SINGULAR RADON TRANSFORMS AND MAXIMAL FUNCTIONS UNDER CONVEXITY ASSUMPTIONS

ANDREAS SEEGER AND STEPHEN WAINGER

ABSTRACT. We prove variable coefficient analogues of results in [5] on Hilbert transforms and maximal functions along convex curves in the plane.

1. Introduction

The purpose of this paper is to prove L^p boundedness results on singular Radon transforms and maximal operators for variable curves in the plane. We shall prove a diffeomorphism invariant extension of the result for translation invariant averages along along convex curves in [5].

To fix our notation let Ω_0, Ω_1, Ω be open sets in $\mathbb{R}^2$ with compact closure, so that $\Omega \subset\subset \Omega_1 \subset\subset \Omega_0$. We assume that for each $x \in \Omega_0$ we are given a curve

$$(1.1) \qquad t \mapsto \Gamma(x,t), \qquad -c_0 \le t \le c_0$$

so that $\Gamma(x,t) \in \Omega_0$ for all x in a neighborhood of the closure of Ω_1 and all $t \in [-c_0, c_0]$. Furthermore assume that Γ satisfies

$$(1.2) \qquad \Gamma(x,0) = x,$$

for all $x \in \Omega_0$. We denote by $\dot{\Gamma}(x,t)$ the t-derivative of Γ and assume that $\dot{\Gamma}$ is an L^∞ function, and that Γ and $\dot{\Gamma}$ depend smoothly on x. We shall assume that for $|t| \le c_0$ the map $x \mapsto \Gamma(x,t)$ is a diffeomorphism on a neighborhood of Ω_1 (for small t this is of course implied by (1.2)). The inverse is denoted by Γ^*; thus $x = \Gamma^*(y,t)$ iff $y = \Gamma(x,t)$.

The two operators under consideration are the maximal operator

$$(1.3) \qquad \mathcal{M}f(x) = \sup_{0 < h < \varepsilon} \frac{1}{2h} \int_{-h}^{h} |f(\Gamma(x,t))| dt$$

and the singular Radon transform

$$(1.4) \qquad \mathcal{R}f(x) = \text{p.v.} \int \omega(x,t) f(\Gamma(x,t)) \frac{dt}{t}$$

where ω is a C_0^∞ function supported in $\Omega_0 \times [-\varepsilon, \varepsilon]$. Here $\varepsilon \le c_0$. Since $\dot{\Gamma}$ is bounded it is not hard to see that for $f \in C^1$ the principal value integral (1.2) is well defined. Our task will be to show that under suitable assumptions the operators $\mathcal{M}$ and $\mathcal{R}$ are L^p bounded. We observe that it suffices to

---

prove L^p estimates under the assumption that $\varepsilon \ll c_0$ as the contribution for t bounded away from 0 is easy to handle.

As we are seeking to generalize the result in [5] we wish to make two assumptions on Γ, namely a *convexity* hypothesis and a *doubling* hypothesis. Since we consider a variable situation our assumptions ought to be invariant under changes of variables (and the usual assumptions of convexity fail to meet this requirement).

In order to introduce an invariant convexity assumption we follow [23] and say that a function h defined on an interval J is *quasi-monotonic* on J if there is a constant $\kappa \geq 0$ so that $h'(t) = a(t) + E(t)$ for $t \in I$ where a has constant sign in I and $|E(t)| \leq \kappa |h(t)|$ (typically h is monotonic modulo a function in the ideal generated by h). A family of functions is *uniformly quasi-monotonic* if in the latter inequality we can choose a universal κ.

The relevant quantities are

$$(1.5) \qquad G(x,t) = \det \left(\dot{\Gamma}(x,t) \quad \dot{\Gamma}^*(w,0) \right)_{w=\Gamma(x,t)}$$

$$(1.6) \qquad G^*(y,t) = \det \left(\dot{\Gamma}^*(y,t) \quad \dot{\Gamma}(z,0) \right)_{z=\Gamma^*(y,t)}$$

We now make the following

Convexity Hypothesis (C.H.). *For all $x \in \Omega_1$, $y \in \Omega_1$ the functions $G(x,\cdot)$ and $G^*(y,\cdot)$ are uniformly quasi-monotonic on $[0,c_0]$ and on $[-c_0,0]$.*

We turn to our doubling hypothesis. We say that a non-negative continuous function g on $[0,c_0]$ is a *doubling function* if $g(0) = 0$, $g(t) > 0$ for $t > 0$ and if there is $A \geq 1$ so that

$$(1.7) \qquad g(t_2) \geq 2g(t_1) \quad \text{if } t_2 \geq At_1.$$

An immediate consequence is that

$$(1.8) \qquad g(t_1) \lesssim (t_1/t_2)^\delta g(t_2), \qquad t_1 \leq A^{-1}t_2, \, t_2 \leq c_0,$$

for some $\delta > 0$.

Doubling Hypothesis (D.H.). *There is $C_0 \geq 1$ and a doubling function g on $[0,c_0]$ so that*

$$(1.9) \qquad C_0^{-1}g(A^{-1}|t|) \leq |G(x,t)| \leq C_0 g(A|t|)$$

and

$$(1.10) \qquad C_0^{-1}g(A^{-1}|t|) \leq |G^*(y,t)| \leq C_0 g(A|t|).$$

for all $x \in \Omega_1$, $y \in \Omega_1$ and $|t| \leq c_0$.

In particular the inequality (1.8) holds for $G(x,\cdot)$ and $G^*(y,\cdot)$ if $t_1 \leq A^{-3}t_2$, $t_2 \leq c_0$.

We can now formulate our main result.

Theorem A. *If the convexity hypothesis (C.H.) and the doubling hypothesis (D.H.) are satisfied then $\mathcal{M}$ is bounded from L^p to $L^p(\Omega)$, for $p > 1$; moreover $\mathcal{H}$ is bounded from L^p to $L^p(\Omega)$ for $1 < p < \infty$.*

Under very general finite type conditions the L^p boundedness of $\mathcal{M}$ and $\mathcal{H}$ has been proved by Christ, Nagel, Stein and Wainger [7] (see also Greenblatt [11]). Thus we are mainly interested in the

flat case. The translation invariant model case of the theorem (where $\Gamma(x,t) = (x_1 + t, x_2 + u(t))$, with u convex) was obtained in Carlsson *et al.* [5] (*cf.* also [9]); the special case $p = 2$ goes back to [19], [20] and in [19] it was also shown that our condition is necessary when u is an even function. See also [9] for a necessary condition in the general case. In the 'semi-translation invariant' case where $\Gamma(x,t) = (x_1 + t, x_2 + s(x_1,t))$ the L^2 result had been obtained by one of the authors in [23]. L^p theorems in somewhat different variable coefficient settings are in [3], [4] and in [2]. More closely related to the setting here is the recent paper by Carbery and Pérez [1] who proved L^p bounds for the semi-translation-invariant case under more restrictive third order assumptions. Optimal results on the Heisenberg group related to Theorem A were obtained by J. Kim [15], [16].

Invariance properties and alternative formulations. The main feature of hypotheses (C.H.) and (D.H.) is the invariance under diffeomorphisms. This is easy to check. Namely if $y = \Gamma(x,t)$, and $y = \Phi(z)$, $x = \Phi(u)$, then $z = \widetilde{\Gamma}(u,t)$ with $\widetilde{\Gamma}(u,t) = \Phi^{-1}(\Gamma(\Phi(u),t))$; moreover $\widetilde{\Gamma}^*(u,t) = \Phi^{-1}(\Gamma^*(\Phi(u),t))$. Hence we get $\dot{\widetilde{\Gamma}}(u,t) = D\Phi^{-1}_{\Gamma(\Phi(u),t)}\dot{\Gamma}(\Phi(u),t)$, and similarly $\dot{\widetilde{\Gamma}}^*(w,0) = D\Phi^{-1}_{\Gamma^*(\Phi(w),0)}\dot{\Gamma}^*(\Phi(w),0)$. The latter we apply for $w = \widetilde{\Gamma}(u,t)$ and notice that $\Gamma^*(\Phi(\widetilde{\Gamma}(u,t)),0) = \Gamma^*(\Gamma(\Phi(u),t),0) = \Gamma(\Phi(u),t)$. Now let $\widetilde{G}$ denote the determinant (1.5) corresponding to the curve $\widetilde{\Gamma}$; then we obtain

$$\widetilde{G}(\Phi(u),t) = \det(D\Phi^{-1}(\Gamma(\Phi(u),t))) \det\left(\dot{\Gamma}(\Phi(u),t) \quad \dot{\Gamma}^*(\Gamma(\Phi(u),t),0)\right).$$

A similar calculation applies to (1.6). From this the invariance property easily follows, with the possible change of the constants A, C_0 (see also the discussion in [23]).

We also note the our assumptions do not depend on the particular parametrization. If $t = u(x,s)$ with $u_s \neq 0$ we have $\partial_s(\Gamma(x,u(x,s))) = u_s(x,s)\dot{\Gamma}(x,u(x,s))$ and the independence of the parametrization is easily verified.

Our hypotheses can also be described in terms of defining functions as in [21], [24]. Namely let $\Sigma = \{(x,y) : y = \Gamma(x,t), \text{ some } t\}$ then if we restrict to small values of t the variety Σ is a smooth hypersurface in $\Omega \times \Omega$ and $\Sigma = \{(x,y) : \Psi(x,y) = 0\}$ where $\Psi'_x \neq 0$ and $\Psi'_y \neq 0$. Our quasimonotonicity and doubling assumptions may be replaced by similar assumptions on the functions

$$(1.11) \qquad\qquad t \mapsto \det(\Psi_y(x,y), \Psi_y(y,y))\big|_{y=\Gamma(x,t)}$$

$$(1.12) \qquad\qquad t \mapsto \det(\Psi_x(x,y), \Psi_x(x,x))\big|_{x=\Gamma^*(y,t)}.$$

If $N^*\Sigma \subset (T^*_L\Omega \setminus 0) \times (T^*_R\Omega \setminus 0)$ denotes the conormal bundle of Σ then

$$N^*\Sigma = \{(x,\xi,y,\eta) : \xi = \tau\Psi'_x, \eta = \tau\Psi'_y, \tau \neq 0, \Psi(x,y) = 0\}$$

and assumptions on (1.11-12) reflect properties of the projections of $N^*\Sigma$ to $T^*_L\Omega$ and $T^*_R\Omega$.

In order to see that the conditions involving (1.11-12) are equivalent to the conditions involving (1.5-6) we first observe that the conditions for (1.11-12) are invariant under changes of variables, moreover they do not depend on the particular choice of defining function. By the above discussion we may without loss of generality assume that

$$(1.13) \qquad\qquad \Gamma(x,t) = (x_1 - t, \gamma(x_1, x_2, t)).$$

Then $\Gamma^*(y,t) = (y_1 + t, \gamma^*(y_1, y_2, t))$ where $\gamma(x_1, x_2, 0) = x_2$, $\gamma^*(y_1, y_2, 0) = y_2$, $\frac{\partial\gamma}{\partial x_2} \neq 0$ and $\frac{\partial\gamma^*}{\partial y_2} \neq 0$. In fact $\frac{\partial\gamma}{\partial x_2}(x_1, x_2, 0) = 1$ and $\frac{\partial\gamma^*}{\partial y_2}(y_1, y_2, 0) = 1$. The equivalence is now obtained by working with

3

the defining functions $\Psi(x,y) = y_2 - \gamma(x_1, x_2, x_1 - y_1)$ or $\widetilde{\Psi}(x,y) = x_2 - \gamma^*(y_1, y_2, y_1 - x_1)$. These are both defining functions and they are related by

$$(1.14) \qquad y_2 - \gamma(x, x_1 - y_1) = \mathfrak{a}(x,y)(x_2 - \gamma^*(y, y_1 - x_1))$$

where

$$\mathfrak{a}(x,y) = \int_0^1 \frac{\partial \gamma}{\partial x_2}(x_1, (1-s)\gamma^*(y, y_1 - x_1) + sx_2, x_1 - y_1)ds.$$

To see this expand $y_2 - \gamma(x, x_1 - y_1)$ about $x_2 = \gamma^*(y, y_1 - x_1)$ and use that

$$y_2 = \gamma(x_1, \gamma^*(y, y_1 - x_1), x_1 - y_1).$$

Note that if ε is chosen small enough we can assume that

$$(1.15) \qquad |\mathfrak{a}(x,y) - 1| \le 1/2 \qquad \text{if } (x,y) \in \Omega_1 \times \Omega_1, \ |x_1 - y_1| \le \varepsilon.$$

For later reference we state that the boundedness of γ_{x_2} and $\nabla \gamma_{x_2}$ (as assumed in Theorem B below) imply that $\mathfrak{a}$ has a bounded gradient.

A change of variable. The invariance under changes of variables allows us to to make a crucial choice of coordinates in order to reduce the situation (1.13) with the additional normalization $\dot{\gamma}(x,0) = 0$. A related change of coordinates was suggested years ago by C. Fefferman, in connection with the problem of differentiation along variable lines. A similar argument was also used in [25].

We set $\Phi(u_1, u_2) = (u_1, \rho(u_1, u_2))$ where the smooth function ρ is to be determined and will satisfy $\rho(0, u_2) = u_2$. This also implies that for small u_1 the function $u_2 \mapsto \rho(u_1, u_2)$ is invertible, with inverse σ, so that $\sigma(u_1, \rho(u_1, u_2)) = u_2$. Now suppose that are already given Φ and we would then have

$$(1.16) \qquad \Phi^{-1}\Gamma(\Phi(u), t) = (u_1 - t, \sigma(u_1 - t, \gamma(u_1, \rho(u_1, u_2), t))).$$

Thus we need to take $\rho(\cdot, u_2)$ to satisfy the equation

$$(1.17) \qquad -\sigma_{x_1}(u_1, \gamma(u_1, \rho(u_1, u_2), 0)) + \sigma_{x_2}(u_1, \gamma(u_1, \rho(u_1, u_2), 0))\dot{\gamma}(u_1, \rho(u_1, u_2), 0) = 0.$$

Now $\gamma(u_1, \rho, 0) = \rho$ and $\sigma_{x_1}(u_1, \rho(u_1, u_2)) + \sigma_{x_2}(u_1, \rho(u_1, u_2))\rho_{u_1}(u_1, u_2) = 0$, and thus (1.17) is implied by $\sigma_{x_2} \ne 0$ and

$$(1.18) \qquad \rho_{u_1}(u_1, u_2) + \dot{\gamma}(u_1, \rho(u_1, u_2), 0) = 0.$$

Thus if we solve the ordinary differential equation (1.18), with parameter u_2, under the initial value condition $\rho(0, u_2) = u_2$ then we have $\rho_{u_2} \ne 0$ and thus $\sigma_{x_2}(u_1, \rho(u)) \ne 0$ for small u_1 and therefore $\widetilde{\gamma}(u, t) = \sigma(u_1 - t, \gamma(u_1, \rho(u_1, u_2), t))$ will satisfy $\dot{\widetilde{\gamma}}(u, 0) = 0$.

From now on we may and shall work with families of curves defined by (1.13) which also satisfy

$$(1.19) \qquad \dot{\gamma}(x, 0) = 0.$$

By implicit differentiation it also follows that

$$(1.20) \qquad \dot{\gamma}^*(y_1, y_2, 0) = 0.$$

In this situation our convexity hypothesis simplifies to

$$(1.21) \qquad \ddot{\gamma}(x, t) = a(t, x) + O(\dot{\gamma}(x, t)),$$

$$(1.22) \qquad \ddot{\gamma}^*(y, t) = a^*(t, y) + O(\dot{\gamma}^*(y, t)),$$

where $a(x, \cdot)$ and $a^*(y, \cdot)$ are of constant sign for $t > 0$ and of constant sign for $t < 0$. Our doubling hypothesis becomes

$$(1.23) \qquad C_0^{-1}g(A^{-1}t) \le |\dot{\gamma}(x, t)| \le C_0 g(A|t|),$$

$$(1.24) \qquad C_0^{-1}g(A^{-1}t) \le |\dot{\gamma}^*(y, t)| \le C_0 g(A|t|)$$

for some doubling function g and suitable constant $A \ge 1$.

We then have the following result:

4

Theorem B. *Assume that γ and γ^* satisfy the hypotheses (1.19-24). Also suppose $|\partial_{x_2}\gamma(x,t)| \geq c_1 > 0$. Then $\mathcal{M}$ is bounded from L^p to $L^p(\Omega)$, for $p > 1$, and $\mathcal{H}$ is bounded from L^p to $L^p(\Omega)$. The operator norms depend only on the cutoff function ω, the doubling function g, the constants A in (1.23-24) and the L^∞ norms of γ_{x_2}, $\nabla\gamma_{x_2}$, $\gamma^*_{y_2}$, $\nabla\gamma^*_{y_2}$.*

With the change of variables discussed above, Theorem B implies Theorem A.

Remark. Note that the operator norms do not explicitly depend on the L^1 norm of $\ddot\gamma$. Thus by limiting arguments Theorem B covers examples such as $\gamma(x,t) = x_2 + u(t)$ where u is even or odd, continuous, linear on $(2^{-j}, 2^{-j+1})$ with $u(2^{-j}) = 2^{-mj}$, as well as variable perturbations.

The organization of the paper is as follows. In Section 2 we introduce some notation and make a preliminary Littlewood Paley decomposition of our operators; moreover we prove the L^p estimates for the 'Calderón-Zygmund part' of the operator. In section 3 we give an outline of the proof of Theorem B, and handle the technical details of the main error estimate in §4.

2. Preliminary decompositions and Calderón-Zygmund estimates

Let $\phi \in C_0^\infty(\mathbb{R})$ be supported in $(1/2, 2) \cup (-2, -1/2)$ and define $\phi_j(s) = 2^j\phi(2^j s)$ for $j > 0$.

Also let $\chi \in C_0^\infty(\Omega \times \Omega)$, $\widetilde\chi$, $\widetilde\phi$ nonnegative so that $0 \leq |\chi| \leq \widetilde\chi$, $0 \leq |\phi| \leq \widetilde\phi$ and $\widetilde\phi_j = 2^j\widetilde\phi(2^j\cdot)$. Define

$$
\begin{aligned}
\mathcal{R}_j f(x) &= \int \phi_j(x_1 - y_1)\chi(x, y_1, x_2 + \gamma(x, x_1 - y_1))f(y_1, x_2 + \gamma(x, x_1 - y_1))dy_1 \\
\mathcal{M}_j f(x) &= \int \widetilde\phi_j(x_1 - y_1)\widetilde\chi(x, y_1, x_2 + \gamma(x, x_1 - y_1))f(y_1, x_2 + \gamma(x, x_1 - y_1))dy_1;
\end{aligned}
$$
(2.1)

here we want to explicitly include the case that the functions $\widetilde\chi$ and χ and the functions $\widetilde\phi$ and ϕ coincide (and are nonnegative), so that some aspects of $\mathcal{R}_j$ and $\mathcal{M}_j$ can be treated at the same time.

The L^p inequality for the maximal function in (1.3) is a simple consequence of the L^p boundedness of the maximal operator $\mathcal{M}$ defined with slight abuse of notation by

$$\mathcal{M}f(x) = \sup_{j \in J}|\mathcal{M}_j f(x)|;$$

here J is a finite set of integers $j > C$ and the bound is not supposed to depend on the cardinality of J. Working with suitable positive cutoff functions we obtain uniform bounds for $\mathcal{M}$ from uniform bounds for the maximal function

$$\sup_{j \in J}|\mathcal{R}_j f|.$$

Notice that since every individual operator $\mathcal{R}_j$ is bounded on L^1 and L^∞ we need to take the supremum over large j only. Similarly the boundedness of the Hilbert transform follows from the L^p boundedness of the operator $\sum_{j>C}\mathcal{R}_j f$ under the additional assumption that the cutoff function satisfies $\int \phi(s)ds = 0$; indeed we can choose ϕ such that $\sum_{j=-\infty}^\infty \phi_j(s) = 1/s$.

Denote by δ_0 the Dirac measure on the real line, at the origin. Following [21] we express $\delta_0(y_2 - \gamma(x,t))$ as an oscillatory integral distribution using the Fourier inversion formula,

$$\delta_0(y_2 - \gamma(x,t)) = (2\pi)^{-1}\int e^{i\tau(y_2 - \gamma(x,t))}d\tau,$$

and then decompose the singular integral operator as in [23] into two parts, a low frequency part where the cancellation of ϕ is crucially used, and a high frequency part where this cancellation

does not play a role. See also [18], [12], [22] for earlier variants of this approach. The analogous decomposition is made for the maximal operator where of course no cancellation of ϕ is needed.

In order to proceed with this decomposition we set $B = 2^{20}A$ where A is the constant in (1.7-10) and define integers a_j, b_j so that

(2.2)
$$2^{-a_j-1} < 2^{-j}g(2^{-5-j}B) \leq 2^{-a_j}$$
$$2^{-b_j-1} < 2^{-j}g(2^{5-j}B^{-1}) \leq 2^{-b_j}$$

For later reference we note that for $2^{j-k} \leq (4A)^{-1}$ we have $2^{a_j-a_k} \lesssim 2^{j-k}$; this does not use the full strength of the doubling assumption as it follows from (1.8) with $\delta = 0$. By the doubling assumption the former estimate can be improved to $2^{a_j-a_k} \lesssim 2^{(j-k)(1+\delta)}$, for some $\delta > 0$.

Let β_0 be an even function in $C_0^\infty(\mathbb{R})$ so that $\beta_0(s) = 1$ if $|s| \leq 1/2$ and $\beta_0(s) = 0$ if $|s| \geq 3/4$ and let $\beta_k(s) = \beta_0(2^{-k}s) - \beta_0(2^{-k+1}s)$, for $k \geq 1$, so that $\beta_k(s)$ can be nonzero for $k \geq 1$ only when $2^{k-1} < s < 2^{k+1}$. Clearly we have $\sum_{k=0}^\infty \beta_k(s) \equiv 1$.

For $k > 1$ we define operators $\mathcal{R}_j^k$ with distribution kernel

$$R_j^k(x,y) = \chi(x,y)\phi_j(x_1 - y_1)\int e^{i\tau(y_2 - \gamma(x,x_1-y_1))}\beta_k(\tau)d\tau$$

as well as operators $\mathcal{H}_j$ with distribution kernel

$$H_j(x,y) = \chi(x,y)\phi_j(x_1 - y_1)\int e^{i\tau(y_2 - \gamma(x,x_1-y_1))}\beta_0(2^{-a_j}\tau)d\tau;$$

then our basic decomposition is given by

(2.3)
$$2\pi\mathcal{R}_j = \mathcal{H}_j + \sum_{k>a_j}\mathcal{R}_j^k$$

In the remainder of this section we shall first prepare further the term $\mathcal{R}_j^k$ which for $k > a_j$ will later be treated as a piece of a singular Fourier integral operator and then deal with the contribution $\sum_j \mathcal{H}_j$ (or the associated maximal function) which corresponds to a kind of Calderón-Zygmund operator.

In (2.3) the decomposition in k corresponds essentially to a Littlewood-Paley decomposition in the variable dual to x_2. To make this precise we introduce a Littlewood-Paley operator $\mathcal{L}^k$ defined by

(2.4)
$$\widehat{\mathcal{L}^k f}(\xi) = \left[\beta_0(2^{-k-10}\xi_2) - \beta_0(2^{-k+10}\xi_2)\right]\widehat{f}(\xi)$$

so that the multiplier is supported where $2^{k-10} \leq |\xi_2| \leq 2^{k+11}$ and equals 1 on $2^{k-9} \leq |\xi_2| \leq 2^{k+9}$.

Lemma 2.1. *For $1 \leq p \leq \infty$*
$$\|\mathcal{R}_j^k - \mathcal{L}^k\mathcal{R}_j^k\mathcal{L}^k\|_{L^p \to L^p} \lesssim 2^{-k}$$

Lemma 2.1 tells us that for $k \geq a_j$ we may replace the operators $\mathcal{R}_j^k$ by $\mathcal{L}^k\mathcal{R}_j^k\mathcal{L}^k$ since

(2.5)
$$\sum_j\sum_{k \geq a_j}\|\mathcal{R}_j^k - \mathcal{L}^k\mathcal{R}_j^k\mathcal{L}^k\|_{L^p \to L^p} \lesssim \sum_j\sum_{k \geq a_j}2^{-k} \lesssim \sum_j 2^{-a_j} \lesssim 1.$$

Proof of Lemma 2.1. We write $\mathcal{R}_j^k - \mathcal{L}^k\mathcal{R}_j^k\mathcal{L}^k = \mathcal{R}_j^k(I - \mathcal{L}^k) + (I - \mathcal{L}^k)\mathcal{R}_j^k\mathcal{L}^k$. Thus it suffices to show that $\mathcal{R}_j^k(I - \mathcal{L}^k)$ and $(I - \mathcal{L}^k)\mathcal{R}_j^k$ satisfy the asserted bounds. Let $\mathcal{P}^{l,2}$ be the convolution

6

operator with Fourier multiplier $\beta_0(2^{-l}\xi_2)$ and let $\mathcal{Q}^{l,2}$ be the convolution operator with multiplier $\beta_1(2^{-l}\xi_2)$. Then by the support properties the symbols we have

$$I - \mathcal{L}^k = (I - \mathcal{L}^k)(\mathcal{P}^{k-5,2} + \sum_{l \geq k+5} \mathcal{Q}^{l,2})$$

and consequently it suffices to show that

(2.6.1) $$\|\mathcal{R}_j^k \mathcal{Q}^{l,2}\|_{L^p \to L^p} + \|\mathcal{Q}^{l,2}\mathcal{R}_j^k\|_{L^p \to L^p} \lesssim 2^{-l} \text{ if } l \geq k+5,$$

(2.6.2) $$\|\mathcal{R}_j^k \mathcal{P}^{k-5,2}\|_{L^p \to L^p} + \|\mathcal{P}^{k-5,2}\mathcal{R}_j^k\|_{L^p \to L^p} \lesssim 2^{-k}.$$

These estimates follow by standard integration by parts arguments (see [13]). For the sake of completeness we include the argument. We first estimate the kernel of the operator $\mathcal{R}_j^k \mathcal{Q}^{l,2}$ which is given by

$$K_{kl}(x,z) = \phi_j(x_1 - z_1) \iiint \chi(x, z_1, y_2) e^{i(\tau(y_2 - \gamma(x, x_1 - z_1)) + \xi_2(y_2 - z_2))} \beta_k(\tau)\beta_l(\xi_2) dy_2 d\tau d\xi_2.$$

Note that on the support of the symbol we have $|\xi_2 + \tau| \approx 2^l$. We integrate by parts once with respect to y_2 and then we integrate by parts with respect to τ and ξ_2. This yields the bound

$$|K_{kl}(x,z)| \lesssim |\phi_j(x_1 - z_1)| 2^{-l} \int \frac{2^k}{(1 + 2^k|y_2 - \gamma(x, x_1 - z_1)|)^N} \frac{2^l}{(1 + 2^l|y_2 - z_2|)^N} dy_2$$

and integration with respect to z yields that $\int |K_{kl}(x,z)|dz \lesssim 2^{-l}$ uniformly in x. If we take into account (1.15) then we also get that $\int |K_{kl}(x,z)|dx \lesssim 2^{-l}$ uniformly in z and the asserted bound (2.6.1) for $\mathcal{R}_j^k \mathcal{Q}^{l,2}$ follows. The proof of (2.6.2) for $\mathcal{R}_j^k \mathcal{P}^{k-5,2}$ is the same.

Next we examine the kernel of the operator $\mathcal{Q}^{l,2}\mathcal{R}_j^k$ which is given by

$$\widetilde{K}_{kl}(x,z) = \phi_j(x_1 - z_1) \iiint \chi(x_1, w_2, z) e^{i(\xi_2(x_2 - w_2) + \tau(z_2 - \gamma(x_1, w_2, x_1 - z_1)))} \beta_k(\tau)\beta_l(\xi_2) dw_2 d\tau d\xi_2.$$

The difference is now the nonlinear dependence on the phase in w_2. To remove this potential difficulty we may again invoke (1.15) and change variables in the oscillatory integral to $\sigma = \tau\mathfrak{a}(x_1, w_2, z)$. Thus we get

$$\widetilde{K}_{kl}(x,z) =$$
$$\phi_j(x_1 - z_1) \iiint \frac{\chi(x_1, w_2, z)}{\mathfrak{a}(x_1, w_2, z)} e^{i(\xi_2(x_2 - w_2) + \sigma(w_2 - \gamma^*(z, x_1 - z_1)))} \beta_k\left(\frac{\sigma}{\mathfrak{a}(x_1, w_2, z)}\right) \beta_l(\xi_2) dw_2 d\tau d\xi_2.$$

With this representation the estimation of $\widetilde{K}_{kl}$ is exactly the same as for K_{kl}. Recall that $|\mathfrak{a} - 1| \leq 1/2$. In the integration by parts with respect to w_2 we shall also need the boundedness of $\partial_{x_2}\mathfrak{a}$ which is guaranteed by our assumption, *cf.* the remark following (1.15). As above we see that the $L^p \to L^p$ bound for $\mathcal{R}_j^k \mathcal{P}^{k-5,2}$ is $O(2^{-k})$. The proof of the bound $\|\mathcal{P}^{k-5,2}\mathcal{R}_j^k\|_{L^p \to L^p} = O(2^{-k})$ is the same. $\square$

Concerning the operators $\mathcal{H}_j$ we make the following simple observation (which is valid without any cancellation property).

7

Lemma 2.2. *The kernels H_j satisfy*

(2.7.1)
$$\left|\partial_{y_1}^{n_1}\partial_{y_2}^{n_2}H_j(x,y)\right| \le A_0 2^{jn_1+a_jn_2}2^{j+a_j}(1+2^j|x_1-y_1|)^{-2}(1+2^{a_j}|x_2-y_2|)^{-2},$$

(2.7.2)
$$\left|\partial_{x_1}^{n_1}\partial_{x_2}^{n_2}H_j(x,y)\right| \le A_0 2^{jn_1+a_jn_2}2^{j+a_j}(1+2^j|x_1-y_1|)^{-2}(1+2^{a_j}|x_2-y_2|)^{-2},$$

for $(n_1,n_2)=(0,0)$, $(0,1)$ or $(1,0)$.

Proof. By integration by parts we have

$$|H_j(x,y)| \lesssim 2^j\chi_{[-2^{1-j},2^{-j+1}]}(x_1-y_1)2^{a_j}(1+2^{a_j}|y_2-\gamma(x,x_1-y_1)|)^{-N}$$

Now if $|x_1-y_1| \le 2^{-j+1}$ then

$$|\gamma(x,x_1-y_1)-x_2| = |x_1-y_1|\left|\int_0^1\dot\gamma(x,s(x_1-y_1))ds\right|$$
$$\lesssim |x_1-y_1|g(A2^{-j+1}) \lesssim 2^{-a_j}$$

and thus $1+2^{a_j}|y_2-\gamma(x,x_1-y_1)| \approx 1+2^{a_j}|y_2-x_2|$ if $|x_1-y_1| \le 2^{-j+1}$. This yields the asserted estimate for $n_1=n_2=0$. The estimates for the derivatives are analogous. $\square$

Let M_{str} be the strong maximal operator (involving averages over rectangles parallel to the coordinate axes). Then the following estimate is an immediate consequence of Lemma 2.2.

Corollary 2.3. *For all $x \in \Omega$,*

(2.8)
$$\sup_j |\mathcal{H}_j f(x)| \lesssim M_{str}f(x).$$

In the case where $\int\phi(s)ds=0$ we get a bound for the sum $\sum_j\mathcal{H}_j$. As in [2] the L^p boundedness is proved by invoking the Hardy space $H^1_{prd} := H^1(\mathbb{R}\times\mathbb{R})$ defined using the two-parameter dilations (see [6], [14], [10]). Recall from [6] that operators which are bounded on H^1_{prd} and bounded on L^2 are also bounded on L^p for $1 < p < 2$.

Proposition 2.4. *Suppose that the cancellation condition $\int\phi(s)ds=0$ holds. Then the operators $\sum_j\mathcal{H}_j$ and $\sum_j\mathcal{H}_j^*$ are both bounded on L^2 and on H^1_{prd}, and consequently on L^p, for $1 < p < \infty$.*

Proof. We first show the L^2 boundedness. By the almost orthogonality lemma of Cotlar and Stein it suffices to verify

(2.9)
$$\|\mathcal{H}_j\mathcal{H}_k^*\|_{L^2\to L^2} + \|\mathcal{H}_j^*\mathcal{H}_k\|_{L^2\to L^2} \lesssim 2^{-|j-k|/2}.$$

By taking adjoints it suffices to show (2.9) for $k \ge j$ and since the operator norms of $\mathcal{H}_j$ are uniformly bounded it suffices to consider the case where $2^{k-j} \ge 2^{10}A$.

We first examine $\mathcal{H}_j\mathcal{H}_k^*$; its kernel is given by

$$\int H_j(x,z)\overline{H_k(y,z)}dz = \sum_{n\ge0}H_{jk}^n(x,y)$$

where

$$H_{jk}^n(x,y) = \iiint \phi_j(x_1 - z_1)\phi_k(y_1 - z_1)e^{i(\tau(z_2 - \gamma(x,x_1 - z_1)) - \sigma(z_2 - \gamma(y,y_1 - z_1)))}$$

(2.10)
$$\times \beta_0(2^{-a_j}\tau)\beta_0(2^{-a_k}\sigma)\beta_n(2^{-a_j}\sigma)dzd\tau d\sigma.$$

Here we used that $\sum_{n=0}^\infty \beta_n(2^{-a_j}\sigma) \equiv 1$. Observe that in view of the support properties of the symbol we have the restriction $a_j + n \le a_k + 1$.

Now let $h(z_1) = \phi_j(x_1 - z_1)e^{i(\tau(z_2 - \gamma(x,x_1 - z_1)) - \sigma(z_2 - \gamma(y,y_1 - z_1)))}$. We use the cancellation of ϕ_k to replace $h(z_1)$ in (2.10) by $h(z_1) - h(y_1) = (z_1 - y_1)\int h'(y_1 + s(z_1 - y_1))ds$; this will be relevant for small n. We write $\zeta_{j,k,n}(\sigma,\tau) = \beta_0(2^{-a_j}\tau)\beta_0(2^{-a_k}\sigma)\beta_n(2^{-a_j}\sigma)$ and obtain

$$H_{jk}^n(x,y) = \int_0^1 [I_s(x,y) + II_s(x,y) + III_s(x,y)]ds$$

where

$$I_s(x,y) = \iiint e^{i(\tau(z_2 - \gamma(x,x_1 - z_1)) - \sigma(z_2 - \gamma(y,y_1 - z_1)))}\zeta_{j,k,n}(\sigma,\tau)\chi_{j,k,1}(x,y,s)dzd\tau d\sigma,$$

$$II_s(x,y) = \iiint e^{i(\tau(z_2 - \gamma(x,x_1 - y_1 + s(y_1 - z_1))) - \sigma(z_2 - \gamma(y,y_1 - z_1)))}\zeta_{j,k,n}(\sigma,\tau)\tau\chi_{j,k,2}(x,y,s)dzd\tau d\sigma,$$

$$III_s(x,y) = \iiint e^{i(\tau(z_2 - \gamma(x,x_1 - z_1)) - \sigma(z_2 - \gamma(y,s(y_1 - z_1))))}\zeta_{j,k,n}(\sigma,\tau)\sigma\chi_{j,k,3}(x,y,s)dzd\tau d\sigma.$$

with

$$\chi_{j,k,1}(x,y,s) = (y_1 - z_1)\phi_j'(x_1 - y_1 + s(y_1 - z_1))\phi_k(y_1 - z_1)$$
$$\chi_{j,k,2}(x,y,s) = \phi_j(x_1 - z_1)\phi_k(y_1 - z_1)(y_1 - z_1)\dot\gamma(x,x_1 - y_1 + s(y_1 - z_1))$$
$$\chi_{j,k,3}(x,y,s) = \phi_j(x_1 - z_1)\phi_k(y_1 - z_1)(y_1 - z_1)\dot\gamma(y,s(y_1 - z_1)).$$

Since $|y_1 - z_1| \le 2^{-k+1}$ we then have

$$|\chi_{j,k,1}(x,y,s)| \le 2^{2j-k}$$
$$2^{a_j}|\chi_{j,k,2}(x,y,s)| \lesssim 2^{a_j}g(2^{1-j}A)2^{-k} \lesssim 2^{j-k}$$
$$2^{a_j+n}|\chi_{j,k,3}(x,y,s)| \lesssim 2^{a_j+n}g(2^{1-k}A)2^{-k} \lesssim 2^{a_j-a_k+n}.$$

These estimates are used after additional integration by parts in τ and σ. For the term I_s we obtain (taking into account the symbol properties of $\zeta_{j,k,n}$)

$$|I(x,y)| \lesssim 2^{2j-k}2^k$$
$$\times \iint_{\substack{|x_1-z_1|\le 2^{-j+1} \\ |y_1-z_1|\le 2^{-k+1}}} \frac{2^{a_j}}{(1 + 2^{a_j}|z_2 - \gamma(x,x_1 - z_1)|)^N}\frac{2^{a_j+n}}{(1 + 2^{a_j+n}|z_2 - \gamma(y,y_1 - z_1)|)^N}dz_1dz_2.$$

Observe that the integral $\int_{x_2} 2^{a_j}(1 + 2^{a_j}|z_2 - \gamma(x,x_1 - z_1)|)^{-N}dx_2$ is $O(1)$ in view of (1.15). Thus in evaluating $\int |I(x,y)|dx$, for fixed y, we perform an x_2 integration first and see that

$$\int |I(x,y)|dx \lesssim 2^{2j} \iint_{\substack{|x_1-z_1|\le 2^{-j+1} \\ |y_1-z_1|\le 2^{-k+1}}} \int \frac{2^{a_j+n}}{(1 + 2^{a_j+n}|z_2 - \gamma(y,y_1 - z_1)|)^N}dz_2\,dz_1dx_1 \lesssim 2^{j-k}$$

9

We argue similarly for the terms II_s and III_s. By integration by parts we get the pointwise estimate

$$|II_s(x,y)| \lesssim 2^{j+k} 2^{a_j - a_k}$$

$$\times \iint_{\substack{|x_1 - z_1| \le 2^{-j+1} \\ |y_1 - z_1| \le 2^{-k+1}}} \frac{2^{a_j}}{(1 + 2^{a_j}|z_2 - \gamma(x, x_1 - z_1 + s(z_1 - y_1))|)^N} \frac{2^{a_j + n}}{(1 + 2^{a_j + n}|z_2 - \gamma(y, y_1 - z_1)|)^N} dz_1 dz_2$$

and

$$|III_s(x,y)| \lesssim 2^{j+k} 2^{a_j - a_k + n}$$

$$\times \iint_{\substack{|x_1 - z_1| \le 2^{-j+1} \\ |y_1 - z_1| \le 2^{-k+1}}} \frac{2^{a_j}}{(1 + 2^{a_j}|z_2 - \gamma(x, x_1 - z_1)|)^N} \frac{2^{a_j + n}}{(1 + 2^{a_j + n}|z_2 - \gamma(y, s(y_1 - z_1))|)^N} dz_1 dz_2.$$

Since $2^{a_j - a_k} \lesssim 2^{j-k}$ we obtain the same bound $O(2^{j-k})$ for $\int |II_s(x,y)| dx$ as above, similarly for $\int |III_s(x,y)| dx$ we obtain the bound $O(2^{a_j - a_k + n})$ which is $O(2^{j-k+n})$. Thus

$$(2.11) \qquad \int |H_{jk}^n(x,y)| dx \le \int_0^1 \int \Big[|I_s| + |II_s| + |III_s| \Big] dx \, ds \lesssim 2^{j-k+n}.$$

The same bound is obtained for $\int |H_{jk}^n(x,y)| dy$, uniformly in x.

For large n the estimate (2.11) is not sufficient but we can now use an integration by parts in z_2, in order to gain a factor $2^{-a_j - n}$; this is followed as above by integration by parts with respect to τ and σ. The result is that for $n \ge 10$

$$\int |H_{jk}^n(x,y)| dx \lesssim 2^{-a_j - n}$$

uniformly in y and again the same bound holds also for $\int |H_{jk}^n(x,y)| dy$, uniformly in x.

We sum in n and obtain the bound

$$\|\mathcal{H}_j \mathcal{H}_k^*\|_{L^2 \to L^2} \le C\Big(2^{j-k} + 2^{a_j - a_k} + \sum_{n \ge 10} \min\{2^{-a_j - n}, 2^{j-k+n}\} \Big).$$

Now if $n \ge 10$ we use the bound $2^{-a_j - n}$ for $n > (k-j)/2$ and the bound 2^{j-k+n} for $n \le (k-j)/2$. We sum in n and obtain the asserted bound (2.9) for the term $\|\mathcal{H}_j \mathcal{H}_k^*\|$.

The estimation of $\|\mathcal{H}_j^* \mathcal{H}_k\|$ is largely analogous. However we first use (1.15) to represent the kernel of $\mathcal{H}_j$ as

$$H_j(v,w) = \phi_j(v_1 - w_1) \chi(v,w) \int e^{i\tau(v_2 - \gamma^*(w, w_1 - v_1))} \beta_0(2^{-a_j} \frac{\tau}{\mathfrak{a}(v,w)}) \frac{d\tau}{\mathfrak{a}(v,w)}.$$

Thus the kernel of $\mathcal{H}_j^* \mathcal{H}_k$ is given by

$$\int \overline{H_j(z,x)} H_k(z,y) dz = \iiint \overline{\phi_j(z_1 - x_1) \chi(z,x)} \phi_k(z_1 - y_1) \chi(z,y) \times$$

$$e^{-i(\tau(z_2 - \gamma^*(x, x_1 - z_1)) - \sigma(z_2 - \gamma^*(y, y_1 - z_1)))} \overline{\beta_0(2^{-a_j} \frac{\tau}{\mathfrak{a}(z,x)})} \beta_0(2^{-a_k} \frac{\sigma}{\mathfrak{a}(z,y)}) \frac{d\tau}{\mathfrak{a}(z,x)} \frac{d\sigma}{\mathfrak{a}(z,y)} dz$$

and by using this expression the above proof for $\mathcal{H}_j \mathcal{H}_k^*$ can be repeated here. Again the only difference is that we have to take into account the limited differentiability of the symbol, but our assumptions on γ_{x_2} and its gradient still allow us to once integrate by parts with respect to z_2. $\quad \square$

In order to complete the proof of the $H_{prd}^1 \to L^1$ boundedness we use the following Lemma which is proved by standard arguments.

Lemma 2.5. *Let $\{H_j\}_{j\in I}$ be a finite family of Schwartz kernels and let $\mathcal{H}_j$ be the associated operators. Assume that $T := \sum_{j\in I} \mathcal{H}_j$ is bounded in L^2 with operator norm A_1, and suppose that the inequalities (2.7.1) hold for $(n_1, n_2) = (0,0)$, $(0,1)$ and $(1,0)$. Then T maps H^1_{prd} to L^1 with norm $\leq C(A_0 + A_1)$ (in particular C is independent of I).*

Proof. This is a straightforward consequence of a theorem by R. Fefferman [10] which says that it suffices to check the operator on rectangle atoms. Suppose that f is supported on a rectangle parallel to the coordinate axes, with center (c_1, c_2) with sidelength $2^{-\ell_1} \times 2^{-\ell_2}$ and that $\|f\|_2 \leq 2^{(\ell_1+\ell_2)/2}$, moreover f satisfies the strong cancellation condition $\int f(x_1, x_2)dx_1 = 0$ and $\int f(x_1, x_2)dx_2 = 0$. Fefferman's theorem states that if T is L^2 bounded and if the estimate

$$(2.12) \qquad \iint_{|x_1-c_1|\geq 2^{-\ell_1+n}} |Tf(x)|dx_1 dx_2 + \iint_{|x_2-c_2|\geq 2^{-\ell_2+n}} |Tf(x)|dx_1 dx_2 \lesssim 2^{-n\varepsilon}$$

holds for some $\varepsilon > 0$ then T maps $H^1_{prd}(\mathbb{R}^2)$ boundedly to $L^1(\mathbb{R}^2)$. Since we assume L^2 boundedness it suffices to prove (2.12).

We estimate the corresponding integrals with T replaced by $\mathcal{H}_j$. We use the size estimate in (2.7.1) obtaining the bound $O(2^{\ell_1-n-j})$ for the L^1 norm in $\{x : |x_1 - c_1| \geq 2^{-\ell_1+n}\}$ and we use the cancellation in y_1 together with the estimate (2.7.1) for the y_1-derivative to get the bound $2^{-\ell_1+j}$. Thus

$$\iint_{|x_1-c_1|\geq 2^{-\ell_1+n}} |\mathcal{H}_j f(x)|dx_1 dx_2 \lesssim \min\{2^{-\ell_1+j}, 2^{\ell_1-j-n}\}.$$

We sum in j and estimate the first term on the left of (2.12) by $C2^{-n/2}$.

Similarly (using now cancellation with respect to the y_2 variable) we obtain

$$\iint_{|x_2-c_2|\geq 2^{-\ell_2+n}} |\mathcal{H}_j f(x)|dx_1 dx_2 \lesssim \min\{2^{-\ell_2+a_j}, 2^{\ell_2-a_j-n}\}.$$

Clearly the right hand side is $O(2^{-n/2})$. Let j_0 be the maximal j with $2^{2a_j} \leq 2^{-n+2\ell_2}$. Then there is an absolute constant C_1 so that $2^{a_j} \lesssim 2^{a_{j_0}} 2^{j-j_0}$ if $j \leq j_0 - C_1$. Thus $\sum_{j\leq j_0-C_1} 2^{-\ell_2+a_j} \lesssim 2^{-\ell_2+a_{j_0}} \lesssim 2^{-n/2}$. Similarly if j_1 denotes the minimal j with $2^{2a_j} \geq 2^{-n+2\ell_2}$ then there is C_2 so that for $j \geq j_1 + C_2$ we have $2^{-a_j} \lesssim 2^{-a_{j_1}} 2^{j_1-j}$ and thus $\sum_{j\geq j_1+C_2} 2^{\ell_2-a_j-n} \lesssim 2^{\ell_2-a_{j_1}-n} \lesssim 2^{-n/2}$. We have only a bounded number of terms with $j_0 - C_1 \leq j \leq j_1 + C_2$; for those we use the bound $O(2^{-n/2})$. Combining the three estimates yields the bound $2^{-n/2}$ for the second term in (2.12). $\square$

3. L^p-boundedness of the Fourier integral contributions

We now give an outline of the proof of Theorem B and consider first the maximal operator. In view of Lemma 2.1 and Corollary 2.3 it suffices to consider the maximal function

$$(3.1) \qquad \sup_j \Big| \sum_{k>a_j} \mathcal{L}^k \mathcal{R}^k_j \mathcal{L}^k f \Big|.$$

where the sup is extended over a finite index set J. We use a familiar square-function technique and dominate

$$\sup_j \Big| \sum_{k>a_j} \mathcal{L}^k \mathcal{R}^k_j \mathcal{L}^k f \Big|$$

$$(3.2) \qquad \leq \Big(\sum_j \Big| \sum_{a_j<k\leq b_j} \mathcal{L}^k \mathcal{R}^k_j \mathcal{L}^k f \Big|^2 \Big)^{1/2} + \Big(\sum_j \Big| \sum_{k>b_j} \mathcal{L}^k \mathcal{R}^k_j \mathcal{L}^k f \Big|^2 \Big)^{1/2}$$

11

We define an operator $\mathfrak{M}$ acting on $F \in L^p(\ell^q)$ and an operator $\widetilde{\mathfrak{M}}$ acting on $G \in L^p(\ell^q(\ell^2))$ by by

$$
\begin{aligned}
(\mathfrak{M}F)_j &= \mathcal{M}_j F_j \\
(\widetilde{\mathfrak{M}}G)_{j,k} &= \mathcal{M}_j G_{j,k};
\end{aligned}
$$

(3.3)

here the ℓ^2 norm is taken with respect to the k variable. We denote by $\|\mathfrak{M}\|_{p,q}$ the $L^p(\ell^q) \to L^p(\ell^q)$ operator norm of $\mathfrak{M}$ and by $\|\widetilde{\mathfrak{M}}\|_{p,q,2}$ the $L^p(\ell^q(\ell^2)) \to L^p(\ell^q(\ell^2))$ operator norm of $\widetilde{\mathfrak{M}}$.

We follow M. Christ [7] (see also Nagel, Stein and Wainger [17] for a closely related earlier argument) and observe

Lemma 3.1. *For $1 \le p \le 2$*

(3.4)
$$\|\mathfrak{M}\|_{p,2} \lesssim (1 + \|\mathcal{M}\|_{L^p \to L^p})^{1-p/2}$$

(3.5)
$$\|\widetilde{\mathfrak{M}}\|_{p,2,2} \lesssim (1 + \|\mathcal{M}\|_{L^p \to L^p})^{1-p/2}$$

Proof. Since the operators $\mathcal{M}_j$ are bounded on L^p, uniformly in j, it is clear that the operator norm of $\mathfrak{M}$ on $L^p(\ell^p)$ is $O(1)$; the same applies to the vector-valued setting by which we see that the operator norm of $\widetilde{\mathfrak{M}}$ on $L^p(\ell^p(\ell^2))$ is finite.

Since $|\mathcal{M}_j f_j| \lesssim \mathcal{M}[\sup_\nu |f_\nu|]$ we see that the operator norm of $\mathfrak{M}$ on $L^p(\ell^\infty)$ and the operator norm of $\widetilde{\mathfrak{M}}$ on $L^p(\ell^\infty(\ell^2))$ are bounded by the L^p norm of $\mathcal{M}$. Interpolation gives the assertion. $\quad\square$

We now consider the first term on the right hand side of (3.2). First observe that there is a pointwise bound

(3.6)
$$|\mathcal{R}_j^k g| \lesssim M_{str}(\mathcal{M}_j(|g|))(x).$$

To see this we use integration by parts in τ to estimate

$$
|\mathcal{R}_j^k g| \lesssim \int \frac{2^k \widetilde{\phi}_j(x_1 - y_1)\widetilde{\chi}(x,y)}{(1 + 2^k|s|)^N} |g(y_1, \gamma(x_1, x_2, x_1 - y_1) + s)| dy_1 ds
$$

and change variables $s = \gamma(x_1, x_2 + u, x_1 - y_1) - \gamma(x_1, x_2, x_1 - y_1)$ which is legitimate since γ_{x_2} is close to 1.

By Littlewood-Paley theory for the operators $\mathcal{L}_k$, the pointwise bound (3.6) and the Fefferman-Stein theorem for the strong maximal function we get

$$
\left\| \left(\sum_j | \sum_{a_j < k \le b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f|^2 \right)^{1/2} \right\|_p \lesssim \left\| \left(\sum_{\substack{j,k: \\ a_j < k \le b_j}} |M_{str}(\mathcal{M}_j(|\mathcal{L}^k f|))|^2 \right)^{1/2} \right\|_p
$$

(3.7)
$$
\lesssim \left\| \left(\sum_{\substack{j,k: \\ a_j < k \le b_j}} |\mathcal{M}_j(\mathcal{L}^k f)|^2 \right)^{1/2} \right\|_p \lesssim \|\widetilde{\mathfrak{M}}\|_{p,2,2} \left\| \left(\sum_{\substack{j,k: \\ a_j < k \le b_j}} |\mathcal{L}^k f|^2 \right)^{1/2} \right\|_p \lesssim \|\widetilde{\mathfrak{M}}\|_{p,2,2} \|f\|_p
$$

where for the last application of Littlewood-Paley theory we have used that for fixed k the cardinality of the set $\{j : a_j < k \le b_j\}$ is bounded.

12

Similar but somewhat more complicated arguments apply to the second term in (3.2). We need to introduce an additional dyadic decomposition in the variable dual to x_1 and define operators $\mathcal{P}_l$, $\mathcal{Q}_l$, Π_m by

$$\widehat{\mathcal{P}_l f}(\xi) = \beta_0(2^{-l}\xi_1)\widehat{f}(\xi)$$
$$\widehat{\mathcal{Q}_l f}(\xi) = \beta_l(\xi_1)\widehat{f}(\xi)$$

and

$$\Pi_m = \mathcal{P}_{m-a_j+10} - \mathcal{P}_{m-b_j-10}$$

and decompose for fixed k the identity operator as

$$(3.8) \qquad I = \mathcal{P}_{j+k-b_j-10} + \Pi_{j+k} + \sum_{l>j+k-a_j+10} \mathcal{Q}_l.$$

Then we change variables $k = b_j + n$, $l = j + k - a_j + m = j + (b_j - a_j) + m + n$ and see that

$$\left\| \left(\sum_j \Big| \sum_{k>b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f \Big|^2 \right)^{1/2} \right\| \lesssim \sum_{n>0} \left\| \left(\sum_j \big| \mathcal{L}^{b_j+n} \mathcal{R}_j^{b_j+n} \mathcal{P}_{j+n-10} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p$$

$$+ \left\| \left(\sum_j \Big| \sum_{k>b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k \Pi_{j+k} f \Big|^2 \right)^{1/2} \right\|_p$$

$$(3.9) \qquad + \sum_{n>0}\sum_{m>0} \left\| \left(\sum_j \big| \mathcal{L}^{b_j+n} \mathcal{R}_j^{b_j+n} \mathcal{Q}_{j+n+m+b_j-a_j} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p.$$

We need to show part (i) of the following proposition (part (ii) will be needed for the singular Radon transform).

Proposition 3.2. *Let $p_0 > 1$, let $p_0 \le p \le 2$ and define $\theta \in [0,1]$ by $(1/p_0 - 1/p) = \theta(1/p_0 - 1/2)$. Then for $n > 0$, $m > 0$*

(i)

$$(3.10) \qquad \left\| \left(\sum_j \big| \mathcal{R}_j^{b_j+n} \mathcal{P}_{j+n-10} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\theta n/2} \|\mathfrak{M}\|_{p,2}^{1-\theta} \|f\|_p$$

$$(3.11) \qquad \left\| \left(\sum_j \Big| \sum_{k>b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k \Pi_{j+k} f \Big|^2 \right)^{1/2} \right\|_p \lesssim \|\widetilde{\mathfrak{M}}\|_{p,2,2}^{1-\theta} \|f\|_p$$

$$(3.12) \qquad \left\| \left(\sum_j \big| \mathcal{R}_j^{b_j+n} \mathcal{Q}_{j+n+m+b_j-a_j} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\theta(n+m)/2} \|\mathfrak{M}\|_{p,2}^{1-\theta} \|f\|_p$$

(ii)

$$(3.13) \qquad \left\| \left(\sum_j \big| \mathcal{P}_{j+n-10} \mathcal{R}_j^{b_j+n} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\theta n/2} \|\mathfrak{M}\|_{p,2}^{1-\theta} \|f\|_p$$

$$(3.14) \qquad \left\| \left(\sum_j \big| \mathcal{Q}_{j+n+m+b_j-a_j} \mathcal{R}_j^{b_j+n} \mathcal{L}^{b_j+n} f \big|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\theta(n+m)/2} \|\mathfrak{M}\|_{p,2}^{1-\theta} \|f\|_p$$

13

If we use the Fefferman-Stein theorem for vector-valued maximal functions we see that (3.9) and Proposition 3.2 imply the bound

$$\left\|\left(\sum_j |\sum_{k>b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f|^2\right)^{1/2}\right\|_p \lesssim (1 + \|\mathfrak{M}\|_{p,2} + \|\widetilde{\mathfrak{M}}\|_{p,2,2})\|f\|_p$$

(3.15)
$$\lesssim (1 + \|\mathcal{M}\|_{L^p \to L^p})^{1-p/2}\|f\|_p$$

where for the last inequality we have used Lemma 3.1. This in conjunction with (2.5), (2.8) and (3.7) shows that

$$(3.16) \qquad \|\mathcal{M}\|_{L^p \to L^p} \lesssim (1 + \|\mathcal{M}\|_{L^p \to L^p}^{1-p/2})$$

which implies of course the L^p boundedness of $\mathcal{M}$ for $1 < p \leq 2$, with bound independent of J. Since $\mathcal{M}$ is bounded on L^∞ the L^p boundedness for $2 < p < \infty$ follows as well. By the monotone convergence theorem this shows the L^p boundedness of the maximal operator in Theorem B.

We turn to the proof of Proposition 3.2. The main technical Lemma used here concerns L^2 estimates for the kernels $\mathcal{R}_j^k \mathcal{P}_{j+k-b_j-10}$ and $\mathcal{R}_j^k \mathcal{Q}_l$.

Lemma 3.3. *We have for $n > 0$, $m > 0$*

$$(3.17) \qquad \|\mathcal{R}_j^k \mathcal{Q}_l\|_{L^2 \to L^2} \lesssim 2^{-n/2}, \quad \text{if } k = b_j + n, l = j + k - a_j + m,$$

and

$$(3.18) \qquad \|\mathcal{R}_j^k \mathcal{P}_{j+k-b_j-10}\|_{L^2 \to L^2} \lesssim 2^{-(n+m)/2} \quad \text{if } k = b_j + n.$$

The estimates (3.17-18) also hold with $\mathcal{R}_j^k$ replaced by its adjoint $(\mathcal{R}_j^k)^*$.

The proof of Lemma 3.3 will be given in the next section. The estimates involving the adjoint operator are only needed for estimating the singular Radon transform. Taking the lemma for granted we can now give the

Proof of Proposition 3.2. The scheme of the proof is the same as for the chain of inequalities in (3.7). For the main term (3.11) we use the Littlewood-Paley inequality

$$(3.19) \qquad \left\|\left(\sum_{j,k} |\mathcal{L}^k \Pi_{j+k} f|^2\right)^{1/2}\right\|_p \lesssim \|f\|_p, \quad 1 < p < \infty,$$

and with (3.19) the proof of (3.11) follows by the argument given in (3.7). The inequality (3.19) in turn follows by the usual argument involving the Marcinkiewicz multiplier theorem and Rademacher functions (see [26]). Here it is necessary to show the L^p boundedness of the operators $\sum_{j,k} \pm \mathcal{L}^k \Pi_{j+k}$ (for any choice of $\pm$) and the doubling assumption on g is crucially used here. We note that (3.19) is essentially a version of the angular Littlewood-Paley theorem used in [17], [5], [7], [25] and elsewhere.

For the terms (3.10), (3.12) we use that for fixed n the $L^p(\ell^2)$ norm of $\{\mathcal{L}^{b_j+n} f\}_{j\in\mathbb{Z}}$ is bounded by $C\|f\|_p$ and the argument in (3.7) shows that the left hand sides of both (3.10) and (3.12) are dominated by $C\|\mathfrak{M}\|_{p,2}\|f\|_p$ if $p > 1$.

For $p = 2$ we have better bounds by Lemma 3.3; indeed the left hand side of (3.10) for $p = 2$ is dominated by $C2^{-\varepsilon n}\|f\|_2$, the left hand side of (3.11) by $C\|f\|_2$ and for (3.12) we obtain the bound $C2^{-\varepsilon(n+m)}\|f\|_2$. Interpolation yields (3.10), (3.11) (3.12). The proof of (3.13) and (3.14) is analogous if we take part (iii) of Lemma 3.3 into account. $\square$

L^p-boundedness of the singular Radon transform. In view of Proposition 2.4 and estimate (2.5) we have to prove the boundedness of the Fourier integral operator $\mathcal{F}$ given by

$$(3.20) \qquad \mathcal{F}f = \sum_j \sum_{k>a_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f.$$

By a Littlewood-Paley estimate in the x_2 variables we see that

$$\Big\| \sum_j \sum_{a_j < k \le b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f \Big\|_p$$

$$\lesssim \Big\| \Big(\sum_k | \sum_{j: a_j < k \le b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f |^2 \Big)^{1/2} \Big\|_p$$

$$(3.21) \qquad \lesssim \|\widetilde{\mathfrak{M}}\|_{p,2,2} \|f\|_p$$

where we argue as in the proof of (3.7). Next we use the decomposition (3.8) and obtain

$$\Big\| \sum_j \sum_{k>b_j} \mathcal{L}^k \mathcal{R}_j^k \mathcal{L}^k f \Big\|_p \lesssim \sum_{n \ge 10} I_n + II + \sum_{n \ge 10} \sum_{m>0} III_{m,n}$$

where

$$I_n = \Big\| \sum_j \mathcal{L}^{b_j+n} \mathcal{P}_{j+n-10} \mathcal{R}_j^{b_j+n} \mathcal{L}^{b_j+n} f \Big\|_p$$

$$II = \Big\| \sum_j \sum_{k>b_j} \mathcal{L}^k \Pi_{j+k} \mathcal{R}_j^k \mathcal{L}^k f \Big\|_p$$

$$III_{m,n} = \Big\| \sum_j \mathcal{L}^{b_j+n} \mathcal{Q}_{j+n+m+b_j-a_j} \mathcal{R}_j^{b_j+n} \mathcal{L}^{b_j+n} f \Big\|_p.$$

Now using for fixed n the Littlewood-Paley decomposition $\{\mathcal{L}^{b_j+n}\}_{j \in \mathbb{Z}}$ we see that I_n is estimated by the left hand side of (3.13) and thus by $C2^{-n\varepsilon(p)}\|f\|_p$ with $\varepsilon(p) > 0$. Similarly $III_{m,n}$ is dominated by $C2^{-(m+n)\varepsilon(p)}\|f\|_p$, by (3.14).

Finally by Littlewood-Paley theory

$$II \lesssim \Big\| \Big(\sum_j \sum_{k>b_j} |\mathcal{R}_j^k \mathcal{L}^k f|^2 \Big)^{1/2} \Big\|_p := \widetilde{II}.$$

Now we decompose as in (3.8), but to the right hand side of $\mathcal{R}_k^j$. Thus

$$\widetilde{II} \lesssim \sum_{n \ge 10} I_n' + II' + \sum_{n \ge 10} \sum_{m>0} III_{m,n}'$$

where I_n' is the left hand side of (3.10), and $III_{m,n}'$ is the left hand side of (3.12). Moreover

$$II' = \Big\| \Big(\sum_j \sum_{k>b_j} |\mathcal{R}_j^k \mathcal{L}^k \Pi_{j+k} f|^2 \Big)^{1/2} \Big\|_p$$

which is dominated by $C\|\widetilde{\mathfrak{M}}\|_{p,2}\|f\|_p$; here we use again (3.19) and the Fefferman-Stein inequality.

Since we have already established the L^p bounds for the maximal operator we know now by (3.4), (3.5) that $\|\mathfrak{M}\|_{p,2}$, $\|\mathfrak{M}_{p,2,2}\|_{p,2,2}$ are $O(1)$ and thus the combination of previous estimates shows the L^p boundedness of the Fourier integral operator $\mathcal{F}$ in (3.20). As pointed out above this yields the L^p boundedness for the singular Radon transform, for $1 < p \le 2$. The estimates can be applied to the adjoint operator which yields the estimates in the complementary range $2 < p < \infty$. $\quad\square$

4. Proof of Lemma 3.3

The kernel of $\mathcal{R}_j^k \mathcal{Q}_l$ is given by

(4.1)
$$K(x,y) = \iiint e^{i(\tau(y_2 - \gamma(x, x_1 - z_1)) + \xi_1(z_1 - y_1))} \chi(x, z_1, y_2)\phi_j(x_1 - z_1)\beta_1(2^{-k}\tau)\beta_1(2^{-l}\xi_1)d\xi_1 d\tau dz_1.$$

By integration by parts with respect to z_1 we obtain

$$K(x,y) = \iiint e^{i(\tau(y_2 - \gamma(x, x_1 - z_1)) + \xi_1(z_1 - y_1))} i^{-1} a(x, z_1, y_2, \tau, \xi_1)d\xi_1 d\tau dz_1$$

where $a = a_1 + a_2$ with

$$a_1(x, z_1, y_2, \tau, \xi_1) = \frac{\ddot{\gamma}(x, x_1 - z_1)}{(\tau\dot{\gamma}(x, x_1 - z_1) + \xi_1)^2}\chi(x, z_1, y_2)\phi_j(x_1 - z_1)\beta_1(2^{-k}\tau)\beta_1(2^{-l}\xi_1)$$

$$a_2(x, z_1, y_2, \tau, \xi_1) = \frac{\partial_{z_1}(\chi(x, z_1, y_2)\phi_j(x_1 - z_1))}{\tau\dot{\gamma}(x, x_1 - z_1) + \xi_1}\beta_1(2^{-k}\tau)\beta_1(2^{-l}\xi_1).$$

The localization of the symbol in (4.1) implies that here $|\xi_1| \gg |\tau\dot{\gamma}(x, x_1 - z_1)|$.

The following fact will be crucial in the estimation of the L^1 norms in x or y.

Sublemma. *For large j we have the estimates*

(i)

(4.2)
$$\int_{|z_1 - x_1| \leq 2^{-j+1}} 2^{-j}|\ddot{\gamma}(x, x_1 - z_1)|dz_1 \lesssim 2^{-a_j},$$

(ii)

(4.3)
$$\int_{2^{-j-1} \leq |z_1 - x_1| \leq 2^{-j+1}} 2^j \frac{|\ddot{\gamma}(x, x_1 - z_1)|}{(\dot{\gamma}(x, x_1 - z_1))^2}dz_1 \lesssim 2^{b_j}.$$

Proof. Let $I_j(x_1) := \{z_1 : 0 \leq x_1 - z_1 \leq 2^{-j+1}\}$. By the quasimonotonicity assumption we have for $t > 0$ that $\ddot{\gamma}(x, t) = a(x, t) + O(\dot{\gamma}(x, t))$ where a does not change sign and thus

$$\int_{I_j(x_1)} |\ddot{\gamma}(x, x_1 - z_1)|dz_1 \lesssim \left|\int_{I_j(x_1)} \ddot{\gamma}(x, x_1 - z_1)dz_1\right| + \int_{I_j(x_1)} |\dot{\gamma}(x, x_1 - z_1)|dz_1$$

$$\lesssim g(A2^{-j-1}) \lesssim 2^{j-a_j}.$$

The same bound holds for the contribution over $x_1 \leq z_1$. This proves (4.2).

Now we turn to (4.3) and let $J_j(x_1) := \{z_1 : 2^{-j-1} \leq x_1 - z_1 \leq 2^{-j+1}\}$. Again by the quasimonotonicity of $\dot{\gamma}$ we get

$$\int_{J_j(x_1)} \frac{|\ddot{\gamma}(x, x_1 - z_1)|}{(\dot{\gamma}(x, x_1 - z_1))^2}dz_1 \leq \left|\int_{J_j(x_1)} \frac{\ddot{\gamma}(x, x_1 - z_1)}{(\dot{\gamma}(x, x_1 - z_1))^2}dz_1\right| + \int_{J_j(x_1)} |\dot{\gamma}(x, x_1 - z_1)|^{-1}dz_1.$$

The first term on the right hand side equals

$$\left|\frac{1}{\dot{\gamma}(x, 2^{-j+1})} - \frac{1}{\dot{\gamma}(x, 2^{-j-1})}\right| \lesssim [g(2^{-j-1}/A)]^{-1} \lesssim 2^{b_j - j}$$

and the second term is estimated by

$$\int_{J_j(x_1)} g(2^{-j-1}/A)dz_1 \lesssim 2^{b_j-2j} \lesssim 2^{b_j-j}.$$

The contributions for $2^{-j-1} \le z_1 - x_1 \le 2^{-j+1}$ are estimated in the same way and (4.3) is proved. $\quad\square$

Proof of Lemma 3.3, cont. Integration by parts yields

$$|K(x,y)| \lesssim \iiint \frac{(I-2^{2l}\partial_{\xi_1}^2)^N(I-2^{2k}\partial_\tau^2)^N a(x,z_1,y_2,\tau,\xi_1)}{(1+2^{2k}|y_2-\gamma(x,x_1-z_1)|^2)^N(1+2^{2l}|y_1-z_1|^2)^N}dz_1 d\xi_1 d\tau.$$

Assume that $|\mathcal{C}| \le 1$, $|\tau|, |\xi_1| \ge 1$ and that either $|\mathcal{C}\tau| \ge 2|\xi_1|$ or $|\xi_1| \ge 2|\mathcal{C}\tau|$. (actually, for the present proof of (3.17) we need this for $|\xi_1| \ge 2|\mathcal{C}\tau|$).

It is easy to verify that under these assumptions we have the product type symbol estimates

$$\left|\partial_{\xi_1}^{\alpha_1}\partial_\tau^{\alpha_2}\big(\beta_1(2^{-k}\tau)\beta_1(2^{-l}\xi_1)\tau(\mathcal{C}\tau+\xi_1)^{-2}\big)\right| \lesssim 2^{-l\alpha_1}2^{-k\alpha_2}2^k(|\mathcal{C}\tau|+|\xi_1|)^{-2},$$

$$\left|\partial_{\xi_1}^{\alpha_1}\partial_\tau^{\alpha_2}\big(\beta_1(2^{-k}\tau)\beta_1(2^{-l}\xi_1)(\mathcal{C}\tau+\xi_1)^{-1}\big)\right| \lesssim 2^{-l\alpha_1}2^{-k\alpha_2}(|\mathcal{C}\tau|+|\xi_1|)^{-1}.$$

We apply this with $\mathcal{C} = \dot\gamma(x, x_1 - z_1)$ and see that

$$\left|\partial_{\xi_1}^{\alpha_1}\partial_\tau^{\alpha_2} a_1(x,z_1,y_2,\tau,\xi_1)\right| \lesssim 2^{-l\alpha_1}2^{-k\alpha_2}\Big[\frac{2^{k+j}}{|\mathcal{C}\tau|+|\xi_1|)^2} + \frac{2^j}{|\mathcal{C}\tau|+|\xi_1|}\Big],$$

$$\left|\partial_{\xi_1}^{\alpha_1}\partial_\tau^{\alpha_2} a_2(x,z_1,y_2,\tau,\xi_1)\right| \lesssim 2^{-l\alpha_1}2^{-k\alpha_2}\frac{2^{2j}}{|\mathcal{C}\tau|+|\xi_1|}.$$

Consequently

$$(4.4)\quad |K(x,y)| \lesssim \iiint_{\substack{|\xi_1|\approx 2^l \\ |\tau|\approx 2^k \\ |x_1-z_1|\le 2^{-j+1}}} \Big[\frac{2^{j+k}|\ddot\gamma(x,x_1-z_1)}{(|\tau\dot\gamma(x,x_1-z_1)|+|\xi_1|)^2} + \frac{2^{2j}}{|\tau\dot\gamma(x,x_1-z_1)|+|\xi_1|}\Big]$$

$$\times (1+2^{2k}|y_2-\gamma(x,x_1-z_1)|^2)^{-N}(1+2^{2l}|y_1-z_1|^2)^{-N}dz_1 d\xi_1 d\tau.$$

We now examine the L^1 norm in y. We interchange the order of integration and first integrate out in the y-variable. We take into account that now $|\xi_1| \ge 2|\tau\dot\gamma(x,x_1-z_1)|$ and obtain

$$(4.5)\qquad \int |K(x,y)|dy \lesssim \iiint_{\substack{|\xi_1|\approx 2^l \\ |\tau|\approx 2^k \\ |x_1-z_1|\le 2^{-j+1}}} \big[2^{j+k}|\ddot\gamma(x,x_1-z_1)||\xi_1|^{-2} + 2^{2j}|\xi_1|^{-1}\big]2^{-k-l}dz_1 d\xi_1 d\tau.$$

By part (i) of the Sublemma this is dominated by a constant times

$$(4.6)\qquad 2^{-l-k}\iint_{\substack{|\xi_1|\approx 2^l \\ |\tau|\approx 2^k}} [2^{2j+k-a_j-2l} + 2^{j-l}]d\xi_1 d\tau \lesssim \big[2^{2j+k-a_j-2l} + 2^{j-l}\big] \lesssim 2^{-m-n}.$$

It is possible to show the same inequality for $\int |K(x,y)|dx$ but we can get away with the bound $O(1)$ for the latter integral and still get (3.18).

We proceed similarly for the estimation of the kernel $\widetilde{K}$ of $\mathcal{R}_j^k \mathcal{P}_{j+k-b_j-10}$. Now however we have the restrictions $|\xi_1| \lesssim 2^{j+k-b_j-9}$ and $|\tau\dot{\gamma}(x,t)| \geq 2^{k-1}g(2^{-j-1}/A) \geq 2^{k-1+j-b_j}$ so that the latter expression is dominant.

The above analysis leads to

$$
|\widetilde{K}(x,y)| \lesssim \iiint_{\substack{|\xi_1| \leq 2^{j+k-b_j-9} \\ 2^{k-1} \leq |\tau| \leq 2^{k+1} \\ |x_1-z_1| \leq 2^{-j+1}}} \Big[\frac{2^{j+k}|\ddot{\gamma}(x,x_1-z_1)}{|\tau\dot{\gamma}(x,x_1-z_1)|^2} + \frac{2^{2j}}{|\tau\dot{\gamma}(x,x_1-z_1)|} \Big]
$$

$$
\times (1+2^{2k}|y_2-\gamma(x,x_1-z_1)|^2)^{-N}(1+2^{2(j+k-b_j)}|y_1-z_1|^2)^{-N} dz_1 d\xi_1 d\tau.
$$

and as above we get

$$
\int |\widetilde{K}(x,y)| dy \lesssim \int_{|x_1-z_1| \leq 2^{-j+1}} \Big[\frac{2^{j+k}|\ddot{\gamma}(x,x_1-z_1)|}{|2^k\dot{\gamma}(x,x_1-z_1)|^2} + \frac{2^{2j}}{|\tau\dot{\gamma}(x,x_1-z_1)|} \Big] dz_1 d\xi_1 d\tau \lesssim 2^{b_j-k} \lesssim 2^{-n}
$$

where for the second to last inequality we use (4.3). Combining this with $\int |\widetilde{K}(x,y)| dx = O(1)$ we obtain (3.17). $\square$

REFERENCES

1. A. Carbery and S. Pérez, *Maximal functions and Hilbert transforms along variable flat curves*, Math. Res. Lett. **6** (1999), 237–249.

2. A. Carbery, A. Seeger, S. Wainger and J. Wright, *Classes of singular integral operators along variable lines*, Journal Geom. Anal. **9** (1999), 584–605.

3. A. Carbery, S. Wainger and J. Wright, *Hilbert transforms and maximal functions associated to flat curves on the Heisenberg group*, J. Amer. Math. Soc. **8** (1995), 141–179.

4. _____, *Hilbert transforms and maximal functions along flat variable plane curves*, J. Fourier analysis and its applications **1995**, no. Special Issue, 119–139.

5. H. Carlsson, M. Christ, A. Córdoba, J. Duoandikoetxea, J. L. Rubio de Francia, J. Vance, S. Wainger and D. Weinberg, L^p *estimates for maximal functions and Hilbert transforms along flat curve in R^2*, Bull. Amer. Math. Soc. **14** (1986), 263–267.

6. S.Y.A. Chang and R. Fefferman, *The Calderón-Zygmund decomposition on product domains*, Amer. J. Math **104** (1982), 445–468.

7. M. Christ, *Endpoint bounds for singular fractional integral operators*, preprint 1988, unpublished.

8. M. Christ, A. Nagel, E. M. Stein and S. Wainger, *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. **150** (1999), 489–577.

9. A. Córdoba and J. L. Rubio de Francia, *Estimates for Wainger's singular integrals along curves*, Rev. Mat. Iberoamericana **2** (1986), 105–117.

10. R. Fefferman, *Harmonic Analysis on product spaces*, Annals of Math. **126** (1987), 109–130.

11. M. Greenblatt, *A method for proving L^p boundedness of singular Radon transforms in codimension 1*, Duke Math. J. **108** (2001), 363–393.

12. A. Greenleaf and G. Uhlmann, *Estimates for singular Radon transforms and pseudo-differential operators with singular symbols*, J. Funct. Anal. **89** (1990), 202–232.

13. L. Hörmander, *Fourier integral operators I*, Acta. Math. **127** (1971), 79–183.

14. J.L. Journé, *Calderón-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), 55–91.

15. J. Kim, *Hilbert transforms along curves in the Heisenberg group*, Proc. London Math. Soc. **80** (2000), 611–642.

16. _____, L^p *estimates for singular integrals and maximal operators associated with flat curves on the Heisenberg group*, Duke Math. J. (to appear).

17. A. Nagel, E. M. Stein and S. Wainger, *Differentiation in lacunary directions*, Proc. Nat. Acad. Sc. USA **75** (1978), 1060–1062.

18. _____ , *Hilbert transforms and maximal functions related to variable curves*, Harmonic Analysis in Euclidean spaces (Proc. Sympos. Pure Math. Williams Coll., Williamstown, Mass., 1978), Part 1,, Proc. Sympos. Pure Math. XXXV, Amer. Math. Soc., 1979, pp. 95-98.

19. A. Nagel, J. Vance, S. Wainger and D. Weinberg, *Hilbert transforms for convex curves*, Duke Math. J. **50** (1983), 735–744.

20. _____ , *Maximal functions for convex curves*, Duke Math. J. **52** (1985), 715–722.

21. D. H. Phong and E. M. Stein, *Hilbert integrals, singular integrals and Radon transforms, I*, Acta Math. **157** (1986), 99–157.

22. _____ , *Singular Radon transforms and oscillatory integrals*, Duke Math. J. **58** (1989), 347–369.

23. A. Seeger, L^2 *estimates for a class of singular oscillatory integrals*, Math. Res. Lett. **1** (1994), 65–71.

24. _____ , *Radon transforms and finite type conditions*, J. Amer. Math. Soc. **11** (1998), 869–897.

25. A. Seeger and S. Wainger, *Bounds for singular fractional integrals and related Fourier integral operators*, preprint (2000), submitted.

26. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, N.J., 1971.

Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA

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

Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA

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