Hence for general $f \in L^q(\mathbb{H}^n)$, $\frac{4n}{2n-1} < q \le \frac{4n+2}{2n-1}$,

$$\|f * \mu_k\|_q = \Big\| \sum_{m \in \mathbb{Z}^{2n+1}} (f\mathbb{1}_{Q_m}) * \mu_k \Big\|_q \lesssim_B \Big(\sum_{m \in \mathbb{Z}^{2n+1}} \|(f\mathbb{1}_{Q_m}) * \mu_k\|_q^q \Big)^{1/q}$$

$$\lesssim_{B,q} 2^{-k(2n-1)/q} \Big(\sum_{m \in \mathbb{Z}^{2n+1}} \|f\mathbb{1}_{Q_m}\|_q^q \Big)^{1/q} \lesssim_{B,q} 2^{-k(2n-1)/q}\|f\|_q.$$

Now convolution with μ_k is uniformly bounded on L^∞ and thus interpolation gives

$$\|f * \mu_k\|_q \le C(q)2^{-k(2n-1)/q}\|f\|_q \text{ for } \frac{4n}{2n-1} < q \le \infty. \tag{5.10}$$

Alternatively one can argue with as in [13] with the Wolff version of decoupling for $q > \frac{4n+2}{2n-1}$. By duality (5.10) also implies

$$\|f * \mu_k\|_p \le C(p)2^{-k(2n-1)(1-\frac{1}{p})}\|f\|_p \text{ for } 1 \le p < \frac{4n}{2n+1}. \tag{5.11}$$

By modifying the definition of γ in (4.4b) we also get the same estimate for μ_k replaced with $2^{-k}\frac{d}{ds}\mathrm{Dil}_s\mu_k$ when $s \approx 1$. This proves Proposition 2.1. A standard Sobolev imbedding argument yields Corollary 2.2.

Remark. Up to this point we worked with a measure μ in $\mathbb{R}^{2n}$ supported on a graph $x_{2n} = g(x_1, \dots, x_{2n-1})$, with D^2g nondegenerate. We must also also consider the case where the surface is given by $x_j = g(x_1, \dots, x_{j-1}, x_{j+1}, \dots)$.

However this situation can be reduced to the former by permuting the variables; one just needs to note that the change of variables argument in §4.2 applies, and that the skew symmetric matrix J in the former case is replaced by $P^\intercal JP$ after a change of variables, with P a suitable permutation matrix.

L^p Sobolev result. Proposition 2.1 can be reformulated as a regularity result in Besov spaces for functions in $\mathbb{R}^{2n+1}$ which are supported on a compact set. However one can combine Proposition 2.1, part (i) with the result in Theorem 1.1 of [12] to show a better result using Sobolev spaces $L^q_\alpha(\mathbb{R}^d)$ for $\alpha = (d-2)/q$, with $d = 2n+1$.

Corollary 5.1. *Let $\mathcal{U}$ be a compact neighborhood of the origin and let $L^q_\alpha(\mathbb{R}^{2n+1})$ be the usual Sobolev space. Let $\mathcal{R}f = f * \mu$ with μ as above, and with the convolution on the Heisenberg group $\mathbb{H}^n$. Then we have for $\frac{4n}{2n-1} < q < \infty$,*

$$\|\mathcal{R}f\|_{L^q_{(2n-1)/q}} \le C(\mathcal{U}, q)\|f\|_q$$

whenever f is supported in $\mathcal{U}$.

See the discussion in §2 of [12] for related examples. Theorem 1.1 of [12] actually gives a better statement using Besov and Triebel-Lizorkin spaces, namely that $\mathcal{R} : (B^0_{q,q})_{\mathrm{comp}} \to F^{(2n-1)/q}_{q,r}$ for all $r > 0$, in the q-range of the corollary.

6. Estimates for the global maximal operator

By the Marcinkiewicz interpolation theorem it suffices to prove a weak type (p,p) estimate for $p > \frac{2n}{2n-1}$, i.e.

$$\operatorname{meas}\big(\{(x,u) : \sup_t |f * \mathrm{Dil}_t\mu(x,u)| > \alpha\}\big) \lesssim \alpha^{-p}\|f\|^p_{L^p(\mathbb{H}^n)} \tag{6.1}$$

for $p > \frac{2n}{2n-1}$.

We use Calderón-Zygmund theory on the Heisenberg group, with respect to the nonisotropic balls

$$\begin{aligned} B((a,b),\delta) &= \{(y,v) : |(y,v)^{-1} \cdot (a,b)| \le \delta\} \\ &= \{(y,v) : |(a-y, b-v-y^\intercal Ja)| \le \delta\} \end{aligned}$$

see for example [3], [4], or [17]. We apply the Calderón-Zygmund decomposition to the function $|f|^p$.

Let

$$\Omega_\alpha \equiv \Omega_\alpha(f) = \{(x,u) : M_{HL}(|f|^p)(x,u) > \alpha^p\}$$

so that

$$\operatorname{meas}(\Omega_\alpha) \lesssim \alpha^{-p}\|f\|^p_p. \tag{6.2}$$

Let $g_1 = f\mathbb{1}_{\Omega_\alpha^\complement}$. Then $|g_1(x)| \le \alpha$ almost everywhere

We have $\Omega_\alpha = \cup_{Q\in\mathcal{Q}_\alpha} Q$ where the sets Q in the family $\mathcal{Q}_\alpha$ are disjoint and measurable and there are constants c_1, c_2 such that $2 \le c_1 < c_2/8$ such that for every Q there is a point P_Q and an $r_Q > 0$ with

$$B(P_Q, r_Q) \subset Q \subset B(P_Q, c_1 r_Q) \subset B(P_Q, 2c_1 r_Q) \subset \Omega_\alpha;$$

moreover

$$\sum_Q \mathbb{1}_{B(P_Q,2c_1 r_Q)}(x) \lesssim C \text{ almost everywhere}$$

and

$$B(P_Q, c_2 r_Q) \cap \Omega_\alpha^{\complement} \neq \emptyset\,.$$

We decompose

$$f\mathbb{1}_{\Omega_\alpha} = \sum_{Q\in\mathcal{Q}_\alpha} f\mathbb{1}_Q.$$

Let ϕ be a C^∞ function supported in the Euclidean ball of radius 1 centered at the origin and such that $\int \phi = 1$. We introduce some cancellation using suitable convolutions as in [2]. Let

$$\phi_m(x,u) = 2^{-m(2n+2)}\phi(2^{-m}x, 2^{-2m}u)$$

let m_Q be such that $c_1 r_Q/2 \le 2^{m_Q} < c_1 r_Q$ and set

$$\begin{aligned} g_Q &= (f\mathbb{1}_Q) * \phi_{m_Q}, \\ b_Q &= f\mathbb{1}_Q - (f\mathbb{1}_Q) * \phi_{m_Q}. \end{aligned}$$

Then g_Q, b_Q are supported in $B(P_Q, 2c_1 r_Q)$. We set $g_2 = \sum_{Q\in\mathcal{Q}_\alpha} g_Q$ and we have $|g_2(x)| \le C\alpha$ almost everywhere. Let $g = g_1 + g_2$ and $b = f - g$. Now

$$\begin{aligned} &\operatorname{meas}\big(\{(x,u) : \sup_t |f * \mathrm{Dil}_t\mu(x,u)| > \alpha\}\big) \\ &\le \operatorname{meas}(\Omega_\alpha) + \operatorname{meas}\big(\{(x,u) : \sup_t |g * \mathrm{Dil}_t\mu(x,u)| > \alpha/2\}\big) \\ &\quad + \operatorname{meas}\big(\{(x,u) \in \Omega_\alpha^{\complement} : \sup_t |b * \mathrm{Dil}_t\mu(x,u)| > \alpha/2\}\big). \end{aligned}$$

Since $n \ge 2$ we know the L^2 boundedness of our maximal operator and the standard argument using $\|g\|_\infty \lesssim \alpha$ gives

$$\begin{aligned} &\operatorname{meas}\big(\{(x,u) : \sup_t |g * \mathrm{Dil}_t\mu(x,u)| > \alpha/2\}\big) \\ &\le (\alpha/2)^{-2}\|\sup_t |g * \mathrm{Dil}_t\mu|\|_2^2 \qquad (6.3) \\ &\le C^2\alpha^{-2}\|g\|_2^2 \le \widetilde{C}^p\alpha^{-p}\|f\|_p^p \end{aligned}$$

It suffices to estimate

$$\operatorname{meas}\big(\{((x,u) : \sup_t |b * \mathrm{Dil}_t\mu(x,u)| > \alpha/2\}\big).$$

For $m \in \mathbb{Z}$ let

$$b^{[m]} = \sum_{\substack{Q\in\mathcal{Q}_\alpha:\\ m_Q = m}} b_Q\,.$$

Observe that

$$\operatorname{supp}\,(b^{[m]}*\mu_t)\subset\Omega_\alpha \text{ if } t\le 2^{m-C_1}$$

where $C_1\in\mathbb{N}$ and C_1 depends only on the support of μ.

If $(x,u)\notin\Omega_\alpha$ we have

$$\begin{aligned}
\sup_{t>0}|b*\mathrm{Dil}_t\mu(x,u)| &\le \Big(\sum_{j\in\mathbb{Z}}\sup_{1/2<s<1}|b*\mathrm{Dil}_{2^js}\mu(x,u)|^p\Big)^{1/p}\\
&=\Big(\sum_{j\in\mathbb{Z}}\sup_{1/2<s<1}\Big|\sum_{m\le j+C_1}b^{[m]}*\mathrm{Dil}_{2^js}\mu(x,u)\Big|^p\Big)^{1/p}\\
&\le\sum_{k\ge0}\Big(\sum_{j\in\mathbb{Z}}\sup_{1/2<s<1}\Big|\sum_{m\le j+C_1}b^{[m]}*\mathrm{Dil}_{2^js}\mu_k(x,u)\Big|^p\Big)^{1/p}. \qquad (6.4)
\end{aligned}$$

We have straightforward estimates

$$\begin{aligned}
&\|\mu_k\|_{L^1}+2^{-k}\|\nabla\mu_k\|_{L^1}\le C,\\
&2^{-k}\Big\|\frac{d}{ds}\mathrm{Dil}_s\mu_k\Big\|_{L^1}+2^{-2k}\Big\|\nabla\frac{d}{ds}\mathrm{Dil}_s\mu_k\Big\|_{L^1}\le C,\quad \frac12\le s\le 1.
\end{aligned}$$

which we use for $m\le j-C_2k$ with sufficiently large C_2, say $C_2=10$. For this range we estimate the corresponding part in (6.4)

$$\begin{aligned}
&\Big\|\sum_{k\ge0}\Big(\sum_{j\in\mathbb{Z}}\sup_{1/2<s<1}\Big|\sum_{m\le j-C_2k}b^{[m]}*\mathrm{Dil}_{2^js}\mu_k\Big|^p\Big)^{1/p}\Big\|_p\\
&\le\sum_{k\ge0}\sum_{l\ge0}\Big\|\Big(\sum_{j\in\mathbb{Z}}\sup_{1/2<s<1}\big|b^{[j-C_2k-l]}*\mathrm{Dil}_{2^js}\mu_k\big|^p\Big)^{1/p}\Big\|_p\\
&=\sum_{k\ge0}\sum_{l\ge0}\Big(\sum_{j\in\mathbb{Z}}\Big\|\sup_{1/2<s<1}\big|b^{[j-C_2k-l]}*\mathrm{Dil}_{2^js}\mu_k\big|\Big\|_p^p\Big)^{1/p}
\end{aligned}$$

Now let $f^{[m]}=\sum_{Q:m_Q=m}f\mathbb{1}_Q$ so that $b^{[m]}=f^{[m]}*(\delta-\phi_m)$ (with δ the Dirac measure at the origin). Then

$$\begin{aligned}
&\Big\|\sup_{1/2<s<1}\big|b^{[j-C_2k-l]}*\mathrm{Dil}_{2^js}\mu_k\big|\Big\|_p\\
&\le\big\|b^{[j-C_2k-l]}*\mathrm{Dil}_{2^{j-1}}\mu_k\big\|_p+\int_{1/2}^1\Big\|b^{[j-C_2k-l]}*\frac{d}{ds}\mathrm{Dil}_{2^js}\mu_k\Big\|_p ds\\
&\le\|f^{[j-C_2k-l]}\|_p\Big[\|(\delta-\phi_{j-C_2k-l})*\mathrm{Dil}_{2^{j-1}}\mu_k\|_{L^1}\\
&\qquad\qquad+\int_{1/2}^1\|(\delta-\phi_{j-C_2k-l})*\frac{d}{ds}\mathrm{Dil}_{2^js}\mu_k\|_{L^1}ds\Big]\\
&\lesssim 2^{2k}2^{-C_2k-l}\|f^{[j-C_2k-l]}\|_p
\end{aligned}$$

and thus

$$\begin{aligned}
&\Big\| \sum_{k\ge 0} \Big(\sum_{j\in\mathbb{Z}} \sup_{1/2<s<1} \big| \sum_{m\le j-C_2k} b^{[m]} * \mathrm{Dil}_{2^js}\mu_k \big|^p \Big)^{1/p} \Big\|_p \\
&\lesssim \sum_{k\ge 0}\sum_{l\ge 0} 2^{2k} 2^{-C_2k-l} \Big(\sum_{j\in\mathbb{Z}} \| f^{[j-C_2k-l]}\|_p^p \Big)^{1/p} \\
&\lesssim \|f\|_p \sum_{k\ge 0} 2^{(2-C_2)k} \sum_{l\ge 0} 2^{-l} \lesssim \|f\|_p
\end{aligned}$$

Note that this estimate holds for all $p \ge 1$.

The main contributions come from the range $j - C_2k \le m \le j + C_1$. Here we use Corollary 2.2 and bound, for fixed k,

$$\begin{aligned}
&\Big\| \Big(\sum_{j\in\mathbb{Z}} \sup_{1/2<s<1} \big| \sum_{j-C_2k<m\le j+C_1} b^{[m]} * \mathrm{Dil}_{2^js}\mu_k \big|^p \Big)^{1/p} \Big\|_p \\
&\le \Big(\sum_{j\in\mathbb{Z}} \sum_{j-C_2k<m\le j+C_1} (C_1 + C_2k)^{p/p'} \| b^{[m]} * \mathrm{Dil}_{2^js}\mu_k\|_p^p \Big)^{1/p} \\
&\lesssim 2^{k(\frac{2n}{p}-2n+1)} \Big(\sum_{j\in\mathbb{Z}} \sum_{j-C_2k<m\le j+C_1} (C_1 + C_2k)^{p/p'} \| b^{[m]}\|_p^p \Big)^{1/p} \\
&\lesssim_{C_1,C_2} (1+k) 2^{k(\frac{2n}{p}-2n+1)} \Big(\sum_{m\in\mathbb{Z}} \| b^{[m]}\|_p^p \Big)^{1/p} \\
&\lesssim (1+k) 2^{k(\frac{2n}{p}-2n+1)} \|f\|_p .
\end{aligned}$$

Thus for $p > \frac{2n}{2n-1}$,

$$\Big\| \sum_{k\ge 0} \Big(\sum_{j\in\mathbb{Z}} \sup_{1/2<s<1} \big| \sum_{\substack{j-C_2k<\\ m\le j+C_1}} b^{[m]} * \mathrm{Dil}_{2^js}\mu_k \big|^p \Big)^{1/p} \Big\|_p \le C_p \|f\|_p \,.$$

We combine the L^p estimates and get $p > \frac{2n}{2n-1}$

$$\big\| \sup_{t>0} |b * \mathrm{Dil}_t \mu| \big\|_{L^p(\Omega_\alpha^{\complement})} \lesssim_p \|f\|_p .$$

Hence by Tshebyshev's inequality,

$$\mathrm{meas}\Big(\big\{ (x,u) \in \Omega_\alpha^{\complement} : \sup_t |b * \mathrm{Dil}_t \mu(x,u)| > \alpha/2 \big\} \Big) \lesssim \alpha^{-p} \|f\|_p^p. \tag{6.5}$$

The desired weak type bound (6.1) follows from (6.2), (6.3) and (6.5). □

Appendix A. *An integration by parts lemma*

To perform the decoupling step we used a familar integration by parts lemma (which has been used many times but is often not found in the precise form needed in an application). We formulate what we need here and include some details of the proof, for convenience.

Let $h \in C_c^\infty$ function on $\mathbb{R}^n$ and let $w \mapsto \psi(w)$ be a real valued C^∞ function such that $\nabla\psi \neq 0$ on the support of h.

We define

$$Lh = \operatorname{div}\big(\frac{h\nabla\psi}{|\nabla\psi|^2}\big).$$

Then the formal adjoint $L^* = -\langle \frac{\nabla\psi}{|\nabla\psi|^2}, \nabla\rangle$ satisfies $i\lambda^{-1}L^* e^{i\lambda\psi} = e^{i\lambda\psi}$.

We let $L^0 h = h$ and define inductively $L^N h = LL^{N-1}h$. We then have by integration by parts

$$\int e^{i\lambda\psi(w)} h(w)dw = \Big(\frac{i}{\lambda}\Big)^N \int e^{i\lambda\psi(w)} L^N h(w)dw$$

for $N = 1, 2, \dots$.

We need to analyze the behavior of $L^N h$ and the following terminology will be helpful.

Definition A.1. (i) The term h is of type $(A, 0)$. A term is of type (A, j) if it is $h_j/|\nabla\psi|^j$ where h_j is a derivative of order j of h.

(ii) A term is of type $(B, 0)$ if it is equal to 1. A term is of type (B, j) for some $j \geq 1$ if it is of the form $\psi_{j+1}/|\nabla\psi|^{j+1}$ where ψ_{j+1} is a derivative of order $j + 1$ of ψ.

Lemma A.2. *Let $N = 0, 1, 2, \dots$. Then*

$$L_N h = \sum_{\nu=1}^{K(N,n)} c_{N,\nu} h_{N,\nu}$$

where each $h_{N,\nu}$ is of the form

$$P(\tfrac{\nabla\psi}{|\nabla\psi|})\beta_A \prod_{l=1}^{M} \gamma_l$$

where P is a polynomial of n variables (independent of h and ψ), β_A is of type (A, j_A) for some $j_A \in \{0, \dots, N\}$ and the terms γ_l are of type (B, κ_l), so that $j_A + \sum_{l=1}^{M} \kappa_l = N$. The terms P, β_A, γ_l depend on ν.

Sketch of proof. The statement holds when $N = 0$. We compute

$$La = \frac{1}{|\nabla\psi|^2} \sum_{j=1}^{n} \frac{\partial\psi}{\partial w_j}\frac{\partial a}{\partial w_j} + \frac{a}{|\nabla\psi|^2} \sum_{j=1}^{n} \frac{\partial^2\psi}{(\partial w_j)^2}$$
$$- \frac{2a}{|\nabla\psi|^4} \sum_{j=1}^{n} \frac{\partial\psi}{\partial w_j} \sum_{k=1}^{n} \frac{\partial\psi}{\partial w_k} \frac{\partial^2\psi}{\partial w_j \partial w_k}.$$

In particular one can check immediately that the assertion holds for $N = 1$.

Now let a be of type (A,m), and γ be of type (B,m). One observes that

$$\frac{1}{|\nabla\psi|^2}\frac{\partial\psi}{\partial w_j}\frac{\partial a}{\partial w_j} = P(\tfrac{\nabla\psi}{|\nabla\psi|})a_{m+1} + a\sum_{i=1}^n P_{A,i}(\tfrac{\nabla\psi}{|\nabla\psi|})b_{1,i}$$

$$\frac{1}{|\nabla\psi|^2}\frac{\partial\psi}{\partial w_j}\frac{\partial \gamma}{\partial w_j} = P(\tfrac{\nabla\psi}{|\nabla\psi|})b_{m+1} + \gamma\sum_{i=1}^n P_{B,i}(\tfrac{\nabla\psi}{|\nabla\psi|})b_{1,i}$$

where a_{m+1} is of type $(A, m+1)$, b_{m+1} is of type $(B, m+1)$, and the $b_{1,i}$ are of type $(B,1)$. The polynomials are given by $P(x) = x_j$ and $P_{A,i}(x) = -mx_ix_j$, $P_{B,i}(x) = -(m+1)x_ix_j$.

Next, if P is a polynomial of n variables then

$$\frac{1}{|\nabla\psi|^2}\frac{\partial\psi}{\partial w_j}\frac{\partial}{\partial w_j}(P(\tfrac{\nabla\psi}{|\nabla\psi|})) = \sum_{i=1}^n \tilde{P}_i(\tfrac{\nabla\psi}{|\nabla\psi|})b_i$$

where the b_i are of type $(B,1)$ and $\tilde{P}_i(x) = x_j(\frac{\partial P}{\partial x_i} - x_i\sum_{k=1}^n x_k\frac{\partial P}{\partial x_k})$.

Concerning the other terms in the definition of Lb (with b equal to a or γ) we note that $b\frac{\partial^2\psi}{(\partial w_j)^2}|\nabla\psi|^{-2}$ is a product of b and a type $(B,1)$ term and

$$\frac{b}{|\nabla\psi|^4}\frac{\partial\psi}{\partial w_j}\frac{\partial\psi}{\partial w_k}\frac{\partial^2\psi}{\partial w_j\partial w_k}$$

is a product of $P_2(\frac{\nabla\psi}{|\nabla\psi|})b$ and a type $(B,1)$ term, with $P_2(x) = x_jx_k$.

One combines these observations and uses them together with the Leibniz rule to carry out the induction step. □

References

[1] J. Bourgain, C. Demeter, *The proof of the ℓ^2 decoupling conjecture,* Ann. of Math. (2) 182 (2015), no. 1, 351–389.

[2] A. Carbery, *Variants of the Calderón-Zygmund theory for L^p-spaces.* Rev. Mat. Iberoamericana 2 (1986), no. 4, 381–396.

[3] R.R. Coifman, G. Weiss, Analyse harmonique non-commutative sur certains espaces homogènes. (French) Étude de certaines intégrales singulières. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971. v+160 pp.

[4] G. Folland, E.M. Stein, Hardy spaces on homogeneous groups. Princeton Univ. Press, Princeton University 1982.

[5] A. Greenleaf, A. Seeger, *Fourier integral operators with fold singularities,* J. reine ang. Math., 455, (1994), 35–56.

[6] ______, *On oscillatory integral operators with folding canonical relations,* Studia Math. 132 (1999), no. 2, 125–139.

[7] L. Hörmander, *Fourier integral operators.* I. Acta Math. 127 (1971), no. 1-2, 79–183.

[8] I. Łaba, T. Wolff, *A local smoothing estimate in higher dimensions.* Dedicated to the memory of Tom Wolff. J. Anal. Math. 88 (2002), 149–171.

[9] D. Müller, A. Seeger, *Singular spherical maximal operators on a class of two step nilpotent Lie groups,* Israel J. Math. 141 (2004), 315–340.

[10] E.K. Narayanan, S. Thangavelu, *An optimal theorem for the spherical maximal operator on the Heisenberg group.* Israel J. Math. 144 (2004), 211–219.

[11] A. Nevo, S. Thangavelu, *Pointwise ergodic theorems for radial averages on the Heisenberg group,* Adv. Math. 127 (1997), 307–334.

[12] M. Pramanik, K.M. Rogers, A. Seeger, *A Calderón-Zygmund estimate with applications to generalized Radon transforms and Fourier integral operators.* Studia Math. 202 (2011), no. 1, 1-15.

[13] M. Pramanik, A. Seeger, *L^p regularity of averages over curves and bounds for associated maximal operators,* Amer. J. Math. 129 (2007), no. 1, 61–103.

[14] ———, *Optimal L^p-Sobolev regularity of a class of generalized Radon transforms.* Manuscript.

[15] D.H. Phong and E.M. Stein, *Hilbert integrals, singular integrals, and Radon transforms.* I. Acta Math. 157 (1986), no. 1-2, 99–157.

[16] C.D. Sogge and E.M. Stein, *Averages over hypersurfaces. Smoothness of generalized Radon transforms.* J. Analyse Math. 54 (1990), 165–188.

[17] E.M. Stein, Harmonic Analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton University Press, Princeton, NJ, 1993.

[18] T. Wolff, *Local smoothing type estimates on L^p for large p,* Geom. Funct. Anal. 10 (2000), no. 5, 1237–1288.

Theresa C. Anderson, Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI 53706, USA

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

Laura Cladek, Department of Mathematics, UCLA, Box 951555, Los Angeles, CA 90095-1555, USA

E-mail address: `cladekl@math.ucla.edu`

Malabika Pramanik, Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver, B.C., Canada V6T 1Z2

E-mail address: `malabika@math.ubc.ca`

Andreas Seeger, Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI 53706, USA

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