The strong maximal function with respect to measures

by

BJÖRN JAWERTH and ALBERTO TORCHINSKY (Bloomington, Indiana)

Abstract. A classical result of Jensen, Marcinkiewicz and Zygmund [12] asserts that the basis \( R = \{ R \} \) of rectangles \( R \) in Euclidean space \( \mathbb{R}^n \) with sides parallel to the coordinate axis differentiates the class \( L(\log^+ L)^{n-1}(\mathbb{R}) \). The quantitative version of this result is the following estimate. Let \( |E| \) denote the Lebesgue measure of \( E \). Associated with \( \mathcal{A} \) we consider the strong maximal operator

\[
Mf(x) = \sup_{x \in \mathbb{R}^n} \left\{ f(y) \right\}_{y \leq x}, \quad x \in \mathbb{R}.
\]

Then for each \( \lambda > 0 \)

\[
(\text{JMZ}) \quad \|x \in \mathbb{R}^n : Mf(x) > \lambda\| \leq C \left( \frac{\|f\|}{\lambda} \left( 1 + \log^+ \frac{\|f\|}{\lambda} \right)^{n-1} \right) dx.
\]

Together with \( |Mf|_{L^p} \leq \|f\|_{L^p} \) (JMZ) obtains the boundedness of \( M \) in the \( L^p(\mathbb{R}^n) \) spaces, \( 1 < p < \infty \).

It is our purpose to extend (JMZ) to the context of maximal operators with respect to measures, including the study of maximal functions on "product basis" in \( \mathbb{R}^n \). Our approach exhibits the close connection existing between iteration and induction techniques and allows us to consider several applications, including a problem of Zygmund [12] recently solved by Córdoba [3], rearrangement inequalities [1] and covering results [5].

Introduction. A classical result of Jensen, Marcinkiewicz and Zygmund [12] asserts that the basis \( R = \{ R \} \) of rectangles \( R \) in Euclidean space \( \mathbb{R}^n \) with sides parallel to the coordinate axis differentiates the class \( L(\log^+ L)^{n-1}(\mathbb{R}) \). The quantitative version of this result is the following estimate. Let \( |E| \) denote the Lebesgue measure of \( E \). Associated with \( \mathcal{A} \) we consider the strong maximal operator

\[
Mf(x) = \sup_{x \in \mathbb{R}^n} \left\{ f(y) \right\}_{y \leq x}, \quad x \in \mathbb{R}.
\]

* Research partially supported by the N.F.R. and N.S.F.
Then for each \( \lambda > 0 \)
\[
(M) \quad | \{ x \in \mathbb{R}^n : M^\lambda (x) > \lambda \} | \leq c \int_{\mathbb{R}^n} \left| \frac{f(x)}{\lambda} \right| \left( 1 + \log^+ \frac{|f(x)|}{\lambda} \right)^{-1} \, dx .
\]

Together with \( \| M^\lambda \|_{L^p} \leq \| f \|_{L^p} \) (JMZ) obtains the boundedness of \( M \) in the \( L^p (\mathbb{R}^n) \) spaces, \( 1 < p < \infty \). The basic idea in proving (JMZ) is to dominate \( M \) by compositions, or iterates, of the better understood one-dimensional Hardy–Littlewood maximal function. Recently Fava [6] proved a weak type inequality for products of sublinear operators from which a simple proof of (JMZ) follows. In addition to Fava’s result the interest in this area was revived by a new proof of (JMZ) due to Córdoba and R. Fefferman [5]. Their proof relies on a deeper understanding of the geometry of rectangles and uses induction. One of their main tools is a selection procedure for families of rectangles, through the notion of sparseness, leading to sharp covering results.

It is our purpose to extend (JMZ) to the context of maximal operators with respect to measures, including the study of maximal functions on “product basis” in \( \mathbb{R}^n \). Our approach exhibits the close connection existing between the iteration and induction techniques and allows us to consider several applications, including a problem of Zygmund recently solved by Córdoba [3], rearrangement inequalities [1] and covering results [5].

To illustrate the character of our results let \( w \) be a positive locally summable function in \( \mathbb{R}^n \) and put
\[
M_w f(x) = \sup_{R \ni x} \frac{1}{w(R)} \int_R |f(y)| w(y) \, dy , \quad x \in \mathbb{R}^n ,
\]
where \( w(R) = \int_R w(y) \, dy \). Under very general conditions, namely that \( w \) be an admissible measure we show that for \( \lambda > 0 \)
\[
(1) \quad w(\{ x \in \mathbb{R}^n : M_w f(x) > \lambda \}) \leq c \int_{\mathbb{R}^n} \left( \frac{|f(x)|}{\lambda} \left( 1 + \log^+ \frac{|f(x)|}{\lambda} \right) \right)^{-1} \, w(x) \, dx .
\]

R. Fefferman [8] initiated the study of the boundedness properties of \( M_w \) and proved that if \( w \) is uniformly \( A_\infty \) in every variable, then
\[
\| M_w \|_{L^p} \leq c \| f \|_{L^p} , \quad 1 < p < \infty .
\]

An estimate similar to (1), but with the loss of a logarithm, holds for general basis \( \mathcal{B} = \text{product of Buseman–Feller basis which differentiate } L \) functions.

Our presentation is organized as follows. In Section 1 we discuss the notion of “admissible measure”. In Section 2 we prove a slight extension of (1) and in Section 3 we discuss the result for the basis \( \mathcal{B} \). In Section 4 we cover the problem of Zygmund and we present applications to rearrangement estimates, covering properties and the rest.

1. Admissible measures. A collection \( \mathcal{B} = \{ B \} \) of bounded, measurable sets of positive Lebesgue measure in \( \mathbb{R}^n \) is said to be a differentiation basis if for each \( x \in \mathbb{R}^n \) there is a subfamily \( \mathcal{B}(x) \) of \( \mathcal{B} \) such that
   \begin{enumerate}
   \item[(i)] if \( B \in \mathcal{B}(x) \), then \( x \in B \);
   \item[(ii)] each \( \mathcal{B}(x) \) contains sets of arbitrarily small diameter.
   \end{enumerate}
\( \mathcal{B} \) is said to be a Buseman–Feller basis, or \( B-F \) basis if in addition
   \begin{enumerate}
   \item[(iii)] each \( B \) is open;
   \item[(iv)] if \( x \in B \), then \( B \in \mathcal{B}(x) \).
   \end{enumerate}
Consider a finite number, say \( N \), of \( B-F \) bases \( \mathcal{B}_i = \mathcal{B}_i (\mathbb{R}^n) \) in \( \mathbb{R}^n \), \( 1 \leq i \leq N \), and put \( n_i = n_1 + \ldots + n_i \). The family \( \mathcal{B} = \{ B \subset \mathbb{R}^n : B = B_1 \times \ldots \times B_N, B_i \in \mathcal{B}_i , i = 1, \ldots , N \} \) is a \( B-F \) basis in \( \mathbb{R}^n \). \( \mathcal{B} \) is called the product basis of the \( \mathcal{B}_i \)’s and is often denoted by \( \mathcal{B}_1 \times \ldots \times \mathcal{B}_N \) or \( \prod_{i=1}^N \mathcal{B}_i \). We identify points \( x \) in \( \mathbb{R}^n \) with \( (x_1 , \ldots , x_N) , x_i \in \mathbb{R}^n \), and, when appropriate, subsets of \( \mathbb{R}^n \) with subsets of \( \mathbb{R}^n \) (by adding a number of coordinates 0) and vice versa.

To each positive locally summable function \( w \) defined in \( \mathbb{R}^n \) we associate the measure
\[
w(E) = \int_E w(x) \, dx , \quad E \subset \mathbb{R}^n ,
\]
and the restriction measures \( w_{x_1} dx_1 \ldots dx_N = dx_1 \ldots dx_N, i = 1, \ldots , N \), given by
\[
w_{x_i} (E) = w(x_i, E) = \int_E w_{x_i} (x_1 , \ldots , x_i , \ldots , x_N) \, dx_1 \ldots dx_i \ldots dx_N ,
\]
for \( E \subset \mathbb{R}^{n_i} \).

As customary, the subscript \( x_i \) denotes the fact that the corresponding variable, or differential, is missing. Thus
\[
dx_1 \ldots dx_i \ldots dx_n = dx_1 \ldots dx_{i-1} \, dx_{i+1} \ldots dx_n ,
\]
etc.

Similarly we define the restrictions \( w_{x_1 x_2} dx_1 \ldots dx_N \), \( w_{x_1 x_2 x_3} dx_1 \ldots dx_N \), etc. We reserve the notation \( \mathcal{B}^\oplus \) for \( \mathcal{B}_{x_1} \times \ldots \times \mathcal{B}_{x_n} \).

To subsets \( E \subset \mathbb{R}^n \) and to each \( x_i \in \mathbb{R}^n \), \( 1 \leq i \leq N \), we assign the sections \( E_{x_i} \subset \mathbb{R}^{n_i} \) given by
\[
E_{x_i} = \{ (x_1 , \ldots , x_i , \ldots , x_N) : (x_1 , \ldots , x_N) \in E \} .
\]
\( E_{x_i} \) is called the section of \( E \) at \( x_i \). One last notation. Let \( \mathcal{B} = \bigcup_{i=1}^N \mathcal{B}_i , N \geq 2 \).
For \( E \subset B_{x_i} \), \( B \in \mathcal{B} \), we set
\[
W(x_i, E, B) = w(x_i, E) / w(x_i, B) .
\]
For a positive locally summable function \( w \) we introduce the maximal operator with respect to \( w \) and \( \mathcal{A}(R^n) \) by

\[
M_{w} f(x) = \sup_{\mathcal{A}(R^n)} \frac{1}{w(R)} \int_{R} |f(y)| w(y) dy, \quad x \in R, \ R \in \mathcal{A}(R^n).
\]

Similarly, for \( \nu(x_1, \ldots, x_j, \ldots, x_k) \) defined on \( R^{*j} \) we set

\[
M_{\nu}^{j-1} f(x) = \sup_{\mathcal{A}(R^n)} \frac{1}{\nu(R)} \int_{R} |f(y)| \nu(y) dy, \quad x \in R, \ R \in \mathcal{A}^{j-1}(R^{*j}).
\]

Finally for \( \nu(x) \) defined on \( R^n \) we put

\[
M_{\nu}^0 f(x) = \sup_{\mathcal{A}(R^n)} \frac{1}{\nu(R)} \int_{R} |f(y)| \nu(y) dy, \quad x \in R, \ R \in \mathcal{A}(R^n),
\]

1 \( \leq \ j \leq k, \ M_{\nu}^{j-1} \) corresponds to the usual Hardy-Littlewood maximal function and as is well known, if \( \nu \) is merely doubling, the \( \nu \)-type estimate

\[
\nu([x \in R^n : M_{\nu}^0 f(x) > \lambda]) \leq \frac{c}{\lambda} \int_{R^n} |f(y)| \nu(y) dy, \quad \lambda > 0,
\]

obtains.

As for the maximal operators \( M_{\nu}^{j-1}, \nu \), for a fixed 1 \( \leq \ j \leq k, \) following Jaworwicz [11] we say that \( \nu \) is a density weight (with respect to \( \mathcal{A}^{j-1}(R^{*j}) \)) if there is a function \( c(\cdot) \) such that

\[
\nu([x \in R^{*j} : M_{\nu}^{j-1} f(x) > \lambda]) \leq c(\lambda) \nu(E)
\]

for each measurable set \( E \subseteq R^{*j} \) and 0 \( \leq \lambda \leq 1. \) This condition is satisfied if, for instance, \( M_{\nu}^{j-1} \) is of restricted weak type \((p, p) \) for some \( p < \infty. \) When \( \nu = w_x \) and \( \nu = \nu_{j}(x_j), \) 1 \( \leq \ j \leq k, \) are the restrictions of a fixed \( w \) defined on \( R^n, \) we say that they are uniformly a density weight and of weak type \( (1, 1), \) respectively, if the quantities \( c(\cdot) \) in (2) and \( c \) in (1) are independent of (almost all) \( x_j \) and \( x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k, \) respectively.

We can now state our first result.

**Theorem 2.1.** Suppose that \( \nu \in \mathcal{A}(R^n) \) and that \( w_x \) is a density weight, uniformly with respect to \( \mathcal{A}^{j-1}(R^{*j}), \) 1 \( \leq \ j \leq k. \) Then there are constants \( c \) and \( c_1, \) such that for all \( f \) and \( \lambda > 0 \)

\[
w([x \in R^n : M_{w} f(x) > \lambda]) \leq \frac{c}{\lambda} \sum_{j=1}^{k} \int_{R^{*j}} M_{w}^{j-1} f(x) w(x) dx.
\]

**Proof.** Let \( \mathcal{E}_n = \{x \in R^n : M_{w} f(x) > \lambda\}; \) \( \mathcal{E}_n \) is an open set in \( R^n. \) If \( F \) is
an arbitrary compact subset of $E$, there is a finite family of rectangles $\{R_i\}$, $0 \leq i \leq M$, such that

\begin{equation}
    w(E) \leq w(\bigcup R_i)
\end{equation}

and

\begin{equation}
    \left\{ \int_{R_i} w(x) dx > \lambda w(R_i), \quad 0 \leq i \leq M. \right\}
\end{equation}

Since $w \in \mathcal{D}(R^d)$ the family $\mathcal{A} = \{R_i\}$ can be divided into $k$ disjoint subfamilies $\mathcal{A}_i$, $\mathcal{A} = \bigcup_{i=1}^k \mathcal{A}_i$, so that for rectangles in $\mathcal{A}_i$ the $i$th direction is admissible. Clearly,

\begin{equation}
    w(\bigcup R_i) = w\left(\bigcup_{i=1}^k \bigcup_{R \in \mathcal{A}_i} R_i\right) \leq \sum_{i=1}^k w\left(\bigcup_{R \in \mathcal{A}_i} R_i\right).
\end{equation}

The rest of the argument is symmetric in $i$, so fix ideas we consider $\mathcal{A}_1$.

We first select a sparse subfamily $\{R_j\}$ of $\mathcal{A}_1$, which satisfies

\begin{equation}
    w\left(\bigcup_{R \in \mathcal{A}_1} R_i\right) \leq c w(\bigcup_{j=1}^j R_j).
\end{equation}

The selection procedure and criteria of sparseness are as follows: since each $R_j \in \mathcal{A}_1(R^d)$, the side lengths corresponding to the $n_i$ block are all equal. Let $R_0$ be an $R_j \in \mathcal{A}_1$ with largest side length in a direction in the $n_i$ block, the $n_i$ direction say. If $R_0, \ldots, R_{j-1}$ have been chosen, let $R_j$ be a rectangle $R_I = I_1 \times \ldots \times I_l$ in $\mathcal{A}_1$ with largest side length in the $n_i$ direction such that

\begin{equation}
    w(x_1, R_{n_j}) \bigcup (R_{n_j}) \geq (1-\gamma) w(x_1, R_{n_j})
\end{equation}

obtains for almost all $x_1$, in $I_1$.

We go on until $\mathcal{A}_1$ is exhausted. Next we show that (6) holds. Let $R^*$ denote the rectangle obtained from $R$ by tripling the sides of $R$ in the $n_i$ block. We observe that

\begin{equation}
    \bigcup_{R \in \mathcal{A}_1} \leq \{M^{[1, \infty]}_{x_1, y_1} (x_1, y_1) \geq \gamma \}
\end{equation}

Indeed, if $R_j \in \mathcal{A}_1$ is one of the $R_j$s, then (8) clearly holds. So assume $R \in \mathcal{A}_1$ is not one of the $R_j$s. Then $R$ was discarded after $R_0, \ldots, R_{j-1}$ had been selected, thus

\begin{equation}
    \{\text{side length in } n_i\} \geq \{\text{side length in } n_i\}
\end{equation}

\begin{equation}
    \{\text{direction of } R\} \geq \{\text{direction of } R_{j_0}\}
\end{equation}

and $j_0$ is the smallest index with this property. By the selection procedure we have

\begin{equation}
    w(x_1, R_{n_j} \bigcup (R_{n_j})) \geq \gamma w(x_1, R_{n_j})
\end{equation}

and consequently also

\begin{equation}
    w(x_1, R_{n_j} \bigcup (R_{n_j})) \geq \gamma w(x_1, R_{n_j})
\end{equation}

for a set of $x_1$s in $I_1$ of positive measure.

Since

\begin{equation}
    \\{\text{side length in } n_i\} \geq \{\text{side length in } n_i\}
\end{equation}

\begin{equation}
    \{\text{direction of } R_{j_0-1}\} \geq \{\text{direction of } R\}
\end{equation}

we readily see that

\begin{equation}
    R_{n_j} \bigcup (R_{n_j}) = \text{fixed set independent of } x_1 \in I_1.
\end{equation}

From (7) and the fact that the $n_1$ direction is admissible it thus follows that

\begin{equation}
    w(x_1, R_{n_j} \bigcup (R_{n_j})) \geq \gamma w(x_1, R_{n_j})
\end{equation}

for almost all $x_1$ in $R_{n_1}$.

Whence

\begin{equation}
    M^{[1, \infty]}_{x_1, y_1} (x_1, y_1, \ldots, y_l) \geq \gamma \text{ for } (x_1, \ldots, x_l) \in R_{n_1} \text{ and } x_1 \in I_1,
\end{equation}

and (8) obtains.

Moreover, since $w_{x_1}$ is a density weight uniformly in $x_1$, from (8) we get that

\begin{equation}
    w(x_1, \bigcup_{R \in \mathcal{A}_1} \leq w(x_1, \{M^{[1, \infty]}_{x_1, y_1} (x_1, y_1) \geq \gamma \})
\end{equation}

\begin{equation}
    \leq c(\gamma) w(x_1, \bigcup_{R \in \mathcal{A}_1} R_j).
\end{equation}

Integrating this inequality with respect to $x_1$ we see that

\begin{equation}
    w\left(\bigcup_{R \in \mathcal{A}_1} R_i\right) \leq c w(\bigcup_{j=1}^j R_j).
\end{equation}

However, it is readily seen that $w(\bigcup_{j=1}^j R_j) \leq c w(\bigcup_{j=1}^j R_j)$ since $w \in \mathcal{D}(R)$. We have thus proved (6).

To complete the proof of our theorem, with $M^{[1, \infty]}_{x_1, y_1}$ on the right-hand side, we will estimate $\sum_{j=1} w(R_j)$. Let $\{E_j\}$ be the disjoint sequence obtained from the $R_j$s by

\begin{equation}
    E_j = R_j \setminus \bigcup_{i<j} R_i, \quad E_0 = R_0.
\end{equation}

Then $\bigcup_{j} E_j$. From (7) we see that

\begin{equation}
    w(x_1, E_j) \geq (1-\gamma) w(x_1, R_j)
\end{equation}

for all $j$. 

Now
\[ \sum_{x \in \mathcal{A}_k} \sum_{j \in J(x)} |f(x, y_1, \ldots, y_n)w_m(x, y_1, \ldots, y_n)dy_2 \ldots dy_n \]
\[ \leq c \sum_{x \in \mathcal{A}_k} \sum_{j \in J(x)} M^{(1)}_{1, w_m} f(x, y_1, \ldots, y_n) \]
\[ \leq c \sum_{x \in \mathcal{A}_k} \left( \inf_{j \in J(x)} M^{(1)}_{1, w_m} f(x, y_1, \ldots, y_n) \right) \]
\[ = c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y_1, \ldots, y_n)w_m(x, y_1, \ldots, y_n)dy_2 \ldots dy_n \]
\[ \leq c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y_1, \ldots, y_n)w_m(x, y_1, \ldots, y_n)dy_2 \ldots dy_n \]
\[ + c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y_1, \ldots, y_n)w_m(x, y_1, \ldots, y_n)dy_2 \ldots dy_n. \]

Hence integrating (10) over \( R^k \) we see that
\[ \sum_{j \in J} \int f(y)w(y)dy \leq c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y)w(y)dy. \]
Also from (4) it follows that
\[ \sum_{j \in J} \int f(y)w(y)dy \leq c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y)w(y)dy. \]
Thus combining (11) and (12) and choosing \( \epsilon \) so that \( c \epsilon < 1/2 \) and \( c_1 < 1/2c \) with \( c \) the constant in (11), we obtain that
\[ \sum_{j \in J} \int f(y)w(y)dy \leq \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y)w(y)dy. \]
Repeating this argument for each family \( \mathcal{A}_k \), and summing over \( h \), from (5) and (13) we get
\[ \sum_{j \in J} \int f(y)w(y)dy \leq c \sum_{x \in \mathcal{A}_k} M^{(1)}_{1, w_m} f(x, y)w(y)dy. \]
Since the compact set \( E \) was arbitrary, by combining (3) and (14) our proof is complete.

The above proof shows in fact that if \( w \) is such that each rectangle \( R \) in \( \mathcal{A}(R^k) \) has an admissible direction in the set \( \{ n_k \}_{k=0}^L, L \leq 1 \}, \) then the sum which appears on the right-hand side of the conclusion of the theorem has only to be extended over \( j \in J \). Moreover, the reader can use Theorem 2.1 with \( w \equiv 1 \), and the Hardy–Littlewood theorem, to obtain (JMZ). More generally, we may also deduce

**Theorem 2.2.** Suppose that \( w \in C(\mathcal{D} R^k), n_k \leq n, \) and that \( w \) is doubling in each \( n_k \) (block) direction uniformly. Then for \( \lambda > 0 \)
\[ w(\{ x \in \mathbb{R}^n : M_{k,w} f(x) > \lambda \}) \leq c \int \frac{|f(x)|}{\lambda} \left( 1 + \log \frac{1 + |f(x)|}{\lambda} \right)^{-1} w(x)dx. \]

**Proof.** If it were true that \( w \in C(\mathcal{D} R^{n+k}) \), we could use Theorem 2.1 repeatedly and at the final step apply the weighted Hardy–Littlewood maximal theorem to obtain the result. Unfortunately, it is not clear whether this property holds in general. Nevertheless, we show next that this strategy almost works.

We begin by dividing \( \mathcal{A}_k \) into \( k! \) disjoint subfamilies \( \mathcal{A}_{k,h} \), one corresponding to each permutation of \( \{ 1, \ldots, k \} \), in such a way that \( \mathcal{A}_{k,h} \) only contains those rectangles \( B = \bigcap_{j=1}^k B_j \) for which we have for each \( s, 1 \leq s \leq k-1, \) and \( E \subseteq \bigcap_{j=1}^k B_j \),
\[ \text{ess sup}_{x \in B \cap \mathcal{A}_{k,h}} W(x, E, B) \geq \gamma \]
implies
\[ \text{ess inf}_{x \in B \cap \mathcal{A}_{k,h}} W(x, E, B) \geq \gamma. \]
Here \( W(x, E, B) \) denotes the quantity introduced in Section 1. Now for each \( \pi \) we introduce the auxiliary operators \( M_{k,w}^\pi \) and \( M_{k,w}^{-w_n} \), \( 1 \leq s \leq k-1, \) by
\[ M_{k,w}^\pi f(x) = \sup_{R} \int_{R} f(y)w(y)dy, \quad x \in R, \quad R \in \mathcal{A}_{k,h}(R^n). \]
and
\[ M_{k,w}^{-w_n} f(x) = \sup_{R} \int_{R} f(y)w_n(y)dy, \quad x \in R, \quad R \subseteq R^n \cap \mathcal{A}_{k,h}(R^n), \quad R = R_{n(1)} \cdots R_{n(k)} \text{ for some } R \in \mathcal{A}_{k,h}(R^n). \]
Since \( w \in C(\mathcal{D} R^n) \) we have
\[ M_{k,w} f(x) \leq \sum_{\pi} M_{k,w}^\pi f(x). \]
We now show that for \( M_{k,w}^\pi \) the iteration argument described above works. Indeed, for \( s \leq k-1, \) \( M_{k,w}^{-w_n} f(x), \) being majorized by a weighted
Hardy-Littlewood maximal operator, is of weak type (1,1). Successive applications of the proof of Theorem 2.1 yield
\[
\left( \int_{\mathbb{R}^n} \left( \frac{1}{\lambda} \int_{\mathbb{R}^{n-1}} M_{\lambda}(x_1, \ldots, x_{n-1}, f(y)) w(x_1, \ldots, x_{n-1}) \, dy \right) \, dx_0 \right)^{\frac{1}{n}}
\]
for 0 ≤ s ≤ k − 1. In this estimate we use x_{n-1} = x and w_{n-1} = w.

It is now easy to complete the argument along the lines of the iteration argument described above; the details are left for the reader to complete.

The reader will also observe that in many cases the conclusion of Theorem 2.2 follows at once from Theorem 2.1. This is the case when, for instance, w is doubling in every direction and in addition every direction, except possibly one not depending on the rectangle, is admissible. In that case the iteration argument works since we have the following result.

**Proposition 2.3.** Let \( \mathcal{A}_i \subseteq \mathbb{R}^n \), \( i = 0, \ldots, k − 1 \), be a partition of \( \mathbb{R}^n \) into \( k + 1 \) parts. Suppose that each of the \( \mathcal{A}_i \) (block) directions, \( 1 ≤ i ≤ k − 1 \), are admissible at every \( R \) in \( \mathcal{A}_i \). Then for almost every \( x \), \( w_0(x) \), has the property that \( w_0 \) on the \( \mathcal{A}_i \), \( 2 ≤ i ≤ k − 1 \), are admissible at every \( R \) in \( \mathcal{A}_i \).

**Proof.** Let us show that, for instance, the \( \mathcal{A}_i \)-direction is admissible. For \( x \in \mathbb{R}^n \) be the cube in \( \mathbb{R}^n \) with side length \( 2 \) and center \( x \). For \( w \) the \( \mathcal{A}_i \)-direction is admissible. Thus there are constants \( 0 < \gamma_i < 1 \), such that for every rectangle \( R \times R' \) in \( \mathcal{A}_i \) and every subset \( E \subseteq \mathbb{R}^n \),
\[
\esssup_{x \in \mathbb{R}^n} \frac{w_t(x, x_2, E)}{w_t(x_1, x_2, x'_2)} \geq \gamma_i
\]
for all \( x \in \mathbb{R}^n \).

Let now \( \varepsilon \) tend to 0. By the (usual) Lebesgue differentiation theorem \( \frac{w_t(x, x_2, E)}{w_t(x_1, x_2, x'_2)} \) tends to \( w(x_1, x_2, E) \) for almost all \( x \). Similarly \( \frac{w_t(x, x_2, R)}{w_t(x_1, x_2, R')} \) tends to \( w(x_1, x_2, R) \), again for almost all \( x \). Consequently (15) implies that for almost all \( x \), the statement of (15) holds with \( x_1 \) replaced by \( x_0 \). This is equivalent to saying that the \( \mathcal{A}_i \)-direction is admissible for \( w_0 \).

An important class of weights for which Proposition 2.3 holds, and consequently also (JMZ), Theorem 2.1, and Theorem 2.2 hold, is the collection of those weights \( w \) which are uniformly in \( \mathcal{A}_0 \) in each variable except possibly one in which they are merely doubling.

3. Maximal functions on product basis with respect to measures. Iteration techniques suffice to deal with the maximal function with respect to measures in the general setting of \( B-F \) bases provided we are dealing with product measures (cf. de Guzmán [10]). In this section we discuss arbitrary measures.

To fix ideas and to simplify notation we restrict ourselves to the basis \( \mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \), the product of (two) \( B-F \) bases \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \). Points \( x \in \mathbb{R}^n \), \( n = n_1 + n_2 \), are denoted by \( x = (x_1, x_2), x_i \in \mathbb{R}^{n_i} \). Let
\[
M_{\mathcal{B}, w}(f)(x) = \sup_{x_i \in \mathcal{B}_i} \frac{1}{w(B)} \int_B f(x) \,dy, \quad x \in B, \mathcal{B} \text{ in } B.
\]

The maximal operators \( M_{\mathcal{B}_1} \) and \( M_{\mathcal{B}_2} \) are defined similarly. We can then prove

**Theorem 3.1.** Suppose that \( \mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \) and that \( w \) is a density weight uniformly with respect to \( \mathcal{B}_i \) \( (\mathbb{R}^{n_i}) \), \( i = 1, 2 \). Then there are constants \( c \) and \( c_i \) such that for \( \lambda > 0 \)
\[
w(\{x \in \mathbb{R}^n : M_{\mathcal{B}, w}(f)(x) > \lambda \}) \leq \frac{c}{\lambda} \left( \int_{\mathbb{R}^{n_1}} M_{\mathcal{B}_1, w_{1_2}} M_{\mathcal{B}_2, w_{1_2}} f(x) \,dx + \int_{\mathbb{R}^{n_2}} M_{\mathcal{B}_1, w_{1_2}} M_{\mathcal{B}_2, w_{1_2}} f(x) \,dx \right),
\]

**Proof.** Let \( \mathcal{B}_1 = \{M_{\mathcal{B}_1, w}(f)(x) > \lambda \} \); then \( \mathcal{B}_1 \) is an open set in \( \mathbb{R}^n \). If \( E \) is an arbitrary compact subset of \( \mathcal{B}_1 \), there is a (finite) family of \( \mathcal{B}_j \subseteq \mathcal{B}(\mathbb{R}^n) \) such that
\[
w(E) \leq cw(\bigcup_j \mathcal{B}_j)
\]
and
\[
\int_{\mathcal{B}_j} f(x) \,dx > \lambda w(\mathcal{B}_j), \quad \text{all } j.
\]

Since we \( \mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \) the family \( \mathcal{B} = \{\mathcal{B}_j \} \) can be divided into two disjoint subfamilies, \( \mathcal{B}_1' \) and \( \mathcal{B}_2' \). Let \( \mathcal{B}_j \) be a family of \( \mathcal{B}_j \) of \( \mathcal{B}_2' \) such that
\[
w(\bigcup \mathcal{B}_j) \leq cw(\bigcup \mathcal{B}_j).
\]
The selection procedure is as follows: choose \( B_0 \) to be the first \( B \). Once \( B_0, \ldots, B_{j-1} \) have been chosen let \( B_j = B_{j-1} \times B_{j-1} \) be the first among the
remaining $B_j$'s with the following property: let $\gamma$ be the constant associated to $w \in \mathcal{S}(\mathcal{B})$. If $E_j \equiv B_{j,1}$ is defined by
\begin{align*}
(4>) \quad E_j = \{x_1 \in B_{j,1}: w(x_1, B_{j,2} \setminus \bigcup_{i < j} B_{i,2}) \geq \frac{1}{2} w(x_1, B_{j,2})\},
\end{align*}
then
\begin{align*}
(5>) \quad w(x_2, E_j) \geq (1 - \gamma) w(x_2, B_{j,1}) \quad \text{for almost all } x_2 \text{ in } \mathbb{R}^2.
\end{align*}
We go on until the family $\mathcal{A}_2$ is exhausted. We now show that (3) holds for $\mathcal{A}_2$. Observe that
\begin{align*}
(6) \quad \bigcup_{\mathcal{x} \in \mathcal{A}_2} B_{i} \subseteq \{(x_1, x_2) \in \mathbb{R}^2: M_{\mathcal{A}_1, \mathcal{A}_2} x_{i_0} (M_{\mathcal{A}_1, \mathcal{A}_2} x_{i_0})^2 \geq \gamma\} \quad = M_{\mathcal{A}_1, \mathcal{A}_2} x_{i_0} \geq \gamma, \quad \text{say}.
\end{align*}
If $B_i$ is among the $B_j$'s (6) clearly holds. If $B_i$ is not one of the $B_j$'s let $j_0 + 1$ be the largest index smaller than that of $B_i$ so that $B_{j_0 + 1}$ is among the $B_j$'s. Since $B_i$ was discarded we have that
\begin{align*}
(5<) \quad w(x_1, B_{j_0 + 1} \setminus E) \geq \gamma w(x_1, B_{j_0 + 1})
\end{align*}
for a set of $x_2$'s of positive measure in $\mathbb{R}^2$.

Moreover, since $B_i \in \mathcal{A}_2$ it follows that
\begin{align*}
(7) \quad w(x_2, B_{j_0 + 1} \setminus E) \geq \gamma w(x_2, B_{j_0 + 1}) \quad \text{for almost all } x_2 \text{ in } \mathbb{R}^2.
\end{align*}
But
\begin{align*}
(8) \quad B_{j_0 + 1} \setminus E_i = \{x_1 \in B_{j_0 + 1}: w(x_1, B_{j_0 + 1} \setminus (\bigcup_{i < j_0} B_{i,2})) \geq \frac{1}{2} w(x_1, B_{j_0 + 1})\} \subseteq \{x_1 \in B_{j_0 + 1}: M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0} (M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0})^2 \geq \gamma\}, \quad \text{say}.
\end{align*}
Thus by (7) and (8)
\begin{align*}
M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0} > \frac{1}{2} \quad \text{for almost all } x_1 \in B_{j_0 + 1} \text{ and } x_1 \in B_{j_0 + 1}.
\end{align*}
This means that $B_i$ is contained in
\begin{align*}
(9) \quad \{x_1 \in B_{j_0 + 1}: \text{for } x_2 \in \mathcal{B}(2), M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0} (M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0})^2 \geq \frac{1}{2}, x_1 \geq \gamma\}.
\end{align*}
This proves (6), for the set in (9) only increases if we replace $j_0$ by the largest $j$. Now (3) follows readily. Indeed,
\begin{align*}
w(x_2, B_{j_0 + 1}) \leq c(\gamma) w(x_2, \mathcal{B})
\end{align*}
whence integrating over $x_2$ we obtain
\begin{align*}
w(\bigcup B_j) \leq c w(\bigcup B_j).
\end{align*}
Thus to complete our proof we only need estimate $\sum_j w(B_j)$. Put $B_j = B_1 \times B_2$. Notice that from (4>) and (5>) we get that
\begin{align*}
B_1 = E \cup (B_1 \setminus E) \quad \text{and} \quad w(x_1, B_1 \setminus (\bigcup_{j < i} B_j)) \geq \frac{1}{2} w(x_1, B_1)
\end{align*}
for almost all $x_1$ in $E$ and
\begin{align*}
w(x_2, E) > (1 - \gamma) w(x_2, B_1), \quad \text{or} \quad w(x_2, B_1 \setminus E) \leq \gamma w(x_2, B_1)
\end{align*}
for almost all $x_2$ in $\mathbb{R}^2$. Therefore
\begin{align*}
w(x_2, B_j) = w(x_2, E) + w(x_2, B_1 \setminus E) \leq w(x_2, E) + \gamma w(x_2, B_1),
\end{align*}
or
\begin{align*}
w(x_2, B_j) \leq \frac{1}{1 - \gamma} w(x_2, E)
\end{align*}
for almost all $x_2$ in $\mathbb{R}^2$.

Consequently,
\begin{align*}
\int_{B_j} \int f(x_1, x_2) w(x_1, x_2) dx_1 dx_2 
\leq \int_{B_2} \int_{B_1} f(y_1, x_2) M_{\mathcal{A}_1, \mathcal{A}_2} x_{i_0} f(M_{\mathcal{A}_1, \mathcal{A}_2} x_{i_0})^2 dx_1 dx_2 
\leq (1 - \gamma)^{-1} \int_{B_2} \int_{B_1} f(y_1, x_2) M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0} f(M_{\mathcal{A}_2, \mathcal{A}_1} x_{i_0})^2 dx_1 dx_2
\leq c \int_{B_2} \int_{B_1} f(x_1, x_2) w(x_1, x_2) dx_1 dx_2.
\end{align*}
Thus by (2) and (11) and with $g = M_{\mathcal{A}_1, \mathcal{A}_2} f$ we see that
\begin{align*}
\sum_j w(B_j) \leq \frac{1}{\lambda} \sum_j \int_{B_2} \int_{B_1} f \int f(x_1, x_2) w(x_1, x_2) dx_1 dx_2 dx_1
\leq \frac{1}{\lambda} \sum_j \int_{B_2} \int_{B_2} g(x_1, x_2) w(x_1, x_2) dx_2 dx_1 dx_2
= \frac{1}{\lambda} \sum_j \int_{B_1} \int_{B_2} g(x_1, x_2) w(x_1, x_2) dx_1 dx_2 dx_2.
\end{align*}
Choose $\varepsilon = \frac{1}{2} c$ to obtain the desired conclusion for $\mathcal{A} \mathfrak{D}$. As we said the argument for $\mathcal{A} \mathfrak{F}$ follows along similar lines and is therefore omitted. This completes our proof.

One of the applications of Theorem 3.1 is to integral inequalities. First some notations. To a Young's function $\psi(t)$ satisfying the $\mathcal{A} \mathfrak{D}$-notation we associate $\psi^*(t)$ defined by

$$
\int \psi^*(t/s) ds = \psi^*(t), \quad t \text{ large}.
$$

Since $\psi(t)/t$ is non-decreasing, $\psi^*(t)$ is always at most of order $\psi(t) \log^+ t$. Nevertheless, if there is a $p > 1$ so that $\psi(t)/t^p$ is non-decreasing, then actually $\psi^*(t)$ is of order $\psi(t)$.

Suppose $\psi_1$ are as above and $\psi^*_1$ are defined with $c = c_1/2$, $c_1$ the constant of Theorem 3.1, $i = 1$. We then have

**Proposition 3.2.** Suppose $w$ is as in Theorem 3.1 and that for some $c > 0$

and all $\lambda > 0$

\[ w(x_1, \{M_{\mathcal{A} \mathfrak{D}, \mathfrak{F}} f > \lambda\}) \leq c \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

and

\[ w(x_1, \{M_{\mathcal{A} \mathfrak{D}, \mathfrak{F}} f > \lambda\}) \leq c \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt.

Then

\[ w(\{M_{\mathcal{A} \mathfrak{D}, \mathfrak{F}} f > \lambda\}) \leq c \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

\[ + c \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt,

where $c_1$ is the constant of Theorem 3.1.

Proof. By homogeneity we may assume that $\lambda = 1$. Also by Theorem 3.1 it will suffice to estimate the integrals on the right-hand side of the conclusion of the theorem. We only do

\[ I = \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

For $t > 0$ put $f^* = f$ when $|f| \geq t$ and 0 otherwise. Clearly

\[ \{M_{\mathcal{A} \mathfrak{D}, \mathfrak{F}} f^* > t\} \equiv \{M_{\mathcal{A} \mathfrak{D}, \mathfrak{F}} f > t^2\}. \]

Now

\[ I \leq \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

\[ \leq c \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

\[ = c \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

\[ + \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

Observe that

\[ I_1 \leq c \int \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

Since $|f| \leq |f|$ and in the range that interests us $t \leq |f|/c_1$ and $s \leq 2|f|/c_1$ by Fubini's theorem we get

\[ I_2 \leq c \int \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt \]

\[ = c \int \int \int \int \psi^*(f(t)/t) w \left( x_1, \left( \frac{f(t)}{t} \right) \right) dt.

"
Similarly for $I_1$ we have

$$I_1 \leq C \int_{R^2} \int_{t_{1/2}}^{t} \int_{t_{1/2}}^{t} \int_{t_{1/2}}^{t} \int_{t_{1/2}}^{t} \frac{w(x_1, y_1) w(y_2, z_2)}{(x_1 - y_1)^2 (y_2 - z_2)^2} d\psi_1(s) d\psi_2(s) dx_1 dy_1 dz_1$$

Thus combining these estimates and since clearly a similar argument applies to the other term, the desired conclusion follows.

4. The problem of Zygmund and other applications. \( R \in \mathcal{A}(R^{k+1}) \) is said to be an \( a \times b \times \varphi(a, b) \) rectangle if the sidelengths in the \( x_i \) and \( x_j \) directions are arbitrary numbers \( a \) and \( b \) and the side length in the \( x_k \) (k-dimensional) block direction is \( \varphi(a, b) \). Here \( \varphi \) is a function which is continuous at the origin, \( \varphi(a, 0) = \varphi(0, 0) = 0 \), and monotone in each variable separately.

Let \( \mathcal{A} \) be the \( B \)-F basis containing all such rectangles \( R \). Recently Córdoba [3] answered affirmatively a problem of Zygmund by showing that for

$$M_{\mathcal{A}} f(x) = \sup_{R \in \mathcal{A}} \frac{1}{|R|} \int_R |f(y)| dy, \quad x \in R, \quad R \in \mathcal{A},$$

the weak type estimate

$$\|M_{\mathcal{A}} f \|_1 = \sup_{a \in \mathbb{R}} \int_{x \in a \mathcal{A}} \frac{|f(x)|}{a} \left( 1 + \log^+ \frac{|f(x)|}{a} \right) dx$$

obtains.

We open this chapter by considering a slight extension of this result to the vector of weighted maximal functions.

In the results we discussed in Sections 2 and 3 it was possible to establish an ordering of the rectangles considered. In fact the \( B \)-F basis \( \mathcal{A}(R^n) \) is basically equivalent to the family of n-fold products of dyadic cubes \( B \) and such cubes have the property:

(D) \( B_1 \cap B_2 \neq \emptyset \), then either \( B_1 \subseteq B_2 \) or \( B_2 \subseteq B_1 \).

It is readily seen that for any \( B \)-F basis \( \mathcal{A} \) for which (D) holds and for any weight \( M_{\mathcal{A} w} \) is of weak type \((1,1)\). Consequently the results of Section 2 remain true with \( \mathcal{A} \) there replaced by \( \mathcal{A} = \prod_{j=1}^{k} \mathcal{A}_j \) provided each \( \mathcal{A}_j \) is a \( B \)-F basis for which (D) holds. One way to construct such basis \( \mathcal{A}_j \) is to let \( \varphi_1(t), \ldots, \varphi_k(t) \) be a collection of monotone, continuous functions of \( t \geq 0 \) with \( \varphi_i(0) = 0, 1 \leq i \leq k \), and let \( \mathcal{A}_j \) be the family of \( t \times \varphi_1(t) \times \ldots \times \varphi_k(t) \) rectangles in \( \mathcal{A}(R^{k+1}) \).

A more general setting in which the above remarks apply is as follows: suppose \( \mathcal{A} \) satisfies (D) and that \( \varphi \) is a set-valued mapping from \( \mathcal{A} \) into subsets of \( R^n \) which verifies

(i) \( \varphi(\mathcal{A}) \) is a \( B \)-F basis verifying (D);

(ii) \( \varphi \) is monotone, i.e., if \( B \subseteq B' \), then \( \varphi(B) \subseteq \varphi(B') \) whenever \( B, B' \in \mathcal{A} \).

It is then immediate to verify that \( \mathcal{A} \times \varphi(\mathcal{A}) = \{ B \times \varphi(B); B \in \mathcal{A} \} \) also satisfies (D) and that \( M_{\mathcal{A} \times \varphi(\mathcal{A})} \) is of weak type \((1,1)\) for any \( w \). Out of the many possible extensions of this result we consider that in the direction of Zygmun's problem.

Let then \( \mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \) be a product of \( B \)-F bases each of which satisfies (D) and consider the \( B \)-F basis \( \mathcal{A} \times \varphi(\mathcal{A}) \). In this case condition (ii) is equivalent to the function \( \varphi \) being monotone in each "variable" separately.

Let us introduce some notations. With \( B = B_1 \times B_2 \times B_3 \in \mathcal{A} \times \varphi(\mathcal{A}) \) put

$$\mathcal{N}_1 f(x) = \sup_{B \in \mathcal{A}} \frac{1}{|B|} \int_{x \in B} \inf_{y \in B} |f(x, y)| dy,$$

and similarly

$$\mathcal{N}_2 f(x) = \sup_{B \in \mathcal{A}} \frac{1}{|B|} \int_{x \in B} \inf_{y \in B} |f(x, y)| dy.$$

Also let \( \mathcal{N}_3 f(x) \) be the supremum over \( B \) containing \( x \) of the expression

$$w(z)^{-1} \inf_{B \in \mathcal{A}} \left\{ \int_{y \in B} |f(z, y)| dy \right\} \left\{ \int_{z \in B} |f(z, x)| dx \right\}.$$

The weights \( w \) we consider satisfy some of the following properties:

there are constants \( 0 < \gamma, \gamma' < 1 \), such that for each \( B = B_1 \times B_2 \times \varphi(B_1, B_2) \)

(1) If \( E \) is a measurable subset of \( B_1 \times \varphi(B_1, B_2) \), then

$$\mathcal{N}_1 f(x) \mathcal{N}_2 f(x) \mathcal{N}_3 f(x),$$

imply

$$\mathcal{N}_1 f(x) \mathcal{N}_2 f(x) \mathcal{N}_3 f(x),$$

(2) If \( E \) is a measurable subset of \( \varphi(B_1, B_2) \), then

$$\mathcal{N}_1 f(x) \mathcal{N}_2 f(x) \mathcal{N}_3 f(x),$$

imply

$$\mathcal{N}_1 f(x) \mathcal{N}_2 f(x) \mathcal{N}_3 f(x),$$

5 = Studia Mathematica 102
Moreover, if \( E_i = B_i \cup \bigcup_{j \neq i} B_j \), then
\[
w(x_i, (B_i)_{x_i}) = w(x_i, (E_i)_{x_i})
\]
for a.e. \( x_i \). Consequently,
\[
w(f, \mathcal{V}, f > \lambda) \leq c \sum \bigwedge_j w(B_j)
\]

\[
\begin{align*}
&\leq \frac{c}{\lambda} \sum_{J} \sum_{x_1, x_2, x_3} \int_{(B_j)_{x_1}} \int_{(B_j)_{x_2}} \int_{(B_j)_{x_3}} |f(x_1, x_2, x_3)| \, w(x_1, (E_i)_{x_1}) \, dx_1 \\
&= \frac{c}{\lambda} \sum_{J} \int \left( \int \int |f(x_1, x_2, x_3)| \, w(x_1, (E_i)_{x_1}) \, dx_1 \right) dx_2 \, dx_3
\end{align*}
\]

\[
\leq \frac{c}{\lambda} \int |f| \, w,
\]

which is what we wanted to show. Clearly a similar result holds for \( \mathcal{V}' \).

In the spirit of Zygmund’s problem we can show

**Theorem 4.2.** Let \( \mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2 \times \varphi(\mathcal{B}_1, \mathcal{B}_2) \) and suppose that \( w \) satisfies properties (1)-(5). Then
\[
w(\{M_{\mathcal{M}} f > \lambda\}) \leq c \frac{1 + \log^* \frac{|f(x)|}{\lambda}}{\lambda} \int w(x) \, dx.
\]

**Proof.** We have just proved that \( \mathcal{N}' \) and \( \mathcal{N} \) are operators of weak type \((1,1)\) and by (5) we readily see that

\[
\mathcal{N} f(x) \leq c \mathcal{N}' f(x) \mathcal{N} f(x).
\]

Thus \( w(\cdot, f > \lambda) \leq c(\lambda) \int |f| \, w \, dx \).

It suffices to show that given any finite family \( \{B_i\} \) in \( \mathcal{M} \) it is possible to extract a subsequence \( \{B_j\} \) which verifies

\[
\int \exp(\sum \lambda_p \mathcal{I}(x) \, w(x)) \, dx \leq c \sum w(B_j),
\]

and

\[
w(\bigcup B_j) \leq c \sum w(B_j).
\]

Once (8) and (9) are established the conclusion follows readily by means of Young’s inequality and the usual choice of \( \{B_i\} \) with the property that

\[
\lambda \leq \frac{1}{w(B_j)} \int |f| \, w \, dx \text{ for all } a.
\]

By renaming the \( B_j \)’s if necessary we may assume that the Lebesgue measure of the \( n_j \)-block is ordered in a non-decreasing fashion. For \( f = 0 \) let
\(B_0\) be the first \(B_j\) and once \(B_0, \ldots, B_{j-1}\) have been chosen let \(B_j\) be the first among the remaining \(B_k's\) such that
\[
\int_{B_j} \exp \left( \sum_{k \leq j} x_k \right) w(x) \, dx \leq 100w(B_j).
\]
Once \(\{B_k\}\) is exhausted we are left with our family \(\{B_j\}\). That (8) holds can be readily checked. Indeed, suppose \(0 \leq j \leq N\), then
\[
\mathcal{F} \leq 100e \left( \sum_{j=0}^N w(B_j) \right).
\]
By repeated use of this argument we get
\[
\mathcal{F} \leq 100e \left( \sum_{j=0}^N w(B_j) \right),
\]
which is precisely (8). As for (9) we may assume that \(B_k\) is not contained in \(\bigcup_{j=0}^N B_j\). Let \(j_0\) be the (smallest) index such that
\[
\int_{B_{j_0}} \exp \left( \sum_{j < j_0} x_j \right) w(x) \, dx > 100w(B_{j_0}).
\]
The collection \(\{B_j\}_{j=0}^{j_0-1}\) can be divided into two disjoint subfamilies: in the first family we collect all those \(B_j's\) with the \(x_i\)-side containing the \(x_j\)-side of \(B_k\); the remaining \(B_j's\) form the second family. By the monotonicity properties of \(w\) the \(x_j\)-side of the \(B_j's\)'s in the second family contains the \(x_j\)-side of \(B_k\). Let \(I = \{j: B_j\text{ is in the first family}\}\), similarly \(II = \{j: B_j\text{ is in the second family}\}\). We write
\[
\int_{B_k} \exp \left( \sum_{j < j_0} x_j \right) w(x) \, dx = \int_{B_k} \exp \left( \sum_{j \in I} x_j \right) w(x) \, dx + \int_{B_k} \exp \left( \sum_{j \in II} x_j \right) w(x) \, dx + \int_{B_k} \exp \left( \sum_{j < j_0 \in I} x_j \right) w(x) \, dx + \int_{B_k} \exp \left( \sum_{j < j_0 \in II} x_j \right) w(x) \, dx
\]
and
\[
\mathcal{F}_1 = \int_{B_k} \exp \left( \sum_{j < j_0} x_j \right) w(x) \, dx.
\]
We consider each \(\mathcal{F}_i, k = 1, 2, 3, \) separately. As far as \(\mathcal{F}_1\) is concerned, on \(B_k\) the expression \(\sum_{j \in I} x_j\) is independent of \(x_i\) and \(x_j\) since the sides in those directions of each \(B_j, j \in I\), contain the corresponding sides of \(B_k\) (the projections actually have this property). Thus if \(B_k = B_{k,1} \times B_{k,2} \times B_{k,3}\) we have
\[
\mathcal{F}_1 = \int_{B_k} \exp \left( \sum_{j \in I} x_j \right) w(x) \, dx \leq w(B_k) \mathcal{N}_2 \left( \exp \left( \sum_{j \in I} x_j \right), y \right).
\]
Similarly
\[
\mathcal{F}_2 \leq w(B_k) \mathcal{N}_2 \left( \exp \left( \sum_{j \in II} x_j \right), y \right).
\]
As for the last term, we have
\[
\mathcal{F}_3 = \int_{B_k} \exp \left( \sum_{j \in I} x_j \right) \exp \left( \sum_{j \in II} x_j \right) w(x) \, dx \leq w(B_k) \mathcal{N}_3 \left( \exp \left( \sum_{j \in I} x_j \right), y \right).
\]
Consequently
\[
\bigcup_{k=1}^N \left\{ y: \mathcal{N}_4 \left( \exp \left( \sum_{j} x_j \right), y \right) > 1 \right\}
\]
and
\[
w \left( \bigcup_{k=1}^N B_k \right) \leq \sum_{k=1}^N \left\{ y: \mathcal{N}_4 \left( \exp \left( \sum_{j} x_j \right), y \right) > 1 \right\}
\]
\[
\leq c \int_{\mathbb{R}^d} \exp \left( \sum_{j} x_j \right) w(x) \, dx \leq c 100e \sum_{j=0}^N w(B_j).
\]
This is precisely what we wanted to show.

In the context of weighed measures the \(L^2(\mathbb{R}^d)\) version of Theorem 4.2 for \(p > 1\) is also of interest. In this case the assumptions on \(w\) can be considerably relaxed. Indeed, let \(w\) satisfy the following property: there are constants \(0 < \gamma < 1\) such that for \(B = B_1 \times B_2 \times \varphi(B_1, B_2),\) we have
\[
\text{ess sup}_{x_1 \in B_1, x_2 \in B_2} w(x, E, x_3)w(x_1, B_2, x_3) > \gamma
\]
for a.e. \( x_2 \), implies
\[
\text{ess inf}_{x_1 \in B_1, x_3 \in W(B_1, x_2)} w(x_1, x_2, x_3) > \gamma
\]
for a.e. \( x_2 \), and
(11) If \( E \) is a measurable subset of \( B_1 \), then
\[
\text{ess sup}_{x_3 \in W(B_1, x_2)} w(E, x_3) > \gamma
\]
for a.e. \( x_2 \), implies
\[
\text{ess inf}_{x_3 \in W(B_1, x_2)} w(E, x_3) > \gamma
\]
for a.e. \( x_2 \).

In this case we have

**Theorem 4.3.** Assume that \( \mathcal{A} \) is as before, that \( 1 < p < \infty \) and that \( w \) verifies properties (10) and (11). Then
\[
\|M_{\mathcal{A},u,f}\|_{L^p} \leq c\|f\|_{L^p}.
\]

Proof. We only sketch a proof here. As usual
\[
w((M_{\mathcal{A},u,f} > \lambda)) \approx w(\bigcup B_\lambda),
\]
where \( \{B_\lambda\} \) is a family in \( \mathcal{A} \) such that
\[
w(B_\lambda) \approx \lambda^p w(\lambda).
\]

Using the selection procedure as in Theorem 3.1 with \((x_1, x_3)\) here in place of the admissible \( x_2 \)-direction there and with
\[
\widehat{M}_{x_2}^{1,3} f(x_1, x_3) = \sup_{c(B)} \left( \frac{1}{c(B)} \int f(y_1, y_3) d\lambda_1(x_1) dy_1 \right),
\]
\[(x_1, x_3) \in B, B = B_1 \times \varphi(B_1, B_2), B_1 \in \mathcal{A}_1, B_2 \in \mathcal{A}_2,
\]
we can find a subfamily \( \{B_\lambda\} \subseteq \{B_\lambda\} \) so that
\[
\bigcup B_\lambda = \left[ M_{\mathcal{A}_2, x_1, x_3} \widehat{M}_{x_2}^{1,3} f \right] > \gamma
\]
(12)

In addition \( \{B_\lambda\} \) satisfies a sparseness property.

As a consequence of (11) and Theorem 2.2 (that is to say, of a similar argument involving \( \widehat{M}_{x_2}^{1,3} \) instead of the strong maximal function; observe that we don’t need doubling since \( \mathcal{A}_1 \) satisfies property (D)), \( \widehat{M}_{x_2}^{1,3} \) is bounded on \( L^p, p > 1 \). Also \( M_{\mathcal{A}_2, x_1, x_3} \) is of weak type \((1,1)\). By (12) we have
\[
w(\bigcup B_\lambda) \approx c w(\bigcup B_\lambda)
\]
for a.e. \( x_1 \), implies
\[
\text{ess inf}_{x_1 \in \bigcup B_\lambda} w(x_1, x_2, x_3) > \gamma
\]
for a.e. \( x_3 \), and
(11) If \( E \) is a measurable subset of \( B_1 \), then
\[
\text{ess sup}_{x_3 \in \bigcup B_\lambda} w(E, x_3) > \gamma
\]
for a.e. \( x_2 \), implies
\[
\text{ess inf}_{x_3 \in \bigcup B_\lambda} w(E, x_3) > \gamma
\]
for a.e. \( x_2 \).

As in the proof of Theorem 3.1 it follows that the sparseness of the \( \{B_\lambda\} \) implies that
\[
w(\bigcup B_\lambda) \leq c \int \widehat{M}_{x_2}^{1,3} \left[ M_{\mathcal{A}_2, x_1, x_3} \varphi \right] f(x_1, x_2, x_3) dy_1 dy_2.
\]

Consequently a similar inequality holds with \( w((M_{\mathcal{A}_2, f} > \lambda)) \) in the left-hand side. Multiplying both sides of this new inequality by \( \mu p^{-1} \) and integrating yields
\[
\|M_{\mathcal{A}_2, u,f}\|_{L^p} \leq c \|\widehat{M}_{x_2}^{1,3} \left[ M_{\mathcal{A}_2, x_1, x_3} \varphi \right] f\|_{L^p}.
\]

The boundedness in \( L^p \) of \( \widehat{M}_{x_2}^{1,3} \) and \( M_{\mathcal{A}_2, x_1, x_3} \varphi \) readily gives us the desired estimate.

**Remark 4.4.** Conditions (10) and (11) are not symmetric in the variables involved. Nevertheless, they can be relaxed in the direction of the definition of \( D \) to overcome this point.

While the proof of Theorem 4.2 does not seem to generalize to, for instance, basis \( \mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{A}_3 \) the generalization is immediate for Theorem 4.3. We point out here one such instance. Suppose that \( \varphi(x_1, \ldots, x_p) \geq 0, i = 1, \ldots, n \), is a continuous function, monotone in each variable separately and satisfying the conditions \( \varphi(x_1, \ldots, x_p) \geq 0, i = 1, \ldots, n \) and \( \varphi(x_1, \ldots, x_p) \approx \varphi(2x_1, \ldots, 2x_p) \). Consider the basis \( \mathcal{A} \) of \( x_1 \times \cdots \times x_n \times \varphi(x_1, \ldots, x_n) \) rectangles and a weight \( w \) which satisfies
\[
(13) \quad \text{There are constants } 0 < \gamma, \gamma < 1 \text{ such that for each } B \in \mathcal{A} \text{ and measurable } E \subseteq B, |E| B > \gamma \text{ implies } w(E) > \gamma.
\]

We then have that \( M_{\mathcal{A}_2, u} \) is bounded in \( L^p, p > 1 \). A few comments about the proof of this result. First, by letting \( \lambda \to 0 \) we find that
\[
\text{ess sup}_{x_3 \in x} w_{x_3, x_1 + 1} (E) w_{x_3, x_1 + 1} (B) > \gamma
\]
implies
\[
\text{ess inf}_{x_3 \in x} w_{x_3, x_1 + 1} (E) w_{x_3, x_1 + 1} (B) > \gamma,
\]
where \( B = B_1 \times \cdots \times B_n \times \varphi(B_1, \ldots, B_n) \).

The proof of Theorem 4.3 carries over almost unchanged in this setting. Nevertheless, since \( \mathcal{A} \) does not necessarily verify (D) we must bring in doubling at some point. Property (13) does not seem to imply doubling in each direction. It is clear, however, that (13) implies
\[
w(B^p) \leq c w(B).
\]
where $B^*$ is the rectangle with same center as $B$ but with $x_n$ side length twice as large and with other sidelengths $\leq c_1$: the corresponding sidelengths of $B$. Here $c_1$ depends on the constant appearing in the relation $\varphi(t_1, \ldots, t_n) \approx \varphi(2t_1, \ldots, 2t_n)$. It then follows that

$$w(\bigcup B_j) \leq c w(\bigcup B_j^*) \leq c \sum w(B_j).$$

This inequality replaces $w(\bigcup B_j) \leq c w(\bigcup B_j)$ in Theorem 2.1 and it is sufficient to complete the rest of the proof as in that theorem.

To close the paper we briefly mention some applications to various topics. In the first place, we would like to discuss the relation between maximal functions and rearrangements, in this context see Bagby [1].

Theorem 2.1 gives

$$(M_{A,w} f^*)(t) \leq c \left( \sum_{j=1}^{N} M_{A,w_j} f^j(t) \right).$$

Similarly Theorem 3.1 gives

$$(M_{A,w} f^*)^j(t) \leq c \left( M_{A,w_2} M_{A,w_1} f^j(t) \right).$$

These estimates can be used to obtain mapping properties of the maximal operators on rearrangement invariant spaces (of index $>1$), for instance.

In another direction Córdoba and R. Fefferman [3] have observed that continuity properties of maximal operators can often be stated in terms of coverings. The same observation applies to our context. For instance, Proposition 3.2 can be thought of as an extension to the case of product basis of Córdoba's result concerning the equivalence between maximal theorems and coverings.

References


DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
Bloomington, Indiana 47405, U.S.A.

Received November 2, 1983