Mathematical Research Letters **8**, 789–800 (2001)

OSCILLATORY INTEGRALS RELATED TO CARLESON'S THEOREM

Elias M. Stein and Stephen Wainger

1. Introduction

The famous theorem of Carleson, in the n-dimensional form given it by Sjölin [Sj] is the following maximal estimate:

Suppose K is an appropriate Calderón-Zygmund kernel in $\mathbb{R}^n$, and let

$$T_\lambda(f)(x) = \int_{\mathbb{R}^n} e^{i\lambda\cdot y}\, K(y)\, f(x-y)\, dy, \tag{1.1}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots \lambda_n) \in \mathbb{R}^n$. Then the mapping $f \mapsto \sup_{\lambda\in\mathbb{R}^n} |T_\lambda(f)(x)|$ is bounded on $L^2(\mathbb{R}^n)$.

It is natural to want to put this result in a broader context. One formulation that suggests itself is to replace the linear form $\lambda \cdot y$ occuring in the exponential above by a real polynomial $P_\lambda(y) = \sum_{1\le|\alpha|\le d} \lambda_\alpha y^\alpha$ of fixed degree d, and now define the corresponding maximal operator by taking the sup over all the coefficients $\lambda = (\lambda_\alpha)$, that is with each λ_α ranging over $\mathbb{R}$.

What are the chances that such a wider result holds? There are a number of specific facts that suggests that this may be true. First is the situation which occurs when, in effect, the "stopping times" involved are themselves polynomials in x. This means we consider operators of the form $T(f)(x) = \int_{\mathbb{R}^n} e^{iP(x,y)}\, K(y)\, f(x-y)\, dy$, where $P(x,y)$ is a real polynomial on $\mathbb{R}^n \times \mathbb{R}^n$. The L^2 boundedness of these operators (with bounds that depend only on the degree of P and not its coefficients) is known ([RS]), and these occur in the study of singular integrals on nilpotent groups.

In another direction, the special case when $n = 1$, $d = 2$, and $P(y) = \lambda y^2$ (with no lower term) was proved in [St1]. the argument given there was based in part on a good asymptotic formula for the Fourier transform of the kernel $e^{i\lambda y^2}/y$; however, this approach would not seem susceptible to easy extension.

To attack the more general problem we need a different approach, and this is what we want to present below.

Received August 20, 2001.
This work was supported by NSF grants DMS-0070692 & DMS-0098757 .

Our results will be of a general character, save for one significant reservation: we shall consider all real polynomials $P_\lambda(y) = \sum_{2\le|\alpha|\le d} \lambda_\alpha y^\alpha$ in $\mathbb{R}^n$, of degree $\le d$, but with the restriction that they have no first-order terms.

With $T_\lambda(f)(x) = \int_{\mathbb{R}^n} e^{iP_\lambda(y)}\, K(y)\, f(x-y)\, dy$, we shall prove that the mapping

$$f \longrightarrow \sup_\lambda |T_\lambda(f)(x)|$$

is bounded on $L^2(\mathbb{R}^n)$.

As we have indicated, the proof of this result will not be based on the analysis of the Fourier transform of $e^{iP_\lambda(y)}\, K(y)$, but instead we will reduce matters to a corresponding maximal oscillatory estimate which would seem to be of interest in its own right. This maximal theorem can be stated as follows.

Let φ be a suitable "bump" function supported in the unit ball, and denote $e^{iP_\lambda(y)}\,\varphi(y)$ by $\Phi^\lambda(y)$. Write Φ^λ_a for $\Phi^\lambda_a(y) = a^{-n}\,\Phi^\lambda(y/a)$, when $a > 0$. Then the maximal operator

$$f \to \sup_{|\lambda|\ge r,\, a>0} |(f * \Phi^\lambda_a)(x)|\,,$$

has a norm (acting on L^2) which decays like $O(r^{-\delta})$, for some $\delta > 0$.

There is no direct analogue of this maximal estimate when $P_\lambda(y)$ has first order terms, (then there is no decay in r), and for this reason our results do not include the Carleson-Sjölin theorem. It might be expected, however, that further work, combining one of the known proofs of the Carleson-Sjölin theorem with our arguments could yield the full result. We hope to return this matter at a future time. In the meanwhile we have learned that M.T. Lacey, combining ideas of his proof with Thiele for Carleson's theorem [LT], with the arguments in [St1], has obtained the desired result in the case when $n = 1$, $d = 2$. (See [L]).

2. Two van der Corput-like propositions in n-dimensions

The following are variants or refinements of known estimates. The particular versions that follow are needed below, but do not seem to be stated explicitly in the literature. Here

$$P(x) = \sum_{1\le|\alpha|\le d} \lambda_\alpha\, x^\alpha$$

is a polynomial in $\mathbb{R}^n$ of degree $\le d$, with real coefficients and no constant term. We denote $|\lambda| = \sum_{1\le|\alpha|\le d} |\lambda_\alpha|$. We also assume that φ is a given $C^{(1)}$ function defined in the unit ball, $B = \{x : |x| \le 1\}$, and let Ω be any convex subset of B.

Proposition 2.1.

$$\left| \int_\Omega e^{iP(x)}\, \varphi(x)\, dx \right| \le c|\lambda|^{-1/d} \sup_{x\in B} \left(|\varphi(x)| + |\nabla\varphi(x)| \right) .$$

The constant c depends on the dimension n and the degree d, but not otherwise on P, φ, or Ω.

The significance of the proposition for us is the uniform decay taken over all polynomials of degree d, as a function of the total size of the coefficients. The exact power $-1/d$, while optimal, is not essential.

Proposition 2.2. *With the same notation as above*

$$|\{x \in B : |P(x)| \le \epsilon\}| \le c\, \epsilon^{1/d}\, |\lambda|^{-1/d}, \ \textit{for every } \epsilon > 0 .$$

Again, c does not depend on the coefficient of P, but only on n and d.

Earlier results, related to these propositions, maybe found in [AKC], [RS], and [CCW]. The proofs require several lemmas.

Lemma 2.1. *Let* $Q(x) = \sum_{|\alpha|=k} \lambda_\alpha\, x^\alpha$ *be a homogeneous polynomial of degree* k. *Then there is a unit vector* ξ, *so that*

$$\left|(\xi\cdot\partial_x)^k\, Q(x)\right| \ge c|\lambda|, \ \textit{with } |\lambda| = \sum_{|\alpha|=k} |\lambda_\alpha|$$

Proof. We have

$$\sum_{|\alpha|=k} \left| \left(\frac{\partial}{\partial x}\right)^\alpha Q(x) \right| = \sum_{|\alpha|=k} \alpha! |\lambda_\alpha| \ge c|\lambda| .$$

On the other hand, there are unit vector $\xi_1, \dots \xi_N$, so that $\{(\xi_j \cdot x)^k\}_{j=1}^N$ form a basis of the homogeneous polynomials of degree k. (See e.g. [St2], p. 343). Hence for appropriate c_j^α, $\left(\frac{\partial}{\partial x}\right)^\alpha = \sum_{j=1}^N c_j^\alpha (\xi_j\cdot\partial_x)^k$, and we need only pick α so that λ_α maximizes $|\lambda_\alpha|$, and then pick $\xi = \xi_j$, so that $(\xi_j\cdot\partial_x)^k\, Q(x)$ achieves a maximum as j varies, $1 \le j \le N$. □

Lemma 2.2. *Suppose* $P(x) = \sum_{1\le|\alpha|\le d} \lambda_\alpha x^\alpha$ *is a given polynomial of degree* $\le d$. *Then there is a* k, $1 \le k \le d$, *and a unit vector* ξ, *so that*

$$|(\xi\cdot\partial_x)^k\, P(x)| \ge c|\lambda|, \ \textit{for all } x \in B ,$$

where $\lambda = \sum_{1\le|\alpha|\le d} |\lambda_\alpha|$.

Proof. Let us write

$$\lambda^{(k)} = \sum_{|\alpha|=k} |\lambda_\alpha|,$$

so that $|\lambda| = \sum_{k=1}^{d} \lambda^{(k)}$. We shall find for some k, $1 \le k \le d$, a "dominant" $\lambda^{(k)}$ among $\lambda^{(1)}, \lambda^{(2)}, \ldots \lambda^{(d)}$; it will have the following properties:

(i) $\lambda^{(k)} \approx |\lambda|$

(ii) $\epsilon\lambda^{(k)} \ge \lambda^{(j)}$, for all $j > k$.

Here ϵ is a (small) preassigned number.

We argue as follows. Fix $\epsilon > 0$, and choose k so that $\epsilon^{-k}\lambda^{(k)}$ is a maximum among $\epsilon^{-1}\lambda^{(1)}$, $\epsilon^{-2}\lambda^{(2)}$, $\ldots \epsilon^{-k}\lambda^{(k)}$, $\ldots \epsilon^{-d}\lambda^{(d)}$. Thus $\epsilon\lambda^{(k)} \ge \epsilon^{k-j+1}\lambda^{(j)}$. As a result, if $j > k$, then $\epsilon\lambda^{(k)} \ge \lambda^{(j)}$, (assuming $\epsilon \le 1$), which verifies property (ii). However, when $j < k$, $\lambda^{(k)} \ge \epsilon^{k-j}\lambda^{(j)} \ge \epsilon^k\lambda^{(j)}$. Altogether then $\lambda^{(k)} \ge c_\epsilon \sum_{j=1}^{k} \lambda^{(j)}$, and therefore $\lambda^{(k)} \approx |\lambda| = \sum_{j=1}^{k} \lambda^{(j)}$, so property (i) is also proved.

Now with k chosen write

$$P(x) = P_0(x) + Q(x) + P_1(x),$$

where

$$P_0(x) = \sum_{1 \le |\alpha| < k} \lambda_\alpha x^\alpha, \qquad Q(x) = \sum_{|\alpha|=k} \lambda_\alpha x^\alpha,$$

and

$$P_1(x) = \sum_{k<|\alpha| \le d} \lambda_\alpha x^\alpha.$$

We apply $(\xi \cdot \partial_x)^k$ to P, where ξ is the unit vector guaranteed by Lemma 2.1. Thus
$(\xi \cdot \partial_x)^k P_0 \equiv 0$, and if $|x| \le 1$, $|(\xi \cdot \partial_x)^k P_1(x)| \le c\sum_{j>k} \lambda^{(j)}$. The latter is $\le c'\epsilon\lambda^{(k)}$ by the fact that $\lambda^{(k)}$ was dominant. However, $|(\xi \cdot \partial_x)^k Q(x)| \ge c\lambda^{(k)}$ by the lemma, and hence we see that our conclusion is proved if we take ϵ to be sufficiently small, when we recall that $\lambda^{(k)} \approx |\lambda|$.

We now prove Proposition 2.1. We pick a coordinate system so that x_1 lies in the direction of the unit vector ξ given by Lemma 2.2, and $x_1, \ldots x_n$ are in orthogonal directions. Then $(\xi \cdot \partial_x)^k = \frac{\partial^k}{\partial x_1^k}$, and we can use the usual van der Corput estimate in the x_1 variable (see e.g [St2]) and then integrate the estimate in $x_2, \ldots x_n$.

Proposition 2.2 is then reduced to a one-variable estimate (which we assume; see [C]) in the same way as in the proof of Proposition 2.1. □

3. A maximal lemma

We shall also need the following easy variant of the standard maximal function.

For any set $E \subset B_2$, with $B_2 = \{x : |x| \leq 2\}$, we denote by χ_E its characteristic function, and $(\chi_E)_a(x) = a^{-n}\,\chi_E(x/a)$, $a > 0$. We define

$$\mathcal{M}_\epsilon(f)(x) \;=\; \sup_{\substack{a>0\\ |E|\leq\epsilon}} \;|f| \,*\, (\chi_E)_a\,(x)\,, \tag{3.1}$$

with the sup taken over all subsets E of B_2 of measure $\leq \epsilon$, and all $a > 0$.*

Proposition 3.1.

$$\| \,\mathcal{M}_\epsilon(f)\, \|_{L^2} \leq \, c\,\epsilon^{1/2} \;\; \|\, f \,\|_{L^2} \;\; .$$

The main point here is a (power) decrease in ϵ, as $\epsilon \longrightarrow 0$. The proof is based on the observation that

$$|\{x :\, \mathcal{M}_\epsilon(f)(x) \,>\, \alpha\}| \,\leq\, (c/\alpha) \int_{|f|\geq 2\alpha/\epsilon} |f|\, dx\,,\;\text{ all } \alpha > 0\,. \tag{3.2}$$

In fact, $\mathcal{M}_\epsilon(1) \,\leq\, \epsilon$, and $\mathcal{M}_\epsilon(f) \,\leq\, cM(f)$ where M is the standard maximal function. Now if f is positive, $f \,\leq\, f_1 + \alpha/2\epsilon$, where $f_1 \,=\, f$ if $f(x) \,>\, \alpha/2\epsilon$, $f_1 = 0$ otherwise. Hence $\{x : \mathcal{M}_\epsilon(f)(x)\} > \alpha\} \subset \{x : cM(f_1) \,>\, \alpha/2\}$ and the latter has measure $\leq \frac{c}{\alpha} \displaystyle\int f_1 dx$ and (3.2) is proved. Therefore,

$$\| \;\mathcal{M}_\epsilon(f) \;\|_{L^p}^p \leq \, cp \int_0^\infty \alpha^{p-2} \left(\int_{|f|\,\geq\,\alpha/2\epsilon} |f|\, dx \right) d\alpha \;=\; c_p'\, \epsilon^{p-1} \;\; \|\; f\; \|_{L^p}^p \;,$$

and the case $p = 2$ is the assertion of Proposition 3.1.

4. The first main proposition

Let us set $\Phi^\lambda(x) \,=\, e^{iP_\lambda(x)}\,\varphi(x)$ where $P_\lambda(x) \,=\, \displaystyle\sum_{2\leq|\alpha|\leq d} \lambda_\alpha\, x^\alpha$ is a real polynomial of degree d; here we assume that the linear terms vanish. The function φ is a fixed $C^{(1)}$ function supported in the unit ball. For each $a > 0$, we set

$$\Phi^\lambda_a(x) \;=\; a^{-n}\,\Phi^\lambda\,(x/a)$$

*It is a slight technical advantage to state the results for $\mathcal{M}_\epsilon$ defined over the ball B_2, instead of the unit ball B.

Theorem 1.

$$\| \sup_{\substack{|\lambda| \geq r \\ a>0}} |(f * \Phi_a^\lambda)(x)| \; \|_{L^2} \leq \; A r^{-\delta} \; \| \, f \, \|_{L^2} \, , r \geq 1 \, . \tag{4.1}$$

Here δ is some fixed small positive number, and as before $\lambda = \sum_{2\leq|\alpha|\leq d} |\lambda_\alpha| \, .$

Remarks. It is important to note that we do not allow first degree terms in P_λ, for otherwise there would be no decay in r. For our applications it suffices to know that $\delta > 0$; the proof below gives $\delta = 1/6d$, but this is most likely not optimal.

We note that in proving the theorem it suffices first to establish it when the sup in (4.1) is restricted to $r \leq |\lambda| \leq 2r$; because then adding the corresponding estimates, when the range $r \leq |\lambda|$ is decomposed dyadically, gives a convergent geometric series, in view of the asserted decay in r.

The proof of this more restricted conclusion uses the method of TT^*, more precisely an adaptation of a Kolmogorov-Seliverstov argument. To carry out this argument we proceed as follows.

We denote by $\widetilde{\Phi}^\lambda$ the "dual" to Φ^λ, i.e. $\widetilde{\Phi}^\lambda(x) \, = \, \overline{\Phi}^\lambda(-x) \, = \, e^{-iP_\lambda(-x)} \, \bar{\varphi}(-x)$. We then claim:

Lemma 4.1.

$$\left| \left(\Phi_h^\nu \, * \, \widetilde{\Phi}_1^\mu \right) (x) \right| \; \leq \; c \left(r^{-2\delta} \, \chi_{B_2}(x) \, + \, \chi_{E_\mu}(x) \right) \tag{4.2}$$

when $r \leq |\nu| \leq 2r$, $r \leq |\mu| \leq 2r$, and $0 < h \leq 1$. Here E_μ is a subset of the ball $B_2 = \{|x| \leq 2\}$, with $|E_\mu| \, \leq \, r^{-4\delta}$. While E_μ depends on μ, it is independent of ν and h. The bound c is independent of ν, μ, r, and h.

Proof. We have

$$\left(\Phi_h^\nu \, * \, \widetilde{\Phi}_1^\mu \right) (x) \; = \; h^{-n} \int_{\mathbb{R}^n} e^{i(P_\nu(y/h) \, - \, P_\mu(y-x))} \varphi(y/h) \, \overline{\varphi(y-x)} \, dy \, .$$

Given that φ is supported in the unit ball, and $h \leq 1$, it is clear that the above convolution is supported in $|x| \, \leq \, 2$. We next make the change of variables $y \longrightarrow hy$ which shows that

$$\left(\Phi_h^\nu \, * \, \widetilde{\Phi}_1^\mu \right) (x) \; = \; \int_{\mathbb{R}^n} e^{i(P_\nu(y) - P_\mu(hy-x)} \, \varphi(y) \overline{\varphi(hy-x)} \, dy \, . \tag{4.3}$$

There are now two case: Case I, h small, $0 < h \leq \eta$, where we choose η below. Case II, h not small, $\eta < h \leq 1$.

For Case I, note that when $|x| \leq 2$,

$$P_\nu(y) \; - P_\mu(hy - x) \; = \; \sum_{2\leq|\alpha|\leq d} \; (\nu_\alpha + O(h|\mu|)) \, y^\alpha \; - P_\mu(-x) \, .$$

while $\sum |\nu_\alpha + O(h|\mu|)| \geq \sum |\nu_\alpha| - c\eta\,|\mu| \geq c\sum |\nu_\alpha| \geq cr$, if η is sufficient small, since $r \leq |\nu| = \sum |\nu_\alpha| \leq 2r$, $r \leq |\mu| \leq 2r$, and $h \leq \eta$.

Also $\varphi(y)\,\varphi(hy - x)$ is supported in the unit ball, and is uniformly in $C^{(1)}$. Thus in this case we can apply Proposition 1 and see that (4.3) is majorized by $c\,r^{-1/d}\,\chi_{B_2}(x) \leq c\,r^{-2\delta}\,\chi_{B_2}(x)$. (At this stage first-degree terms in P could have been allowed).

We now assume that $\eta < h \leq 1$, with η now fixed, $\eta > 0$. We examine the terms of degree 1 in y in the phase $P_\nu(y) - P_\mu(hy - x)$.

For this we recall that we assumed we had no first order terms in y in $P_\nu(y)$, so it follows that the first order term of the above are

$$-h\sum_{j=1}^{n} P_\mu^{(j)}(x)\,\cdot\,y_j\,, \quad \text{where}$$

$$P_\mu^{(j)}(x) \;=\; \sum_\alpha \alpha_j\,\mu_\alpha\,x^{\alpha - e_j}\,,$$

and $e_j = (0, \dots 1, 0 \dots)$ with 1 in the j^{th} component.

It then follows from Proposition 1 that (4.3) is majorized by

$$c\left(\sum_{j=1}^{n}\left|P_\mu^{(j)}\,(x)\right|\right)^{-1/d}\,. \tag{4.4}$$

We now divide the ball B_2 into two sets, E_μ and its complement. We define

$$E_\mu \;=\; \{x \in B_2 : \sum_{j=1}^{n}|P_\mu^{(j)}(x)| \;\leq\; \rho\},$$

and ρ will be chosen in terms of r momentarily.

In the complement of the set $x \in E_\mu$ we get (in view of 4.4) $c\rho^{-1/d}$ as a bound for (4.3). So for those x we estimate (4.3) by $c\rho^{-1/d}\chi_{B_2}(x)$. Note however that by Proposition 2,

$$|E_\mu| \;\leq\; c\left(\sum_{j}\sum_{\substack{\alpha \\ |\alpha - e_j| \geq 1}} \alpha_j|\mu_\alpha|\right)^{-1/d}\rho^{1/d}$$

and

$$\sum_{\substack{j \\ |\alpha - e_j| \geq 1}} \sum_{\alpha} \alpha_j |\mu_\alpha| \;\geq\; \sum_{\alpha} |\mu_\alpha| \;\geq r\,.$$

Thus for $x \in E_\mu$ we have as an estimate for (4.3), $c\chi_{E_\mu}(x)$, with $|E_\mu| \leq c(\rho/r)^{1/d}$.

Now we only need take $\delta = 1/6d$, $\rho = \bar{c}r^{1/3}$ with $\bar{c}$ appropriately small. Then $|E_\mu| \leq r^{-4\delta}$, and $\rho^{-1/d} =$ constant $r^{-2\delta}$ so (4.2) is completely proved. □

As an immediate consequence we have:

Corollary 4.1.

$$\begin{aligned}\left|\left(\Phi^{\nu}_{a_1} * \widetilde{\Phi}^{\mu}_{a_2}\right)(x)\right| \;&\leq\; c\,r^{-2\delta}\,\{a_1^{-n}\,\chi_{B_2}(x/a_1) \;+\; a_2^{-n}\,\chi_{B_2}(x/a_2)\} \\ &+ c\{a_1^{-n}\chi_{E\nu}(x/a_1) \;+\; a_2^{-n}\,\chi_{E_\mu}(x/a_2)\}\end{aligned} \tag{4.5}$$

We still assume that $r \leq |\nu| \leq 2r$, and $r \leq |\mu| \leq 2r$, but now a_1 and a_2 are any two positive numbers. Here both $|E_\nu|$ and $|E_\mu|$ are $\leq r^{-4\delta}$.

Proof. Consider the case $a_2 \geq a_1$. Then by rescaling by a_2^{-1} we may reduce matters to the case $a_2 = 1$, $a_{1/a_2} = h \leq 1$, which we had considered previously, but where the terms depending on a_1 and ν in the right-side are not present. The situation is symmetric in a_1 and a_2 and thus the same inequality hold when $a_1 \geq a_2$.

We now pass to the proof of Theorem 1. We pick $x \to \lambda(x) = \{\lambda_\alpha(x)\}$, and $x \longrightarrow a(x)$ arbitrary "stopping times," i.e. finite-valued measurable functions, with each $\lambda_\alpha(x)$ being real, and $a(x)$ positive. We assume also that $r \leq |\lambda(x)| \leq 2r$, for every x. Denote by T the linear operator $f \to T(f)$ given by

$$T(f)(x) \;=\; \int_{\mathbb{R}^n} f(x-y)\,\Phi^{\lambda(x)}_{a(x)}(y)\,dy.$$

It suffices to prove that

$$\|\,T\,\|_{L^2 \to L^2} \leq\; c\,r^{-\delta},$$

with the bound independent of the choice of the functions $\lambda(x)$ and $a(x)$.

Now $\|\,T\,\| = \|\,TT^*\,\|^{1/2}$, and TT^* is an operator with kernel $K(x,y)$

$$(TT^*f)(x) \;=\; \int K(x,y)\,f(y)\,dy\,, \text{ where}$$

$$K(x,y) \;=\; \left(\Phi^{\nu}_{a_1} * \widetilde{\Phi}^{\mu}_{a_2}\right)(x-y),$$

with

$$\nu = \lambda(x)\,, \quad \mu = \lambda(y)$$

$$a_1 = a(x)\,, \quad a_2 = a(y)\,.$$

Then by (4.5) it follows that

$$\left|\int K(x,y)\, f(y)\, \bar{g}(x)\, dx\, dy\right| \leq$$

$$c\, r^{-2\delta} \left(\int_{\mathbb{R}^n} M(f)(x)|g(x)|\, dx + \int_{\mathbb{R}^n} M(g)(y)|f(y)|\, dy\right) +$$

$$c\left(\int_{\mathbb{R}^n} \mathcal{M}_\epsilon(f)(x)|g(x)|\, dx + \int_{\mathbb{R}^n} \mathcal{M}_\epsilon(g)(y)|f(y)|\, dy\right)\,.$$

Here M is the standard n-dimensional maximal function, and $\mathcal{M}_\epsilon$ is the maximal function appearing in Section 3, with $\epsilon = r^{-4\delta}$. Thus by the usual L^2 estimates for M and the estimates given in Section 3 for $\mathcal{M}_\epsilon$ we have

$$|(TT^* f, g)| \leq c\, r^{-2\delta} \, \| f \|_{L^2} \| g \|_{L^2}$$

As a result $\| TT^* \| \leq c\, r^{-2\delta}$, and $\| T \| \leq c'\, r^{-\delta}$, and the theorem is proved. □

For the application below we need a slight restatement of the theorem which we state in two stages. We replace the isotropic norm $|\lambda| = \sum |\lambda_\alpha|$ by the non-isotropic norm $N(\lambda) = \sum |\lambda_\alpha|^{1/|\alpha|}$. Note that since $N(\lambda) \leq c|\lambda|$, when $N(\lambda) \geq 1$, then as a consequence of (4.1) we have

$$\| \sup_{\substack{N(\lambda)\geq r \\ a>0}} |(f * \Phi_a^\lambda)(x)\ \|_{L^2} \leq A r^{-\delta}\ \| f \|_{L^2}\,, \quad r \geq 1 \tag{4.6}$$

An immediate implication of (4.6) is

$$\| \sup_{\substack{N(\lambda)\geq 1 \\ a>0}} N(\lambda)^{\delta_1}\, |\,(f * \Phi_a^\lambda)\,(x)|\ \|_{L^2} \leq A\ \| f \|_{L^2} \tag{4.7}$$

whenever $\delta_1 < \delta$. Indeed,

$$\sup_{\substack{N(\lambda)\geq 1 \\ a>0}} N(\lambda)^{\delta_1}\, |(f * \Phi_a^\lambda\,(x)|$$

$$\leq \sum_{j=0}^{\infty} 2^{j\delta_1} \sup_{N(\lambda)\,\geq 2^j} |(f * \Phi_a^\lambda)(x)|\,,$$

and so (4.7) follows from (4.6) whenever $\delta_1 < \delta$, since each term in the sum has norm $\leq c2^{j(\delta_1 - \delta)}\ \| f \|_{L^2}$

5. The second main proposition

Let K be a Calderón-Zygmund kernel in $\mathbb{R}^n$. For our purposes we shall assume that K satisfies the following properties.

(a) K is a tempered distribution which agrees with a $C^{(1)}$ function $K(x)$, for $x \neq 0$

(b) $K^\wedge$, the Fourier transform of K, is an L^∞ function

(c) $|\partial_x^\alpha K(x)| \leq A|x|^{-n+|\alpha|}$, for $0 \leq |\alpha| \leq 1$.

Write as before $P_\lambda(x)$ for the real polynomial of degree d with coefficients $\lambda = (\lambda_\alpha)$, i.e. $P_\lambda(x) = \sum_{2\leq|\alpha|\leq d} \lambda_\alpha x^\alpha$; here again we assume that no first-degree terms are present. Considering the distribution which arises as the product of the function e^{iP_λ} with K, we can define the operator

$$T_\lambda(f)(x) = \int_{\mathbb{R}^n} e^{iP_\lambda(y)}\, K(y)\, f(x-y)\, dy$$

(at least for test functions f). We then have,

Theorem 2.

$$\| \sup_\lambda |T_\lambda(f)(x)| \, \| \leq A \| f \|_{L^2} .$$

The sup *is taken over all the real coefficients of* P_λ.

To prove the theorem we decompose the kernel K as

$$K = \sum_{j=-\infty}^{\infty} 2^{-nj}\, \varphi^{(j)}\, (2^{-j} \cdot x) \tag{5.1}$$

where the $\varphi^{(j)}$ are each $C^{(1)}$ functions supported in $1/4 < |x| \leq 1$; they satisfy $|\partial_x^\alpha \varphi^{(j)}| \leq A$, $0 \leq |\alpha| \leq 1$, uniformly in j; and $\int_{\mathbb{R}^n} \varphi^{(j)}(x)dx = 0$, all j. (For this decomposition see e.g. [St2], Chapter 13)

Now for each λ, write $K_\lambda = K_\lambda^+ + K_\lambda^-$, where K_λ^- is the sum in (5.1) of all terms where $2^j < 1/N(\lambda)$, and K_λ^+ is the sum of the term when $2^j \geq 1/N(\lambda)$. We also write $T_\lambda = T_\lambda^+ + T_\lambda^-$, with

$$T_\lambda^\pm(f)(x) = \int e^{iP_\lambda(y)}\, K_\lambda^\pm\, (y)\, f(x-y)\, dy\,.$$

We estimate $\sup_\lambda |T_\lambda^+(f)|$ and $\sup_\lambda |T_\lambda^-(f)|$ separately. The majorization of T_λ^- is easily handled by standard estimates.

Notice that $K_\lambda^-(x)$ is supported where $|x| \leq 1/N(\lambda)$, (and agrees with $K(x)$ when $|x| \leq 1/4N(\lambda)$). Thus on the support of $K_\lambda^-(x)$ we have

$$|e^{iP_\lambda(x)} - 1| \leq c \sum_{2\leq|\alpha|\leq d} |\lambda_\alpha|\,|x^\alpha| \leq c' \sum_{2\leq|\alpha|\leq d} N(\lambda)^{|\alpha|}\,|x|^{|\alpha|} \leq c\,N(\lambda)|x|\,,$$

since $N(\lambda)|x| \leq 1$, and $|\lambda_\alpha| \leq N(\lambda)^{|\alpha|}$. Hence,

$$T_\lambda^-(f)(x) = \int K_\lambda^-(y)\,f(x-y)\,dy \,+\, O(N(\lambda)\int_{|y|\,\leq\,1/N(\lambda)} |f(x-y)|\,|y|^{-n+1}\,dy)\,.$$

Next, we observe that the sup in λ of the first term on the right-side is dominated by the truncated-singular-integral maximal function, and the second term by the standard maximal function. Therefore we obtain

$$\|\sup_\lambda |T_\lambda^-(f)(x)|\,\|_{L^2} \leq A\,\|\,f\,\|_{L^2}\ .$$

Turning to T_λ^+ we have

$$T_\lambda^+(f)(x) = \sum_{2^j>1/N(\lambda)} \int e^{iP_\lambda(y)}\,2^{-nj}\,\varphi^{(j)}(2^{-j}y)\,f(x-y)\,dy\,. \tag{5.2}$$

We introduce the notation $2^j\circ\lambda$ to denote $2^j\circ\lambda = (2^{j|\alpha|}\,\lambda_\alpha)$ when $\lambda = (\lambda_\alpha)$. Then clearly $P_\lambda(y) = P_{2^j\circ\lambda}\,(2^{-j}\cdot y)$, and thus $e^{iP_\lambda(y)}\,2^{-nj}\,\varphi^{(j)}\,(2^{-j}\cdot y)$ can be written as

$$e^{iP_{2^j\circ\lambda}(2^{-j}y)}\,2^{-nj}\,\varphi^{(j)}\,(2^{-j}\cdot y) = {}^{(j)}\Phi^{2^j\circ\lambda}_{2^j},$$

with ${}^{(j)}\Phi^\mu(x) = e^{iP_\mu(x)}\,\varphi^{(j)}(x)$. We now apply Theorem 1 (that is inequality (4.7)) to each of these terms in (5.2) and we see that $\|\sup_\lambda |T_\lambda^+(f)(x)|\,\|_{L^2}$ is dominated by

$$c\sum_{2^j>1/N(\lambda)} N(2^j\circ\lambda)^{-\delta_1}\ \|\,f\,\|_{L^2}\ .$$

However, $\sum_{2^j>1/N(\lambda)} N(2^j\circ\lambda)^{-\delta_1} = N(\lambda)^{-\delta_1} \sum_{2^j>1/N(\lambda)} 2^{-j\delta_1} \leq c < \infty$, because $N(2^j\circ\lambda) = 2^j N(\lambda)$, and $\delta_1 > 0$. With this our desired estimate for $\sup_\lambda |T_\lambda(f)|$ is achieved.

Remark. A simple consequence of the above arguments is that the maximal operator in Theorem 2 is also bounded on L^p, $1 < p < \infty$.

References

[AKC] G.I. Arkhipov, A.A. Karacuba, and V.N.Čubarikov, *Multiple Trigonometric Sums*, Proc. of the Steklov Institute Math. **151**, (1980).

[C] M. Christ, *Hilbert transform along curves I. Nilpotent groups.* Ann. of Math. (2) **122** (1985), no. 3, 575–596.

[CCW] A. Carbery, M. Christ, and J. Wright, *Multi-dimensional van der Corput and sub-level set estimates,* J. Amer. Math. Soc. **12** (1999), no. 4, 981–1015.

[L] M.T. Lacey, *Carleson's theorem with quadratic phase functions,* preprint.

[LT] M. Lacey and C. Thiele, *A proof of the boundedness of the Carleson operator,* Math. Res. Lett. **7** (2000), no. 4, 361–370.

[RS] F. Ricci and E. Stein, *Harmonic analysis on nilpotent groups and singular integrals I. Oscillatory integrals.* J. Funct. Anal. **73** (1987), no. 1, 179–194.

[Sj] P. Sjölin, *Convergence almost everywhere of certain singular integrals and multiple Fourier series,* Ark. Mat. **9**, 65–90. (1971).

[St1] E.M. Stein, *Oscillatory integrals related to Radon-like transforms.* Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993). J. Fourier Anal. Appl. **1995**, Special Issue, 535–551.

[St2] E.M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA.
E-mail address: `stein@math.princeton.edu`

Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA.
E-mail address: `wainger@math.wisc.edu`