ISRAEL JOURNAL OF MATHEMATICS 141 (2004), 315–340

SINGULAR SPHERICAL MAXIMAL OPERATORS ON A CLASS OF TWO STEP NILPOTENT LIE GROUPS

BY

DETLEF MÜLLER

Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel
Ludewig-Meyn-Str. 4, 24098 Kiel, Germany
e-mail: mueller@math.uni-kiel.de

AND

ANDREAS SEEGER*

Department of Mathematics, University of Wisconsin
480 Lincoln Drive, Madison, WI 53706, USA
e-mail: seeger@math.wisc.edu

ABSTRACT

Let $H^n \cong \mathbb{R}^{2n} \ltimes \mathbb{R}$ be the Heisenberg group and let μ_t be the normalized surface measure for the sphere of radius t in $\mathbb{R}^{2n}$. Consider the maximal function defined by $Mf = \sup_{t>0} |f * \mu_t|$. We prove for $n \geq 2$ that M defines an operator bounded on $L^p(H^n)$ provided that $p > 2n/(2n-1)$. This improves an earlier result by Nevo and Thangavelu, and the range for L^p boundedness is optimal. We also extend the result to a more general class of surfaces and to groups satisfying a nondegeneracy condition; these include the groups of Heisenberg type.

1. Introduction

Let G be a finite-dimensional two step nilpotent group which we may identify with its Lie algebra $\mathfrak{g}$ by the exponential map. We assume that $\mathfrak{g}$ splits as a direct sum $\mathfrak{g} = \mathfrak{w} \oplus \mathfrak{z}$ so that

$$[\mathfrak{w}, \mathfrak{w}] \subset \mathfrak{z}, \quad [\mathfrak{w}, \mathfrak{z}] = \{0\},$$

and that $\dim(\mathfrak{w}) = d$, $\dim(\mathfrak{z}) = m$.

Throughout we shall make the following

* The second author was supported in part by the National Science Foundation.
Received March 27, 2003

NONDEGENERACY HYPOTHESIS. For every nonzero linear functional $\omega \in \mathfrak{z}^*$ the bilinear form

$$\mathcal{J}_\omega\colon \begin{array}{l} \mathfrak{w} \times \mathfrak{w} \to \mathbb{R} \\ (X, Y) \mapsto \omega([X, Y]) \end{array}$$

is nondegenerate.

Note that the skew symmetry of $\mathcal{J}_\omega$ and the nondegeneracy hypothesis imply that d is even.

There is a natural dilation structure relative to $\mathfrak{w}$ and $\mathfrak{z}$, namely for $X \in \mathfrak{w}$ and $U \in \mathfrak{z}$ we consider the dilations

$$\delta_t\colon (X, U) \mapsto (tX, t^2U).$$

With the identification of the Lie algebra with the group δ_t becomes an automorphism of the group.

In exponential coordinates (x, u), $x \in \mathbb{R}^d$, $u \in \mathbb{R}^m$, the group multiplication is given by

$$(x, u) \cdot (y, v) = (x + y, u + v + x^t J y) \tag{1.1}$$

where $x^tJy = (x^tJ_1y, \ldots, x^tJ_my) \in \mathbb{R}^m$ and the J_i are skew-symmetric matrices acting on $\mathbb{R}^d$ (i.e. $J_i^t = -J_i$). For $u \in \mathbb{R}^m$ we also form the skew-symmetric matrices $J_u = \sum_{i=1}^m u_iJ_i$ and the nondegeneracy hypothesis is equivalent with the invertibility of J_u for all $u \neq 0$.

The most prominent examples are the Heisenberg groups H^n which arise when $d = 2n$, $m = 1$ and $J = J_1$ is the standard symplectic matrix on $\mathbb{R}^{2n}$. These belong to the class of Heisenberg-type groups (termed H-type groups in [9]), for which $J_u^2 = -4|u|^2I$, so that the nondegeneracy hypothesis is clearly satisfied in this case. Note that in general m has to be small compared to d (see [9] where the connection with Radon–Hurwitz numbers is pointed out). The class considered here has been introduced by Métivier [10] in his study of analytic hypoellipticity; the nondegeneracy assumption is termed "Condition (H)" in [10]. There are many groups which satisfy the nondegeneracy condition but which are not isomorphic to a Heisenberg-type group; we give an example in §7.

Let Σ be a smooth convex hypersurface in $\mathfrak{w}$ and let μ be a compactly supported smooth density on Σ. We make the following

CURVATURE HYPOTHESIS. The Gaussian curvature of Σ does not vanish on the support of μ.

Define the dilate μ_t by

$$\langle \mu_t, f\rangle = \int f(tx,0)d\mu(x). \tag{1.2}$$

We recall the definition of convolution

$$\begin{aligned} f * g(x,u) &= \int f(y,v)g((y,v)^{-1}\cdot(x,u))dydv \\ &= \int f(y,v)g(x-y,u-v+x^tJy)dydv \end{aligned} \tag{1.3}$$

and define for Schwartz functions the maximal operator M by

$$Mf(x,u) = \sup_{t>0} |f*\mu_t(x,u)|.$$

We prove the following sharp result.

THEOREM: *Suppose $d > 2$. Then M extends to a bounded operator on $L^p(G)$ if and only if $p > d/(d-1)$.*

Remarks: (i) Other more "regular" spherical maximal functions on the Heisenberg group have been considered in [2], [15]. In these papers the maximal functions are generated by measures on hypersurfaces and the averaging operators are Fourier integral operators associated to local canonical graphs. In our work the maximal functions are generated by measures on surfaces of codimension $m+1$, and the associated canonical relations project with fold singularities.

(ii) A previous result is due to Nevo and Thangavelu [12] who considered the case of spherical means on the noncentral part of the Heisenberg groups ($m=1$) and obtained L^p boundedness in the smaller range $p > (d-1)/(d-2)$, $d > 2$.

(iii) Our theorem is an analogue of Stein's theorem [16] in the Euclidean case. The necessity of the condition $p > d/(d-1)$ follows from the example in [16]; one tests M on the function given by $f(y,v) = |y|^{1-d}(\log|y|)^{-1}\chi(y,v)$ with a suitable cutoff function χ. The L^2 methods in this paper are not sufficient to establish L^p boundedness for $p>2$ for the case $d=2$ (that is, for an extension of Bourgain's result [1] in the Euclidean case).

(iv) The result should remain true for any nilpotent Lie group of step ≤ 2; i.e. the nondegeneracy hypothesis should not be necessary. This is currently an open problem.

(v) As a corollary of the L^p estimate for the maximal operator one obtains the pointwise convergence result $\lim_{t\to 0}\mu_t * f(x) = cf(x)$ almost everywhere, if $f \in L^p$ and $c = \int d\mu$. Moreover, the L^p bounds of the maximal operator are relevant for certain results in ergodic theory, where one needs to have pointwise control for large t.

(vi) We use in an essential way the invariance of the subspace $\mathfrak{w}$ under the dilation group $\{\delta_t\}$. Namely, this implies a favorable bound for the principal symbol of $(d/dt)\mu_t$ on the fold surface of the associated canonical relation. A similar phenomenon was observed in [11] for averages along light rays.

(vii) One can replace the measure on $\mathfrak{w}$ by a measure supported on a perturbed subspace $\mathfrak{W}$ which is transversal to the center but no longer invariant under $\{\delta_t\}$; then the phenomenon in the last remark does not occur. In the above coordinates $\mathfrak{W}$ is given as

$$\mathfrak{W} = \{(x, \Lambda x), x \in \mathbb{R}^d\}, \tag{1.4}$$

where $\Lambda = (\Lambda_{ij})$ is a $m \times d$ matrix. Define a measure μ_t^Λ by

$$\langle \mu_t^\Lambda, f\rangle = \int f(tx, t^2\Lambda x) d\mu(x);$$

we also set $\mu^\Lambda := \mu_1^\Lambda$. Consider the maximal operator M^Λ defined by

$$M^\Lambda f = \sup_{t>0} |f * \mu_t^\Lambda|. \tag{1.5}$$

For general Λ we then prove the partial result that M^Λ is bounded for $p > (3d-1)/(3d-4)$. We conjecture that boundedness holds for $p > d/(d-1)$ which by our theorem holds true for $\Lambda = 0$.

Structure of the paper: In §2 we shall give the basic decompositions of the operator. Almost orthogonality arguments are used in §2 to reduce matters to a "local" maximal operator (where the dilation parameter is ≈ 1). In order to estimate the local maximal operator it is necessary to understand the precise regularity properties of the averages. It turns out that these are Fourier integral operators with folding canonical relations and our main decomposition is in terms of the (scaled) distance to the surface of degeneracy. In §4 we state the main (known) estimates for oscillatory integral operators associated to fold singularities. In §5 we first reduce the estimate for the averages to estimates for oscillatory integral operators; this argument is rather standard and similar to calculations in [5]. The main part of §5 is concerned with showing that the uniform assumptions (4.3)–(4.5) on the phase in the case of folding canonical relations are indeed satisfied. We then conclude that section discussing the L^2 estimates for $\partial_t[f * \mu_t]$; here we take advantage of the fact that the principal symbol of this operator vanish on the surface of degeneracy. In §6 we complete the proof of the main theorem by deriving appropriate weak type (1, 1) bounds. In the appendix §7 we give an example of a two step nilpotent group which is not isomorphic to a Heisenberg-type group but satisfies the nondegeneracy hypothesis.

Notation: Given two quantities A and B we write $A \lesssim B$ if there is a positive constant C, such that $A \leq CB$.

Note: After a preprint version of our article had been circulated, S. Thangavelu informed us that he and E. K. Narayanan had obtained another proof of the sharp L^p inequality for the case of the spherical maximal function on the Heisenberg group H^n, $n \geq 2$, shortly before they became aware of our preprint. Their argument extends ideas from [12] and is based on estimates for Laguerre functions. It is contained in a preprint entitled "An optimal theorem for the spherical maximal operator on the Heisenberg group".

ACKNOWLEDGEMENT: We thank the referee for suggestions concerning the exposition.

2. Preliminary decompositions

We shall present the argument for the maximal operator M^Λ in (1.5). We shall denote by Λ_j the j^{th} column of Λ and by $\|\Lambda\|$ the matrix norm of Λ with respect to the Euclidean norms on $\mathbb{R}^d$ and $\mathbb{R}^m$. In what follows we shall always assume that $\|\Lambda\| \leq C_1$ for some fixed C_1 (and various bounds may depend on C_1). If $\|\Lambda\|$ occurs explicitly in an estimate then we are interested in the behavior for $\Lambda \to 0$, as the case of our Theorem corresponds to $\Lambda = 0$.

We note that by localizations and rotations in $\mathbb{R}^d$ one can assume that μ has small support and that the projection of Σ to $\mathfrak{w}$ is given as a graph $x_d = \Gamma(x')$, $x' = (x_1, \dots, x_{d-1})$, so that $\nabla_{x'}\Gamma(0) = 0$ and so that μ is supported in a small neighborhood of $(0, \Gamma(0))$ (we may assume that $|\nabla_{x'}\Gamma(x')| \leq C_0^{-1}c_0/100$ where c_0, C_0 are defined in (5.10) below). Note that a rotation has the effect of replacing the matrices J_i in the group law by $Q^t J_i Q$ with $Q \in SO(d)$. We thus will need to prove an estimate which is uniform in these rotations.

Using the Fourier inversion formula for Dirac measures we may write

$$\mu^\Lambda(x, u) = \chi(x, u) \iint e^{i(\sigma(x_d - \Gamma(x')) + \tau \cdot (u - \Lambda x))} d\sigma d\tau$$

where χ is a smooth compactly supported function and the integral converges in the sense of oscillatory integrals (thus in the sense of distributions).

We split the integrals by introducing dyadic decompositions in (σ,τ) and then also in σ, when $|\sigma|<|\tau|$.

Let $\zeta_0\in C_0^\infty(\mathbb{R})$ be an even function so that $\zeta_0(s)=1$ if $|s|\le 1/2$ and $\operatorname{supp}(\zeta_0)\subset(-1,1)$. Also define $\zeta_1(s)=\zeta_0(s/2)-\zeta_0(s)$ and for $k\ge 1$, $1\le l<k/3$,

$$\beta_0(\sigma,\tau)=\zeta_0(\sqrt{\sigma^2+|\tau|^2}),\tag{2.1.1}$$

$$\beta_{k,0}(\sigma,\tau)=\zeta_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})(1-\zeta_0(2^{-k}\sigma)),\tag{2.1.2}$$
$$\beta_{k,l}(\sigma,\tau)=\zeta_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})\zeta_1(2^{l-k}\sigma),$$

$$\widetilde{\beta}_k(\sigma,\tau)=\zeta_1(2^{-k}\sqrt{\sigma^2+|\tau|^2})\zeta_0(2^{[k/3]-k-1}\sigma).\tag{2.1.3}$$

Then observe that

$$\beta_0+\sum_{k\ge1}\Big(\beta_{k,0}+\sum_{1\le l<k/3}\beta_{k,l}+\widetilde{\beta}_k\Big)=1,$$

and for $k>0$ the function $\beta_{k,0}$ is supported where $|\sigma|\approx 2^k$ and $|\tau|\lesssim 2^k$, $\beta_{k,l}$ is supported where $|\tau|\approx 2^k$ and $|\sigma|\approx 2^{k-l}$ and $\widetilde{\beta}_k$ is supported where $|\tau|\approx 2^k$ and $|\sigma|\lesssim 2^{2k/3}$.

Define

$$K^0(x,u)=\chi(x,u)\iint e^{i(\sigma(x_d-\Gamma(x'))+\tau\cdot(u-\Lambda x))}\beta_0(\sigma,\tau)d\sigma d\tau,\tag{2.2.1}$$

$$K^{k,l}(x,u)=\chi(x,u)\iint e^{i(\sigma(x_d-\Gamma(x'))+\tau\cdot(u-\Lambda x))}\beta_{k,l}(\sigma,\tau)d\sigma d\tau,\quad 0\le l<k/3,\tag{2.2.2}$$

$$\widetilde{K}^k(x,u)=\chi(x,u)\iint e^{i(\sigma(x_d-\Gamma(x'))+\tau\cdot(u-\Lambda x))}\widetilde{\beta}_k(\sigma,\tau)d\sigma d\tau;\tag{2.2.3}$$

moreover, for $t>0$ define the dilates

$$[K_t^0,K_t^{k,l},\widetilde{K}_t^k](x,u)=t^{-(d+2m)}[K^0,K^{k,l},\widetilde{K}^k](t^{-1}x,t^{-2}u).$$

Note that $\mu_t^\Lambda=K_t^0+\sum_{k\ge1}(K_t^{k,0}+\sum_{1\le l<k/3}K_t^{k,l}+\widetilde{K}_t^k)$.

Since K^0 is a bounded compactly supported function the associated maximal function is controlled by the appropriate variant of the Hardy–Littlewood maximal function and therefore ([17]) we have the inequality

$$\|\sup_t|f*K_t^0|\|_p\le C_p\|f\|_p$$

for $1<p\le\infty$.

Using known estimates for oscillatory integral operators with fold singularities and additional almost orthogonality estimates we shall derive in §3 and §5 the following L^2 estimates.

PROPOSITION 2.1: *Suppose $k > 0$. Then for $0 \le l < k/3$*

$$\|\sup_t |f * K_t^{k,l}|\|_2 \lesssim \sqrt{k} 2^{-k(d-2)/2}(1 + \|\Lambda\| 2^l)^{1/2}\|f\|_2; \tag{2.3}$$

moreover,

$$\|\sup_t |f * \widetilde{K}_t^k|\|_2 \lesssim \sqrt{k} 2^{-k(d-2)/2}(1 + \|\Lambda\| 2^{k/3})^{1/2}\|f\|_2. \tag{2.4}$$

To obtain L^p results we shall interpolate with weak type inequalities proved in §6.

LEMMA 2.2: *Let $k > 0$. For all $\alpha > 0$ we have*

$$\operatorname{meas}(\{(x,u) : \sup_{t>0} |f * K_t^{k,l}(x,u)| > \alpha\}) \lesssim k 2^{k-l}(1 + \|\Lambda\| 2^l)\alpha^{-1}\|f\|_1 \tag{2.5}$$

for $0 \le l < k/3$ and

$$\operatorname{meas}(\{(x,u) : \sup_{t>0} |f * \widetilde{K}_t^k(x,u)| > \alpha\}) \lesssim k 2^{2k/3}(1 + \|\Lambda\| 2^{k/3})\alpha^{-1}\|f\|_1. \tag{2.6}$$

We interpolate by the real method and obtain

COROLLARY 2.3: *Suppose $1 < p \le 2$ and $k > 0$. Then for $0 \le l < k/3$*

$$\|\sup_t |f * K_t^{k,l}|\|_p \le C_p k^{1/p} 2^{-k(d-1-d/p)} 2^{-l(2/p-1)}(1 + \|\Lambda\| 2^l)^{1/p}\|f\|_p; \tag{2.7}$$

moreover,

$$\|\sup_t |f * \widetilde{K}_t^k|\|_p \le C_p k^{1/p} 2^{-k(d-4/3-d/p+2/3p)}(1 + \|\Lambda\| 2^{k/3})^{1/p}\|f\|_2. \tag{2.8}$$

Now if $p < 2$ we may sum in k and l and see that M^Λ is L^p bounded if $d-4/3-d/p+1/(3p) > 0$ which is equivalent to $p > (3d-1)/(3d-4)$ (showing the estimate mentioned in remark (vii) in the introduction). If $\Lambda = 0$ we get a better bound, namely that L^p boundedness holds if $d - 1 - d/p > 0$ or $p > d/(d-1)$. This proves our main Theorem.

3. Square functions and almost orthogonality

It is advantageous to introduce cancellation in the above kernels, modulo small acceptable errors. Indeed

$$\left|\iint K^{k,l}(x,u)dxdu\right| + \left|\iint \widetilde{K}^k(x,u)dxdu\right| \le C_N 2^{-kN},$$

for all $N = 0, 1, \dots$, and this estimate follows by an integration by parts in the (x, u) variables. Thus there is a C_0^∞ function b which is equal to 1 on $\mathrm{supp}(\chi)$, and constants $\gamma_{k,l}$, γ_k so that

$$\tag{3.1} \begin{gathered} \iint K^{k,l}(x,u)dxdu = \gamma_{k,l}\iint b(x,u)dxdu, \\ \iint \widetilde{K}^k(x,u)dxdu = \gamma_k \iint b(x,u)dxdu, \end{gathered}$$

where

$$\tag{3.2} |\gamma_k| + |\gamma_{k,l}| \le C_N 2^{-kN}.$$

We define

$$\tag{3.3.1} \mathcal{K}^{k,l}(x,u) = K^{k,l}(x,u) - \gamma_{k,l} b(x,u),$$

$$\tag{3.3.2} \widetilde{\mathcal{K}}^k(x,u) = \widetilde{K}^k(x,u) - \gamma_k b(x,u),$$

and denote by $\mathcal{K}_t^{k,l}$, $\widetilde{\mathcal{K}}_t^k$ their dilates, as before. Then the functions $\mathcal{K}_t^{k,l}$, $\mathcal{K}_t^k$ have integral zero.

Since the maximal operator generated by the kernel b (with nonisotropic dilations) is bounded by the nonisotropic Hardy–Littlewood maximal operator we see that for $1 < p \le \infty$

$$\|\sup_t |f * (\mathcal{K}_t^{k,l} - K_t^{k,l})|\|_p \le C_{N,p} 2^{-kN} \|f\|_p.$$

Now in order to deal with the main term we shall use the following standard lemma in the subject which is an immediate consequence of a similar one stated in [17, p. 499].

LEMMA 3.1: *Suppose that*

$$\sup_{s\in[1,2]} \Big(\sum_{n\in\mathbb{Z}} \|F_n(\cdot,s)\|_2^2\Big)^{1/2} \le A_1 \quad \textit{and} \quad \sup_{s\in[1,2]} \Big(\sum_{n\in\mathbb{Z}} \Big\|\frac{\partial F_n}{\partial s}(\cdot,s)\Big\|_2^2\Big)^{1/2} \le A_2.$$

Then

$$\Big\|\sup_n \sup_{s\in[1,2]} |F_n(\cdot,s)|\Big\|_2 \le C(A_1 + \sqrt{A_1 A_2}).$$

We omit the proof. Using Lemma 3.1 one sees that the estimates

$$\|\sup_t |f * \mathcal{K}_t^{k,l}|\|_2 \lesssim \sqrt{k}2^{-k(d-2)/2}(1+\|\Lambda\|2^l)^{1/2}\|f\|_2,$$
$$\|\sup_t |f * \widetilde{K}_t^k|\|_2 \lesssim \sqrt{k}2^{-k(d-2)/2}(1+\|\Lambda\|2^{k/3})^{1/2}\|f\|_2$$

follow from the following estimates which are uniform in $s \in [1,2]$:

$$\Big(\sum_n \|f * \mathcal{K}_{2^n s}^{k,l}\|_2^2\Big)^{1/2} \lesssim \sqrt{k}2^{-k(d-1)/2}2^{l/2}\|f\|_2, \tag{3.4}$$

$$\Big(\sum_n \Big\|f * \Big[t\frac{\partial}{\partial t}\mathcal{K}_t^{k,l}\Big]_{t=2^n s}\Big\|_2^2\Big)^{1/2} \lesssim \sqrt{k}2^{-k(d-3)/2}2^{-l/2}(1+\|\Lambda\|2^l)\|f\|_2, \tag{3.5}$$

for $l < k/3$, and

$$\Big(\sum_n \|f * \widetilde{\mathcal{K}}_{2^n s}^{k}\|_2^2\Big)^{1/2} \lesssim \sqrt{k}2^{-k(d-1)/2+k/6}\|f\|_2, \tag{3.6}$$

$$\Big(\sum_n \Big\|f * \Big[t\frac{\partial}{\partial t}\widetilde{\mathcal{K}}_t^{k}\Big]_{t=2^n s}\Big\|_2^2\Big)^{1/2} \lesssim \sqrt{k}2^{-k(d-3)/2-k/6}(1+\|\Lambda\|2^{k/3})\|f\|_2. \tag{3.7}$$

Note by scaling that it suffices to prove these estimates for $s = 1$. We shall first use the cancellation of the kernels $\mathcal{K}_{2^n s}^{k,l}$ and $\widetilde{\mathcal{K}}_{2^n s}^{k}$ to show certain almost orthogonality properties (for the sums in n) and then we use stronger estimates for oscillatory integrals to establish decay estimates for fixed n.

An almost orthogonality lemma. We first state a simple and presumably well known consequence of the Cotlar–Stein Lemma.

Lemma 3.2: *Suppose $0 < \varepsilon < 1$, $A \le B/2$ and let $\{T_n\}_{n=1}^{\infty}$ be a sequence of bounded operators on a Hilbert space H so that the operator norms satisfy*

$$\|T_n\| \le A \tag{3.8}$$

and

$$\|T_n T_{n'}^*\| \le B^2 2^{-\varepsilon|n-n'|}. \tag{3.9}$$

Then for all $f \in H$

$$\Big(\sum_{n=1}^{\infty} \|T_n f\|^2\Big)^{1/2} \le CA\sqrt{\varepsilon^{-1}\log(B/A)}\|f\|. \tag{3.10}$$

Proof: For $N \geq 1$ consider the operator

$$\mathcal{T}_N \colon H \to \ell^2(H)$$

which maps f to the sequence $(T_1 f, \dots, T_N f, 0, 0, \dots)$. Now $\|\mathcal{T}_N\| = \|\mathcal{T}_N^* \mathcal{T}_N\|^{1/2}$ where $\mathcal{T}_N^* \mathcal{T}_N \colon H \to H$ is given by

$$\mathcal{T}_N^* \mathcal{T}_N f = \sum_{n=1}^{N} T_n^* T_n f.$$

We let $S_n = T_n^* T_n$ and observe that

$$\begin{aligned} \|S_k^* S_l\| = \|S_k S_l^*\| &= \|T_k^* T_k T_l^* T_l\| \\ &\leq \|T_k^*\| \|T_k T_l^*\| \|T_l\| \leq A^2 \min\{A^2, B^2 2^{-|k-l|\varepsilon}\}. \end{aligned}$$

The standard Cotlar–Stein Lemma [17] gives

$$\|\mathcal{T}_N^* \mathcal{T}_N\| \leq \sum_{m=\infty}^{\infty} \max\{ \sup_{k-l=m} \|S_k^* S_l\|^{1/2}, \sup_{k-l=m} \|S_k S_l^*\|^{1/2}\}$$

and thus

$$\begin{aligned} \|\mathcal{T}_N\|^2 &\leq A \sum_{m=-\infty}^{\infty} \min\{A, B 2^{-|m|\varepsilon}\} \\ &\leq C^2 \varepsilon^{-1} A^2 \log(B/A). \end{aligned}$$

Thus $\|\mathcal{T}_N f\|_{\ell^2(H)}$ is dominated by the right hand side of (3.10), and the assertion follows by taking the limit as $N \to \infty$. ∎

Remark: We proved Lemma 3.2 by using the statement of the Cotlar–Stein Lemma. Using the proof of the Cotlar–Stein Lemma one can also show the following more general fact: If $\|T_n T_{n'}^*\| \leq \alpha^2(n - n')$ then

$$\Big(\sum_{n=1}^{N} \|T_n f\|^2 \Big)^{1/2} \lesssim \Big(\sum_{j \in \mathbb{Z}} |\alpha(j)|^2 \Big)^{1/2} \|f\|.$$

Of course, Lemma 3.2 is an immediate consequence of this inequality.

Almost orthogonality estimates. Here we wish to apply Lemma 3.2 to convolutions on groups. If $Tf = f * g$ we first note that its adjoint is given by $T^* f = f * g^*$ where $g^* = \overline{g(\cdot^{-1})}$. Moreover, using Minkowski's inequality and the unimodularity of nilpotent Lie groups one obtains the standard convolution inequality

$$\|f * g\|_2 \leq \|g^*\|_1 \|f\|_2 = \|g\|_1 \|f\|_2.$$

We now fix k, l and $s \in [1,2]$ and derive almost orthogonality properties for the operators of convolution with $\mathcal{K}^{k,l}_{2^n s}$.

Notice that for $n \le 0$ the function $\mathcal{K}^{k,l}_{2^n s}$ is supported in a (small) ball of radius $C2^n$ (in fact in a smaller nonisotropic ball). Moreover, we have $|\nabla_{y,v}\mathcal{K}^{k,l}_s(y,v)| \le 2^{k(m+2)}$ and using the cancellation of $\mathcal{K}^{k,l}_{2^n s}$ we obtain

$$|\mathcal{K}^{k,l}_s * (\mathcal{K}^{k,l}_{2^n s})^*(x,u)| \lesssim 2^{k(m+2)}2^n \quad \text{if } n \le 0.$$

By scaling and applying Schur's Lemma we obtain

$$\|f * \mathcal{K}^{k,l}_{2^{n'} s} * (\mathcal{K}^{k,l}_{2^n s})^*\|_2 \lesssim 2^{k(m+2)}2^{-|n-n'|}\|f\|_2 \tag{3.11}$$

first for $n \le n'$ and then by taking adjoints also for $n < n'$. This and the following estimates are uniform in $s \in [1,2]$.

Similarly we get

$$\Big\|f * s\frac{\partial\mathcal{K}^{k,l}_{2^{n'} s}}{\partial s} * s\frac{\partial(\mathcal{K}^{k,l}_{2^{n'} s})^*}{\partial s}\Big\|_2 \lesssim 2^{k(m+4)}2^{-|n-n'|}\|f\|_2 \tag{3.12}$$

and also

$$\|f * \widetilde{\mathcal{K}}^k_{2^{n'} s} * (\widetilde{\mathcal{K}}^k_{2^n s})^*\|_2 \lesssim 2^{k(m+2)}2^{-|n-n'|}\|f\|_2, \tag{3.13}$$

$$\Big\|f * s\frac{\partial\widetilde{\mathcal{K}}^k_{2^n s}}{\partial s} * s\frac{\partial(\widetilde{\mathcal{K}}^k_{2^{n'} s})^*}{\partial s}\Big\|_2 \lesssim 2^{k(m+4)}2^{-|n-n'|}\|f\|_2. \tag{3.14}$$

In §5 we shall prove the inequalities

$$\|f * K^{k,l}\|_2 \lesssim 2^{-k(d-1)/2}2^{l/2}\|f\|_2, \tag{3.15}$$

$$\Big\|f * \Big[\frac{\partial K^{k,l}_s}{\partial s}\Big]_{s=1}\Big\|_2 \lesssim 2^{-k(d-3)/2}2^{-l/2}(1+\|\Lambda\|2^l)\|f\|_2 \tag{3.16}$$

for $l < k/3$, and

$$\|f * \widetilde{K}^k\|_2 \lesssim 2^{-k(d-1)/2}2^{k/6}\|f\|_2, \tag{3.17}$$

$$\Big\|f * \frac{\partial\widetilde{K}^k_s}{\partial s}|_{s=1}\Big\|_2 \lesssim 2^{-k(d-3)/2}2^{-k/6}(1+\|\Lambda\|2^{k/3})\|f\|_2. \tag{3.18}$$

By scaling and by (3.2) the same inequalities hold with $K^{k,l}$ and $\widetilde{K}^k$ replaced by $\mathcal{K}^{k,l}_t$ and $\widetilde{\mathcal{K}}^k_t$ and with $\partial_s K^{k,l}$, $\partial_s\widetilde{K}^k$ replaced by $\partial_s\mathcal{K}^{k,l}_{2^n s}$, $\partial_s\widetilde{\mathcal{K}}^k_{2^n s}$, for $1 \le s \le 2$.

Now the inequality (3.4) follows from (3.15) and (3.11) if we apply Lemma 3.2 with $A = 2^{-k(d-1)/2}2^{l/2}$ and $B = 2^{k(m+4)}$. Similarly (3.5) follows from (3.16) and (3.12), (3.6) from (3.17) and (3.13), and (3.7) from (3.18) and (3.14).

The next two sections are concerned with the derivation of inequalities (3.15–18).

4. Preliminaries on oscillatory integral operators with folding canonical relations

We shall reduce matters to estimates for oscillatory integral operators whose canonical relations have two-sided fold singularities. We consider localizations near the fold surface and the estimate goes back to Phong and Stein [13] for certain conormal operators in the plane; the general case is implicit in Cuccagna's paper [3]. For the version needed here we refer to [6].

Let $\Omega \in \mathbb{R}^n \times \mathbb{R}^n$ be an open set and let Γ be an open set in some finite dimensional space. We consider phases $\varphi(x,y,\gamma)$ and amplitudes $a_\lambda(x,y,\gamma)$, $(x,y,\gamma) \in \Omega \times \Omega \times \Gamma$, and assume that

$$|\partial_x^\alpha \partial_y^\beta \varphi(x,y,\gamma)| \le C, \tag{4.1}$$

$$|\partial_x^\alpha \partial_y^\beta a_\lambda(x,y,\gamma)| \le C\lambda^{(|\alpha|+|\beta|)/3} \tag{4.2}$$

say, for all multiindices α, β with $|\alpha|, |\beta| \le 10n$, with uniform bounds in $\Omega \times \Gamma$; we also assume that all derivatives depend continuously on the parameter γ.

We shall assume that

$$\mathcal{C}_\varphi = \{(x, \varphi_x, y, -\varphi_y)\}$$

is a folding canonical relation, i.e. for each point $P_0 = (x_0, y_0, \gamma_0)$ we have

$$\text{rank}\varphi''_{xy}(P_0) \ge n-1, \tag{4.3}$$

and for unit vectors U, V

$$\varphi''_{xy}(P_0)V = 0 \Longrightarrow |\langle V, \nabla_y\rangle \det \varphi''_{xy}| \ge c, \tag{4.4}$$

$$U^t\varphi''_{xy}(P_0) = 0 \Longrightarrow |\langle U, \nabla_x\rangle \det \varphi''_{xy}| \ge c, \tag{4.5}$$

for some $c > 0$.

We consider the oscillatory integral operator $T_\lambda[b]$ defined by

$$T_\lambda[b]f(x) = \int e^{i\lambda\varphi(x,y,\gamma)} b(x,y,\gamma) f(y)dy$$

which is bounded on all L^p if b is bounded and compactly supported. We shall take for b certain localizations of the symbol in terms of the size of $\det \varphi''_{xy}$. Let ϑ be smooth and compactly supported in $(-1,1)$ so that $\vartheta(s) = 1$ for $|s| \le 1/2$ and set

$$\vartheta_l(x,y,\gamma) = \vartheta(2^l \det \varphi''_{xy}(x,y,\gamma)) - \vartheta(2^{l+1} \det \varphi''_{xy}(x,y,\gamma)),$$

so that ϑ_l localizes to the set where $|\det\varphi''_{xy}| \approx 2^{-l}$. We also define

$$\zeta^\lambda(x,y) = 1 - \sum_{2^l<\lambda^{1/3}} \vartheta_l(x,y)$$

so that $|\det\varphi''_{xy}| \lesssim \lambda^{-1/3}$ on $\mathrm{supp}(\zeta^\lambda)$.

Then there is a neighborhood $\mathcal{U}$ of (x_0, y_0, γ_0) so that for all a_λ satisfying (4.2), supported in $\mathcal{U}$ the following estimates hold for the operator norms:

$$\|T_\lambda[a_\lambda\vartheta_l]\|_{L^2\to L^2} \le C_1 2^{l/2}\lambda^{-n/2}, \quad 2^l \le \lambda^{1/3} \tag{4.6}$$

and

$$\|T_\lambda[a_\lambda\zeta^\lambda]\|_{L^2\to L^2} \le C_1\lambda^{1/6-n/2}. \tag{4.7}$$

These estimates are a consequence of Theorem 2.1 in [6].

5. Reduction to oscillatory integral operators

We now consider the operator of convolution with $K^{k,l}$ and give the proof of the bound (3.15). The operator $\partial_s K^{k,l}$ is more singular, but its estimation is rather analogous, so we shall point out the modifications needed for (3.16) at the end of this section. The estimations for $\widetilde{K}^k$ and $\partial_s\widetilde{K}^k_s$ will be similar.

Since $K^{k,l}$ is compactly supported in a fixed neighborhood we may use the translation invariance to reduce to the case that f is also compactly supported in a fixed neighborhood of the origin. Thus it suffices to show the desired bound for the operator with Schwartz kernel

$$\chi_1(x,u)K^{k,l}(x-y, u-v+x^tJy)\chi_2(y,v), \tag{5.1}$$

for suitable compactly supported smooth functions χ_1 and χ_2. In what follows we set $\lambda = 2^k$ and then by a change of variables the kernel (5.1) can be written as

$$H^{\lambda,l}(x,u,y,v) = \lambda^{m+1}\iint e^{i\lambda\phi(x,u,y,v,\sigma,\tau)}\chi_0(x,u,y,v)\eta_l(\sigma,\tau)d\sigma d\tau \tag{5.2}$$

where

$$\phi(x,u,y,v,\sigma,\tau) = \sigma(x_d - y_d - \Gamma(x'-y')) + \tau\cdot(u - v + x^tJy - \Lambda(x-y))$$

and where $|\tau| \approx 1$ and $|\sigma| \approx 2^{-l}$ on the support of η_l; specifically

$$\eta_l(\sigma,\tau) = \zeta_1(\sqrt{\sigma^2+|\tau|^2})\zeta_1(2^l\sigma),$$

and $\chi_0(x,u,y,v) = \chi_1(x,u)\chi(x-y, u-v+x^tJy)\chi_2(y,v)$.

Notation: We let $P\colon \mathbb{R}^d \to \mathbb{R}^{d-1}$ be the linear map with $Pe_i = e_i$, $i = 1, \dots, d-1$ and $Pe_d = 0$. We also use the notation P for the $(d-1) \times d$ matrix

$$P = (\, I \quad 0\,)$$

and P^t for its transpose.

STATIONARY PHASE CALCULATIONS. We wish to apply stationary phase arguments to reduce matters to the estimation of an oscillatory integral operators without frequency variables (see e.g. the general discussion in [5]).

We shall apply a scaled Fourier transform on $\mathbb{R}^{m+1}$, in the (x_d, u) variables. Define

$$\mathcal{F}_\lambda g(x', x_d, u) = \iint e^{-i\lambda(x_d z_d + u\cdot w)} g(x', z_d, w)\, dz_d dw;$$

then $(\lambda/2\pi)^{(m+1)/2}\mathcal{F}_\lambda$ is a unitary operator and thus, if $\mathcal{H}^{\lambda,l}$ denotes the operator with Schwartz kernel $H^{\lambda,l}$, we have to prove that $\mathcal{F}_\lambda \mathcal{H}^{\lambda,l}$ maps L^2 to itself with operator norm $O(\lambda^{-(d+m)/2} 2^{l/2})$. Let $\chi_3(x_d, u)$ denote a smooth compactly supported function which is equal to one whenever $|x_d| + |u| \le 10$, and define $\mathcal{F}_{\lambda,1}$ by

$$\mathcal{F}_{\lambda,1} g(x', x_d, u) = \chi_3(x_d, u) \iint e^{-i\lambda(x_d z_d + u\cdot w)} g(x', z_d, w)\, dz_d dw;$$

moreover, let $\mathcal{F}_{\lambda,2} = \mathcal{F}_\lambda - \mathcal{F}_{\lambda,1}$. Then the Schwartz kernel of $\mathcal{F}_{\lambda,1} H^{\lambda,l}$ is given by

$$\lambda^{m+1} \int e^{i\lambda\Psi(x,u,y,v,\theta)} b_l(x, u, y, v, \theta)\, d\theta, \tag{5.3}$$

where with

$$\theta = (z_d, w, \sigma, \tau)$$

the phase function Ψ is given by

$$\begin{aligned}\Psi(x, u, y, v, \theta) = &- x_d z_d - u\cdot w + \sigma(z_d - y_d - \Gamma(x' - y')) \\ &+ \tau^t(w - v + \Lambda P^t(x' - y') + \Lambda_d(z_d - y_d) + ({x'}^t, z_d) J y),\end{aligned}$$

and the amplitude is given by

$$b_l(x, u, y, v, \theta) = \chi_3(x_d, u)\chi_0(x', z_d, y, w)\eta_l(\sigma, \tau).$$

For the error term $\mathcal{F}_{\lambda,2} H^{\lambda,l}$ we have a similar formula, only with χ_3 replaced by $1 - \chi_3$. Then in view of the support properties of $(1 - \chi_3)$ we see that

$|\nabla_{z_d,w}\Psi| \geq |x_d| + |u|$ on $\operatorname{supp}(1-\chi_3)$ and by integration by parts with respect to the (z_d, w) variables we see that the kernel of $\mathcal{F}_{\lambda,2}H^{\lambda,l}$ is bounded by $C_N\lambda^{m+1-N}(|x_d| + |u|)^{-N}$. Moreover, this kernel is supported on a set where $|x_d| + |u| \geq 1$ and where $|x'| + |y| + |v| \leq C$. Thus, with an obvious application of Schur's Lemma we conclude that the operator $\mathcal{F}_{\lambda,2}H^{\lambda,l}$ is bounded on L^2 with operator norm $O(\lambda^{-N})$ for any N.

We return to the main term $\mathcal{F}_{\lambda,1}H^{\lambda,l}$ and it remains to be shown that

$$\|\mathcal{F}_{\lambda,1}\mathcal{H}^{\lambda,l}\| \lesssim 2^{l/2}\lambda^{-(d+m)/2}. \tag{5.4}$$

Note that for fixed (x, u, y, v) the phase function Ψ is a polynomial of degree ≤ 2 in the θ variables and that the Hessian $\Psi''_{\theta\theta}$ is nondegenerate.

Indeed,

$$\begin{aligned} \Psi'_{z_d} &= -x_d + e_d^t J_\tau y + \sigma + \tau^t \Lambda_d, \\ \Psi'_w &= \tau - u, \\ \Psi'_\tau &= w - v + (x'^t, z_d)Jy + \Lambda P^t(x'-y') + \Lambda_d(z_d - y_d), \\ \Psi'_\sigma &= z_d - y_d - \Gamma(x'-y'), \end{aligned} \tag{5.5}$$

and with Ξ denoting the column vector in $\mathbb{R}^m$ with coordinates $\Xi_i = e_d^t J_i y + \Lambda_{id}$ we have

$$\Psi''_{\theta\theta} = \begin{pmatrix} 0 & 0 & \Xi^t & 1 \\ 0 & 0 & I & 0 \\ \Xi & I & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Clearly the linear equations $\Psi_\theta = 0$ have a unique solution

$$\theta_{\text{crit}} = [z_d, w, \tau, \sigma]_{\text{crit}}(x, u, y, v),$$

with

$$\begin{aligned} (z_d)_{\text{crit}} &= y_d + \Gamma(x'-y'), \\ (w_i)_{\text{crit}} &= v_i - (x'^t, y_d + \Gamma(x'-y'))J_i y - e_i^t \Lambda P^t(x'-y') - \Lambda_{id}\Gamma(x'-y'), \\ (\tau_i)_{\text{crit}} &= u_i, \\ \sigma_{\text{crit}} &= x_d - \sum_{i=1}^m u_i(e_d^t J_i y + \Lambda_{id}), \end{aligned}$$

and we can apply the method of stationary phase (with respect to the $2(m+1)$

frequency variables θ). Setting

(5.6)
$$\Phi(x,u,y,v):=\Psi(x,u,y,v,\theta_{\rm crit}(x,u,y,v))=-x_d(y_d+\Gamma(x'-y'))$$
$$-\sum_{i=1}^{m}u_i(v_i-(x'^t,y_d+\Gamma(x'-y'))J_iy-\Lambda_{id}\Gamma(x'-y')-e_i^t\Lambda P^t(x'-y'))$$

we obtain that

(5.7)
$$\lambda^{m+1}\int e^{i\lambda\Psi(x,u,y,v,\theta)}b_l(x,u,y,v,\theta)d\theta$$
$$=e^{i\lambda\Phi(x,u,y,v)}\sum_{j=0}^{N-1}\mathcal{E}_j^l(x,u,y,v)\lambda^{-j}+R_N^{\lambda,l}(x,u,y,v)$$

where

$$(5.8)\qquad \begin{aligned}\mathcal{E}_j^l(x,u,y,v)=&(2i)^{-j}(\det(\Psi_{\theta\theta}(x,y,u,v,\theta_{\rm crit}(x,u,y,v))/2\pi i)^{-1/2}\\ &\times\frac{1}{j!}\langle\Psi_{\theta\theta}^{-1}D_\theta,D_\theta\rangle^j b_l(x,u,y,v,\theta)|_{\theta=\theta_{crit}(x,u,y,v)}\end{aligned}$$

and

$$(5.9)\qquad |R_N^{\lambda,l}(x,u,y,v)|\le C_N\|b_l\|_{L^2_{m+2+2N}}\lambda^{-N}\le C_N'2^{l(m+2+2N)}\lambda^{-N}.$$

Here we have applied Lemma 7.7.3 in [7].

Since $2^l\le\lambda^{1/3}$, the error term $R_N^{\lambda,l}$ (which is compactly supported) defines a bounded operator on L^p with norm $O(\lambda^{-(2m+1+N)/3})$ which for large N is much better than the desired bound in (5.4).

Claim 5.1: *The operators with kernels $\lambda^{-j}\mathcal{E}_j^l(x,u,y,v)e^{i\lambda\Phi(x,u,y,v)}$ have L^2 operator norm $O(\lambda^{-(d+m)/2-j/3}2^{l/2})$.*

This clearly implies (5.4).

Geometry of the canonical relation. We consider the canonical relation $\mathcal{C}_\Phi=(x,u,\Phi_x,\Phi_u;y,v,-\Phi_y,-\Phi_v)$ and the singularities of the maps p_L: $(y,v)\mapsto(\Phi_x,\Phi_u)$, $p_R:(x,u)\mapsto(\Phi_y,\Phi_v)$. It is our objective to check the analogues of (4.3–4.5) and we will have to verify a few elementary linear algebra facts.

Let A denote the $(d-1)\times(d-1)$ matrix $\Gamma''(x'-y')$ and let B denote the column vector $\Gamma'(x'-y')\in\mathbb{R}^{d-1}$; recall that we may assume that $\|B\|$ is small. Indeed if

$$(5.10.1)\qquad c_0=\min_{u\in S^{m-1}}\|J_u^{-1}\|^{-1},$$

$$(5.10.2)\qquad C_0=\max_{u\in S^{m-1}}\|J_u\|,$$

we may assume that

$$\|B\| \le C_0^{-1} c_0/100.$$

Now p_L is explicitly given by

$$\begin{aligned}
\Phi_{x'} &= -x_d\Gamma'(x'-y') + PJ_u y + \Gamma'(x'-y')e_d^t J_u y + u^t\Lambda_d\Gamma'(x'-y') + u^t\Lambda P^t,\\
\Phi_{x_d} &= -y_d - \Gamma(x'-y'),\\
\Phi_{u_i} &= -(v_i - ({x'}^t, y_d + \Gamma(x'-y'))J_i y - e_i^t\Lambda P^t(x'-y') - \Lambda_{id}\Gamma(x'-y')).
\end{aligned}$$

We compute the differential Dp_L as

(5.11)

$$\Phi''_{(x,u),(y,v)} = \begin{pmatrix} (x_d - e_d^t J_u y - u^t\Lambda_d)A + PJ_uP^t + Be_d^tJ_uP^t & PJ_ue_d & 0\\ B^t & -1 & 0\\ C & c & I\end{pmatrix}$$

where I is an $m\times m$ identity matrix and C is an $m\times(d-1)$ matrix with rows $C_i = {x'}^t PJ_iP^t + y_d e_d^t J_i P^t - (e_d^t J_i y + \Lambda_{id})B^t + e_i^t\Lambda P^t + \Gamma(x'-y')e_d^t J_i P^t$ and c is the column in $\mathbb{R}^m$ with $c_i = ({x'}^t, 0)J_i e_d + e_d^t J_i y$. In this calculation the skew symmetry of the J_i is used.

We now compute the determinant of (5.11) and obtain

$$\det\Phi''_{(x,u),(y,v)} = (-1)^d\det((x_d - e_d^tJ_uy - u^t\Lambda_d)A + PJ_uP^t + E(B)) \tag{5.12}$$

where

$$E(B) = Be_d^tJ_uP^t + PJ_ue_dB^t. \tag{5.13}$$

Here we used the factorization

$$\begin{pmatrix}\sigma A + PJ_uP^t + Be_d^tJ_uP^t & PJ_ue_d\\ B^t & -1\end{pmatrix} = \begin{pmatrix}\sigma A + PJ_uP^t + E(B) & PJ_ue_d\\ 0 & -1\end{pmatrix}\begin{pmatrix} I & 0\\ -B^t & 1\end{pmatrix}.$$

Note that $E(B)$ is a skew-symmetric $(d-1)\times(d-1)$ matrix and so is $PJ_uP^t + E(B)$. Thus, since $d-1$ is odd, the rank of $PJ_uP^t + E(B)$ is at most $d-2$, and the following lemma shows that for small B the rank is equal to $d-2$.

LEMMA 5.2: *Suppose that*

$$\|B\| \le c_0/4C_0.$$

Then the following holds:

(i) *If* $W \in \operatorname{Ker}(PJ_uP^t + E(B))$ *then*

$$|e_d^tJ_uP^tW| \ge \frac{c_0}{2}\|W\|. \tag{5.14}$$

(ii) $\dim \operatorname{Ker}(PJ_uP^t + E(B)) = 1$.

(iii) *If X belongs to the orthogonal complement of* $\operatorname{Ker}(PJ_uP^t + E(B))$ *then*

$$\|(PJ_uP^t + E(B))X\| \geq \frac{c_0}{2}\|X\|. \tag{5.15}$$

Proof: Observe that

$$\|E(B)\| \leq 2C_0\|B\|.$$

Thus if $W \in \operatorname{Ker}(PJ_uP^t + E(B))$ and $\|W\| = 1$ then

$$\begin{aligned} 1 = \|P^tW\| &\leq \|J_u^{-1}\|\|J_uP^tW\| \\ &\leq \|J_u^{-1}\|(|e_d^tJ_uP^tW| + \|PJ_uP^tW\|) \\ &= \|J_u^{-1}\|(|e_d^tJ_uP^tW| + \|E(B)W\|) \\ &\leq c_0^{-1}(|e_d^tJ_uP^tW| + 2C_0\|B\|) \end{aligned}$$

and thus, if $\|B\| \leq c_0/4C_0$, we obtain $|e_d^tJ_uP^tW| \geq c_0/2$ which is (5.14).

Let $S_u = J_u + E(B)$. Since S_u is skew symmetric, it can be diagonalized over $\mathbb{C}$, and the eigenvalues are imaginary. The bounds (5.10.1/2) are still valid if J_u^{-1} is acting as a linear transformation on $\mathbb{C}^d$. Let $\eta \in \mathbb{C}^d$ be a unit eigenvector of S_u so that $S_u\eta = i\lambda\eta$ and $\|\eta\| = 1$; then

$$|\lambda| = \|S_u\eta\| \geq \|J_u\eta\| - \|E(B)\eta\| \geq c_0 - \|E(B)\| \geq c_0 - 2C_0\|B\| \geq c_0/2$$

by assumption on B. Hence $|\lambda| \geq c_0/2$ for every eigenvalue $i\lambda$ of S_u. In particular S_u is nondegenerate. But then $PS_uP^t = PJ_uP^t + E(B)$ has rank $d-2$ and therefore a one-dimensional kernel and all nontrivial eigenvalues of S_u are also eigenvalues of PS_uP^t. This implies for vectors X orthogonal to the kernel of PS_uP^t that

$$PS_uP^tX \geq \frac{c_0}{2}\|X\|$$

which is (5.15).

LEMMA 5.3: *Let $\mathfrak{A}$ be a symmetric positive definite matrix on $\mathbb{R}^n$ and let S be a skew-symmetric matrix on $\mathbb{R}^n$. Then:*

(i) *For all $\sigma \neq 0$, the matrix $\sigma\mathfrak{A} + S$ is invertible and the inverse satisfies the bounds*

$$\|(\sigma\mathfrak{A} + S)^{-1}\| \leq |\sigma|^{-1}\|\mathfrak{A}^{-1}\|. \tag{5.16}$$

(ii) *If S is invertible then $\sigma\mathfrak{A} + S$ is invertible for all σ and we have the bound*

$$\|(\sigma\mathfrak{A} + S)^{-1}\| \leq 2\|S^{-1}\| \quad \text{if } |\sigma| \leq (2\|\mathfrak{A}\|\|S^{-1}\|)^{-1}. \tag{5.17}$$

Proof: For a unit vector e in $\mathbb{R}^n$ we get

$$\|(\sigma\mathfrak{A}+S)e\| \geq |\langle(\sigma\mathfrak{A}+S)e,e\rangle| = |\langle\sigma\mathfrak{A}e,e\rangle| \geq |\sigma|\|\mathfrak{A}^{-1}\|^{-1}.$$

Here we have used that by the skew symmetry of S we have $\langle Se,e\rangle = 0$, and also that $\|\mathfrak{A}^{-1}\| = 1/\lambda_{\min}$, where $\lambda_{\min}$ is a minimal eigenvalue of $\mathfrak{A}$. This establishes invertibility and the bound (5.16).

If in addition S is invertible and σ is small, we may simply use the Neumann series to get invertibility of $\sigma\mathfrak{A}+S$. Namely, if $|\sigma| \leq (2\|\mathfrak{A}\|\|S^{-1}\|)^{-1}$ we get $(\sigma\mathfrak{A}+S)^{-1} = S^{-1}(I+\sum_{j=1}^{\infty}(-1)^j\sigma^j(\mathfrak{A}S^{-1})^j)$ and the bound (5.17) is immediate. ∎

LEMMA 5.4: *Let $\ell \geq 1$ be an odd integer, let Ω_1 be the cone of real symmetric positive definite $\ell \times \ell$ matrices and let Ω_2 be the set of all skew symmetric $\ell \times \ell$ matrices with rank $\ell - 1$.*

For $S \in \Omega_2$ choose a unit vector e_S in the kernel of S and let π_S be the orthogonal projection to the orthogonal complement of e_S.

Then for $A \in \Omega_1$, $S \in \Omega_2$, $\sigma \in \mathbb{R}$ we have

$$\det(\sigma A+S) = \sigma\langle Ae_S,e_S\rangle \det(\pi_S(\sigma A+S)\pi_S^*) + \sigma^2F(A,S,\sigma) \tag{5.18}$$

where F is a smooth function on $\Omega_1 \times \Omega_2 \times \mathbb{R}$.

Proof: Let $Q = Q(S)$ be an orthogonal transformation with $e_S^tQ = (0,\ldots,1)$. Then

$$Q^t(\sigma A+S)Q = \begin{pmatrix} \sigma A_0+S_0 & \sigma a \\ \sigma a^t & \sigma\eta \end{pmatrix}$$

where S_0 is a skew-symmetric invertible $(\ell-1)\times(\ell-1)$ matrix, A_0 is positive definite, $a \in \mathbb{R}^{\ell-1}$ and $\eta = \langle Ae_S,e_S\rangle$. We apply Lemma 5.3 to σA_0+S_0 and factor

$$\begin{pmatrix} \sigma A_0+S_0 & \sigma a \\ \sigma a^t & \sigma\eta \end{pmatrix} = \begin{pmatrix} I & 0 \\ \sigma a^t(\sigma A_0+S_0)^{-1} & 1 \end{pmatrix} \begin{pmatrix} \sigma A_0+S_0 & \sigma a \\ 0 & \sigma\eta - \sigma^2a^t(\sigma A_0+S_0)^{-1}a \end{pmatrix}$$

and conclude that

$$\det(\sigma A+S) = \det(\sigma A_0+S_0)(\sigma\eta - \sigma^2a^t(\sigma A_0+S_0)^{-1}a).$$

The assertion follows since $\det(\sigma A_0+S_0) = \det(\pi_S(\sigma A+S)\pi_S^*)$. ∎

VERIFICATION OF (4.3)–(4.5). We now use the above lemmata to verify the analogues of conditions (4.3–5) for the phase function Φ in (5.6). By Lemma 5.3 the determinant of $\Phi''_{(x,u),(y,v)}$ can only vanish when $\sigma := \sigma_{cr} \equiv x_d - e_d^t J_u y - u^t \Lambda_d$ vanishes. In this case the dimension of the kernel $\Phi''_{(x,u),(y,v)}$ is equal to the dimension of the kernel of $PJ_uP^t + E(B)$ with $B = \Gamma'(x' - y')$, thus equal to 1. Thus $\operatorname{rank}(\Phi''_{(x,u),(y,v)}) \geq d + m - 1$ everywhere.

In order to verify (4.4) let V_L be a nonvanishing vector field which is in the kernel of Dp_L when the mixed Hessian (5.11) becomes singular (i.e. when $x_d - e_d^t J_u y - u^t \Lambda_d = 0$). Then

$$V_L = \sum_{j=1}^{d-1} W_{L,j} \frac{\partial}{\partial y_j} + g_L \frac{\partial}{\partial y_d} + \sum_{i=1}^{m} h_{L,i} \frac{\partial}{\partial v_i}, \tag{5.19}$$

and with $A = \Gamma''(x' - y')$, we have $g_L = B^t W_L$ and

$$(\sigma A + PJ_uP^t + Be_d^t J_u P^t + PJ_u e_d B^t) W_L = 0; \tag{5.20}$$

moreover, the functions $h_{L,i}$ are in the ideal generated by the $W_{L,j}$ (and the coefficients can be computed from (5.11)). To get a nontrivial kernel (when $\sigma = 0$) we must choose a nonvanishing vector W_L satisfying (5.20). Notice that then $|e_d^t J_u P^t W_L|$ is bounded below, by (5.14). By Lemma 5.4 we have

$$\begin{aligned} &V_L(\det \Phi''_{(x,u),(y,v)}) \\ &\qquad = (-1)^d F_1(x,y,u) e_d^t J_u P^t W_L + F_2(x,y,u,v)(x_d - e_d^t J_u y - u^t \Lambda_d) \end{aligned}$$

where F_1 and F_2 are smooth and F_1 does not vanish. Thus $|V_L(\det \Phi''_{(x,u),(y,v)})| \geq c$ on the zero set of $\det \Phi''_{(x,u),(y,v)}$.

Next we consider the map p_R and let V_R be a nonvanishing vector field which is in the kernel of Dp_R (or the cokernel of (5.11)) when $x_d - e_d^t J_u y - u^t \Lambda_d = 0$. Then

$$V_R = \sum_{j=1}^{d-1} W_{R,j} \frac{\partial}{\partial x_j} + g_R \frac{\partial}{\partial x_d} + \sum_{i=1}^{m} h_{R,i} \frac{\partial}{\partial u_i}$$

where by (5.11) the functions $h_{R,i}$ vanish when $x_d - e_d^t J_u y - u^t \Lambda_d = 0$ and

$$\begin{aligned} &W_R^t[\sigma A + PJ_uP^t + Be_d^t J_u P^t] + g_R B^t = 0, \\ &W_R^t P J_u e_d - g_R = 0; \end{aligned}$$

thus since A is symmetric and J_u skew symmetric we have essentially the same equation for W_L above, except that J_u is replaced by $-J_u$:

$$(\sigma A - PJ_uP^t - PJ_u e_d B^t - e_d^t J_u P^t) W_R = 0. \tag{5.21}$$

Moreover, $g_R = e_d^t J_u P^t W_R$ does not vanish by (5.14). As $x_d - e_d^t J_u y - u^t \Lambda_d$ does not depend on x' we get

$$V_R(\det \Phi''_{(x,u),(y,v)}) = \widetilde{F}_1(x,y,u) e_d^t J_u P^t W_R + \widetilde{F}_2(x,y,u,v)(x_d - e_d^t J_u y - u^t \Lambda_d)$$

with smooth functions $\widetilde{F}_1$, $\widetilde{F}_2$ and nonvanishing $\widetilde{F}_1$. Thus $|V_R(\det \Phi''_{(x,u),(y,v)})|$ is bounded below on the zero set of $\det \Phi''_{(x,u),(y,v)}$ and we have verified the statements analogous to (4.3-5).

Proof of Claim 5.1, conclusion. For small l the bound is immediate from Hörmander's standard L^2 estimate for nondegenerate oscillatory integrals ([8], cf. (5.12) and Lemma 5.3 above). For large l we can, by Lemma 5.4, rewrite the amplitude $\mathcal{E}_j^l$ as a finite sum

$$\mathcal{E}_j^l(x,y,u,v) = 2^{2jl} \sum_{|i| \le C} \zeta_1(2^{l+i} \det \Phi''_{(x,u,y,v)}) q_{l+i}(x,u,y,v)$$

where the q_{l+i} are compactly supported and smooth and satisfy the estimates $\partial^\alpha_{x,y,u,v} q_{l+i} = O(2^{l\alpha})$. Since $2^l \le \lambda^{1/3}$, this type of blowup is covered by (4.2) and we can apply the estimate (4.6) and see that the operator with kernel $\lambda^{-j}\mathcal{E}_j^l$ has L^2 operator norm $\lesssim 2^{2jl}\lambda^{-j}\lambda^{-(d+m)/2}2^{l/2}$. This implies our claim.

Modifications for the proof of (3.16). By scaling we need to consider the operator of convolution with $\partial_s K_s^{k,l}|_{s=1}$.

Let ϕ be as in (5.2) and

$$\begin{aligned} \rho(x', x_d, u, y, v, \sigma, \tau) &= \frac{\partial}{\partial s} \phi(\frac{x}{s}, \frac{u}{s^2}, \frac{y}{s}, \frac{v}{s^2}, \sigma, \tau)\Big|_{s=1} \\ &= \sigma(-x_d + y_d + (x' - y') \cdot \nabla_{x'} \Gamma(x' - y')) \\ &\quad + 2\sum_{i=1}^m \tau_i(-u_i + v_i - x^t J_i y) + \sum_{i=1}^m \tau_i e_i^t \Lambda(y - x). \end{aligned} \tag{5.22}$$

As before we set $\lambda = 2^k$ and observe that our operator is a sum of an operator $\mathcal{G}^{\lambda,l}$ with Schwartz kernel

$$\begin{aligned} &G^{\lambda,l}(x,u,y,v) \\ &\quad = \lambda^{m+2} \iint e^{i\lambda\phi(x,u,y,v,\sigma,\tau)} \rho(x', x_d, u, y, v, \sigma, \tau) \chi_0(x,u,y,v) \eta_l(\sigma,\tau) d\sigma d\tau \end{aligned}$$

and an operator which has similar properties as $H^{\lambda,l}$ above (thus satisfies estimates which are better than claimed in (3.16)).

We now need to carry out the stationary phase calculations as before for the kernel $\mathcal{F}_{\lambda,1}\mathcal{G}^{\lambda,l}$ (since the contribution from $\mathcal{F}_{\lambda,2}\mathcal{G}^{\lambda,l}$ is again negligible). It has the form of (5.3), except that b_l is replaced by λc_l where c_l is given by

$$c_l(x,u,y,v,\theta) = b_l(x,u,y,v,z_d,w,\sigma,\tau)\rho(x',z_d,w,y,v,\sigma,\tau).$$

Then by stationary phase the Schwartz kernel of $\mathcal{F}_{\lambda,1}\mathcal{G}^{\lambda,l}$ can be expanded as

(5.23)
$$\lambda^{m+2}\int e^{i\lambda\Psi(x,u,y,v,\theta)}c_l(x,u,y,v,\theta)d\theta = e^{i\lambda\Phi(x,u,y,v)}\sum_{j=0}^{N-1}\widetilde{\mathcal{E}}^l_j(x,u,y,v)\lambda^{1-j} + \widetilde{R}^{\lambda,l}_N(x,u,y,v)$$

where again the error term $\widetilde{R}^{\lambda,l}_N$ is easy to handle for large N and $\widetilde{\mathcal{E}}^\lambda_j$ is defined as in (5.8) but with b_j replaced by c_j.

In order to finish the proof of (3.16) it is now sufficient to establish that the operator $\mathcal{T}^{\lambda,l}_j$ with kernel $\lambda^{1-j}\widetilde{\mathcal{E}}^l_j e^{i\lambda\Phi(x,u,y,v)}$ satisfies the bound

$$\|\mathcal{T}^{\lambda,l}_j\|_{L^2\to L^2} \lesssim \lambda^{1-(d+m)/2}2^{-l/2}(1+\|\Lambda\|2^l). \tag{5.24}$$

The differentiation in s causes a blowup by not more than λ and by our previous analysis it follows that

$$\|\mathcal{T}^{\lambda,l}_j\|_{L^2\to L^2} \lesssim 2^{l/2}\lambda^{1-(d+m)/2}(2^{2l}\lambda^{-1})^j. \tag{5.25}$$

If $j=1,2,\dots$ this estimate is sufficient for (5.24), since then $2^{l/2}(2^{2l}\lambda^{-1})^j \lesssim 2^{-l/2}$ by our restriction $2^l \le \lambda^{1/3}$.

This crude estimate does not suffice for the leading term in the asymptotic expansion when $\|\Lambda\|$ is small (or zero).

However, note that when $\Lambda = 0$ the coefficient of τ_i in (5.22) vanishes on the critical set where $\theta = \theta_{\text{crit}}(x,u,y,v)$, since $\partial\Psi/\partial\tau = 0$ on that set. We get

$$\begin{aligned}\rho(x',&z_{d,\text{crit}},w_{\text{crit}},y,v,\sigma_{\text{crit}},\tau_{\text{crit}})\\ &=(x_d - e_d^tJ_uy - u^t\Lambda_d)((x'-y')\cdot\nabla_{x'}\Gamma(x'-y') - \Gamma(x'-y'))\\ &\quad + 2\sum_{i=1}^m u_i(e_i^t\Lambda P^t(x'-y') + e_i^t\Lambda_d\Gamma(x'-y')).\end{aligned}$$

Since $|x_d - e_d^tJ_uy - u^t\Lambda_d| \approx 2^{-l}$ on the support of c_l and since the coefficients of u_i are $O(\|\Lambda\|)$, we now gain an additional factor of $O(2^{-l}+\|\Lambda\|)$ in the estimate (5.25) for $j=0$ and thus establish (5.24) also for $j=0$.

MODIFICATIONS FOR THE PROOF OF (3.17), (3.18). The only reason for the modified definition (2.2.3) (replacing (2.2.2) for $l > k/3$) is the preservation of the symbol estimates (4.2), needed for the validity of (4.6), (4.7). The estimation for $\widetilde{K}^k$ is exactly analogous to the estimation of $K^{k,l}$ when $l < k/3$, and the same statement applies to the s-derivatives. Only notational modifications are needed.

6. Weak type (1,1) estimates

We are now proving the weak type inequality (2.5). The proof of (2.6) is omitted since it is exactly analogous.

We apply standard Calderón–Zygmund arguments (with respect to nonisotropic families of balls on nilpotent Lie groups, see [4], [17]). Cf. also [14] and related papers on singular Radon transforms.

Let

$$B_\delta = \{(x,u) : |x| \le \delta, |u| \le \delta^2\}$$

and denote by B_δ^c its complement.

Since we have already checked the L^2 bounds for the maximal function it suffices to check the following Hörmander-type condition for $L^\infty(\mathbb{R}^+)$ valued kernels:

$$\sup_{\delta>0}\sup_{(y,v)\in B_\delta}\int_{B^c_{10\delta}}\sup_{t>0}|K_t^{k,l}((y,v)^{-1}(x,u)) - K_t^{k,l}(x,u)|dxdu \lesssim k2^{k-l}(1+\|\Lambda\|2^l)$$

which follows from the two estimates

$$\sup_{(y,v)\in B_\delta}\int_{B^c_{10\delta}}\sup_{s\in[1,2]}|K_{2^n s}^{k,l}((y,v)^{-1}(x,u)) - K_{2^n s}^{k,l}(x,u)|dxdu \lesssim \begin{cases} 2^{k-l}(1+\|\Lambda\|2^l), \\ 2^{k(m+2)}\min\{2^{-n}\delta, 2^n\delta^{-1}\}. \end{cases}$$

Indeed we use the first bound for the $O(k)$ terms with $2^{-2k(m+1)} \le 2^{-n}\delta \le 2^{2k(m+1)}$ and the second bound for the remaining terms. We then sum the series in n. Using scaling we see that the latter estimates are equivalent to

(6.1)

$$\sup_{(y,v)\in B_r}\int_{B^c_{10r}}\sup_{s\in[1,2]}|K_s^{k,l}(x-y, u-v+x^tJy) - K_s^{k,l}(x,u)|dxdu \lesssim \begin{cases} 2^{k-l}(1+\|\Lambda\|2^l), \\ 2^{k(m+2)}\min\{r^{-1}, r\}. \end{cases}$$

Because of the support properties of the kernel the integral on the left hand side is zero if $r \gg 1$. Now assume that $r \lesssim 1$. Since $|\nabla K_s^{k,l}(x,u)| \lesssim 2^{k(m+2)}$

the bound $2^{k(m+2)}r$ in (6.1) is immediate. It remains to show that

$$\Big\| \sup_{s\in[1,2]} |K_s^{k,l}| \Big\|_1 \lesssim 2^{k-l}(1+\|\Lambda\|2^l),$$

and this follows from

$$\|K^{k,l}\|_1 \lesssim 1, \tag{6.2}$$

$$\|\partial_s K_s^{k,l}\|_1 \lesssim 2^{k-l}(1+\|\Lambda\|2^l). \tag{6.3}$$

By an integration by parts in σ, τ we see that

$$|K^{k,l}(x,u)| \le C_N \frac{2^{k-l}}{(1+2^{k-l}|x_d-\Gamma(x')|)^N} \frac{2^{km}}{(1+2^k|u-\Lambda x|)^N} \tag{6.4}$$

from which (6.2) immediately follows. Moreover, from (5.22) one obtains by the same argument $|\partial_s K_s^{k,l}(x,u)|$ is bounded by $C'_N 2^{k-l}(1+\|\Lambda\|2^l)$ times the right hand side of (6.4). Consequently we obtain (6.3). This finishes the proof of the weak type inequality (2.5). ∎

7. Appendix

In this section we give the example of a two-step nilpotent Lie group G, with 10-dimensional Lie algebra, which satisfies the nondegeneracy condition but which is not isomorphic to a group of Heisenberg type.

For $\mu=(\mu_1,\mu_2)\in\mathbb{R}^2$ let

$$E_\mu = \begin{pmatrix} \mu_1 & 0 & 0 & -\mu_2 \\ \mu_2 & \mu_1 & 0 & 0 \\ 0 & \mu_2 & \mu_1 & 0 \\ 0 & 0 & \mu_2 & \mu_1 \end{pmatrix}$$

and define the 8×8 matrix

$$J_\mu = \begin{pmatrix} 0 & E_\mu \\ -E_\mu^t & 0 \end{pmatrix};$$

then

$$\det J_\mu = (\mu_1^4+\mu_2^4)^2. \tag{7.1}$$

Let $\mathfrak{g}$ be the Lie algebra which is $\mathbb{R}^8\oplus\mathbb{R}^2$ as a vector space, with Lie bracket

$$[X+U, Y+V] = 0 + (X^t J_{(1,0)} Y, X^t J_{(0,1)} Y).$$

By (7.1) the group identified with $\mathfrak{g}$ satisfies our nondegeneracy condition. We now prove by contradiction that $\mathfrak{g}$ is not isomorphic to a Heisenberg-type Lie algebra.

Assume that there is a Lie algebra isomorphism α: $\widetilde{\mathfrak{g}} \to \mathfrak{g}$ where $\widetilde{\mathfrak{g}}$ is a Heisenberg-type algebra. Then $\widetilde{\mathfrak{g}} = \mathfrak{w} \oplus \mathfrak{z}$ where $\mathfrak{z}$ is the center and α is a linear isomorphism from $\mathfrak{z}$ to $\mathbb{R}^2$.

Now with respect to orthonormal bases $u_1, \dots, u_8$ on $\mathfrak{w}$ and u_9, u_{10} on $\mathfrak{z}$ and $e_1, \dots, e_8$ on $\mathbb{R}^8$ and e_9, e_{10} on $\mathbb{R}^2$ the map α is given by the 10×10 matrix

$$\begin{pmatrix} A & 0 \\ L & B \end{pmatrix}$$

where A is an invertible 8×8 matrix and B an invertible 2×2 matrix.

Now let $X = \sum_{i=1}^{8} x_i u_i$, $Y = \sum_{i=1}^{8} y_i u_i$, and express $\omega \in \mathfrak{z}^*$ in terms of the dual basis as $\omega = w_1 u_9^* + w_2 u_{10}^*$. Then, since $\widetilde{\mathfrak{g}}$ is of Heisenberg type we have $\omega([X,Y]) = x^t \widetilde{J}_w y$ with $\widetilde{J}_w^2 = -(w_1^2 + w_2^2)I$; in particular

$$|\det \widetilde{J}_w| = (w_1^2 + w_2^2)^4. \tag{7.2}$$

Now if $\omega = \alpha^t \mu$ (thus $B^t \mu = (w_1, w_2)^t$) then

$$x^t \widetilde{J}_{B^t\mu} y = \omega([X,Y]) = (\alpha^t)^{-1}\omega(\alpha[X,Y]) = \langle \mu, [\alpha X, \alpha Y] \rangle = (Ax)^t J_\mu (Ay)$$

so that $A^t J_\mu A = \widetilde{J}_{B^t\mu}$ and therefore

$$\det \widetilde{J}_{B^t\mu} = (\det A)^2 \det J_\mu.$$

Thus by (7.1) and (7.2) we obtain $|B^t\mu|^8 = (\det A)^2(\mu_1^4 + \mu_2^4)^2$ and therefore, if (a,b) and (c,d) are the rows of the matrix $|\det A|^{-1/4} B^t$,

$$\mu_1^4 + \mu_2^4 = ((a\mu_1 + b\mu_2)^2 + (c\mu_1 + d\mu_2)^2)^2,$$

for all $\mu \in \mathbb{R}^2$. Thus

$$\mu_1^4 + \mu_2^4 = ((a^2 + c^2)\mu_1^2 + (b^2 + d^2)\mu_2^2 + 2(ab + cd)\mu_1\mu_2))^2$$

for all $\mu \in \mathbb{R}^2$. This implies $a^2 + c^2 = b^2 + d^2 = 1$ and setting $\rho = ab + cd$ we obtain after a little algebra that

$$(4\rho^2 + 2)\mu_1\mu_2 + 4\rho(\mu_1^2 + \mu_2^2) = 0$$

for all $\mu \in \mathbb{R}^2$. This implies both $2\rho^2 + 1 = 0$ and $\rho = 0$, thus a contradiction. ∎

References

[1] J. Bourgain, *Averages in the plane over convex curves and maximal operators,* Journal d'Analyse Mathématique **47** (1986), 69–85.

[2] M. Cowling, *On Littlewood-Paley-Stein theory,* Rendiconti del Circolo Matematico di Palermo **1** (1981), 21–55.

[3] S. Cuccagna, *L^2 estimates for averaging operators along curves with two-sided k-fold singularities,* Duke Mathematical Journal **89** (1997), 203–216.

[4] G. Folland and E. M. Stein, *Hardy Spaces on Homogeneous Groups,* Princeton University Press, 1982.

[5] A. Greenleaf and A. Seeger, *Fourier integral operators with fold singularities,* Journal für die reine und angewandte Mathematik **455** (1994), 35–56.

[6] A. Greenleaf and A. Seeger, *On oscillatory integrals with folding canonical relations,* Studia Mathematica **132** (1999), 125–139.

[7] L. Hörmander, *The Analysis of Linear Partial Differential Operators, Vol. I,* Springer-Verlag, New York, Berlin, 1983.

[8] L. Hörmander, *Oscillatory integrals and multipliers on FL^p,* Archiv der Mathematik **11** (1973), 1–11.

[9] A. Kaplan, *Fundamental solutions for a class of hypoelliptic PDE generated by composition of quadratic forms,* Transactions of the American Mathematical Society **258** (1980), 147–153.

[10] G. Métivier, *Hypoellipticité analytique sur des groupes nilpotents de rang* 2, Duke Mathematical Journal **47** (1980), 195–221.

[11] G. Mockenhaupt, A. Seeger and C. D. Sogge, *Local smoothing of Fourier integral operators and Carleson–Sjölin estimates,* Journal of the American Mathematical Society **6** (1993), 65–130.

[12] A. Nevo and S. Thangavelu, *Pointwise ergodic theorems for radial averages on the Heisenberg group,* Advances in Mathematics **127** (1997), 307–334.

[13] D. H. Phong and E. M. Stein, *Radon transforms and torsion,* International Mathematics Research Notices, 1991, pp. 49–60.

[14] F. Ricci and E. M. Stein, *Harmonic analysis on nilpotent groups and singular integrals II: Singular kernels supported on submanifolds,* Journal of Functional Analysis **78** (1988), 56-84.

[15] O. Schmidt, *Maximaloperatoren zu Hyperflächen in Gruppen vom homogenen Typ,* Diplomarbeit, Universität Kiel, 1998.

[16] E. M. Stein, *Maximal functions: spherical means,* Proceedings of the National Academy of Sciences of the United States of America **73** (1976), 2174–2175.

[17] E. M. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality and Oscillatory Integrals,* Princeton University Press, 1993.