STUDIA MATHEMATICA, T. LXXX. (1984)

The strong maximal function with respect to measures*

by

BJÖRN JAWERTH and ALBERTO TORCHINSKY (Bloomington, Indiana)

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

$$Mf(x) = \sup_R \frac{1}{|R|} \int_R |f(y)|\, dy, \qquad x \in R.$$

Then for each $\lambda > 0$

$$\text{(JMZ)} \qquad |\{x \in R^n\colon\ Mf(x) > \lambda\}| \leqslant c \int \frac{|f(x)|}{\lambda} \left(1 + \log^+ \frac{|f(x)|}{\lambda}\right)^{n-1} dx.$$

Together with $\|Mf\|_{L^\infty} \leqslant \|f\|_{L^\infty}$, (JMZ) obtains the boundedness of M in the $L^p(R^n)$ 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 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 Jessen, Marcinkiewicz and Zygmund [12] asserts that the basis $\mathscr{R} = \{R\}$ of rectangles R in Euclidean space R^n with sides parallel to the coordinate axis differentiates the class $L(\log^+ L)^{n-1}_{\text{loc}}(R^n)$. The quantitative version of this result is the following estimate. Let $|E|$ denote the Lebesgue measure of E. Associated with $\mathscr{R}$ we consider the strong maximal operator

$$Mf(x) = \sup_R \frac{1}{|R|} \int_R |f(y)|\, dy, \qquad x \in R.$$

* Research partially supported by the N.F.R. and N.S.F.

Then for each $\lambda > 0$

$$\text{(JMZ)}\qquad |\{x \in R^n \colon Mf(x) > \lambda\}| \leqslant c \int \frac{|f(x)|}{\lambda}\left(1+\log^+ \frac{|f(x)|}{\lambda}\right)^{n-1} dx.$$

Together with $\|Mf\|_{L^\infty} \leqslant \|f\|_{L^\infty}$, (JMZ) obtains the boundedness of M in the $L^p(R^n)$ 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 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 R^n and put

$$M_w f(x) = \sup \frac{1}{w(R)} \int_R |f(y)|\, w(y)\, dy, \qquad x \in R,$$

where $w(R) = \int_R w(y)\, dy$. Under very general conditions, namely that w be an *admissible measure* we show that for $\lambda > 0$

$$\text{(1)}\qquad w(\{x \in R^n \colon M_w f(x) > \lambda\}) \leqslant c \int \frac{|f(x)|}{\lambda}\left(1+\log^+ \frac{|f(x)|}{\lambda}\right)^{n-1} w(x)\, dx.$$

R. Fefferman [8] initiated the study of the boundedness properties of M_w and proved that if w is uniformly A_∞ in each variable, then

$$\|M_w f\|_{L^p_w} \leqslant c \|f\|_{L^p_w}, \qquad 1 < p < \infty.$$

An estimate similar to (1), but with the loss of a logarithm, holds for general basis $\mathscr{B}$ = product of Buseman–Feller basis which differentiate L^1 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 $\mathscr{B}$. 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 $\mathscr{B} = \{B\}$ of bounded, measurable sets of positive Lebesgue measure in R^n is said to be a *differentiation basis* if for each $x \in R^n$ there is a subfamily $\mathscr{B}(x)$ of $\mathscr{B}$ such that

(i) if $B \in \mathscr{B}(x)$, then $x \in B$;

(ii) each $\mathscr{B}(x)$ contains sets of arbitrarily small diameter.

$\mathscr{B}$ is said to be a *Buseman–Feller basis*, or *B–F basis* if in addition

(iii) each B is open;

(iv) if $x \in B$, then $B \in \mathscr{B}(x)$.

Consider a finite number, say N, of B–F bases $\mathscr{B}_i = \mathscr{B}_i(R^{n_i})$ in R^{n_i}, $1 \leqslant i \leqslant N$, and put $n = n_1 + \ldots + n_N$. The family $\mathscr{B} = \{B \subseteq R^n \colon B = B_1 \times \ldots \times B_N,\ B_i \in \mathscr{B}_i,\ i = 1, \ldots, N\}$ is a B–F basis in R^n. $\mathscr{B}$ is called the *product basis* of the $\mathscr{B}_i$'s and is often denoted by $\mathscr{B}_1 \times \ldots \times \mathscr{B}_N$ or $\prod_{i=1}^{N} \mathscr{B}_i$. We identify points x in R^n with $(x_1, \ldots, x_N)$, $x_i \in R^{n_i}$, and, when appropriate, subsets of R^{n_i} with subsets of R^n (by adding a number of coordinates 0) and vice versa.

To each positive locally summable function w defined in R^n we associate the measure

$$w(E) = \int \chi_E(x)\, w(x)\, dx, \qquad E \subseteq R^n,$$

and the *restriction measures* $w_{x_i} dx_1 \ldots d\hat{x}_i \ldots dx_N$, $i = 1, \ldots, N$, given by

$$w_{x_i}(E) = w(x_i, E) = \int \chi_E(x_1, \ldots, \hat{x}_i, \ldots, x_n)\, w(x)\, dx_1 \ldots d\hat{x}_i \ldots dx_n, \qquad \text{for}\quad E \subseteq R^{n-n_i}.$$

As customary, the check $\hat{}$ denotes the fact that the corresponding variable, or differential, is missing. Thus

$$dx_1 \ldots d\hat{x}_i \ldots dx_n = dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_n,$$

etc.

Similarly we define the restrictions $w_{x_{i_1} x_{i_2}}$, $w_{x_{i_1} x_{i_2} x_{i_3}}$, etc. We reserve the notation $w^{(j)}$ for $w_{x_1, \ldots, \hat{x}_j, \ldots, x_n}$.

To subsets E of R^n and to each $x_i \in R^{n_i}$, $1 \leqslant i \leqslant N$, we assign the sections $E_{x_i} \subseteq R^{n-n_i}$ given by

$$E_{x_i} = \{(x_1, \ldots, \hat{x}_i, \ldots, x_n) \colon (x_1, \ldots, x_n) \in E\},$$

E_{x_i} is called the *section of* E *at* x_i. One last notation. Let $\mathscr{B} = \prod_{i=1}^{N} \mathscr{B}_i$, $N \geqslant 2$. For $E \subseteq B_{x_i}$, $B \in \mathscr{B}$, we set

$$W(x_i, E, B) = w(x_i, E)/w(x_i, B_{x_i}).$$

We say that w is an *admissible measure with respect to* $\mathscr{B}$, or plainly that w is an *admissible measure*, and we denote this by $w \in \mathscr{D}(\mathscr{B})$, if there are constants $0 < \gamma,\ \bar{\gamma} < 1$, such that to each $B = \prod_{j=1}^{N} B_j$ in $\mathscr{B}$ there corresponds an index $i = i(B)$, $1 \leqslant i \leqslant N$ such that for $E \subseteq B_{x_i}$

$$\operatorname*{ess\,sup}_{x_i \in B_i} W(x_i, E, B_{x_i}) \geqslant \gamma$$

implies

$$\operatorname*{ess\,inf}_{x_i \in B_i} W(x_i, E, B_{x_i}) \geqslant \bar{\gamma}.$$

The index $i(B)$ is called an *admissible direction* at B. Analogously we introduce a "higher order" variant of $\mathscr{D}(\mathscr{B})$, namely the class $\bar{\mathscr{D}}(\mathscr{B})$. Let π be a permutation of the set of N indices $\{1, \ldots, N\}$. To $1 \leqslant s \leqslant N-1$ and

$$E \subseteq \prod_{j \neq \pi(1), \ldots, \pi(s)} B_j = B_{x_{\pi(1)}, \ldots, x_{\pi(s)}}$$

we associate

$$W(\pi, E, B) = w_{x_{\pi(1)}, \ldots, x_{\pi(s)}}(E) / w_{x_{\pi(1)}, \ldots, x_{\pi(s)}}(B_{x_{\pi(1)}, \ldots, x_{\pi(s)}}).$$

We say that $w \in \bar{\mathscr{D}}(\mathscr{B})$ if there are constants $0 < \gamma,\ \bar{\gamma} < 1$ such that to each B in $\mathscr{B}$ there corresponds a fixed permutation $\pi = \pi(\mathscr{B})$ of $\{1, \ldots, N\}$ such that for each $1 \leqslant s \leqslant N-1$ and $E \subseteq B_{x_{\pi(1)}, \ldots, x_{\pi(s)}}$

$$\operatorname*{ess\,sup}_{x_{\pi(s)} \in B_{\pi(s)}} W(\pi, E, B) \geqslant \gamma$$

implies

$$\operatorname*{ess\,inf}_{x_{\pi(s)} \in B_{\pi(s)}} W(\pi, E, B) \geqslant \bar{\gamma}.$$

for almost all $x_{\pi(1)}, \ldots, x_{\pi(s-1)}$.

The reader will observe that weights w of the form $w(x) = \prod_{i=1}^{N} w_i(x_i)$, corresponding to product measures, are in $\bar{\mathscr{D}}(\mathscr{B})$. However, if only $w(x) = u(x_j)\, v(x_1, \ldots, \hat{x}_j, \ldots, x_n)$, $1 \leqslant j \leqslant N$, then $w \in \mathscr{D}(\mathscr{B})$.

2. The strong maximal function with respect to measures. We start by considering a certain subfamily of $\mathscr{R}(R^n)$. More specifically, for $1 \leqslant k \leqslant n$ and $n = n_1 + \ldots + n_k$ we let $\mathscr{R}^k(R^n)$ be the family of those R's in $\mathscr{R}(R^n)$ such that the length of n_1 of its sides (we assume those corresponding to the first n_1 directions in R^n, otherwise we can rename the directions) are equal, the next n_2 are equal, and so on until the last n_k side-lengths. We call the directions corresponding to equal side-lengths *blocks*. $\mathscr{R}^1(R^n)$ is thus the collection of cubes in R^n and $\mathscr{R}^n(R^n) = \mathscr{R}$.

For a positive locally summable function w we introduce the maximal operator with respect to w and $\mathscr{R}^k(R^n)$ by

$$M_{k,w} f(x) = \sup \frac{1}{w(R)} \int_R |f(y)|\, w(y)\, dy, \qquad x \in R,\ R \in \mathscr{R}^k(R^n).$$

Similarly, for $v(x_1, \ldots, \hat{x}_j, \ldots, x_k)$ defined on R^{n-n_j} we set

$$M_{k-1,v}^{(j)} f(x) = \sup \frac{1}{v(R)} \int_R |f(y)|\, v(y)\, dy, \qquad x \in R,\ R \in \mathscr{R}^{k-1}(R^{n-n_j}).$$

Finally for $\bar{v}(x_j)$ defined on R^{n_j} we put

$$M_{1,\bar{v}}^{(j)} f(x) = \sup \frac{1}{\bar{v}(R)} \int_R |f(y)| \bar{v}(y)\, dy, \qquad x \in R,\ R \in \mathscr{R}^1(R^{n_j}),$$

$1 \leqslant j \leqslant k$. $M_{1,\bar{v}}^{(j)}$ corresponds to the usual Hardy–Littlewood maximal function and as is well known, if $\bar{v}$ is merely doubling, the weak-type estimate

$$\text{(1)} \qquad \bar{v}\big(\{x \in R^{n_j}:\ M_{1,\bar{v}}^{(j)} f(x) > \lambda\}\big) \leqslant \frac{c}{\lambda} \int_{R^{n_j}} |f(y)|\, \bar{v}(y)\, dy, \qquad \lambda > 0,$$

obtains.

As for the maximal operators $M_{k-1,v}^{(j)}$, for a fixed $1 \leqslant j \leqslant k$, following Jawerth [11] we say that v is a *density weight* (with respect to $\mathscr{R}^{k-1}(R^{n-n_j})$) if there is a function $c(\lambda)$ such that

$$\text{(2)} \qquad v\big(\{x \in R^{n-n_j}:\ M_{k-1,v}^{(j)} \chi_E(x) > \lambda\}\big) \leqslant c(\lambda)\, v(E)$$

for each measurable set $E \subseteq R^{n-n_j}$ and $0 < \lambda < 1$. This condition is satisfied if, for instance, $M_{k-1,v}^{(j)}$ is of restricted weak type (p, p) for some $p < \infty$. When $v = w_{x_j}$ and $\bar{v} = w^{(j)}(x_j)$, $1 \leqslant j \leqslant k$, 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 $c(\lambda)$ in (2) and c in (1) are independent of (almost all) x_j and $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$, respectively.

We can now state our first result.

THEOREM 2.1. *Suppose that $w \in \mathscr{D}(R^k)$ and that w_{x_j} is a density weight, uniformly with respect to $\mathscr{R}^{k-1}(R^{n-n_j})$, $1 \leqslant j \leqslant k$. Then there are constants c and c_1 such that for all f and $\lambda > 0$*

$$w\big(\{x \in R^n:\ M_{k,w} f(x) > \lambda\}\big) \leqslant \frac{c}{\lambda} \sum_{j=1}^{N} \int_{\{M_{k-1,w_{x_j}}^{(j)} f > c_1 \lambda\}} M_{k-1,w_{x_j}}^{(j)} f(x)\, w(x)\, dx.$$

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

(3) $$w(E) \leqslant w(\bigcup R_i)$$

and

(4) $$\int_{R_i} |f(x)|\, w(x)dx > \lambda w(R_i), \qquad 0 \leqslant i \leqslant M.$$

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

(5) $$w(\bigcup R_i) = w(\bigcup_{h=1}^{k} \bigcup_{R_i \in \mathscr{A}_h} R_i) \leqslant \sum_{h=1}^{k} w(\bigcup_{R_i \in \mathscr{A}_h} R_i).$$

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

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

(6) $$w(\bigcup_{R_i \in \mathscr{A}_1} R_i) \leqslant cw(\bigcup_j R_j).$$

The selection procedure and criteria of sparseness are as follows: since each $R_i \in \mathscr{R}^k(R^n)$, the side lengths corresponding to the n_1 block are all equal. Let R_0 be an R_i in $\mathscr{A}_1$ with largest sidelength in a direction in the n_1 block, the n_1 direction say. If $R_0, \ldots, R_{j-1}$ have been chosen, let R_j be a rectangle $R = I_1 \times \ldots \times I_k$ in $\mathscr{A}_1$ with largest side length in the n_1 direction such that

(7 >) $$w\big(x_1, R_{x_1} \setminus \bigcup_{h<j} (R_h)_{x_1}\big) \geqslant (1-\gamma)\, w(x_1, R_{x_1})$$

obtains for almost all x_1 in I_1.

We go on until $\mathscr{A}_1$ 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_1 block. We observe that

(8) $$\bigcup_{R_i \in \mathscr{A}_1} R_i \subseteq \{M^{(1)}_{k-1, w_{x_1}}(\chi_{\bigcup_j (R_j^*)_{x_1}}) \geqslant \bar{\gamma}\}.$$

Indeed, if $R_i \in \mathscr{A}_1$ is one of the R_j's, then (8) clearly holds. So assume $R \in \mathscr{A}_1$ is not one of the R_j's. Then R was discarded after $R_0, \ldots, R_{j_0-1}$ had been selected, thus

(9) $$\begin{Bmatrix}\text{side length in } n_1 \\ \text{direction of } R\end{Bmatrix} \geqslant \begin{Bmatrix}\text{side length in } n_1 \\ \text{direction of } R_{j_0}\end{Bmatrix}$$

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

$$w\big(x_1, R_{x_1} \cap \big(\bigcup_{h<j_0} (R_h)_{x_1}\big)\big) \geqslant \gamma w(x_1, R_{x_1})$$

and consequently also

(7 <) $$w\big(x_1, R_{x_1} \cap \big(\bigcup_{h<j_0} (R_h^*)_{x_1}\big)\big) \geqslant \gamma w(x_1, R_{x_1})$$

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

Since

$$\begin{Bmatrix}\text{side length in } n_1 \\ \text{direction of } R_{j_0-1}\end{Bmatrix} \geqslant \begin{Bmatrix}\text{side length in } n_1 \\ \text{direction of } R\end{Bmatrix}$$

we readily see that

$$R_{x_1} \cap \big(\bigcup_j (R_j^*)_{x_1}\big) = \text{fixed set independent of } x_1 \in I_1.$$

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

$$w\big(x_1, R_{x_1} \cap \big(\bigcup_j (R_j^*)_{x_1}\big)\big)/w(x_1, R_{x_1}) \geqslant \bar{\gamma} \quad \text{for almost all } x_1 \text{ in } R^{n_1}.$$

Whence

$$M^{(1)}_{k-1, w_{x_1}}(\chi_{\bigcup (R_j^*)_{x_1}})(x_2, \ldots, x_n) \geqslant \bar{\gamma} \quad \text{for } (x_2, \ldots, x_k) \in R_{x_1} \quad \text{and} \quad x_1 \in I_1,$$

and (8) obtains.

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

$$\begin{aligned} w\big(x_1, \bigcup_{R_i \in \mathscr{A}_1} (R_i)_{x_1}\big) &\leqslant w\big(x_1, \{M^{(1)}_{k-1, w_{x_1}}(\chi_{\bigcup (R_j^*)_{x_1}}) \geqslant \bar{\gamma}\}\big) \\ &\leqslant c(\bar{\gamma})\, w\big(x_1, \bigcup (R_j^*)_{x_1}\big). \end{aligned}$$

Integrating this inequality with respect to x_1 we see that

$$w(\bigcup_{R_i \in \mathscr{A}_1} R_i) \leqslant cw(\bigcup_j R_j^*).$$

However, it is readily seen that $w(\bigcup_j R_j^*) \leqslant cw(\bigcup_j R_j)$ since $w \in \mathcal{D}(\mathscr{B})$. We have thus proved (6).

To complete the proof of our theorem, with $M^{(1)}_{k-1, w_{x_1}}$ on the right-hand side, we will estimate $\sum_j w(R_j)$. Let $\{E_j\}$ be the disjoint sequence obtained from the R_j's by

$$E_j = R_j \setminus (\bigcup_{i<j} R_i), \qquad E_0 = R_0.$$

Then $\bigcup E_j = \bigcup R_j$. From (7 >) we see that

$$w\big(x_1, (E_j)_{x_1}\big) \geqslant (1-\gamma)\, w\big(x_1, (R_j)_{x_1}\big) \quad \text{for all } j.$$

Now

(10)
$$\begin{aligned}\sum_j \int_{(R_j)_{x_1}} &|f(x_1, y_2, \dots, y_n)|\, w_{x_1}(y_2, \dots, y_n)\, dy_2 \dots dy_n\\ &\leqslant c\sum_j w(x_1, (E_j)_{x_1}) \inf_{(y_2,\dots,y_n)\in (R_j)_{x_1}} M^{(1)}_{k-1,w_{x_1}} f(x_1, y_2, \dots, y_n)\\ &\leqslant c\sum_j \int_{(E_j)_{x_1}} M^{(1)}_{k-1,w_{x_1}} f(x_1, y_2, \dots, y_n)\, w_{x_1}(y_2, \dots, y_n)\, dy_2 \dots dy_n\\ &= c \int_{\bigcup (R_j)_{x_1}} M^{(1)}_{k-1,w_{x_1}} f(x_1, y_2, \dots, y_n)\, w_{x_1}(y_2, \dots, y_n)\, dy_2 \dots dy_n\\ &\leqslant c\varepsilon\lambda w(x_1, \bigcup (R_j)_{x_1}) +\\ +c &\int_{\{M^{(1)}_{k-1,w_{x_1}} f > \varepsilon\lambda\}} M^{(1)}_{k-1,w_{x_1}} f(x_1, y_2, \dots, y_n)\, w_{x_1}(y_2, \dots, y_n)\, dy_2 \dots dy_n.\end{aligned}$$

Whence integrating (10) over R^{n_1} we see that

(11)
$$\sum_j \int_{R_j} |f(y)|\, w(y)\, dy \leqslant c\varepsilon\lambda w(\bigcup R_j) + c \int_{\{M^{(1)}_{k-1,w_{y_1}} f > \varepsilon\lambda\}} M^{(1)}_{k-1,w_{y_1}} f(y)\, w(y)\, dy.$$

Also from (4) it follows that

(12)
$$w(\bigcup R_j) \leqslant \sum w(R_j) \leqslant \lambda^{-1} \sum_j \int_{R_j} |f(y)|\, w(y)\, dy.$$

Thus combining (11) and (12) and choosing ε so that $c\varepsilon < 1/2$, and $c_1 < 1/2c$ with c the constant in (11), we obtain that

(13)
$$w(\bigcup R_j) \leqslant \frac{c}{\lambda} \int_{\{M^{(1)}_{k-1,w_{y_1}} f > c_1\lambda\}} M^{(1)}_{k-1,w_{y_1}} f(y)\, w(y)\, dy.$$

Repeating this argument for each family $\mathscr{A}_h$, and summing over h, from (5) and (13) we get

(14)
$$w(\bigcup R_i) \leqslant \frac{c}{\lambda} \sum_{j=1}^{k} \int_{\{M^{(j)}_{k-1,w_{x_j}} f > c_1\lambda\}} M^{(j)}_{k-1,w_{x_j}} f(x)\, w(x)\, dx.$$

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 $\mathscr{R}^k(R^n)$ has an admissible direction in the set $\{n_j\}_{j\in J}$, $J \subseteq \{1, \dots, k\}$, 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 $w \in \mathscr{D}(R^k)$, $1 \leqslant k \leqslant n$, and that w is doubling in each n_i (block) direction uniformly. Then for $\lambda > 0$*

$(\mathrm{JMZ})_w$
$$w(\{x \in R^n\colon M_{k,w} f(x) > \lambda\}) \leqslant c \int \frac{|f(x)|}{\lambda} \left(1 + \log^+ \frac{|f(x)|}{\lambda}\right)^{k-1} w(x)\, dx.$$

Proof. If it were true that $w \in \mathscr{D}(R^k)$ implied that each section $w_{x_j} \in \mathscr{D}(R^{k-1})$, then 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 $\mathscr{R}^k$ into $k!$ disjoint subfamilies $\mathscr{R}^k_\pi$, one corresponding to each permutation π of $\{1, \dots, k\}$, in such a way that $\mathscr{R}^k_\pi$ only contains those rectangles $B = \prod_{j=1}^{k} B_j$ for which we have for each s, $1 \leqslant s \leqslant k-1$, and $E \subseteq \prod_{j \neq \pi(1),\dots,\pi(s)} B_j$,
$$\operatorname*{ess\,sup}_{x_{\pi(s)} \in B_{\pi(s)}} W(\pi, E, B) \geqslant \gamma$$
implies
$$\operatorname*{ess\,inf}_{x_{\pi(s)} \in B_{\pi(s)}} W(\pi, E, B) \geqslant \bar{\gamma}.$$

Here $W(\pi, E, B)$ denotes the quantity introduced in Section 1.

Now for each π we introduce the auxiliary operators $M^\pi_{k,w}$ and $M_{k-s,w_{x_{\pi(1)},\dots,x_{\pi(s)}}}$, $1 \leqslant s \leqslant k-1$, by
$$M^\pi_{k,w} f(x) = \sup \frac{1}{w(R)} \int_R |f(y)|\, w(y)\, dy, \qquad x \in R,\ R \in \mathscr{R}^k_\pi(R^n).$$

and
$$M_{k-s,w_{x_{\pi(1)},\dots,x_{\pi(s)}}} f(x) = \sup \frac{1}{w_{x_{\pi(1)},\dots,x_{\pi(s)}}(r)} \int_r |f(y)|\, w_{x_{\pi(1)},\dots,x_{\pi(s)}}(y)\, dy,$$

$x \in r$ and $r \subseteq R^{n-(n_{\pi(1)}+\dots+n_{\pi(s)})}$, $r = R_{x_{\pi(1)},\dots,x_{\pi(s)}}$ for some $R \in \mathscr{R}^k_\pi(R^n)$.

Since $w \in \mathscr{D}(R^k)$ we have
$$M_{k,w} f(x) \leqslant \sum_\pi M^\pi_{k,w} f(x).$$

We now show that for $M^\pi_{k,w}$ the iteration argument described above works. Indeed, for $s \leqslant k-1$, $M^\pi_{1,w_{x_{\pi(1)},\dots,x_{\pi(s)}}} f(x)$, 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

$$w_{x_{\pi(1)},\dots,x_{\pi(s)}}(\{M^{\pi}_{k-s,w_{x_{\pi(1)},\dots,x_{\pi(s)}}}f>\lambda\})$$

$$\leqslant \frac{c}{\lambda}\int\limits_{\{M^{\pi}_{k-s-1,w_{x_{\pi(1)},\dots,x_{\pi(s-1)}}}f>\lambda\}} M^{\pi}_{k-s-1,w_{x_{\pi(1)},\dots,x_{\pi(s-1)}}}f(y)\,w_{x_{\pi(1)},\dots,x_{\pi(s)}}(y)(dy)$$

for $0\leqslant s\leqslant k-1$. In this estimate we use $x_{\pi(0)}=x$ and $w_{x_{\pi(0)}}=w$.

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 $w\in\mathscr{D}(R^k)$, $k\geqslant 3$. Suppose that each of the n_i (block) directions, $1\leqslant i\leqslant k-1$, are admissible at every R in $\mathscr{R}^k$. Then for almost every x_1, w_{x_1} has the property that the n_i-directions, $2\leqslant i\leqslant k-1$, are admissible at every R in $\mathscr{R}^{k-1}$.*

Proof. Let us show that, for instance, the n_2-direction is admissible. For $\varepsilon>0$ let $r_\varepsilon(x_1)$ be the cube in R^{n_1} with side length 2ε and center x_1. For w the n_2-direction is admissible. Thus there are constants $0<\gamma,\bar{\gamma}<1$, such that for every rectangle $R_2\times R'$ in $\mathscr{R}^{k-1}$ and every subset $E\subseteq R'$

$$\operatorname*{ess\,sup}_{x_2\in R_2} w(r_\varepsilon(x_1),x_2,E)/w(r_\varepsilon(x_1),x_2,R')\geqslant\gamma,$$

(15) implies

$$\operatorname*{ess\,inf}_{x_2\in R_2} w(r_\varepsilon(x_1),x_2,E)/w(r_\varepsilon(x_1),x_2,R')\geqslant\bar{\gamma}$$

for all x_1.

Let now ε tend to 0. By the (usual) Lebesgue differentiation theorem $w(r_\varepsilon(x_1),x_2,E)/(2\varepsilon)^{n_1}$ tends to $w(x_1,x_2,E)$ for almost all x_1. Similarly $w(r_\varepsilon(x_1),x_2,R')/(2\varepsilon)^{n_1}$ tends to $w(x_1,x_2,R')$, again for almost all x_1. Consequently (15) implies that for almost all x_1, the statement of (15) holds with $r_\varepsilon(x_1)$ replaced by x_1 there. This is equivalent to saying that the n_2-direction is admissible for w_{x_1}.

An important class of weights for which Proposition 2.3 holds, and consequently also $(\mathrm{JMZ})_w$, Theorem 2.1, and Theorem 2.2 hold, is the collection of those weights w which are uniformly in A_∞ 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 notations we restrict ourselves to the basis $\mathscr{B}=\mathscr{B}_1\times\mathscr{B}_2$, the product of (two) B–F bases $\mathscr{B}_1(R^{n_1})$ and $\mathscr{B}_2(R^{n_2})$. Points x in R^n, $n=n_1+n_2$, are denoted by $x=(x_1,x_2)$, $x_i\in R^{n_i}$. Let

$$M_{\mathscr{B},w}f(x)=\sup\frac{1}{w(B)}\int_B|f(y)|\,w(y)\,dy,\qquad x\in B,\ B \text{ in } \mathscr{B}.$$

The maximal operators $M_{\mathscr{B}_1,v}$ and $M_{\mathscr{B}_2,\bar{v}}$ are defined similarly. We can then prove

THEOREM 3.1. *Suppose that $w\in\mathscr{D}(\mathscr{B})$ and that w_{x_i} is a density weight uniformly with respect to $\mathscr{B}_i(R^{n_i})$, $i=1,2$. Then there are constants c and c_1 such that for $\lambda>0$*

$$w(\{x\in R^n\colon M_{\mathscr{B},w}f(x)>\lambda\})$$

$$\leqslant\frac{c}{\lambda}\Big(\iint\limits_{\{M_{\mathscr{B}_1,w_{x_2}}M_{\mathscr{B}_2,w_{x_1}}f>c_1\lambda\}} M_{\mathscr{B}_1,w_{x_2}}M_{\mathscr{B}_2,w_{x_1}}f(x)\,w(x)\,dx+$$

$$+\iint\limits_{\{M_{\mathscr{B}_2,w_{x_1}}M_{\mathscr{B}_1,w_{x_2}}f>c_1\lambda\}} M_{\mathscr{B}_2,w_{x_1}}M_{\mathscr{B}_1,w_{x_2}}f(x)\,w(x)\,dx\Big).$$

Proof. Let $\mathscr{O}_\lambda=\{M_{\mathscr{B},w}f>\lambda\}$; then $\mathscr{O}_\lambda$ is an open set in R^n. If E is an arbitrary compact subset of $\mathscr{O}_\lambda$ there is a (finite) family of $\{B_i\}\subseteq\mathscr{B}(R^n)$ such that

(1) $$w(E)\leqslant cw(\bigcup B_i)$$

and

(2) $$\int_{B_i}|f(x)|\,w(x)\,dx>\lambda w(B_i),\qquad \text{all } i.$$

Since $w\in\mathscr{D}(\mathscr{B})$ the family $\mathscr{A}=\{B_i\}$ can be divided into two disjoint subfamilies, $\mathscr{A}_1$ and $\mathscr{A}_2$ say, so that for all rectangles in $\mathscr{A}_i$ the ith direction is admissible, $i=1,2$. We will consider the family $\mathscr{A}_2$, the argument for $\mathscr{A}_1$ being similar is omitted. We first introduce a notion of *sparseness* and select a subfamily $\{B_j\}$ of $\mathscr{A}_2=\{B_i\}$ so that

(3) $$w(\bigcup B_i)\leqslant cw(\bigcup B_j).$$

The selection procedure is as follows: choose B_0 to be the first B_i. Once $B_0,\dots,B_{j-1}$ have been chosen let $B_j=B_{j,1}\times B_{j,2}$ be the first among the remaining B_i's with the following property: let γ be the constant associated to $w \in \mathcal{D}(\mathcal{B})$. If $E_j \subseteq B_{j,1}$ is defined by

$$E_j = \{x_1 \in B_{j,1}\colon\ w(x_1, B_{j,2} \setminus \bigcup_{i<j} B_{i,2}) \geqslant \tfrac{1}{2} w(x_1, B_{j,2})\}, \tag{4>}$$

then

$$w(x_2, E_j) \geqslant (1-\gamma) w(x_2, B_{j,1}) \quad \text{for almost all } x_2 \text{ in } R^{n_2}. \tag{5>}$$

We go on until the family $\mathcal{A}_2$ is exhausted. We now show that (3) holds for $\mathcal{A}_2$. Observe that

$$\bigcup_{B_i \in \mathcal{A}_2} B_i \subseteq \{(x_1, x_2) \in R^n\colon\ M_{\mathcal{B}_1, w_{x_2}} \chi_{\{M_{\mathcal{B}_2, w_{x_1}}(\chi \bigcup_j B_{j,2}) \geqslant 1/2\}} \geqslant \bar{\gamma}\} \tag{6}$$

$$\equiv \{M_{\mathcal{B}_1, w_{x_2}} \chi_{\mathcal{U}} \geqslant \bar{\gamma}\}, \quad \text{say.}$$

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

$$w(x_2, B_{i,1} \setminus E_i) \geqslant \gamma w(x_2, B_{i,1}) \tag{5<}$$

for a set of x_2's of positive measure in R^{n_2}.

Moreover, since $B_i \in \mathcal{A}_2$ it follows that

$$w(x_2, B_{i,1} \setminus E_i) \geqslant \bar{\gamma} w(x_2, B_{i,1}) \quad \text{for almost all } x_2 \text{ in } R^{n_2}. \tag{7}$$

But

$$B_{i,1} \setminus E_i = \{x_1 \in B_{i,1}\colon\ w(x_1, B_{i,2} \cap (\bigcup_{i<j_0} B_{i,2})) > \tfrac{1}{2} w(x_1, B_{i,2})\} \tag{8}$$
$$\subseteq \{x_1 \in B_{i,1}\colon\ M_{\mathcal{B}_2, w_{x_1}}(\chi_{\bigcup_{i<j_0} B_{i,2}}) > \tfrac{1}{2}\} = \mathcal{U}, \quad \text{say.}$$

Thus by (7) and (8)

$$M_{\mathcal{B}_1, w_{x_2}}(\chi_{\mathcal{U}}, x_1) > \tfrac{1}{2} \quad \text{for almost all } x_2 \in B_{i,2} \text{ and } x_1 \in B_{i,1}.$$

This means that B_i is contained in

$$\{M_{\mathcal{B}_1, w_{x_2}}(\chi\{x_1 \in B_{i,1}\colon \text{ for } x_2 \in \mathcal{B}_{i,2},\ M_{\mathcal{B}_2, w_{x_1}}(\chi_{\bigcup_{i<j_0} B_{i,2}}, x_2) > \tfrac{1}{2}\}, x_1) \geqslant \bar{\gamma}\}. \tag{9}$$

This proves (6), for the set in (9) only increases if we replace j_0 by the largest j. Now (3) follows readily. Indeed,

$$w(x_2, \bigcup_i B_{i,1}) \leqslant c(\bar{\gamma}) w(x_2, \mathcal{U})$$

whence integrating over x_2 we obtain

$$w(\bigcup B_i) \leqslant cw(\bigcup_j B_j).$$

Thus to complete our proof we only need estimate $\sum_j w(B_j)$. Put $B_j = B = B_1 \times B_2$. Notice that from (4>) and (5>) we get that

$$B_1 = E \cup (B_1 \setminus E) \quad \text{and} \quad w(x_1, B_2 \setminus (\bigcup_{i<j} B_{i,2})) > \tfrac{1}{2} w(x_1, B_{i,2})$$

for almost all x_1 in E and

$$w(x_2, E) > (1-\gamma) w(x_2, B_1), \quad \text{or} \quad w(x_2, B_1 \setminus E) \leqslant \gamma w(x_2, B_1)$$

for almost all x_2 in R^{n_2}. Therefore

$$w(x_2, B_1) = w(x_2, E) + w(x_2, B_1 \setminus E) \leqslant w(x_2, E) + \gamma w(x_2, B_1), \quad \text{or}$$

$$w(x_2, B_1) \leqslant \frac{1}{1-\gamma} w(x_2, E) \tag{10}$$

for almost all x_2 in R^{n_2}.

Consequently,

$$\int_{B_2} \int_{B_1} |f(x_1, x_2)| w(x_1, x_2)\, dx_1\, dx_2 \tag{11}$$
$$\leqslant \int_{B_2} w(x_2, B_1) \inf_{y_1 \in B_1} M_{\mathcal{B}_1, w_{x_2}} f(y_1, x_2)\, dx_2$$
$$\leqslant (1-\gamma)^{-1} \int_{B_2} w(x_2, E) \inf_{y_1 \in E} M_{\mathcal{B}_1, w_{x_2}} f(y_1, x_2)\, dx_2$$
$$\leqslant c \int_{B_2} \int_E M_{\mathcal{B}_1, w_{x_2}} f(x_1, x_2) w(x_1, x_2)\, dx_1\, dx_2.$$

Thus by (2) and (11) and with $g = M_{\mathcal{B}_1, w_{x_2}} f$ we see that

$$\sum_j w(B_j) \leqslant \frac{1}{\lambda} \sum \int_{B_{j,2}} \int_{B_{j,1}} |f(x_1, x_2)| w(x_1, x_2)\, dx_2\, dx_1$$
$$\leqslant \frac{c}{\lambda} \sum \int_{B_{j,2}} \int_{E_j} g(x_1, x_2) w(x_1, x_2)\, dx_2\, dx_1$$
$$= \frac{1}{\lambda} \sum \int_{E_j} \int_{B_{j,2}} g(x_1, x_2) w(x_1, x_2)\, dx_1\, dx_2$$

$$\leqslant \frac{c}{\lambda}\sum \int \chi_{E_j}(x_1)\, w(x_1, B_{j,2}) \inf_{x_2 \in B_{j,2}} M_{\mathscr{B}_2, w_{x_1}} g(x_1, x_2)\, dx_1$$

$$\leqslant \frac{c}{\lambda}\sum \iint \chi_{E_j}(x_1)\, \chi_{B_{j,2}}(x_2)\, M_{\mathscr{B}_2, w_{x_1}} g(x_1, x_2)\, w(x_1, x_2)\, dx_2\, dx_1$$

$$\leqslant \frac{c}{\lambda}\iint \chi_{\bigcup_j E_j \times B_{j,2}}(x_1, x_2)\, M_{\mathscr{B}_2, w_{x_1}} g(x_1, x_2)\, w(x_1, x_2)\, dx_1\, dx_2$$

$$\leqslant \frac{c}{\lambda}\iint \chi_{\bigcup_j B_{j,1} \times B_{j,2}}(x_1, x_2)\, M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f w\, dx_1\, dx_2$$

$$\leqslant \frac{c}{\varepsilon\lambda} w\Big(\bigcup_j B_j\Big) + \frac{c}{\lambda} \iint_{\{M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f > \varepsilon\lambda\}} M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f(x)\, w(x)\, dx.$$

Choose $\varepsilon = \frac{1}{2}c$ to obtain the desired conclusion for $\mathscr{A}_2$. As we said the argument for $\mathscr{A}_1$ 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 Δ_2-notation we associate $\psi^*(t)$ defined by

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

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

PROPOSITION 3.2. *Suppose w is as in Theorem 3.1 and that for some $\varepsilon > 0$ and all $\lambda > 0$*

$$w(x_2, \{M_{\mathscr{B}_1, w_{x_2}} f > \lambda\}) \leqslant c \int_{\{|f| > \varepsilon\lambda\}} \psi_1\left(\frac{|f(x_1)|}{\lambda}\right) w_{x_2}(x_1)\, dx_1$$

and

$$w(x_1, \{M_{\mathscr{B}_2, w_{x_1}} f > \lambda\}) \leqslant c \int_{\{|f| > \varepsilon\lambda\}} \psi_2\left(\frac{|f(x_2)|}{\lambda}\right) w_{x_1}(x_2)\, dx_2.$$

Then

$$w(\{M_{\mathscr{B}, w} f > \lambda\}) \leqslant c \int_{R^n} \int_{\varepsilon/2}^{2|f(x)|/\varepsilon c_1 \lambda} \psi_2^*(|f(x)|/s\lambda)\, d\psi_1(s)\, w(x)\, dx +$$
$$+ c \int_{R^n} \int_{\varepsilon/2}^{2|f(x)|/\varepsilon c_1 \lambda} \psi_1^*(|f(x)|/s\lambda)\, d\psi_2(s)\, w(x)\, dx,$$

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

$$I = \int_{\{M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f > c_1\}} M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f(x)\, w(x)\, dx.$$

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

$$\{M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f > t\} \subseteq \{M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f^{t/2} > t/2\}.$$

Now

$$I \leqslant \int_{R^{n_1}} \int_{c_1}^{\infty} w_{x_1}(\{M_{\mathscr{B}_2, w_{x_1}} M_{\mathscr{B}_1, w_{x_2}} f^{t/2} > t/2\})\, dt\, dx_1$$

$$\leqslant c \int_{R^{n_1}} \int_{c_1/2}^{\infty} \int_{\{M_{\mathscr{B}_1, w_{x_2}} f^t > \varepsilon t\}} \psi_2(M_{\mathscr{B}_1, w_{x_2}} f^t/t)\, w_{x_1}(x_2)\, dx_2\, dt\, dx_1$$

$$= c \int_{R^{n_2}} \int_{c_1/2}^{\infty} \{\psi_2(\varepsilon)\, w_{x_2}(\{M_{\mathscr{B}_1, w_{x_2}} f^t > \varepsilon t\}) +$$
$$+ \int_{\varepsilon}^{\infty} w_{x_2}(\{M_{\mathscr{B}_1, w_{x_2}} f^t > st\})\, d\psi_2(s)\}\, dt\, dx_2 \equiv I_1 + I_2, \qquad \text{say}.$$

Observe that

$$I_2 \leqslant c \int_{R^{n_2}} \int_{\varepsilon}^{\infty} \int_{c_1/2}^{\infty} \int_{\{|f^t| > \varepsilon s t\}} \psi_1\left(\frac{|f^t|}{st}\right) w_{x_2}(x_1)\, dx_1\, d\psi_2(s)\, dt\, dx_2.$$

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

$$I_2 \leqslant c \int_{R^n} \int_{\varepsilon}^{2|f|/\varepsilon c_1} \int_{c_1/2}^{|f|/\varepsilon s} \psi_1\left(\frac{|f(x)|}{st}\right) dt\, d\psi_2(s)\, w(x)\, dx$$

$$= c \int_{R^n} \int_{\varepsilon}^{2|f|/\varepsilon c_1} \psi_1^*(|f(x)|/s)\, d\psi_2(s)\, w(x)\, dx.$$

Similarly for I_1 we have

$$I_1 \leqslant c \int_{R^{n_2}} \int_{c_1/2}^{\infty} \int_{\varepsilon/2}^{\infty} w_{x_2}(\{M_{\mathscr{B}_1, w_{x_2}} f^t > st\})\, d\psi_2(s)\, dt\, dx_2$$

$$\leqslant c \int_{R^n} \int_{\varepsilon/2}^{2|f|/\varepsilon c_1} \psi_1^*(|f(x)|/s)\, d\psi_2(s)\, w(x)\, dx.$$

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. $R \in \mathscr{R}(R^{k+2})$ is said to be an $a \times b \times \varphi(a, b)$ rectangle if the sidelengths in the x_1 and x_2 directions are arbitrary numbers a and b and the sidelength in the x_3 (k-dimensional) block direction is $\varphi(a, b)$. Here φ is a function which is continuous at the origin, $\varphi(a, 0) = \varphi(0, b) = 0$, and monotone in each variable separately.

Let $\mathscr{B}$ be the B–F basis containing all such rectangles R. Recently Córdoba [3] answered affirmatively a problem of Zygmund by showing that for

$$M_{\mathscr{B}} f(x) = \sup \frac{1}{|R|} \int_R |f(y)|\, dy, \quad x \in R,\ R \in \mathscr{B},$$

the weak type estimate

$$|\{M_{\mathscr{B}} f > \lambda\}| \leqslant c \int_{R^{k+2}} \frac{|f(x)|}{\lambda} \left(1 + \log^+ \frac{|f(x)|}{\lambda}\right) dx$$

obtains.

We open this chapter by considering a slight extension of this result to the context 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 $\mathscr{R}(R^n)$ is basically equivalent to the family of n-fold products of dyadic cubes B and such cubes have the property:

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

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

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

(i) $\varphi(\mathscr{B})$ is a B–F basis verifying (D);

(ii) φ is *monotone*, i.e., if $B \subseteq B'$, then $\varphi(B) \subseteq \varphi(B')$ whenever $B, B' \in \mathscr{B}$.

It is then immediate to verify that $\mathscr{B} \times \varphi(\mathscr{B}) = \{\Omega = B \times \varphi(B)\colon B \in \mathscr{B}\}$ also satisfies (D) and that $M_{\mathscr{B} \times \varphi(\mathscr{B}), w}$ 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 Zygmund's problem.

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

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

$$\mathscr{N}_1 f(x) = \sup_{x \in B} w(B)^{-1} \int_{B_1} \inf_{y_2 \in B_2, y_3 \in B_3} |f(z, y_2, y_3)|\, w(z, B_2, B_3)\, dz,$$

and similarly

$$\mathscr{N}_2 f(x) = \sup_{x \in B} w(B)^{-1} \int_{B_2} \inf_{y_1 \in B_1, y_3 \in B_3} |f(y_1, z, y_3)|\, w(B_1, z, B_3)\, dz.$$

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

$$w(B)^{-1} \int_B \inf_{\substack{y_1 \in B_1 \\ y_3 \in B_3}} |f(y_1, z_2, y_3)| \left\{ \int_{B_1} \inf_{\substack{y_2 \in B_2 \\ y_3 \in B_3}} |f(z_1, y_2, y_3)|\, w(z_1, z_2, B_3)\, dz_1 \right\} dz_2.$$

The weights w we consider satisfy some of the following properties: there are constants $0 < \gamma, \bar{\gamma} < 1$, 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

$$\operatorname*{ess\,sup}_{x_2 \in B_2} w(x_2, E, B)/w(x_1, x_2, \varphi(B_1, B_2)) > \gamma \quad \text{for a.e. } x_1$$

implies

$$\operatorname*{ess\,inf}_{x_2 \in B_2} w(x_2, E, B)/w(x_1, x_2, \varphi(B_1, B_2)) > \bar{\gamma} \quad \text{for a.e. } x_1.$$

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

$$\operatorname*{ess\,sup}_{x_1 \in B_1} w(x_1, x_2, E)/w(x_1, x_2, \varphi(B_1, B_2)) > \gamma \quad \text{for a.e. } x_2$$

implies

$$\operatorname*{ess\,inf}_{x_1 \in B_1} w(x_1, x_2, E)/w(x_1, x_2, \varphi(B_1, B_2)) > \bar{\gamma} \quad \text{for a.e. } x_2.$$

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

$$\operatorname*{ess\,sup}_{x_1 \in B_1} W(x_1, E, B) > \gamma \quad \text{implies} \quad \operatorname*{ess\,inf}_{x_1 \in B_1} W(x_1, E, B) > \bar{\gamma};$$

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

$$\operatorname*{ess\,sup}_{x_2 \in B_2} w(x_1, x_2, E) > \gamma \quad \text{for a.e. } x_1$$

implies

$$\operatorname*{ess\,inf}_{x_2 \in B_2} w(x_1, x_2, E) > \bar{\gamma} \quad \text{for a.e. } x_1;$$

and finally there is a constant C such that

(5) $$\sup_{y_1, y_2} \frac{w(B_1, B_2, B_3)}{w(y_1, B_2, B_3)} \frac{w(y_1, y_2, B_3)}{w(B_1, y_2, B_3)} \leqslant C.$$

We can now prove

PROPOSITION 4.1. *Assume w satisfies properties* (3) *and* (4) *above. Then*

$$w(\{\mathcal{N}_1 f > \lambda\}) \leqslant \frac{c}{\lambda} \|f\|_{1,w}, \quad \lambda > 0,$$

i.e. $\mathcal{N}_1$ maps $L^1_w(R^n)$ into weak $L^1_w(R^n)$.

Sketch of the proof. Let B_α be a subfamily of $\mathcal{B}$ such that $\{\mathcal{N}_1 f > \lambda\} = \bigcup_\alpha B_\alpha$. We will select a sparse subfamily $\{B_j\}$ of $\{B_\alpha\}$ as follows: order the B_α's according to *decreasing* x_1 side-length; by property (D) this actually means by decreasing Lebesgue measure. Choose now inductively those B_j's which verify property (7 >) of Theorem 2.1, that is, after $B_0, \ldots, B_{j-1}$ have been chosen let B_j be a rectangle $B_\alpha = B_{\alpha,1} \times B_{\alpha,2} \times \varphi(B_{\alpha,1}, B_{\alpha,2})$ such that it has largest sidelength in the n_1 direction and

(6) $$w(x_1, B_{\alpha,1} \setminus \bigcup_{h<j} (B_h)_{x_1}) \geqslant (1-\gamma) w(x_1, B_{\alpha,x_1})$$

obtains for almost all x_1 in $B_{\alpha,1}$. We follow this procedure until $\{B_\alpha\}$ is exhausted. By (3) it readily follows that if we put

$$\tilde{M}^{(1)}_{2,v} f(x) = \sup_{x \in B} v(B)^{-1} \int_B |f(y)|\, v(y)\, dy, \quad B = B_2 \times \varphi(B_2, B_2),$$

then

(7) $$\bigcup B_\alpha \subseteq \{\tilde{M}^{(1)}_{2,w_{x_1}} (\chi_{\bigcup (B_j)_{x_1}}) > \bar{\gamma}\}.$$

By (4), $\tilde{M}^{(1)}_{2,w_{x_1}}$ is bounded operator on every L^p_w space, $p > 1$, cf. Section 2. In particular

$$w(\bigcup B_\alpha) \leqslant c w(\bigcup B_j).$$

Moreover, if $E_j = B_j \setminus \bigcup_{i<j} B_i$, then

$$w(x_1, (B_j)_{x_1}) \approx w(x_1, (E_j)_{x_1})$$

for a.e. x_1. Consequently,

$$\begin{aligned} w(\{\mathcal{N}_1 f > \lambda\}) &\leqslant c \sum_j w(B_j) \\ &\leqslant \frac{c}{\lambda} \sum_j \int_{(B_j)_{x_1}} \inf_{y_1 \in (B_j)_{x_2}, y_2 \in (B_j)_{x_3}} |f(x_1, y_2, y_3)|\, w(x_1, (E_j)_{x_1})\, dx_1 \\ &\leqslant \frac{c}{\lambda} \sum_j \int \Big(\iint_{(E_j)_{x_1}} |f(x_1, x_2, x_3)|\, w(x_1, x_2, x_3)\, dx_2\, dx_3 \Big) dx_1 \\ &\leqslant \frac{c}{\lambda} \int |f|\, w, \end{aligned}$$

which is what we wanted to show. Clearly a similar result holds for $\mathcal{N}_2$.

In the spirit of Zygmund's problem we can show

THEOREM 4.2. *Let $\mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \times \varphi(\mathcal{B}_1, \mathcal{B}_2)$ and suppose that w satisfies properties* (1)–(5). *Then*

$$w(\{M_{\mathcal{B}} f > \lambda\}) \leqslant c \int \frac{|f(x)|}{\lambda} \Big(1 + \log^+ \frac{|f(x)|}{\lambda}\Big) w(x)\, dx.$$

Proof. We have just proved that $\mathcal{N}_1$ and $\mathcal{N}_2$ are operators of weak type (1, 1) and by (5) we readily see that

$$\mathcal{N}_3 f(x) \leqslant c \mathcal{N}_1 f(x) \mathcal{N}_2 f(x).$$

Thus $w(\{\mathcal{N}_3 f > \lambda\}) \leqslant c(\lambda) \int |f|\, w\, dx$.

It suffices to show that given any finite family $\{B_\alpha\}$ in $\mathcal{B}$ it is possible to extract a subfamily $\{B_j\}$ which verifies

(8) $$\int \exp\big(\sum \chi_{B_j}(x)\big) w(x)\, dx \leqslant c \sum w(B_j),$$

and

(9) $$w(\bigcup B_\alpha) \leqslant c \sum w(B_j).$$

Once (8) and (9) are established the conclusion follows readily by means of Young's inequality and the usual choice of $\{B_\alpha\}$ with the property that $\lambda < \frac{1}{w(B_\alpha)} \int_{B_\alpha} |f|\, w dx$ for all α.

By renaming the B_α's if necessary we may assume that the Lebesgue measure of the n_3-block is ordered in a non-decreasing fashion. For $j = 0$ let

B_0 be the first B_α and once $B_0, \ldots, B_{j-1}$ have been chosen let B_j be the first among the remaining B_α's such that

$$\int_{B_j} \exp\Big(\sum_{k<j} \chi_{B_k}(x)\Big) w(x)\,dx \leqslant 100\, w(B_j).$$

Once $\{B_\alpha\}$ is exhausted we are left with our family $\{B_j\}$. That (8) holds can be readily checked. Indeed, suppose $0 \leqslant j \leqslant N$, then

$$\begin{aligned}
\mathscr{I} &= \int_{\bigcup_{j=0}^{N} B_j} \exp\Big(\sum_{j=0}^{N} \chi_{B_j}(x)\Big) w(x)\,dx \\
&= e \int_{B_N} \exp\Big(\sum_{j=0}^{N-1} \chi_{B_j}(x)\Big) w(x)\,dx + \int_{(\bigcup_{j=0}^{N-1} B_j)\setminus B_N} \exp\Big(\sum_{j=0}^{N-1} \chi_{B_j}(x)\Big) w(x)\,dx \\
&\leqslant 100 e\, w(B_N) + \int_{\bigcup_{j=0}^{N-1} B_j} \exp\Big(\sum_{j=0}^{N-1} \chi_{B_j}(x)\Big) w(x)\,dx.
\end{aligned}$$

By repeated use of this argument we get

$$\mathscr{I} \leqslant 100 e \sum_{j=0}^{N} w(B_j),$$

which is precisely (8). As for (9) we may assume that B_α is not contained in $\bigcup_{j=0}^{N} B_j$. Let j_0 be the (smallest) index such that

$$\int_{B_\alpha} \exp\Big(\sum_{j<j_0} \chi_{B_j}(x)\Big) w(x)\,dx > 100 w(B_\alpha).$$

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_1-side containing the x_1-side of B_α; the remaining B_j's form the second family. By the monotonicity properties of φ the x_2-side of the B_j's in the second family contains the x_2-side of B_α. Let $\mathrm{I} = \{j\colon B_j \text{ is in the first family}\}$, similarly $\mathrm{II} = \{j\colon B_j \text{ is in the second family}\}$. We write

$$\begin{aligned}
\int_{B_\alpha} \exp\Big(\sum_{j<j_0} \chi_{B_j}(x)\Big) w(x)\,dx &= \int_{B_\alpha \cap (\bigcup_{j\in \mathrm{I}} B_j \setminus \bigcup_{j\in \mathrm{II}} B_j)} \exp\Big(\sum_{j\in \mathrm{I}} \chi_{B_j}(x)\Big) w(x)\,dx + \\
&\quad + \int_{B_\alpha \cap (\bigcup_{j\in \mathrm{II}} B_j \setminus \bigcup_{j\in \mathrm{I}} B_j)} \exp\Big(\sum_{j\in \mathrm{II}} \chi_{B_j}(x)\Big) w(x)\,dx + \\
&\quad + \int_{B_\alpha \cap (\bigcup_{j\in \mathrm{I}} B_j) \cap (\bigcup_{j\in \mathrm{II}} B_j)} \exp\Big(\sum_{j<j_0} \chi_{B_j}(x)\Big) w(x)\,dx \\
&= \mathscr{I}_1 + \mathscr{I}_2 + \mathscr{I}_3, \quad \text{say.}
\end{aligned}$$

We consider each $\mathscr{I}_k$, $k = 1, 2, 3$, separately. As far as $\mathscr{I}_1$ is concerned, on B_α the expression $\sum_{j\in \mathrm{I}} B_j$ is independent of x_1 and x_3 since the sides in those directions of each B_j, $j \in \mathrm{I}$, contain the corresponding sides of B_α (the projections actually have this property). Thus if $B_\alpha = B_{\alpha,1} \times B_{\alpha,2} \times B_{\alpha,3}$ we have

$$\begin{aligned}
\mathscr{I}_1 &= \int_{B_{\alpha,2}} \inf_{x_1 \in B_{\alpha,1},\, x_3 \in B_{\alpha,3}} \exp\Big(\sum_{j\in \mathrm{I}} \chi_{B_j}(x_1, x_2, x_3)\Big) w(x_2, B_{\alpha,1}, B_{\alpha,3})\,dx_2 \\
&\leqslant w(B_\alpha) \inf_{y \in B_\alpha} \mathscr{N}_2\Big(\exp\Big(\sum_{j\in \mathrm{I}} \chi_{B_j}\Big), y\Big).
\end{aligned}$$

Similarly

$$\mathscr{I}_2 \leqslant w(B_\alpha) \inf_{y \in B_\alpha} \mathscr{N}_1\Big(\exp\Big(\sum_{j\in \mathrm{II}} \chi_{B_j}\Big), y\Big).$$

As for the last term, we have

$$\begin{aligned}
\mathscr{I}_3 &= \int_{B_\alpha \cap (\bigcup_{j\in \mathrm{I}} B_j) \cap (\bigcup_{j\in \mathrm{II}} B_j)} \exp\Big(\sum_{j\in \mathrm{I}} \chi_{B_j}\Big) \exp\Big(\sum_{j\in \mathrm{II}} \chi_{B_j}\Big) w(x)\,dx \\
&\leqslant w(B_\alpha) \inf_{y \in B_\alpha} \mathscr{N}_3\Big(\exp\Big(\sum_j \chi_{B_j}\Big), y\Big).
\end{aligned}$$

Consequently

$$\bigcup B_\alpha \subseteq \bigcup_{k=1}^{3} \Big\{y\colon \mathscr{N}_k\Big(\exp\Big(\sum_j \chi_{B_j}\Big), y\Big) > 1\Big\}$$

and

$$\begin{aligned}
w\Big(\bigcup B_\alpha\Big) &\leqslant \sum_{k=1}^{3} w\Big(\Big\{y\colon \mathscr{N}_k\Big(\exp\Big(\sum_j \chi_{B_j}\Big), y\Big) > 1\Big\}\Big) \\
&\leqslant c \int \exp\Big(\sum_j \chi_{B_j}(x)\Big) w(x)\,dx \leqslant c\, 100 e \sum_j w(B_j).
\end{aligned}$$

This is precisely what we wanted to show.

In the context of weighed measures the $L^p_w(R^n)$ 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 $0 < \gamma$, $\bar{\gamma} < 1$ such that for $B = B_1 \times B_2 \times \varphi(B_1, B_2)$,

(10) If E is a measurable subset of B_2, then

$$\operatorname*{ess\,sup}_{x_1 \in B_1,\, x_3 \in \varphi(B_1, B_2)} w(x_1, E, x_3)/w(x_1, B_2, x_3) > \gamma$$

for a.e. x_2, implies

$$\operatorname*{ess\,inf}_{x_1\in B_1, x_3\in\varphi(B_1,B_2)} w(x_1, E, x_3)/w(x_1, B_2, x_3) > \bar{\gamma}$$

for a.e. x_2, and

(11) If E is a measurable subset of B_1, then

$$\operatorname*{ess\,sup}_{x_3\in\varphi(B_1,B_2)} w(E, x_2, x_3)/w(B_1, x_2, x_3) > \gamma \quad \text{for a.e. } x_2,$$

implies

$$\operatorname*{ess\,inf}_{x_3\in\varphi(B_1,B_2)} w(E, x_2, x_3)/w(B_1, x_2, x_3) > \bar{\gamma} \quad \text{for a.e. } x_2.$$

In this case we have

THEOREM 4.3. *Assume that $\mathscr{B}$ is as before, that $1 < p < \infty$ and that w verifies properties* (10) *and* (11). *Then*

$$\|M_{\mathscr{B},w} f\|_{L_w^p} \leqslant c\|f\|_{L_w^p}.$$

Proof. We only sketch a proof here. As usual

$$w(\{M_{\mathscr{B},w} f > \lambda\}) \approx w(\bigcup B_\alpha),$$

where $\{B_\alpha\}$ is a family in $\mathscr{B}$ such that

$$w(B_\alpha)^{-1} \int_{B_\alpha} |f|\, w dy > \lambda.$$

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

$$\tilde{M}_v^{1,3} f(x_1, x_3) = \sup \frac{1}{v(B)} \int_B |f(y_1, y_3)|\, v(y_1, y_3)\, dy_1\, dy_3,$$

$$(x_1, x_3)\in B,\ B = B_1\times\varphi(B_1, B_2),\ B_1\in\mathscr{B}_1,\ B_2\in\mathscr{B}_2,$$

we can find a subfamily $\{B_j\}\subseteq\{B_\alpha\}$ so that

$$\bigcup B_\alpha \subseteq \{M_{\mathscr{B}_2, w_{x_1,x_3}} \chi_{\{\tilde{M}_{w_{x_2}}^{1,3}\chi_{\bigcup B_j}\geqslant 1/2\}} > \bar{\gamma}\}. \tag{12}$$

In addition $\{B_j\}$ satisfies a sparseness property.

As a consequence of (11) and Theorem 2.2 (that is to say, of a similar argument involving $\tilde{M}_{w_{x_2}}^{1,3}$ instead of the strong maximal function; observe that we don't need doubling since $\mathscr{B}_1$ satisfies property (D)), $\tilde{M}_{w_{x_2}}^{1,3}$ is bounded on L_w^p, $p > 1$. Also $M_{\mathscr{B}_2, w_{x_1,x_3}}$ is of weak type (1,1). By (12) we have that

$$w(\bigcup B_\alpha) \leqslant cw(\bigcup B_j).$$

As in the proof of Theorem 3.1 it follows that the sparseness of the $\{B_j\}$ implies that

$$w(\bigcup B_\alpha) \leqslant \frac{c}{\lambda} \int_{\{\tilde{M}_{w_{x_2}}^{1,3} M_{\mathscr{B}_2, w_{x_1,x_3}} f > \lambda\}} \tilde{M}_{w_{x_2}}^{1,3} M_{\mathscr{B}_2, w_{x_1,x_3}} f(y_1, y_2, y_3)\, w(y_1, y_2, y_3)\, dy_1\, dy_2\, dy_3.$$

Consequently a similar inequality holds with $w(\{M_{\mathscr{B},w} f > \lambda\})$ in the left-hand side. Multiplying both sides of this new inequality by $p\lambda^{p-1}$ and integrating yields

$$\|M_{\mathscr{B},w} f\|_{L_w^p}^p \leqslant c\|\tilde{M}_{w_{x_2}}^{1,3} M_{\mathscr{B}_2, w_{x_1,x_3}} f\|_{L_w^p}^p.$$

The boundedness in L_w^p of $\tilde{M}_{w_{x_2}}^{1,3}$ and $M_{\mathscr{B}_2, w_{x_1,x_3}}$ 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 $\mathscr{B} = \mathscr{B}_1\times\ldots\times\mathscr{B}_n\times\varphi(\mathscr{B}_1, \ldots, \mathscr{B}_n)$, this generalization is immediate for Theorem 4.3. We point out here one such instance. Suppose that $\varphi(s_1, \ldots, s_n)$, $s_i\geqslant 0$, $i = 1, \ldots, n$, is a continuous function, monotone in each variable separately and satisfying the conditions $\varphi(s_1, \ldots, s_{i-1}, 0, s_{i+1}, \ldots, s_n) = 0$, $i = 1, \ldots, n$ and $\varphi(s_1, \ldots, s_n)\approx\varphi(2s_1, \ldots, 2s_n)$. Consider the basis $\mathscr{B}$ of $s_1\times\ldots\times s_n\times\varphi(s_1, \ldots, s_n)$ rectangles and a weight w which satisfies

(13) There are constants $0 < \gamma$, $\bar{\gamma} < 1$ such that for each $B\in\mathscr{B}$ and measurable $E\subset B$, $|E|/|B| > \gamma$ implies $w(E)/w(B) > \bar{\gamma}$.

We then have that $M_{\mathscr{B},w}$ is bounded in L_w^p, $p > 1$. A few comments about the proof of this result. First, by letting $s_i\to 0$ we find that

$$\operatorname*{ess\,sup}_{\substack{x_i\in B_i\\ x_{n+1}\in\varphi(B_1,\ldots,B_n)}} w_{x_i,x_{n+1}}(E)/w_{x_i,x_{n+1}}(B) > \gamma$$

implies

$$\operatorname*{ess\,inf}_{\substack{x_i\in B_i\\ x_{n+1}\in\varphi(B_1,\ldots,B_n)}} w_{x_i,x_{n+1}}(E)/w_{x_i,x_{n+1}}(B) > \bar{\gamma},$$

where $B = B_1\times\ldots\times B_n\times\varphi(B_1, \ldots, B_n)$.

The proof of Theorem 4.3 carries over almost unchanged in this setting. Nevertheless, since $\mathscr{B}$ 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

$$w(B^*) \leqslant cw(B),$$

where B^* is the rectangle with same center as B but with x_{n+1}-side length twice as large and with other sidelengths $\leqslant c_1 \cdot$ the corresponding sidelengths of B. Here c_1 depends on the constant appearing in the relation $\varphi(s_1, \ldots, s_n) \approx \varphi(2s_1, \ldots, 2s_n)$. It then follows that

$$w(\bigcup B_\alpha) \leqslant cw(\bigcup B_j^*) \leqslant c\sum w(B_j).$$

This inequality replaces $w(\bigcup B_\alpha) \leqslant cw(\bigcup B_j)$ 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

$$(M_{k,w} f)^*(t) \leqslant c\Big(\sum_{j=1}^{N} M^{(j)}_{k-1,w_{x_j}} f\Big)^{**}(t).$$

Similarly Theorem 3.1 gives

$$(M_{\mathscr{B},w} f)^*(t) \leqslant c\big((M_{\mathscr{B}_1,w_{x_2}} M_{\mathscr{B}_2,w_{x_1}} + M_{\mathscr{B}_2,w_{x_1}} M_{\mathscr{B}_1,w_{x_2}}) f\big)^{**}(t).$$

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 [5] 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

[1] R. J. Bagby, *Maximal functions and rearrangements: some new proofs*, Indiana Univ. Math. J. 32 (1983), 879–891.

[2] S. Y. Chang and R. Fefferman, *Some weight norm inequalities for Córdoba's maximal function* (preprint).

[3] A. Córdoba, *Maximal functions, covering lemmas and Fourier multipliers*, Proc. Symp. Pure Math. 35 (1979), 29–50.

[4] –, *Cinco teoremas de Análisis Armónico Clásico*, vol. 15, Memoirs of the Royal Academy of Sciences, Madrid, Spain, 1982.

[5] – and R. Fefferman, *A geometric proof of the strong maximal theorem*, Ann. of Math. 102 (1975), 95–100.

[6] N. Fava, *Weak type inequalities for product operators*, Studia Math. 42 (1972), 271–278.

[7] R. Fefferman, *Covering lemmas, maximal functions and multiplier operators in Fourier analysis*, Proc. Symp. Pure Math. 35 (1979), 51–60.

[8] –, *Strong differentiation with respect to measures*, Amer. J. Math. 103 (1981), 33–40.

[9] M. de Guzmán, *Real variable methods in Fourier Analysis*, Mathematical Studies 46, North-Holland, 1981.

[10] M. de Guzmán, *An inequality for the Hardy–Littlewood maximal operator with respect to a product of differentiation bases*, Studia Math. 49 (1972), 185–194.

[11] B. Jawerth, *A note on local maximal operators* (preprint).

[12] B. Jessen, J. Marcinkiewicz and A. Zygmund, *Note on the differentiability of multiple integrals*, Fund. Math. 25 (1935), 217–234.

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
Bloomington, Indiana 47 405, U.S.A.

Received November 2, 1983 (1939)