ANNALES DE L'INSTITUT FOURIER

FULVIO RICCI
ELIAS M. STEIN

# Multiparameter singular integrals and maximal functions

*Annales de l'institut Fourier*, tome 42, n° 3 (1992), p. 637-670
<http://www.numdam.org/item?id=AIF_1992__42_3_637_0>

Ann. Inst. Fourier, Grenoble
42, 3 (1983), 637-670.

# MULTIPARAMETER SINGULAR INTEGRALS AND MAXIMAL FUNCTIONS

by F. RICCI and E.M. STEIN

## 0. Introduction.

The purpose of this paper is to study singular integrals and maximal functions in $\mathbf{R}^n$ which reflect a $k$-parameter homogeneity. Before stating our main results, we wish to describe briefly some of the earlier developments relevant to our work.

*Background*

There is first the well-known case corresponding to one-parameter dilations. Here we are considering the mappings

$$(x_1, \ldots, x_n) \mapsto (\delta^{\lambda_1}x_1, \ldots, \delta^{\lambda_n}x_n) = \delta \cdot x$$

for $\delta > 0$, where $\lambda_j$ are fixed positive exponents.

In the standard Calderón-Zygmund theory for this setting we consider operators of the form $Tf = f * K$, where, to begin with, $K$ is homogeneous of the critical degree, i.e. $K(\delta \cdot x) = \delta^{-\lambda}K(x)$, with $\lambda = \lambda_1 + \cdots + \lambda_n$ ; $K$ satisfies some smoothness condition and an appropriate cancellation condition.

This theory can be recast in the more general setting where $K$ is not necessarily homogeneous but can be written in the form

$$K(x) = \sum_{i \in \mathbf{Z}} \mu_i^{(i)}(x), \tag{0.1}$$

*Key words :* Singular integrals - Maximal functions - Calderón-Zygmund kernels - Product domains.

*A.M.S. Classification :* 42B20 - 42B25.

where the functions $\mu_i^{(i)}$ arise as $\mu_i^{(i)}(x) = 2^{-\lambda i}\mu^{(i)}(2^{-i}\cdot x)$; the $\mu^{(i)}$ satisfy a uniform $L^1$-Dini condition and the cancellation condition

$$\widehat{\mu^{(i)}}(0) = 0. \tag{0.2}$$

A second development of this one-parameter theory deals with singular integrals carried on orbits of the dilations. When the orbits are « curved » (i.e. if $\lambda_i \neq \lambda_j$ for some $i$ and $j$), then certain decay estimates of the Fourier transform come into play. In the present context these take the place of the previous regularity assumptions on the $\mu^{(i)}$, and are of the form

$$|\widehat{\mu^{(i)}}(\xi)| \leqslant A(1+|\xi|)^{-\varepsilon}. \tag{0.3}$$

An additional interpolation argument using the standard theory then gives the $L^p$-results. For this see [9], [18].

Consider next the $n$-parameter theory in $\mathbf{R}^n$ (sometimes called the "product theory"). Here we deal with dilations

$$(x_1, \ldots, x_n) \mapsto (\delta_1 x_1, \ldots, \delta_n x_n)$$

where $\delta = (\delta_1, \ldots, \delta_n) \in \mathbf{R}_+^n = (\mathbf{R}_+)^n$. Again we consider $K$ decomposed analogously to (0.1),

$$K(x) = \sum_{I \in \mathbf{Z}^n} \mu_I^{(I)}(x), \tag{0.4}$$

where $I = (i_1, \ldots, i_n)$, $\mu_I^{(I)}(x) = 2^{-|I|}\mu^{(I)}(2^{-i_1}x_1, \ldots, 2^{-i_n}x_n)$ and the cancellation conditions are

$$\widehat{\mu^{(I)}}(\xi) = 0, \qquad \text{whenever } \xi_j = 0 \text{ for some } j. \tag{0.5}$$

There are two levels of generality here depending on the regularity required of the $\mu^{(I)}$. The initial case, treated in [11], allows one to treat the situation when the $\mu^{(I)}$ satisfy uniform Hölder estimates. A more refined approach, developed in [12] (see also [2], [10], [15]) makes it possible to require less regularity of the $\mu^{(I)}$.

*Results of this paper.*

Turning now to the subject proper of this paper, we consider a general family of $k$-parameter dilations, with $1 \leqslant k \leqslant n$. The theory in this setting was initiated in [14]; some other works which have a bearing on our paper are [6], [8], [19], [21].

To describe these dilations we assume we are given fixed exponents $\{\lambda_{ij}\}$, with $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant k$. The dilations are then

$$(0.6) \qquad (x_1, \ldots, x_n) \mapsto (\delta_1^{\lambda_{11}} \ldots \delta_k^{\lambda_{1k}} x_1, \ldots, \delta_1^{\lambda_{n1}} \ldots \delta_k^{\lambda_{nk}} x_n),$$

where $\delta \in \mathbf{R}_+^k$. We associate to the exponents $\{\lambda_{ij}\}$ the matrix $\Lambda = \{\lambda_{ij}\}$ and write $\delta^\Lambda x$ for the action given by (0.6). We also set $2^{\Lambda I} = \delta^\Lambda$ when $\delta = (2^{i_1}, \ldots, 2^{i_k})$ and $I = (i_1, \ldots, i_k)$.

The kernels we shall be considering are of the form

$$(0.7) \qquad K(x) = \sum_{I \in \mathbf{Z}^k} \mu_I^{(I)}(x),$$

where $\mu^{(I)}$ are appropriate distributions and

$$\mu_I^{(I)}(x) = \det(2^{-\Lambda I})\mu^{(I)}(2^{-\Lambda I}x).$$

From our point of view the two main issues are:

What are the natural cancellation conditions to be required of the $\mu^{(I)}$?

What are the appropriate regularity conditions to be imposed on the $\mu^{(I)}$?

To answer the first question we consider the orbits of the dilations in the $\xi$-space. Generically, the maximum dimension of an orbit is $k$. The cancellation condition then becomes the requirement that $\widehat{\mu^{(I)}}(\xi) = 0$, whenever $\xi$ belongs to a subspace which is the union of lower-dimensional orbits.

Appropriate regularity properties are, first, the uniform decay of the Fourier transforms,

$$(0.8) \qquad |\widehat{\mu^{(I)}}(\xi)| \leqslant A(1+|\xi|)^{-\varepsilon}.$$

The second regularity condition is of a new type. It is not required in the usual one-parameter theory, nor did it appear explicitely in the product theory.

It is that there is a finite measure $\sigma$ so that

$$(0.9) \qquad |\mu^{(I)}| \leqslant \sigma, \qquad \forall I.$$

Under these cancellation and regularity conditions on the $\mu^{(I)}$ our main result is that when $K$ is given by (0.7), the operator $f \mapsto f * K$ is bounded from $L^p(\mathbf{R}^n)$ to itself, if $1 < p < \infty$.

As a consequence we can deal with kernels $K$ which are homogeneous of the critical degree under our dilations $\delta^{\Lambda}$; as a special case the kernels may be carried by single orbits of these dilations.

The relevance of the cancellation conditions in this theory can be seen from the following example in $\mathbf{R}^3$. The operator

$$(Tf)(x,y,z) = \text{p.v.} \int f(x-s, y-t, z-|st|) \frac{ds}{s} \frac{dt}{t}$$

is bounded on every $L^p(\mathbf{R}^3)$, $1 < p < \infty$ [8], [14], [21], whereas the operator

$$(Uf)(x,y,z) = \text{p.v.} \int f(x-s, y-t, z-st) \frac{ds}{s} \frac{dt}{t}$$

is not even bounded on $L^2(\mathbf{R}^3)$.

Another interesting situation arises if we consider $\mathbf{R}^2$ with the one-parameter dilations $(x,y) \mapsto (\delta x, \delta^{-1} y)$. It is important to note that one of the exponents is negative, and so what follows is not deducible from the previously described one-parameter theory. A basic singular integral here is the Hilbert transform along the hyperbola

$$f \mapsto \int_{-\infty}^{+\infty} f\left(x-t, y-\frac{1}{t}\right) \frac{dt}{t} = \lim_{\varepsilon \to 0} \int_{\varepsilon}^{1/\varepsilon} \left(f\left(x-t, y-\frac{1}{t}\right) - f\left(x+t, y+\frac{1}{t}\right)\right) \frac{dt}{t}.$$

The cancellation conditions are $\widehat{\mu^{(I)}}(0,0) = 0$, and so we can envisage other variants of this Hilbert transform, e.g.

$$f \mapsto \lim_{\varepsilon \to 0} \int_{\varepsilon}^{1/\varepsilon} \left(f\left(x-at, y-\frac{1}{t}\right) - f\left(x+bt, y+\frac{1}{t}\right)\right) \frac{dt}{t},$$

where $a$ and $b$ are arbitrary non-zero real numbers. Other variants, involving stronger cancellations, have been treated previously and the results were limited to $8/5 < p < 8/3$ [19].

We describe now the maximal functions which can be treated by our methods. Experience shows that there are two ways one can build up maximal functions from simpler elements. The first is by majorization, and is illustrated in $\mathbf{R}^1$ by the fact that if $(Mf)(x) = \sup_{\delta > 0} |(f * \varphi^{\delta})(x)|$, where $\varphi^{\delta}(x) = \delta^{-1} \varphi(x/\delta)$, we have the usual results for $M$ if $\varphi$ is majorized by an integrable symmetric decreasing function. The second is by the « method of rotations », which allows one to obtain results in $\mathbf{R}^n$ from maximal functions associated to lower-dimensional varieties. Our general theorem subsumes both, and is suggested by the uniformity condition (0.9).

Suppose $\mu$ is a finite positive measure on $\mathbf{R}^n$. With $\delta^A$ denoting our $k$-dimensional dilations, we write $\mu_\delta$ for the measure given by

$$\int f\, d\mu_\delta = \int f(\delta^A x)\, d\mu(x).$$

Define

$$(Mf)(x) = \sup_{\delta \in \mathbf{R}^k_+} |f * \mu_\delta(x)|.$$

The condition we impose on $\mu$ is that there is another positive measure $\nu$ so that

$$\sup_{1 \leqslant \delta_j < 2} \mu_\delta \leqslant \nu. \tag{0.10}$$

Under these assumptions, the maximal function $M$ is bounded on $L^p(\mathbf{R}^n)$, $1 < p \leqslant \infty$.

Note that in one dimension, if $d\mu = \varphi(x)\, dx$ and $\varphi$ has a non-decreasing symmetric majorant in $L^1$, then $\mu$ satisfies (0.10). Conversely, if (0.10) holds in one dimension, then $\mu$ is of the above kind. Observe also that a measure $\mu$ supported on a compact portion of an orbit, with bounded density, automatically satisfies (0.10). Previous results for related maximal functions can be found in [3], [4], [6], [8].

We now indicate the main features of the proof of our principal result (Theorem 5.1). The argument, which is somewhat complicated, proceeds in three steps.

First, in the setting of the $n$-parameter dilations, a variant of the conclusion is proved as a consequence of the Journé's theory. Here a modification of the uniformity condition (0.9) already enters.

Next, the decay estimates of the Fourier transform are incorporated by interpolating an $L^2$ with the previously proved $L^p$-estimates. This gives our desired conclusion in the context of the full $n$-parameter dilations (Theorem 2.4).

Finally, the $k$-parameter result is deduced from the $n$-parameter case. This cannot be done directly, because the cancellation conditions required in the former case are not as restrictive as those required in the latter case. So a further decomposition of the elements $\mu^{(I)}$ appearing in (0.7) is needed. This is carried out in Section 5.

Once our general result has been established, the theorem for homogeneous kernels is deduced from it in Section 6. The maximal theorem is proved in Section 7, and is based on an $n$-parameter maximal theorem proved in Section 3.

## 1. Singular kernels adapted to the $n$-parameter dilations on $\mathbf{R}^n$.

On $\mathbf{R}^n$ we consider the $n$-parameter dilations

$$(x_1, \dots, x_n) \mapsto (\delta_1 x_1, \dots, \delta_n x_n)$$

with $\delta_1, \dots, \delta_n > 0$. We will often consider only dyadic dilations, defined by multi-indices $I = (i_1, \dots, i_n) \in \mathbf{Z}^n$ and corresponding to taking $\delta_j = 2^{i_j}$. To abbreviate, we write $2^I x$ for $(2^{i_1}x_1, \dots, 2^{i_n}x_n)$.

If $f$ is a function on $\mathbf{R}^n$, we write $f_I(x) = 2^{-|I|} f(2^{-I}x)$, where $|I| = i_1 + \cdots + i_n$; consistent with this, if $\mu$ is a distribution on $\mathbf{R}^n$, we define $\mu_I$ by the identity

$$\langle \mu_I, f \rangle = \langle \mu, f(2^I \cdot) \rangle,$$

for every $C^\infty$-function $f$ with compact support.

Proposition 1.1. – *For every $I \in \mathbf{Z}^n$* let $\mu^{(I)}$ *be a distribution supported on the unit cube in $\mathbf{R}^n$, and assume that*

(i) $\widehat{\mu^{(I)}}(\xi) = 0$ *if some coordinate of $\xi$ is zero;*

(ii) $|\widehat{\mu^{(I)}}(\xi)| \leqslant C(1+|\xi|)^{-\varepsilon}$ *for some $C$, $\varepsilon > 0$.*

*Then the series $\sum_{I \in \mathbf{Z}^n} \mu_I^{(I)}$ converges in the sense of distributions to a distribution $K$ and the convolution operator with kernel $K$ is bounded on $L^2(\mathbf{R}^n)$ with a norm that depends only on the constants $C$ and $\varepsilon$ in* (ii).

*Proof.* – It is sufficient to prove that the series

$$\sum_{I \in \mathbf{Z}^n} |\widehat{\mu^{(I)}}(2^I \xi)| \tag{1.1}$$

converges almost everywhere to a bounded function. Let $\varphi(x)$ be a Schwartz function on $\mathbf{R}^n$ that equals one on a neighborhood of $\operatorname{supp} \mu^{(I)}$

for every $I$. Then

$$\widehat{\mu^{(I)}}(\xi) = (\widehat{\mu^{(I)}} * \hat{\varphi})(\xi) \qquad \text{and} \qquad \partial^\alpha \widehat{\mu^{(I)}}/\partial\xi^\alpha = \widehat{\mu^{(I)}} * (\partial^\alpha \hat{\varphi}/\partial\xi^\alpha).$$

So for every $\alpha$

$$(1.2) \qquad |(\partial^\alpha \widehat{\mu^{(I)}}/\partial\xi^\alpha)(\xi)| \leqslant C_\alpha(1+|\xi|)^{-\varepsilon}.$$

Since $\widehat{\mu^{(I)}}$ vanishes on each coordinate subspace, (1.2) implies that for each choice of the coordinates $j_1, \ldots, j_k$ we have

$$|\widehat{\mu^{(I)}}(\xi)| \leqslant C|\xi_{j_1}| \cdots |\xi_{j_k}|(1+|\xi|)^{-\varepsilon}.$$

Take now $\xi = (\xi_1, \ldots, \xi_n)$ with $\xi_j \neq 0$ for every $j$. By simply relabeling the terms in (1.1), we can assume that $|\xi_j| \sim 1$ for all $j$. Let $I \in \mathbf{Z}^n$ and assume that $i_{j_1}, \ldots, i_{j_k} \leqslant 0$ and all the other components of $I$ are positive. Then

$$\begin{aligned} |\widehat{\mu^{(I)}}(2^I\xi)| &\leqslant C 2^{i_{j_1}+\cdots i_{j_k}}(1+|2^I\xi|)^{-\varepsilon} \\ &\leqslant C 2^{-\frac{\varepsilon}{n}|I|}. \end{aligned}$$

Q.E.D.

*Remark.* – In Proposition 1.1 we assume that the measures $\mu^{(I)}$ are supported on the unit cube. This is only a normalization introduced in order to simplify its statement. If one asssumes that the supports of the $\mu^{(I)}$ are all contained in a fixed cube, the size of the cube will simply affect the norm of the resulting operator. Similarly, the integer lattice $\mathbf{Z}^n$ can be replaced by any co-compact lattice in $\mathbf{R}^n$.

These remarks apply to most results in this paper, and will be used without further comment.

## 2. $L^p$-boundedness in the $n$-parameter case.

Following [12] and [15], we define the Calderón-Zygmund norms for kernels on $\mathbf{R}^n$ by the following inductive procedure.

(i) Given $\delta > 0$, the $CZ_\delta$-norm of a kernel $K(x)$ on the real line is given by

$$(2.1) \qquad \|K\|_{CZ_\delta(\mathbf{R})} = \|K\|_{2,2} + \sup_{\substack{\gamma>2 \\ h\neq 0}} \gamma^\delta \int_{|x|>\gamma|h|} |K(x+h)-K(x)|\, dx$$

($\|\ \|_{p,p}$ denotes the operator norm on $L^p$).

(ii) For a kernel $K(x_1, \dots, x_n)$ on $\mathbf{R}^n$ $(n \geqslant 2)$,

$$(2.2) \quad \|K\|_{CZ_\delta(\mathbf{R}^n)} = \|K\|_{2,2} + \sum_{j=1}^{n} \sup_{\substack{\gamma > 2 \\ h \neq 0}} \gamma^\delta \int_{|x_j| > \gamma |h|} \|K(\cdot, x_j + h, \cdot) - K(\cdot, x_j, \cdot)\|_{CZ_\delta(\mathbf{R}^{n-1})}\, dx_j .$$

We will also use the following result [12], [15].

Proposition 2.1. — *Assume that* $\|K\|_{CZ_\delta(\mathbf{R}^n)} < +\infty$ *for some* $\delta > 0$. *Then the operator* $Tf = K * f$ *is bounded on* $L^p(\mathbf{R}^n)$, *for* $1 < p < \infty$, *with a norm that depends only on* $p$ *and on* $\|K\|_{CZ_\delta(\mathbf{R}^n)}$.

Dealing with kernels defined as sums of dyadically scaled terms, the following remark is particularly useful.

Lemma 2.2. — *For each* $j \in \mathbf{Z}$ let $\psi^{(j)}$ *be a function supported on the interval* $[-1, 1]$ *taking values in a normed space* $B$, *and assume that for some positive constants* $C$, $\delta$ *and every* $h \in \mathbf{R}$

$$(2.3) \qquad \int \|\psi^{(j)}(x+h) - \psi^{(j)}(x)\|_B\, dx \leqslant C|h|^\delta .$$

Let $K(x) = \sum_{j \in \mathbf{Z}} 2^{-j}\psi^{(j)}(2^{-j}x)$. *Then, if* $\gamma > 2$ *and* $h \neq 0$,

$$\int_{|x| > \gamma |h|} \|K(x+h) - K(x)\|_B\, dx \leqslant A_\delta C \gamma^{-\delta} .$$

*Proof.* — Without loss of generality, we can assume that $|h| \sim 1$ (this may require relabeling the $\psi^{(j)}$). Then

$$\begin{aligned} &\int_{|x| > \gamma} \|K(x+h) - K(x)\|_B\, dx \\ &\qquad \leqslant \sum_{2^{j+1} > \gamma} \int_{|x| > \gamma} \|2^{-j}\psi^{(j)}(2^{-j}(x+h)) - 2^{-j}\psi^{(j)}(2^{-j}x)\|_B\, dx \\ &\qquad \leqslant C \sum_{2^{j+1} > \gamma} 2^{-j\delta} \\ &\qquad = A_\delta C \gamma^{-\delta} . \end{aligned}$$

Q.E.D.

We prove now our basic lemma.

We set

$$\Delta_h^{(k)} f = f(x_1, \ldots, x_k + h, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n).$$

For $h = (h_1, \ldots, h_n)$ we also set $\Delta_h = \Delta_{h_1}^{(1)} \ldots \Delta_{h_n}^{(n)}$.

LEMMA 2.3. – *For each $I \in \mathbf{Z}^n$, let $\varphi^{(I)}$ be a function supported on the unit cube in $\mathbf{R}^n$. Assume that*

(i) $\widehat{\varphi^{(I)}}(\xi) = 0$ *if some coordinate of $\xi$ is zero;*

(ii) *for every permutation $(k_1, \ldots, k_n)$ of $(1, \ldots, n)$ the quantity*

$$(2.4) \qquad \sup_{\substack{i_{k_1} \in \mathbf{Z} \\ 0<|h_{k_1}|<2}} |h_{k_1}|^{-\delta} \int \Bigg( \ldots \Bigg( \sup_{\substack{i_{k_n} \in \mathbf{Z} \\ 0<|h_{k_n}|<2}} |h_{k_n}|^{-\delta} \int |\Delta_h \varphi^{(I)}(x)| \, dx_{k_n} \Bigg) \ldots \Bigg) dx_{k_1}$$

*is bounded by a constant $C$ for some $\delta > 0$. Then, if $K(x) = \sum_{I \in \mathbf{Z}^n} \varphi_I^{(I)}(x)$, the norm $\|K\|_{CZ_\delta(\mathbf{R}^n)}$ is bounded by a constant that depends only on $C$ and $\delta$.*

*Proof.* – We proceed by induction on the dimension $n$. For $n = 1$, the functions $\varphi^{(i)}$ are uniformly in $L^{1,\varepsilon}(\mathbf{R})$ and have mean value zero. The conclusion then follows from Proposition 1.1 and from Lemma 2.2.

Assume that the statement is true in $n - 1$ dimensions. Since (2.4) implies that the $\varphi^{(I)}$ are uniformly in $L^{1,\varepsilon}(\mathbf{R}^n)$, the boundedness of $\|K\|_{2,2}$ follows from Proposition (1.1). We then estimate $\int_{|x_1|>\gamma|h_1|} \|K(x_1 + h_1, \cdot) - K(x_1, \cdot)\|_{CZ_\delta(\mathbf{R}^{n-1})} \, dx_1$ by means of Lemma 2.2. Observe that

$$K(_1, \cdot) = \sum_{i_1 \in \mathbf{Z}} \sum_{I' \in \mathbf{Z}^{n-1}} \varphi_{i_1, I'}^{(i_1, I')}(x_1, \cdot),$$

so that we can take as our function $\psi^{(i_1)}(x_1)$ with values in the space of $CZ_\delta$-kernels the function

$$\psi^{(i_1)}(x_1) = \sum_{I' \in \mathbf{Z}^{n-1}} \varphi_{0,I'}^{(i_1, I')}(x_1, \cdot).$$

The integral in (2.3) is

$$\int \|\psi^{(i_1)}(x_1+h_1)-\psi^{(i_1)}(x_1)\|_{CZ_\delta(\mathbf{R}^{n-1})}\,dx_1$$
$$= \int \Big\| \sum_{I'\in\mathbf{Z}^{n-1}} \Delta^{(1)}_{h_1}\varphi^{(i_1,I')}_{0,I'}(x_1,\cdot)\Big\|_{CZ_\delta(\mathbf{R}^{n-1})} dx_1 .$$

By the inductive hypothesis, this integral is smaller than a constant times the sum of the integrals

$$\int\Bigg( \sup_{\substack{i_{k_2}\in\mathbf{Z}\\ 0<|h_{k_2}|<2}} |h_{k_2}|^{-\delta}\int\dots \sup_{\substack{i_{k_n}\in\mathbf{Z}\\ 0<|h_{k_n}|<2}} |h_{k_n}|^{-\delta}\int |\Delta_h\varphi^{(i_1,I')}(x)|\,dx_{k_n}\dots dx_{k_2}\Bigg)dx_1$$

over all permutations $(k_2,\dots,k_n)$ of $(2,\dots,n)$. By (2.4), this expression is bounded by a constant times $|h_1|^\delta$. Q.E.D.

By Proposition 2.1, the kernel $K$ in the statement of Lemma 2.3 defines a bounded convolution operator on $L^p(\mathbf{R}^n)$ for $1<p<\infty$. We next prove that the same conclusion holds under less restrictive hypotheses.

We will use the following notation: if $\nu$ is a measure on the coordinate subspace of $\mathbf{R}^n$ with coordinates $(x_{k_1},\dots,x_{k_\ell})$, we denote by $p_{k_l}(\nu)$ the push-forward of $\nu$ onto the subspace with coordinates $(x_{k_1},\dots,x_{k_{\ell-1}})$, i.e.

$$\int_{\mathbf{R}^{\ell-1}} f(x_{k_1},\dots,x_{k_{\ell-1}})\,dp_{k_\ell}(\nu) = \int_{\mathbf{R}^\ell} f(x_{k_1},\dots,x_{k_{l\ell}})\,d\nu .$$

In particular, if $\ell=1$, $p_{k_\ell}(\nu)=\int d\nu$. We also note that if $\{\nu_\alpha\}_{\alpha\in A}$ is a set of positive measures, bounded from above by a measure $\mu$, then $\sup_\alpha\nu_\alpha$ is a well-defined positive measure.

THEOREM 2.4. – *For each $I\in\mathbf{Z}^n$, let $\mu^{(I)}$ be a measure supported on the unit sube, such that*

(i) $\widehat{\mu^{(I)}}(\xi)=0$ *if some coordinate of $\xi$ is zero;*

(ii) $|\widehat{\mu^{(I)}}(\xi)|\leqslant C(1+|\xi|)^{-\varepsilon}$ for some $C$, $\varepsilon>0$;

(iii) *for every permutation $(k_1,\dots,k_n)$ of $(1,\dots,n)$ we have*

$$\sup_{i_{k_1}\in\mathbf{Z}} p_{k_1}(\sup_{i_{k_2}\in\mathbf{Z}} p_{k_2}(\dots \sup_{i_{k_n}\in\mathbf{Z}} p_{k_n}(|\mu^{(I)}|)\dots))\leqslant C. \tag{2.5}$$

*Then the convolution operator defined by the kernel* $K = \sum_{I \in \mathbf{Z}^n} \mu_I^{(I)}$ *is bounded on* $L^p(\mathbf{R}^p)$ *for* $1 < p < \infty$, *with a norm that only depends on* $p$, $C$, $\varepsilon$.

In order to understand condition (iii), consider the case of absolutely continuous measures $d\mu^{(i,j)} = f^{(i,j)}(x,y)\,dx\,dy$ in the plane. For each $i_0 \in \mathbf{Z}$, we take all the functions $f^{(i_0,j)}$, i.e. those that will be dilated above the interval where $x \sim 2^{i_0}$, and we require that the functions $g^{(i_0,j)}(x) = \int |f^{(i_0,j)}(x,y)|\ dy$ are majorized for $j \in \mathbf{Z}$ by a common function $g^{(i_0)}(x)$ with $\|g^{(i_0)}\|_1 \leqslant C$. A similar condition is imposed on $f^{(i,j_0)}(x,y)$, for fixed $j_0$, when the rôles of the two coordinates are interchanged.

*Proof.* – Let $\varphi(t)$ be a $C^\infty$-function with compact support on the line such that $\int \varphi(t)\,dt = 1$. For $\ell \geqslant 1$ define

$$\eta_{-\ell}(t) = 2^\ell\varphi(2^\ell t) - 2^{\ell-1}\varphi(2^{\ell-1}t),$$

so that the series $\varphi_0 + \sum_{\ell=1}^{\infty} \eta_{-\ell}$ converges to $\delta_0$ in $\mathscr{S}'(\mathbf{R})$.

Similarly on $\mathbf{R}^n$

$$\begin{aligned}\delta_0 &= \left(\varphi_0(x_1) + \sum_{\ell=1}^{\infty} \eta_{-\ell}(x_1)\right) \cdots \left(\varphi_0(x_n) + \sum_{\ell=1}^{\infty} \eta_{-\ell}(x_n)\right)\\ &= \sum_{L \in \mathbf{N}^n} \psi_L(x).\end{aligned}$$

If $L = (\ell_1, \ldots, \ell_n)$, then $\int \psi_L(x)\,dx_j = 0$ for every $x_1, \ldots, x_{j-1}$, $x_{j+1}, \ldots, x_n$ if and only if $\ell_j > 0$.

We then write

$$\mu^{(I)} = \sum_{L \in \mathbf{N}^n} \mu^{(I)} * \psi_L$$

and for each $L \in \mathbf{N}^n$ we define

$$K_L = \sum_{I \in \mathbf{Z}^n} (\mu^{(I)} * \psi_L)_I.$$

Like in [5], [16] the proof reduces to showing that

(a) $\|K_K\|_{2,2} \leqslant C2^{-\alpha|L|}$ for some $\alpha > 0$;

(b) if $1 < p < \infty$, then $\|K_L\|_{p,p} \leqslant C_{\delta,p}2^{\delta|L|}$ for every $\delta > 0$.

To prove (a) we use Proposition 1.1. The support of $\mu^{(I)} * \psi_L$ is contained in a fixed cube independently of $I$ and $L$, and its Fourier transform vanishes at each point with some zero coordinate. We also have

$$
\begin{aligned}
|(\mu^{(I)}*\psi_L)\hat{}(\xi)| &= |\widehat{\mu^{(I)}}(\xi)|\,|\widehat{\psi_L}(\xi)| \\
&\leqslant C(1+|\xi|)^{-\varepsilon}\prod_{j=1}^{n}|\hat{\varphi}(2^{-\ell_j}\xi_j)-\hat{\varphi}(2^{-\ell_j+1}\xi_j)|
\end{aligned}
$$

where the factor $|\hat{\varphi}(2^{-\ell_j}\xi_j)-\hat{\varphi}(2^{-\ell_j+1}\xi_j)|$ must be replaced by $|\hat{\varphi}(\xi_j)|$ if the corresponding $\ell_j$ is zero. By the smoothness of $\hat{\varphi}$, we have

$$
|\hat{\varphi}(2^{-\ell_j}\xi_j)-\hat{\varphi}(2^{-\ell_j+1}\xi_j)| \leqslant C2^{-\frac{\varepsilon}{2n}\ell_j}|\xi_j|^{\frac{\varepsilon}{2n}};
$$

therefore

$$
\begin{aligned}
|(\mu^{(I)}*\psi_L)\hat{}(\xi)| &\leqslant C(1+|\xi|)^{-\varepsilon}2^{-\frac{\varepsilon}{2n}|L|}\prod_{j=1}^{n}(1+|\xi_j|)^{\frac{\varepsilon}{2n}} \\
&\leqslant C(1+|\xi|)^{-\varepsilon/2}2^{-\frac{\varepsilon}{2n}|L|}.
\end{aligned}
$$

By Proposition 1.1, $\|K_L\|_{2,2} \leqslant 2^{-\frac{\varepsilon}{2n}|L|}$ and this proves (a).

To prove (b) we apply Lemma 2.3. We have

$$
\Delta_h(\mu^{(I)}*\psi_L) = \mu^{(I)} * (\Delta_h\psi_L).
$$

Since $\psi_L$ is a tensor product of functions in one variable,

$$
\Delta_h\psi_L(x) = \Delta_{h_1}^{(1)}\ldots\Delta_{h_n}^{(n)}\psi_L(x) = (\Delta_{h_1}\eta_{-\ell_1}(x_1))\ldots(\Delta_{h_n}\eta_{-\ell_n}(x_n))
$$

where $\eta_{-\ell_j}$ has to be replaced by $\varphi$ if $\ell_j = 0$. For each $\delta > 0$,

$$
|\Delta_{h_j}\eta_{-\ell_j}(x_j)| \leqslant C2^{\delta\ell_j}|h_j|^{\delta}\chi_{[-3,3]}(x_j)
$$

if we assume $|h_j| < 2$. The same estimate holds for $\varphi(x_j)$ when $\ell_j = 0$.

Therefore

$$
\sup_{0<|h_j|<2}|h_1\ldots h_n|^{-\delta}|\Delta_h\psi_L(x)| \leqslant C2^{\delta|L|}\chi_{Q_n}(x) \tag{2.6}
$$

where the cube $Q_n = [-3,3]^n$.

We can then majorize the expression in (2.4) by bringing all the sup's with respect to the various $h_j$ inside the inner integral and then use (2.6). So, by Lemma 2.3 and Proposition 2.1, the norm $\|K_L\|_{p,p}$ is controlled by the supremum, taken over all permutations $(k_1, \ldots, k_n)$ of $(1, \ldots, n)$, of the quantities

$$2^{\delta|L|} \sup_{i_{k_1} \in \mathbf{Z}} \int \left( \ldots \left( \sup_{i_{k_1} \in \mathbf{Z}} \int (|\mu^{(I)}| * \chi_{Q_n})(x)\, dx_{k_n} \right) \ldots \right) dx_{k_1}.$$

Now

$$\begin{aligned}(|\mu^{(I)}| * \chi_{Q_n})(x)\, dx_{k_n} &= 6 p_{k_n}(|\mu^{(I)}|) * \chi_{Q_{n-1}} \\ &\leqslant 6 (\sup_{i_{k_n} \in \mathbf{Z}} p_{k_n}(|\mu^{(I)}|)) * \chi_{Q_{n-1}}.\end{aligned}$$

Iterating this argument $n$ times, we conclude that

$$\|K_L\|_{p,p} \leqslant C_{p,\delta} 2^{\delta|L|}$$

for each $\delta > 0$, when $1 < p < \infty$. This concludes the proof.

Q.E.D.

## 3. Two maximal theorems in the $n$-parameter case.

As before, we let $\{\nu^{(I)}\}_{I \in \mathbf{Z}^n}$ be measures supported on the unit cube, but this time we do not impose any cancellation condition on them. We consider the maximal operator

$$(3.1) \qquad (Mf)(x) = \sup_{I \in \mathbf{Z}^n} |(f * \nu_I^{(I)})(x)|.$$

In order to prove an $L^p$-boundedness theorem for $M$ in the same spirit as Theorem 2.4, it seems necessary to impose more restrictive conditions on the $\nu^{(I)}$ than just (ii) and (iii) in the statement of Theorem 2.4.

We prove two results along these lines. In the first theorem, (ii) is replaced by the stronger assumptions that the $\nu^{(I)}$ are uniformly in $L^{1,\varepsilon}(\mathbf{R}^n)$ for some $\varepsilon > 0$. In the second theorem, we assume that $\nu^{(I)} = \nu > 0$ independent of $I$.

THEOREM 3.1. – Let $\{\nu^{(I)}\}_{I \in \mathbf{Z}^n}$ *be functions supported on the unit cube, such that*

(i) $\|\nu^{(I)}\|_{L^{1,\varepsilon}(\mathbf{R}^n)} \leqslant C$ *for some* $\varepsilon > 0$ ;

(ii) *for each permutation* $(k_1, \dots, k_n)$ *of* $(1, \dots, n)$,

(3.2)

$$\sup_{i_{k_1}\in\mathbf{Z}} \int \left( \sup_{i_{k_2}\in\mathbf{Z}} \int \dots \left( \sup_{i_{k_n}\in\mathbf{Z}} \int |\nu^{(I)}(x_1, \dots, x_n)|\, dx_{k_n} \right) \dots dx_{k_2} \right) dx_{k_1} \leqslant C.$$

*Then for every* $p$, $1 < p \leqslant \infty$, *the maximal operator* $M$ *in* (3.1) *is bounded on* $L^p(\mathbf{R}^n)$, *with a norm that only depends on* $p$, $C$, $\varepsilon$.

*Proof.* – Let $\varphi$, $\eta_{-\ell}$, $\psi_L$ be as in the proof of Theorem 2.4, and write

$$\nu^{(I)} = \sum_{L\in\mathbf{N}_n} \nu^{(I)} * \psi_L$$

with corresponding maximal operators

$$(M_L f)(x) = \sup_{I\in\mathbf{Z}^n} |(f*(\nu^{(I)}*\psi_L)_I)(x)|.$$

It follows from (i) that $\|\nu^{(I)}*\psi_L\|_1 \leqslant C2^{-\varepsilon|L|}$, so that

$$\|M_L\|_{\infty,\infty} \leqslant C2^{-\varepsilon|L|}.$$

It is then sufficient to prove that, for $1 < p < \infty$ and $\delta > 0$ arbitrarily small,

$$\|M_L\|_{p,p} \leqslant C_{p,\delta}{}^{\delta|L|}.$$

The conclusion will then follow by the Marcinkiewicz interpolation theorem.

If all the components $\ell_1, \dots, \ell_n$ of $L$ are positive, then $\mu^{(I)} = \nu^{(I)} * \psi_L$ satisfies assumption (i) in Theorem 2.4. We also have

$$\begin{aligned} |\widehat{\mu^{(I)}}(\xi)| &= |\widehat{\nu^{(I)}}(\xi)|\ |\hat{\psi}_L(\xi)| \\ &\leqslant C\|\nu^{(I)}\|_{L^{1,\varepsilon}(\mathbf{R}^n)}(1+|\xi|)^{-\varepsilon} \\ &\leqslant C(1+|\xi|)^{-\varepsilon}. \end{aligned}$$

It is also clear, using (3.2), that (2.5) holds independently of $L$. So when no component of $L$ is zero, we can conclude that for every choice of $\pm$ signs, the kernel

$$K = \sum_{I\in\mathbf{Z}^n} \pm\, (\nu^{(I)}*\psi_L)_I$$

satisfies $\|K\|_{p,p} \leqslant C_p$ for $1 < p < \infty$. From this, using the standard square-function argument, we conclude that $\|M_L\|_{p,p} \leqslant C_p$ for $1 < p < \infty$.

Considering now $L = 0 = (0, \ldots, 0)$, we observe that

$$|(\nu^{(I)} * \psi_0)(x)| \leqslant C\chi_{Q_n}(x),$$

where $Q_n = [-2,2]^n$, so that $M_0 f$ is controlled by the $n$-parameter Hardy-Littlewood maximal function of $f$.

It remains to consider those $L \neq 0$ having some components equal to zero. We first observe that we have already completed the proof in dimension $n = 1$. We then proceed by induction, assuming that the theorem has been proved in dimensions strictly less than $n$.

Assume for simplicity that $L = (\ell_1, \ldots, \ell_k, 0, \ldots, 0)$, with $\ell_1, \ldots, \ell_k$ different from zero. We then define, for $(i_1, \ldots, i_k) \in \mathbf{Z}^k$,

$$\sigma^{(i_1, \ldots i_k)}(x_1, \ \ldots, x_k) = \sup_{i_{k+1} \in \mathbf{Z}} \int \ldots (\sup_{i_n \in \mathbf{Z}} \int |\nu^{(I)}(x)|\, dx_n) \ldots dx_{k+1}.$$

By (ii), $\|\sigma^{(i_1, \ldots, i_k)}\|_1 \leqslant C$. We have

$$\begin{aligned} &|(\nu^{(I)} * \psi_L)(x)| \\ &\quad \leqslant C(\sigma^{(i_1, \ldots, i_k)} * [|\eta_{-\ell_1}| \otimes \ldots \otimes |\eta_{-\ell_k}|])(x_1, \ldots, x_k)\chi_{Q_{n-k}}(x_{k+1}, \ldots, x_n), \end{aligned}$$

so that $M_L$ is dominated by the composition of the $(n-k)$-dimensional Hardy-Littlewood maximal operator with a $k$-dimensional maximal operator whose norm is estimated by means of the inductive hypothesis.

In fact, for $\delta > 0$ and small,

$$\begin{aligned} &\|\sigma^{(i_1, \ldots, i_k)} * [|\eta_{-\ell_1}| \otimes \ldots \otimes |\eta_{-\ell_k}|]\|_{L^{1,\delta}(\mathbf{R}^k)} \\ &\qquad \leqslant \|\sigma^{(i_1, \ldots, i_k)}\|_1 \||\eta_{-\ell_1}| \otimes \ldots \otimes |\eta_{-\ell_k}|\|_{L^{1,\delta}(\mathbf{R}^k)} \\ &\qquad \leqslant C 2^{\delta|L|} \end{aligned}$$

and (3.2) follows from the definition of $\sigma^{(i_1, \ldots, i_k)}$ and from the boundedness of the $L^1$-norms of the various $\eta_{-\ell}$.

Therefore for these indices $L$, $\|M_L\|_{p,p} \leqslant C_{p,\delta} 2^{\delta|L|}$, and this concludes the proof. Q.E.D.

We want to point out here that some uniformity of the type (3.2) is actually needed when $n \geqslant 2$.

In fact, take $n = 2$ and define $\nu^{(I)} = \nu^{(i_1, i_2)} = \varphi^{(i_2)}(x)\psi(y)$, when $1 \leqslant i_2 \leqslant N$, $\nu^{(I)} = 0$ otherwise. Here $\varphi^{(i_2)}(x) = N$ if $(i_2 - 1)/N \leqslant x \leqslant i_2/N$, $\varphi^{(i_2)}(x) = 0$ otherwise.

We take $\psi$ to be the characteristic function of the interval $[0,1]$ and set $f(x,y) = f_1(x)f_2(y)$, where $f_1(x) = 1$ when $0 \leqslant x \leqslant 1/N$, $f_1(x) = 0$ otherwise; we also set $f_2(y) = \psi(y)$. Then

$$\begin{aligned}(Mf)(x,y) &= \sup_{I \in \mathbf{Z}^2} (f * \nu_I^{(I)})(x,y) \\ &\geqslant \sup_{1 \leqslant i_2 \leqslant N} (f * \nu_{0,i_2}^{(0,i_2)})(x,y) \\ &\geqslant \frac{1}{2}\end{aligned}$$

on the unit square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. So $\|Mf\|_p \geqslant 12$. Also $\|f\|_p = N^{-1/p}$. Because $\|\nu^{(I)}\|_{L^{1,\varepsilon}(\mathbf{R}^2)} \leqslant cN^{-\varepsilon}$, taking $N$ large, this shows that the $L^p$-inequality cannot hold when $p < 1/\varepsilon$.

THEOREM 3.2. — *Let $\nu$ be a positive measure supported on the unit cube, such that $|\hat{\nu}(\xi)| \leqslant C(1+|\xi|)^{-\varepsilon}$ for some $\varepsilon > 0$. Then the maximal operator $M$ in* (3.1) *(with $\nu^{(I)} = \nu$) is bounded on $L^p(\mathbf{R}^n)$, for $1 < p \leqslant \infty$, with a norm that only depends on $p$, $C$, $\varepsilon$.*

*Proof.* — The statement is trivial for $p = \infty$, so we will assume $1 < p < \infty$.

Let $S = \{k_1, \dots, k_\ell\} \subseteq \{1, \dots, n\}$ and define $p_S(\nu) = p_{k_1}(\dots p_{k_\ell}(\nu)\dots)$, $p_\varnothing(\nu) = \nu$. For $\xi \in \mathbf{R}^n$, let $\xi_S \in \mathbf{R}^n$ be the element obtained by replacing the coordinates $\xi_k$, $k \in S$, by zero. Let

$$\mu = \sum_{S \subseteq \{1,\dots,n\}} (-1)^{|S|} p_S(\nu) \prod_{k \in S} \varphi(x_k),$$

where $\varphi$ is a non-negative smooth function supported on $[-1,1]$ with integral equal to one. Observing that

$$(p_S(\nu))\hat{\ }(\xi) = \hat{\nu}(\xi_S),$$

on easily sees that

(i) $\hat{\mu}(\xi) = 0$ if some coordinate of $\xi$ is zero;

(ii) $|\hat{\mu}(\xi)| \leqslant C(1+|\xi|)^{-\varepsilon}$.

Since (iii) in Theorem 2.4 is trivially verified, the kernels $K = \sum_{I \in \mathbf{Z}^n} \pm \mu_I$ satisfy $\|K\|_{p,p} \leqslant C$ independently of the choice of the signs.

Applying again a square-function argument, we conclude that

$$\|\sup_{I \in \mathbf{Z}^n} |f * \mu_I(x)|\|_p \leqslant C\|f\|_p .$$

It remains to prove that if $S \neq \emptyset$, then

$$\left\| \sup_{I \in \mathbf{Z}^n} \left| f * \left( p_S(\nu) \prod_{k \in S} \varphi(x_k) \right)_I \right| \right\|_p \leqslant C \|f\|_p . \tag{3.3}$$

As in the proof of Theorem 3.1, this follows from an inductive argument, observing that the maximal function in (3.3) is dominated by the superposition of an $|S|$-dimensional one and an $(n-|S|)$-dimensional one, the first being a Hardy-Littlewood maximal function.
Q.E.D.

## 4. Dilations on $\mathbf{R}^n$ depending on $k$ parameters.

The formal setting for $k$-parameter dilations on $\mathbf{R}^n$ is the following : we assume that $\mathbf{R}^k$ acts on $\mathbf{R}^n$ continuously by diagonalizable linear transformations. Then, writing this action in an appropriate coordinate system, the element $t = (t_1, \ldots, t_k) \in \mathbf{R}^k$ acts on $(x_1, \ldots, x_n) \in \mathbf{R}^n$ by

$$(x_1, \ldots, x_n) \mapsto (2^{\sum_j \lambda_{1j} t_j} x_1, \ldots, 2^{\sum_j \lambda_{nj} t_j} x_n) . \tag{4.1}$$

In the more usual multiplicative notation, setting $\delta_j = 2^{t_j}$, we have

$$(x_1, \ldots, x_n) \mapsto (\delta_1^{\lambda_{11}} \ldots \delta_k^{\lambda_{1k}} x_1, \ldots, \delta_1^{\lambda_{n1}} \ldots \delta_k^{\lambda_{nk}} x_n) . \tag{4.2}$$

We call

$$\Lambda = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1k} \\ \vdots & \ddots & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nk} \end{pmatrix} \tag{4.3}$$

the *dilation matrix*. We say that two dilation matrices $\Lambda$ and $\Lambda'$ of the same dimensions are equivalent if $\Lambda' = \Lambda P$, for some non-singular $k \times k$ matrix $P$. In particular $\Lambda \sim c\Lambda$ for every $c \neq 0$. It is quite clear that two equivalent matrices induce the same set of dilations.

Since part of the existing literature assumes that the exponents $\lambda_{ij}$ are non-negative, it is worth pointing out that this assumption is not invariant under the just described notion of equivalence. However, a necessary and sufficient condition for a matrix $\Lambda$ to be equivalent to one with non-negative entries is obviously that the range of the associated linear map from $\mathbf{R}^k$ to $\mathbf{R}^n$ contains $k$ independent vectors with non-negative entries. This shows that dilations with non-negative exponents form a proper subclass.

We will often refer to the following two examples of dilations.

*Example* $(a)$ : $(x_1, x_2) \mapsto (\delta x_1, \delta^{-1} x_2)$ on $\mathbf{R}^2$, corresponding to

$$\Lambda = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

*Example* $(b)$ : $(x_1, x_2, x_3) \mapsto (\delta_1 x_1, \delta_2 x_2, \delta_1 \delta_2 x_3)$ on $\mathbf{R}^3$, corresponding to

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

We will consider convolution kernels adapted to the dilations (4.2) and constructed starting from a family $\{\mu^{(I)}\}_{I \in \mathbf{Z}^k}$ of uniformly bounded measures supported on a fixed cube. If we define the dilated $\mu_I^{(I)}$ of $\mu^{(I)}$ by

$$\int_{\mathbf{R}^n} f \, d\mu_I^{(I)} = \int_{\mathbf{R}^n} f(2^{\Sigma \lambda_{1j} i_j} x_1, \ldots, 2^{\Sigma \lambda_{nj} i_j} x_n) \, d\mu^{(I)}(x_1, \ldots, x_n),$$

we want to define

$$K = \sum_{I \in \mathbf{Z}^k} \mu_I^{(I)}.$$

To begin with, we determine what cancellations it is natural to impose on the measures $\mu^{(I)}$ in order to make (4.4) meaningful. More precisely, we want the series

$$\sum_{I \in \mathbf{Z}^k} \widehat{\mu_I^{(I)}}(\xi)$$

to converge to a bounded function. It is natural to consider the simpler case where $\mu^{(I)} = c(I)\mu$, with $\{c(I)\}_{I \in \mathbf{Z}^k}$ a bounded sequence and $\mu$ a fixed measure with compact support.

If $S = \{k_1, \ldots, k_\ell\} \subseteq \{1, \ldots, n\}$, we denote by $\Lambda_S$ the $\ell \times k$ matrix whose $j$-th row is the $k_j$-th row of $\Lambda$, and by $V_S$ the coordinate subspace of $\mathbf{R}^n$ spanned by the elements $e_{k_1}, \ldots, e_{k_\ell}$ of the canonical basis.

We also set $2^{\Lambda I} x = (2^{\Sigma \lambda_{1j} i_j} x_1, \ldots, 2^{\Sigma \lambda_{nj} i_j} x_n)$, for $I \in \mathbf{Z}^k$.

PROPOSITION 4.1. — *Let $\mu$ be a measure with compact support. If*

$$\left|\sum_{I\in \mathbf{Z}^k} c(I)\hat{\mu}(2^{\Lambda I}\xi)\right| \leqslant C \sup_{I\in \mathbf{Z}^k} |c(I)| \tag{4.5}$$

*for every bounded sequence $\{c(I)\}$ and almost every $\xi \in \mathbf{R}^n$, then $\hat{\mu}$ vanishes on each subspace $V_S$ for which* $\operatorname{rank}(\Lambda_S) < k$. *In particular if* $\operatorname{rank}(\Lambda) < k$ *and* (4.5) *holds, then* $\mu = 0$.

*Proof.* — Since $\hat{\mu}$ is continuous, (4.5) is equivalent to

$$\sum_{I\in \mathbf{Z}^k} |\hat{\mu}(2^{\Lambda I}\xi)| \leqslant C$$

uniformly in $\xi$.

Assume that $\operatorname{rank}(\Lambda_S) < k$, and let $\xi \in V_S$. Then the matrix $\operatorname{diag}(\xi_1, \ldots, \xi_n)\Lambda$ has rank smaller than $k$, so there is $\tau \in \mathbf{R}^k\backslash\{0\}$ such that

$$\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_n \end{pmatrix} \Lambda \begin{pmatrix} \tau_1 \\ \vdots \\ \tau_k \end{pmatrix} = 0 .$$

This means that

$$((\Sigma\, \lambda_{1j}\tau_j)\xi_1, \ldots, (\Sigma\, \lambda_{nj}\tau_j)\xi_n) = (0, \ldots, 0),$$

so that for every $s \in \mathbf{R}$

$$(2^{s\Sigma\lambda_{1j}\tau_j}\xi_1, \ldots, 2^{s\Sigma\lambda_{nj}\tau_j}\xi_n) = (\xi_1, \ldots, \xi_n).$$

If we had $\hat{\mu}(\xi) \neq 0$, we would have $|\hat{\mu}(\xi')| \geqslant \alpha > 0$ for $\xi'$ in a neighborhood of $\xi$. On the other hand, infinitely many points $I \in \mathbf{Z}^k$ are arbitrarily close in $\mathbf{R}^k$ to the line generated by $\tau$. We would then have $|\hat{\mu}(2^{\Lambda I}\xi)| \geqslant \alpha$ for infinitely many indices $I$, which contradicts the hypothesis. Q.E.D.

It follows that in order to have a non-vacuous theory, we must assume that $k \leqslant n$ and $\operatorname{rank}(\Lambda) = k$. Also, if $k = n$, then $\Lambda \sim I$, and we are in the $n$-parameter theory discussed in the previous sections.

If we assume, as we will do in the sequel, that $\operatorname{rank}(\Lambda) = k$, then the generic orbit in $\mathbf{R}^n$ under the action of $\mathbf{R}^k$ is a $k$-dimensional manifold. The lower-dimensional orbits are precisely the orbits contained in some subspace $V_S$ corresponding to $\operatorname{rank}(\Lambda_S) < k$. We can then restate Proposition 4.1 by saying that $\hat{\mu}$ must vanish on the lower-dimensional orbits.

Observe that in Example (a) the required cancellation is $\hat{\mu}(0) = 0$. In Example (b) $\hat{\mu}$ must vanish on the three coordinate axes.

## 5. $L^p$-boundedness in the $k$-parameter case.

We consider a set of dilations on $\mathbf{R}^n$ induced by an $n \times k$ matrix $\Lambda$ of rank $k$. For $S \subseteq \{1, \dots, n\}$, we keep the notation $V_S$ and $\Lambda_S$ introduced in the previous section.

THEOREM 5.1. – *Let $\{\mu^{(I)}\}_{I \in \mathbf{Z}^k}$ be a family of measures supported on the unit cube in $\mathbf{R}^n$, such that*

(i) $\widehat{\mu^{(I)}}(\xi) = 0$ *if* $\xi \in V_S$ *and* $\operatorname{rank}(\Lambda_S) < k$;

(ii) $|\widehat{\mu^{(I)}}(\xi)| \leqslant C(1+|\xi|)^{-\varepsilon}$ for some $C, \varepsilon > 0$;

(iii) $|\mu^{(I)}| \leqslant \sigma \in M(\mathbf{R}^n)$.

*Then the series* $\sum_{I \in \mathbf{Z}^k} |\mu^{(I)}(2^{\Lambda I}\xi)|$ *converges to a bounded function and the convolution operator defined by the kernel* $K = \sum_{I \in \mathbf{Z}^k} \mu_I^{(I)}$ *is bounded on* $L^p(\mathbf{R}^n)$, $1 < p < \infty$, *with a norm that only depends on* $p$, $C$, $\varepsilon$ *and* $\|\sigma\|_1$.

The proof consists in a reduction to the $n$-parameter dilations, and assumes Theorem 2.4. Before going into the proof, we explain the main ideas in the setting of Example (a), assuming all the $\mu^{(I)}$ to be equal.

We are given a measure $\mu$ on $\mathbf{R}^2$, and the kernel $K$ is constructed by subjecting $\mu$ to the dilations $(x_1, x_2) \mapsto (2^i x_1, 2^{-i} x_2)$. If it were not for the extra cancellations required by the 2-parameter dilations, we could simply say that we are dealing with a special double sequence $\{\mu^{(i,j)}\}$, with $\mu^{(i,j)} = 0$ if $i + j \neq 0$, and $\mu^{(i,-i)} = \mu$. The problem is that the 2-parameter theory requires $\hat{\mu}(\xi_1, 0) = \hat{\mu}(0, \xi_2) = 0$ and we only assume that $\hat{\mu}(0,0) = 0$.

Therefore we first decompose $\mu$ as $\mu_1 + \mu_2$, with $\hat{\mu}_1(\xi_1, 0) = 0$ and $\hat{\mu}_2(0, \xi_2) = 0$. One way to do this consists in taking a bump function $\varphi$ on the line with integral one, and setting $\mu_1 = \varphi(x_1) \otimes p_2(\mu)$. This gives $\hat{\mu}_1(\xi_1, \xi_2) = \hat{\varphi}(\xi_1)\hat{\mu}(0, \xi_2)$ and $\hat{\mu}_2(\xi_1, \xi_2) = \hat{\mu}(\xi_1, \xi_2) - \hat{\varphi}(\xi_1)\hat{\mu}(0, \xi_2)$.

Correspondingly, we write $K = K_1 + K_2$.

Since $\mu_1$ does not have the required cancellations, we expand it telescopically as

$$\mu_1(x_1,x_2) = \sum_{j=0}^{\infty} (2^{-j}\mu_1(x_1,2^{-j}x_2) - 2^{-j-1}\mu_1(x_1,2^{-j-1}x_2))$$

(with slight abuse of notation in treating $\mu_1$ as if it were a function). This identity can be checked on the Fourier transform side, using the decay of $\hat{\mu}$ at infinity.

If we call $\nu(x_1,x_2) = \mu_1(x_1,x_2) - \frac{1}{2}\mu_1\left(x_1,\frac{1}{2}x_2\right)$, then $\nu$ has the required cancellations, and we have reduced matters to the assumptions of Theorem 2.4 if we write

$$\begin{aligned} K_1(x_1,x_2) &= \sum_{i\in\mathbf{Z}} \mu_1(2^{-i}x_1,2^{i}x_2) \\ &= \sum_{\substack{i\in\mathbf{Z}\\ j\geqslant 0}} 2^{-j}\nu(2^{-i}x_1,2^{i-j}x_2) \\ &= \sum_{i+j\geqslant 0} 2^{-i-j}\nu(2^{-i}x_1,2^{-j}x_2)\,. \end{aligned}$$

Observe that this argument also proves that $K$ is a well-defined distribution and that $\Sigma|\hat{\mu}(2^{i}\xi_1,2^{-i}\xi_2)|$ is bounded.

The proof of Theorem 5.1 requires of course more technical details and an induction argument. The reader can test his comprehension of the proof on Example (b).

*Proof of Theorem* 5.1. – We argue by induction on $n-k$. The case $k=n$ is covered by Theorem 2.4. We assume therefore that $k<n$ and begin by introducing some notation.

For $1\leqslant j\leqslant n$, let $\Lambda^{(j)}$ be the $(n-1)\times k$ matrix obtained from $\Lambda$ by removing its $j$-th row. We can assume, without loss of generality, that rank $(\Lambda^{(n)}) = k$. Let

$$\mathscr{S} = \{S\subseteq\{1,\ldots,n-1\} : S \text{ maximal w.r. to the property rank } (\Lambda_S)=k-1\}$$

$$\mathscr{S}_1 = \{S\in\mathscr{S} : 1\notin S\}$$

and $$\mathscr{S}_2 = \{S\in\mathscr{S}\backslash\mathscr{S}_1 : 2\notin S\}$$

$$\cdots\cdots$$

$$\mathscr{S}_{n-1} = \{S\in\mathscr{S}\backslash(\mathscr{S}_1\cup\ldots\cup\mathscr{S}_{n-2}) : n-1\notin S\}.$$

Then $\mathscr{S}$ is the disjoint union of $\mathscr{S}_1,\ldots,\mathscr{S}_{n-1}$.

For $\xi \in \mathbf{R}^n$ and $S \subseteq \{1, \ldots, n\}$, we denote by $\xi_S$ the projection of $\xi$ on $V_S$, i.e. $\xi_S = (\xi'_1, \ldots, \xi'_n)$ where $\xi'_j = \xi_j$ if $j \in S$ and $\xi'_j = 0$ otherwise.

Finally let $\varphi$ be a smooth function on the line, supported on $[-1,1]$ and with integral equal to one.

For each $j$, $1 \leqslant j \leqslant n$, we introduce the $(k+1)$-parameter set of dilations on $\mathbf{R}^n$ induced by the dilation matrix

$$\Lambda(j) = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1k} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \lambda_{j1} & \cdots & \lambda_{jk} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nk} & 0 \end{pmatrix}. \tag{5.1}$$

We also decompose $\mu^{(I)}$ as

$$\mu^{(I)} = \sum_{j=1}^{n} {}^j\mu^{(I)},$$

where, for $1 \leqslant j \leqslant n-1$,

$$\widehat{{}^j\mu^{(I)}}(\xi) = \sum_{S \in \mathscr{S}_j} \widehat{\mu^{(I)}}(\xi_{S \cup \{n\}}) \prod_{k \notin S \cup \{n\}} \hat{\varphi}(\xi_k),$$

and

$${}^n\mu^{(I)} = \mu^{(I)} - \sum_{j=1}^{n-1} {}^j\mu^{(I)}.$$

We take now $j \leqslant n-1$. Since each summand in ${}^j\mu^{(I)}$ is obtained by pushing forward $\mu^{(I)}$ to some coordinate subspace and then tensoring it with some copies of $\varphi$, it is clear thar supp $({}^j\mu^{(I)})$ is contained in the unit cube. It is also clear that $\widehat{{}^j\mu^{(I)}}$ inherits a decay at infinity from the decay of $\widehat{\mu^{(I)}}$ and from that of $\hat{\varphi}$. We check now that ${}^j\mu^{(I)}$ satisfies some of the cancellation conditions required by the $(k+1)$-parameter dilations induced by $\Lambda(j)$.

A preliminary remark is necessary, since the correct induction argument only works if rank $(\Lambda(j)) = k+1$. If it happens that rank $(\Lambda(j)) < k+1$, then necessarily rank $(\Lambda^{(j)}) < k$. Then the set $\mathscr{S}_j$ contains $S = \{1, \ldots, j-1, j+1, \ldots, n-1\}$ as its unique element, and $\Lambda_{S \cup \{n\}} = \Lambda^{(j)}$ has rank strictly smaller than $k$. Therefore $\widehat{\mu^{(I)}}(\xi_{S \cup \{n\}}) = 0$ for every $\xi \in \mathbf{R}^n$ by hypothesis, and hence ${}^j\mu^{(I)} = 0$.

We can then assume that $\operatorname{rank}(\Lambda(j)) = k+1$, and we determine the subsets $S \subseteq \{1, \ldots, n\}$ such that $\operatorname{rank}(\Lambda(j)_S) < k+1$. It is clear from (5.1) that any $S$ such that $j \notin S$ has this property. It is also clear that if $j \in S$, then

$$\operatorname{rank}(\Lambda(j)_S) < k+1 \Leftrightarrow \operatorname{rank}(\Lambda_{S\setminus\{j\}}) < k.$$

Assume that $j \in S$ and that $\operatorname{rank}(\Lambda_{S\setminus\{j\}}) < k$. If $\xi \in V_S$ (i.e. if $\xi = \xi_S$) we have

$$\begin{aligned} \widehat{{}^j\mu^{(I)}}(\xi) &= \sum_{S' \in \mathscr{S}_j} \widehat{\mu^{(I)}}(\xi_{S' \cup \{n\}}) \prod_{k \notin S' \cup \{n\}} \hat{\varphi}(\xi_k) \\ &= \sum_{S' \in \mathscr{S}_j} \widehat{\mu^{(I)}}(\xi_{(S' \cup \{n\}) \cap S}) \prod_{k \notin S' \cup \{n\}} \hat{\varphi}(\xi_k). \end{aligned}$$

Since $j \notin S'$ when $S' \in \mathscr{S}_j$, then $(S' \cup \{n\}) \cap S \subseteq S\setminus\{j\}$. By assumption, $\widehat{\mu^{(I)}}(\xi_{(S' \cup \{n\}) \cap S}) = 0$ and then $\widehat{{}^j\mu^{(I)}}(\xi) = 0$.

In general, however, $\widehat{{}^j\mu^{(I)}}$ does not vanish on the subspaces $V_S$ with $j \notin S$, i.e. it does not vanish whenever $\xi_j = 0$. For this reason we introduce the measure ${}^j\nu^{(I)}$ given by

$$\text{(5.3)} \quad \widehat{{}^j\nu^{(I)}}(\xi) = \widehat{{}^j\mu^{(I)}}(\xi_1, \ldots, \xi_j, \ldots, \xi_n) - \widehat{{}^j\mu^{(I)}}(\xi_1, \ldots, 2\xi_j, \ldots, \xi_n).$$

Then the ${}^j\nu^{(I)}$ are supported on the double of the unit cube and

$$\text{(5.4)} \quad \widehat{{}^j\mu^{(I)}}(\xi_1, \ldots, \xi_j, \ldots, \xi_n) = \sum_{\ell \geqslant 0} \widehat{{}^j\nu^{(I)}}(\xi_1, \ldots, 2^\ell \xi_j, \ldots, \xi_n)$$

boundedly.

For $L \in \mathbf{Z}^{k+1}$, define ${}^j\omega^{(L)}$ to be equal to ${}^j\nu^{(I)}$ if $L = (I, \ell)$ with $\ell \geqslant 0$, and ${}^j\omega^{(L)} = 0$ otherwise. Then the $\{{}^j\omega^{(L)}\}_{L \in \mathbf{Z}^{k+1}}$ satisfy the assumptions (i), (ii) and (iii) relative to the dilations induced by the matrix $\Lambda(j)$. By the inductive hypothesis,

$$\sum_{L \in \mathbf{Z}^{k+1}} |\widehat{{}^j\omega^{(L)}}(2^{\Lambda(j)L}\xi)| \leqslant C$$

and the kernel

$${}^jK = \sum_{L \in \mathbf{Z}^{k+1}} {}^j\omega_L^{(L)}$$

defines a bounded convolution operator on $L^p(\mathbf{R}^n)$, for $1 < p < \infty$.

By (5.4), for fixed $I \in \mathbf{Z}^k$,

$$\sum_{\ell \in \mathbf{Z}} {}^j\omega_{I,\ell}^{(I,\ell)} = \sum_{\ell \geqslant 0} {}^j\nu_{I,\ell}^{(I)} = {}^j\mu_I^{(I)}$$

and

$$|\widehat{{}^j\mu^{(I)}}(2^{\Lambda I}\xi)| \leqslant \sum_{\ell \in \mathbf{Z}} |\widehat{{}^j\omega^{(I,\ell)}}(2^{\Lambda(j)(I,\ell)}\xi)|.$$

This implies that

$${}^jK = \sum_{I \in \mathbf{Z}^k} {}^j\mu_I^{(I)}$$

and that

$$\sum_{I \in \mathbf{Z}^k} |\widehat{{}^j\mu^{(I)}}(2^{\Lambda I}\xi)| \leqslant C.$$

We take now $j = n$. In order to apply the same argument used for $j \leqslant n - 1$, we only have to show that $\widehat{{}^n\mu^{(I)}}$ vanishes on the subspaces $V_S$ with $n \in S$ and rank $(\Lambda_{S\setminus\{n\}}) < k$. It is sufficient to assume that $S$ is maximal with respect to these two properties, i.e. that $S\setminus\{n\} \in \mathcal{S}$. We have

$$\widehat{{}^n\mu^{(I)}}(\xi) = \widehat{\mu^{(I)}}(\xi) - \sum_{S' \in \mathcal{S}} \widehat{\mu^{(I)}}(\xi_{S'\cup\{n\}}) \prod_{k \notin S'\cup\{n\}} \hat{\varphi}(\xi_k). \tag{5.5}$$

If $\xi \in V_S$ and $S\setminus\{n\} \in \mathcal{S}$, then $\xi_{S'\cup\{n\}} = \xi_{(S'\cup\{n\})\cap S}$. We claim that if $S' \neq S\setminus\{n\}$, then rank $(\Lambda_{(S'\cup\{n\})\cap S}) < k$, so that $\widehat{\mu^{(I)}}(\xi_{S'\cup\{n\}}) = 0$. In fact, if we had rank $(\Lambda_{(S'\cup\{n\})\cap S}) = k$, we would have rank $(\Lambda_{S'\cap(S\setminus\{n\})}) \geqslant k-1$, and actually rank $(\Lambda_{S'\cap(S\setminus\{n\})}) = k - 1$, since both $\Lambda_{S'}$ and $\Lambda_{S\setminus\{n\}}$ have rank $k-1$. But this would imply that also rank $(\Lambda_{S'\cup(S\setminus\{n\})}) = k-1$, contradicting the maximality of $S'$ and of $S\setminus\{n\}$.

So in the right-hand side of (5.5) only the term corresponding to $S' = S\setminus\{n\}$ is different from zero and then

$$\widehat{{}^n\mu^{(I)}}(\xi) = \widehat{\mu^{(I)}}(\xi) - \widehat{\mu^{(I)}}(\xi) \prod_{k \notin S} \hat{\varphi}(\xi_k).$$

But since $\xi \in V_S$, for $k \notin S$, $\hat{\varphi}(\xi_k) = \hat{\varphi}(0) = 1$, hence $\widehat{{}^n\mu^{(I)}}(\xi) = 0$.

Q.E.D.

## 6. Homogeneous kernels.

We assume that $\Lambda$ is an $n \times k$ dilation matrix, with rank $(\Lambda) = k$. If $x \in \mathbf{R}^n$ and $\delta \in \mathbf{R}^k_+$ (i.e. $\delta = (\delta_1, \ldots, \delta_k)$ with $\delta_j > 0$ for every $j$), we denote by $\delta^\Lambda x$ the element $(\delta_1^{\lambda_{11}} \ldots \delta_k^{\lambda_{1k}} x_1, \ldots, \delta_1^{\lambda_{n1}} \ldots \delta_k^{\lambda_{nk}} x_n) \in \mathbf{R}^n$.

A convolution operator $f \mapsto f * K$ commutes with all the dilations $\delta^\Lambda$ if and only if $K$ is homogeneous, in the sense that

$$\langle K, f(\delta^\Lambda \cdot) \rangle = \langle K, f \rangle$$

for every $f \in \mathscr{S}(\mathbf{R}^n)$ and every $\delta \in \mathbf{R}^k_+$. If $K$ coincides with a function on an open set, then this function satisfies the identity $K(x) = \det(\delta^\Lambda) K(\delta^\Lambda x)$. Clearly $\det(\delta^\Lambda) = \delta_1^{\sum_j \lambda_{j1}} \ldots \delta_1^{\sum_j \lambda_{jk}}$.

The class of homogeneous distributions is quite rich when $k < n$. We will deal with a subclass of it, consisting of « principal value kernels » located away from the coordinate hyperplanes. To stay away from these hyperplanes may look as an awkward restriction in certain cases (like the ordinary isotropic one-parameter dilations), but it is quite reasonable in the general context.

Let $D$ be the set of points with non-vanishing coordinates in $\mathbf{R}^n$. Under the action of $\mathbf{R}^k$, $D$ decomposes as a disjoint union of orbits, each orbit being a $k$-dimensional manifold. We say that an $(n-k)$-dimensional manifold $\Sigma \subset D$ is a *representative set* if $\Sigma$ contains exactly one point from each orbit in $D$ and it is transversal to each orbit.

This is equivalent to saying that the map $\Phi : \Sigma \times \mathbf{R}^k \to D$ given by $\Phi(x,\delta) = \delta^\Lambda x$ is a diffeomorphism. Explicit constructions of representative sets for any family of dilations can be found in [14].

The representative set plays the rôle of (a dense open set in) the unit sphere for the one-parameter isotropic dilations. The choice of $\Sigma$ is not unique. For the dilations in Example (a), one can take $\Sigma = \{(\pm 1, y) : y \neq 0\}$ or $\Sigma = \{(x, \pm x) : x \neq 0\}$.

A choice of $\Sigma$ for Example (b), which is different from the one used in [14], is $\Sigma = \{(\pm 1, \pm 1, z\} : z \neq 0\}$.

For $r, s \in \mathbf{R}^k_+$, with $r < s$ (in the sense that $r_j < s_j$ for each $j$) and $F \subseteq \Sigma$, we set

$$F_{r,s} = \{\delta^\Lambda x : x \in F, r \leqslant \delta < s\}.$$

Let $\mu$ be a measure on $D$ that is homogeneous under the dilations $\delta^\Lambda$. If $F$ is a Borel subset of $\Sigma$, then

$$\mu(F_{r,s}) = \sigma(F)\prod_{j=1}^{k}\log(s_j/r_j) \tag{6.1}$$

for some constant $\sigma(F)$. Clearly, $\sigma$ is a measure on $\Sigma$ and if $f$ is a continuous function with compact support in $D$,

$$\int f(x)\,d\mu(x) = \int_{\mathbf{R}_+^k}\int_\Sigma f(\delta^\Lambda x)\,d\sigma(x)\,\det(\delta^\Lambda)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k}. \tag{6.2}$$

Conversely, if $\sigma$ is a measure on $\Sigma$, then (6.2) defines a homogeneous measure $\mu$ on $D$ (see Theorem 2.1 in [14]).

We want to define principal value distributions $K$ on $\mathbf{R}^n$ by the formula

$$\begin{aligned} \text{(6.3)}\quad \text{p.v.}\int f(x)\,dK(x) \\ = \lim_{\varepsilon\to 0}\int_\varepsilon^{1/\varepsilon}\cdots\int_\varepsilon^{1/\varepsilon}\int_\Sigma f(\delta^\Lambda x)\,d\sigma(x)\,\det(\delta^\Lambda)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k}, \end{aligned}$$

where $f$ is a smooth function with compact support in $\mathbf{R}^n$.

In order that (6.3) makes sense, some condition must be imposed on $\sigma$. Once $K$ is defined by (6.3), we want to discuss $L^p$-boundedness for the corresponding convolution operator. This analysis leads to a very simple statement if we assume that the orbits in $D$ are « well-curved ».

For $x \in D$, we call $O_x$ its orbit under the dilations $\delta^\Lambda$ and denote by $\omega_x$ the surface measure on $O_x$.

LEMMA 6.1. — *Let $\Lambda$ be an $n \times k$ dilation matrix of rank. The following are equivalent :*

(i) *for each $x \in D$, the orbit $O_x$ is not contained in any proper affine subspace of $\mathbf{R}^n$ ;*

(ii) *if $\mu$ is a measure supported on an orbit $O_x \subset D$ and $d\mu(y) = \varphi(y)\,d\omega_x(y)$ with $\varphi \in C_0^\infty(O_x)$, then there is $\varepsilon > 0$ such that $|\hat\mu(\xi)| \leqslant C(1+|\xi|)^{-\varepsilon}$ ;*

(iii) *the $n$ rows of $\Lambda$ are all different and none of them is equal to zero.*

*Proof.* — The implication (i) $\Rightarrow$ (ii) is known (see e.g. [17]). Conversely, if a measure $\mu$ is supported in a proper affine subspace, then $\hat{\mu}$ has no decay in the orthogonal directions.

The equivalence of (i) and (iii) is a consequence of the fact that the $n+1$ functions on $\mathbf{R}^k_+ : f_0(\delta) = 1$, $f_1(\delta) = \delta_1^{\lambda_{11}} \dots \delta_k^{\lambda_{1k}}$, $\dots, f_n(\delta) = \delta_k^{\lambda_{n1}} \dots \delta_k^{\lambda_{nk}}$ are linearly independent if and only if they are all different.

Q.E.D.

THEOREM 6.2. — *Assume that the rows of* $\Lambda$ *are all different and non-zero.* Let $\sigma$ *be a measure supported on a compact subset of* $\Sigma$, *such that* $\hat{\sigma}$ *vanishes on each coordinate subspace* $V_S$ *of* $\mathbf{R}^n$ *corresponding to a submatrix* $\Lambda_S$ *of* $\Lambda$ *with rank strictly smaller than* $k$. *Then* (6.3) *defines a homogeneous distribution* $K$ *and the associated convolution operator is bounded on* $L^p(\mathbf{R}^n)$ *for* $1 < p < \infty$.

*Proof.* — Let $\varphi(\delta)$ be a smooth function on $\mathbf{R}^k$ supported on the cube $Q = \left\{\delta : \frac{1}{2} \leqslant \delta_j \leqslant 4, 1 \leqslant j \leqslant k\right\}$ and such that $\sum_{I \in \mathbf{Z}^k} \varphi(2^I \delta) = 1$ for every $\delta \in \mathbf{R}^k_+$.

Let also $\mu$ be the measure on $\mathbf{R}^n$ given by

$$\text{(6.4)} \quad \int f(x)\, d\mu(x) = \int_Q \int_\Sigma f(\delta^\Lambda x)\, d\sigma(x) \varphi(\delta) \det(\delta^\Lambda) \frac{d\delta_1}{\delta_1} \cdots \frac{d\delta_k}{\delta_k}.$$

If we set $\mu^{(I)} = \mu$ for every $I$, we see that the hypotheses of Theorem 5.1 are satisfied. Actually (iii) is trivial and (i), (ii) follow from the fact that they are satisfied by $\hat{\sigma}$.

If we call $K = \sum_{I \in \mathbf{Z}^k} \mu_I$, then $K$ defines a bounded convolution operator on $L^p(\mathbf{R}^n)$, $1 < p < \infty$. We prove now that $K$ is the principal value distribution given by (6.3).

It follows from Theorem 5.1 that

$$\sum_{I \in \mathbf{Z}^k} |\hat{\sigma}(2^{\Lambda I} \xi)| \leqslant C$$

for some constant $C$ and every $\xi \in \mathbf{R}^n$. This easily implies that

$$\int_{\mathbf{R}^k_+} |\hat{\sigma}(\delta^\Lambda \xi)| \frac{d\delta_1}{\delta_1} \cdots \frac{d\delta_k}{\delta_k} \leqslant C'$$

and that

$$\hat{K}(\xi) = \sum_{I \in \mathbf{Z}^k} \int_Q \hat{\sigma}((2^I\delta)^\wedge\xi)\varphi(\delta)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k}$$
$$= \int_{\mathbf{R}_+^k} \hat{\sigma}(\delta^\wedge\xi)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k}.$$

If $K_\varepsilon$ is the distribution given by

$$\int_\varepsilon^{1/\varepsilon}\cdots\int_\varepsilon^{1/\varepsilon}\int_\Sigma f(\delta^\wedge x)\,d\sigma(x)\det(\delta^\wedge)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k},$$

then

$$\hat{K}_\varepsilon(\xi) = \int_\varepsilon^{1/2}\cdots\int_\varepsilon^{1/\varepsilon}\hat{\sigma}(\delta^\wedge\xi)\frac{d\delta_1}{\delta_1}\cdots\frac{d\delta_k}{\delta_k},$$

and $\hat{K}_\varepsilon(\xi)$ tends to $\hat{K}(\xi)$ boundedly as $\varepsilon \to 0$. This ends the proof.

Q.E.D.

## 7. A general $k$-parameter maximal theorem.

We would like now to formulate and prove a general $k$-parameter maximal theorem. Suppose $\mu$ is a finite positive measure on $\mathbf{R}^n$, and define the dilated measures $\mu_\delta$ by

$$\int f(x)\,d\mu_\delta(x) = \int f(\delta^\wedge x)\,d\mu(x),$$

where $x \mapsto \delta^\wedge x$ are the $k$-parameter dilations introduced in Section 4.

We shall be concerned with the $L^p$-boundedness of the operator $M$ defined by

$$(7.1) \qquad (Mf)(x) = \sup_{\delta \in \mathbf{R}_+^k} |(f * \mu_\delta)(x)|.$$

THEOREM 7.1. — *Suppose that the measure $\mu$ satisfies the majorization*

$$(7.2) \qquad \sup_{1 \leqslant \delta_j < 2} \mu_\delta \leqslant \nu,$$

*where $\nu$ is a finite measure. Then the operator $M$ defined by* (7.1) *is bounded from $L^p(\mathbf{R}^n)$ to itself for all $p$, $1 < p \leqslant \infty$.*

*Remark.* – Note that the condition (7.2) is easily seen to be equivalent with the condition that $\sup_{\delta\in\Delta} \mu_\delta$ is a finite measure, whenever $\Delta$ is a non-empty relatively compact open set in $\mathbf{R}_+^k$.

*Proof.* – We shall first prove the Theorem under the assumption that the measures $\mu$ and $\nu$ are concentrated on $D = \{x \in \mathbf{R}^n : x_j \neq 0 \ \forall j\}$ (i.e. they assign no mass to $\mathbf{R}^n \backslash D$).

For every $x_0 \in D$ we define the measure $\mu^{x_0}$ supported on the orbit of dilations through $x_0$ and given by

$$\int f \, d\mu^{x_0} = \int_{1 \leqslant \varepsilon_j < 2} f(\varepsilon^\Lambda x_0) \frac{d\varepsilon_1}{\varepsilon_1} \cdots \frac{d\varepsilon_k}{\varepsilon_k}. \tag{7.3}$$

Next, we define the corresponding maximal operator $M_{x_0}$, given by

$$(M_{x_0} f)(x) = \sup_{\delta \in \mathbf{R}_+^k} |f * \mu_\delta^{x_0}(x)|, \tag{7.4}$$

and this clearly equals

$$\sup_{\delta \in \mathbf{R}_+^k} \left| \int_{\delta_j \leqslant \varepsilon_j < 2\delta_j} f(x - \varepsilon^\Lambda x_0) \frac{d\varepsilon_1}{\varepsilon_1} \cdots \frac{d\varepsilon_k}{\varepsilon_k} \right|.$$

The following lemma, in the case where the exponents of the dilations are non-negative, is essentially contained in [6].

LEMMA 7.2. – *For each $x_0 \in D$, the operator $M_{x_0}$ is bounded from $L^p(\mathbf{R}^n)$ to itself, when $1 < p \leqslant \infty$, with a bound independent of $x_0$.*

*Proof.* – For a given $x_0 \in D$, the boundedness of $\mathrm{M}_{x_0}$ follows easily from Lemma 6.1 (part (ii)) and Theorem 3.2, when the orbit is well-curved (i.e. satisfies one of the equivalent statements of that lemma).

In the case the orbit is not well-curved, we can, by part (iii) of Lemma 6.1, decompose $\mathbf{R}^n$ as $\mathbf{R}^\ell \times \mathbf{R}^{n-\ell}$, so that $x_0 = (x_0', 0)$, with $\delta^\Lambda x_0 = (\delta^{\Lambda'} x_0', 0)$, and the orbit of $x_0'$ under $\delta^{\Lambda'}$ is well-curved in $\mathbf{R}^\ell$. This allows us to assert the boundedness of $M_{x_0}$ in the general case.

To prove the uniformity of the estimates in $x_0$, we consider the $n$-parameter family of scalings $x = (x_1, \ldots, x_n) \mapsto (\rho_1 x_1, \ldots, \rho_n x_n) = \rho x$, with $\rho_j \neq 0$ for all $j$. Observe that these maps commute with the dilatations $\delta^\Lambda$, where $\delta \in \mathbf{R}^k_+$. Moreover, as is easily verified, if we define $R_\rho$ by $R_\rho f(x) = f(\rho_1 x_1, \ldots, \rho_n x_n)$, then $R_\rho^{-1} M_{x_0} R_\rho = M_{\rho x_0}$.

It therefore follows that the inequality $\|M_{x_0} f\|_p \leqslant A_p \|f\|_p$ implies the inequality $\|M_{\rho x_0} f\|_p \leqslant A_p \|f\|_p$, with the same bound $A_p$. Since the $\rho$-scalings act transitively on $D$, the lemma is established. Q.E.D.

We shall next use the « polar coordinates » for $D$ defined in Section 6. Thus if $x \in D$, we write $x = (\delta, z) = \delta^\Lambda z$, where $\delta \in \mathbf{R}^k_+$ and $z \in \Sigma$, and so identify $D$ with $\mathbf{R}^k_+ \times \Sigma$.

LEMMA 7.3. – *Suppose* $\nu$ *is a positive measure concentrated on* $D$. *If*

$$\tilde{\nu} = \int_{1 \leqslant \varepsilon_j < 2} \nu_\varepsilon \frac{d\varepsilon_1}{\varepsilon_1} \cdots \frac{d\varepsilon_k}{\varepsilon_k}, \tag{7.5}$$

*then* $\tilde{\nu}$ *is majorized as*

$$d\tilde{\nu}(x) \leqslant \sum_{I \in \mathbf{Z}^k} \chi_I(\delta) \frac{d\delta_1}{\delta_1} \cdots \frac{d\delta_k}{\delta_k} d\nu^I(z), \tag{7.6}$$

*where* $\chi_I(\delta)$ *is the characteristic function of the set where* $2^{i_j} \leqslant \delta_j < 2^{i_j+1}$, $I = (i_1, \ldots, i_k)$, *and* $\nu^I$ *are positive measures on* $\Sigma$, *with*

$$\sum_{I \in \mathbf{Z}^k} \int d\nu^I(z) < \infty. \tag{7.7}$$

*Proof.* – It is convenient to change variables, via $\delta_j = 2^{y_j}$, $j = 1, \ldots, k$, identifying $\mathbf{R}^k_+$ with $\mathbf{R}^k$. With this (and a slight abuse of notation), we are reduced to the following : suppose $d\nu = d\nu(y,z)$ is a finite measure on $\mathbf{R}^k \times \Sigma$. We write $\nu^h$ for the $y$-translate (by $h \in \mathbf{R}^k$) of $\nu$ ; i.e. symbolically, $d\nu^h(y,z) = d\nu(y-h,z)$. We then define $\tilde{\nu}$ by

$$\tilde{\nu} = \int_{0 \leqslant h_j < 1} \nu^h \, dh_1 \ldots dh_k.$$

Then

$$\tilde{\nu} = \sum_{I \in \mathbf{Z}^k} \int_{0 \leqslant h_j < 1} \chi_I(y) \nu^h \, dh,$$

where $\chi_I$ is the characteristic function of the cube $Q_I = \{y \in \mathbf{R}^k : i_j \leqslant y_j < i_j + 1\}$. However if $y \in Q_I$ and $Q_I^* = \{y \in \mathbf{R}^k : i_j - 1 \leqslant y_j < i_j + 1\}$, then as measures on $Q_I \times \Sigma$,

$$(7.8) \qquad \int_{0 \leqslant h_j < 1} d\nu^h(y, \cdot)\, dh \leqslant \left(\int \chi_{Q_I^*}(y')\, d\nu(y', \cdot)\right) dy.$$

We first prove (7.8) when $\nu$ is absolutely continuous, i.e. $d\nu = f(y,z)\, dy\, dz$. Then $d\nu^h = f(y-h,z)\, dy\, dz$ and (7.8) becomes

$$\int_{0 \leqslant h_j < 1} f(y-h,z)\, dh \leqslant \int_{y' \in Q_I^*} f(y',z)\, dy'$$

when $y \in Q_I$.

The proof of (7.8) in the general case now follows by approximating $\nu$ by absolutely continuous measures in the weak-$*$ topology.

Next define

$$d\nu^I(z) = \int \chi_{Q_I^*}(y')\, d\nu(y', z).$$

Then clearly

$$\sum_I \int_\Sigma d\nu^I(z) = 2^k \int_{\mathbf{R}^k \times \Sigma} d\nu(y,z) < \infty.$$

Restating this result in terms of $\delta_j = 2^{y_j}$ gives us (7.6) and (7.7).

Q.E.D.

*End of the proof of Theorem* 7.1. – Suppose that $\mu$ is concentrated on $D$. We can assume that $\mu_\varepsilon \leqslant \nu$, for $\frac{1}{2} < \varepsilon_j \leqslant 1$, where $\nu$ is a finite measure which is also concentrated on $D$. Then $\mu \leqslant \nu_\varepsilon$, whenever $1 \leqslant \varepsilon_j < 2$, and therefore

$$\mu \leqslant C \int_{1 \leqslant \varepsilon_j \leqslant 2} \nu_\varepsilon \frac{d\varepsilon_1}{\varepsilon_1} \cdots \frac{d\varepsilon_k}{\varepsilon_k} = C\tilde{\nu}.$$

It thus suffices to prove the maximal theorem with $\mu$ replaced by $\tilde{\nu}$. However if $f \geqslant 0$, by (7.6),

$$(f * \tilde{\nu})(x) \leqslant \sum_{I \in \mathbf{Z}^k} \int_{\mathbf{R}^k \times \Sigma} f(x - \delta^\wedge z)\chi_I(\delta) \frac{d\delta_1}{\delta_1} \cdots \frac{d\delta_k}{\delta_k}\, d\nu^I(z),$$

so

$$\sup_{\delta \in \mathbf{R}_+^k} (f * \tilde{\nu}_\delta)(x) \leqslant \sum_I \int_\Sigma M_z(f)\, d\mu_I(z).$$

Applying Lemma 7.2 and the inequality (7.7) then gives $\|Mf\|_p \leqslant A_p\|f\|_p$ as desired.

To lift the restriction that our measures $\mu$ and $\nu$ are concentrated on $D$ we can argue as follows. Decompose $\mathbf{R}^n$ as a disjoint union, $\mathbf{R}^n = \bigcup_S \mathbf{R}^n_S$, where $S$ ranges over the subsets of $\{1, \dots, n\}$ and

$$\mathbf{R}^n_S = \{x \in \mathbf{R}^n : x_j = 0 \Leftrightarrow j \in S\}.$$

Then write $d\mu = \sum_S \chi_S\, d\mu = \sum_S d\mu_S$, $d\nu = \sum_S \chi_S\, d\nu = \sum_S d\nu_S$ where $\chi_S$ is the characteristic function of $\mathbf{R}^n_S$. Since each $\mathbf{R}^n_S$ is a union of orbits, our assumption implies that $\sup_{1 \leqslant \delta_j < 2} (\mu_S)_\delta \leqslant \nu_S$.

Note that when $S = \emptyset$, then $\mathbf{R}^n_S = D$, and this corresponds to the result proved above. The result for $S = \emptyset$ then follows from the case already proved, in the setting of $\mathbf{R}^m$ with $m = n - |S|$.

As a final remark, we point out that using Theorem 3.2 one can prove that the following maximal operator is bounded from $L^p(\mathbf{R}^3)$ to itself, for $1 < p \leqslant \infty$:

$$(Mf)(x,y,z) = \sup_{\substack{r,s>0 \\ k \in \mathbf{Z}}} \int_{-1}^{1} \int_{-1}^{1} |f(x-ru, y-sv, z-2^k rsuv)|\, du\, dv.$$

A general class of maximal operators of the same type can be obtained as follows: let $M_{x_0}$ be the maximal operator in (7.4) associated with a well-curved $k$-parameter orbit $O_{x_0}$ in $\mathbf{R}^n$; considering the full $n$-parameter dilations, define

$$Mf = \sup_{I \in \mathbf{Z}^n} M_{2^I x_0} f.$$

Then $M$ is bounded on $L^p(\mathbf{R}_n)$, $1 < p \leqslant \infty$.

We have assumed the orbit to be well-curved in order to have the decay of the Fourier transform required by Theorem 3.2. On the other hand this hypothesis may not be necessary, since it is known that a result of this kind holds when the orbit is a half-line [2], [7], [13], [20].

## BIBLIOGRAPHY


[1] A. Carbery, Differentiation in lacunary directions and an extension of the Marcinkiewicz multiplier theorem, Ann. Inst. Fourier, Grenoble, 38-1 (1988), 157-168.

[2] A. Carbery, A. Seeger, $H^p$- and $L^p$- variants of multiparameter Calderón-Zygmund theory, preprint.

[3] H. Carlsson, P. Sjögren, Estimates for maximal functions along hypersurfaces, Ark. För Math., 25 (1987), 1-14.

[4] H. Carlsson, P. Sjögren, J.-O. Strömberg, Multiparameter maximal functions along dilation-invariant hypersurfaces, Trans. Amer. Math. Soc., 292 (1985), 335-343.

[5] M. Christ, Hilbert transforms along curves. I. Nilpotent groups, Ann. of Math., 122 (1985), 575-596.

[6] M. Christ, The strong maximal function on a nilpotent group, to appear in Trans. Amer. Math. Soc.

[7] A. Córdoba, R. Fefferman, On the equivalence between the boundedness of certain classes of maximal operators and multiplier operators in Fourier analysis, Proc. Nat. Acad. Sci. USA, 74 (1977), 423-425.

[8] J. Duoandikoetxea, Multiple singular integrals and maximal functions along hypersurfaces, Ann. Inst. Fourier, Grenoble, 36-4 (1986), 185-206.

[9] J. Duoandikoetxea, J.L. Rubio de Francia, Maximal and singular integral operators via Fourier transform estimates, Inv. Math., 84 (1986), 541-561.

[10] R. Fefferman, Calderón-Zygmund theory for product domains. $H^p$ spaces, Proc. Nat. Acad. Sci. USA, 83 (1986), 840-843

[11] R. Fefferman, E.M. Stein, Singular integrals on product spaces, Adv. in Math., 45 (1982), 117-143.

[12] J.L. Journé, Calderón-Zygmund operators on products spaces, Rev. Mat. Ibero-Amer., 3 (1985), 55-91.

[13] A. Nagel, E.M. Stein, S. Wainger, Differentiation in lacunary directions, Proc. Nat. Acad. Sci., USA, 75 (1978), 1060-1062.

[14] A. Nagel, S. Wainger, $L^2$-boundedness of Hilbert transforms along surfaces and convolution operators homogeneous with respect to a multiple parameter group, Amer. J. Math., 99 (1977), 761-785.

[15] J. Pipher, Journé's covering lemma and its extension to higher dimensions, Duke Math. J., 53 (1986), 683-690.

[16] F. Ricci, E.M. Stein, Harmonic analysis on nilpotent groups and singular integrals. II. Singular kernels supported on submanifolds, J. Funct. Anal., 78 (1988), 56-84.

[17] E.M. Stein, Oscillatory integrals in Fourier analysis, in Bejing Lectures in Harmonic Analysis, Princeton Univ. Press, Princeton, 1986.

[18] E.M. Stein, S. Wainger, Problems in harmonic analysis related to curvature, Bull. Amer. Math. Soc., 84 (1978), 1239-1295.

[19] R. Strichartz, Singular integrals supported on submanifolds, Studia Math., 74 (1982), 137-151.

[20] J.-O. Strömberg, Dissertation (1976), Mittag-Leffler Inst., Djursholm, Sweden.

[21] J.T. Vance, $L^p$-boundedness of the multiple Hilbert transform along a surface, Pac. J. Math., 108 (1983), 221-241.

Manuscrit reçu le 20 novembre 1990.

F. Ricci,
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino (Italy).

&

E.M. Stein,
Department of Mathematics
Princeton University
Fine Hall
Princeton, NJ 08540 (USA).