Summation Conditions on Weights

Stephen M. Buckley

1. Introduction and Notation

In this paper, we introduce a class of summation conditions on weights which are equivalent to the dyadic weight conditions A^d_∞, A^d_p, and B^d_p, and provide a useful alternative way of thinking of these weight conditions. We then use this equivalence result to find a new proof of the boundedness of the dyadic square function on $L^p(w)$ for any A^d_p weight w. (Usually one shows, as in [4], that singular integrals, square functions, and related operators are bounded on weighted $L^p(w)$ spaces by using a good-λ inequality, but we avoid such methods entirely.)

Our first task (Section 2) is to state and prove the main equivalence theorem. The summation conditions we introduce here are related to the conditions introduced by R. Fefferman, Kenig, and Pipher in [6], but the methods employed are completely different. In Section 3, we utilize the results and ideas of Section 2 to prove the boundedness of the dyadic square function on weighted $L^p(w)$ spaces.

Harmonic analysis on "product spaces" has been the subject of much scrutiny in recent years (an overview of this field can be found in [3]), and so we finish, in Section 4, by defining analogs of our summation conditions on product spaces and by showing that they are related to the product A^d_p and B^d_p conditions.

Throughout this paper, we will use "C" to indicate a constant that depends only on p and the dimension n. $\mathfrak{D} = \mathfrak{D}(\mathbf{R}^n)$ indicates the set of all dyadic cubes in $\mathbf{R}^n$. For any $Q \in \mathfrak{D}$, $\mathfrak{D}(Q)$ is the collection of proper dyadic subcubes of Q, and $\tilde{Q}$ is the dyadic double of Q (the smallest dyadic cube properly containing Q). For any weight w and set S, $w(S)$ denotes the integral of w over S, $|S|$ denotes the Lebesgue measure of S, and $w_S = w(S)/|S|$. Unless otherwise specified, $1 < p < \infty$, but p is otherwise arbitrary.

2. A^d_p, B^d_p, and Summation Conditions

In this section, we shall examine conditions on a weight w involving the sum

Received April 11, 1991.

The material in this work was drawn from the author's Ph.D. thesis, which was done under the supervision of Robert Fefferman at the University of Chicago.

Michigan Math. J. 40 (1993).

$$S_r(Q_0, w) = \sum_{Q \in \mathfrak{D}(Q_0)} w_{\tilde{Q}}^r \left(\frac{\Delta_Q w}{w_{\tilde{Q}}}\right)^2 |Q|, \tag{1}$$

where $\Delta_Q w = w_Q - w_{\tilde{Q}}$. Before we state the main theorem, let us make the following definitions.

DEFINITION. We say w is a *dyadic doubling weight* (written $w \in \mathbf{Db}^d$) if $w(\tilde{Q}) \leq Cw(Q)$ for all dyadic cubes Q, where $\tilde{Q}$ is the dyadic double of Q (the smallest dyadic cube properly containing Q).

DEFINITION. We say w is an A_p^d *weight* (written $w \in A_p^d$) if

$$\left(\frac{1}{|Q|}\int_Q w\right)\left(\frac{1}{|Q|}\int_Q w^{-1/(p-1)}\right)^{p-1} \leq K \quad \text{for all } Q \in \mathfrak{D}.$$

The smallest such K is referred to as the A_p^d *norm* of w and will be denoted $K_{w,p}$ or simply K_w. We say w is an A_∞^d *weight* if there exists $K', \epsilon > 0$ such that, for all $Q \in \mathfrak{D}$ and all $E \subset Q$, we have

$$\frac{w(E)}{w(Q)} \leq K'\left(\frac{|E|}{|Q|}\right)^\epsilon.$$

DEFINITION. We say w is a *weak-B_p^d weight* (written $w \in B_p^{d,\mathrm{wk}}$) if

$$\left(\frac{1}{|Q|}\int_Q w^p\right)^{1/p} \leq K\frac{1}{|Q|}\int_Q w \quad \text{for all } Q \in \mathfrak{D}.$$

If, in addition, $w \in \mathbf{Db}^d$, we say that w is a B_p^d weight. The smallest such K is referred to as the B_p^d norm of w and will be denoted $K_{w,p}^*$ or simply K_w^*.

REMARK 2.1. Note that, unlike the nondyadic case, the above reverse Hölder inequality does not automatically imply doubling. We require B_p^d weights to be dyadic doubling in our definition, since this is necessary for the theory of such weights to closely mirror the nondyadic case (for example, if $w \in B_p^d$ then $w \in A_\infty^d$).

THEOREM 2.2. *Suppose w is a weight. Then*

(i) $w \in B_p^d \Leftrightarrow w \in \mathbf{Db}^d$ *and* $S_r(Q, w) \leq Kw_Q^r|Q|$ $\forall Q \in \mathfrak{D}(\mathbf{R}^n)$, *where* $r = p$.

(ii) $w \in A_p^d \Leftrightarrow w \in \mathbf{Db}^d$ *and* $S_r(Q, w) \leq Kw_Q^r|Q|$ $\forall Q \in \mathfrak{D}(\mathbf{R}^n)$, *where* $r = -1/(p-1)$.

(iii) $w \in A_\infty^d \Leftrightarrow w \in \mathbf{Db}^d$ *and* $S_r(Q, w) \leq Kw_Q^r|Q|$ $\forall Q \in \mathfrak{D}(\mathbf{R}^n)$, *where* $r = 0$ *or* $r = 1$ (*each is separately equivalent to* $w \in A_\infty^d$).

(iv) $S_r(Q, w) \leq Kw_Q^r|Q|$ $\forall Q \in \mathfrak{D}(\mathbf{R}^n)$, *for any* $0 < r < 1$.

The constant K in (i) *and* (ii) *is equivalent to the* $|r|$th *power of the B_p^d or A_p^d norm of w (up to a constant dependent on r). In fact,*

$$\frac{1}{r(r-1)}c_n^{|r|+1}C_w^{|r|} \leq K \leq \frac{1}{r(r-1)}C_n^{|r|+1}C_w^{|r|},$$

where c_n and C_n are dimensional constants, and $C_w = K_w^*$ for (i) *and $C_w = K_w$ for* (ii). *In* (iv), *$K \leq C_n/r(r-1)$, C_n being a dimensional constant.*

REMARK 2.3. We can actually prove the following, which clearly implies (i):

(i′) $w \in B_p^{d,\mathrm{wk}} \Leftrightarrow S_p(Q,w) \le K w_Q^p |Q| \ \forall Q \in \mathfrak{D}(\mathbf{R}^n)$.

REMARK 2.4. Part (iii) (for the case $r=0$) and nondyadic versions of (i), (ii), and (iii) (for the case $r=0$) were found by R. Fefferman, Kenig and Pipher in [6], using different methods. It is natural to have two different summation conditions equivalent to A_∞^d, since A_∞^d is a limiting version of both A_q^d ($q \to \infty$) and B_q^d ($q \to 1$). For some purposes, the condition involving $S_1(Q,w)$ (the "S_1 condition", for short) has advantages over the S_0 condition. For example, Muckenhoupt's C_p condtion [13] is more similar to a limiting B_q condition than a limiting A_q condition; in [1], we were able to get a condition involving $S_1(Q,w)$ equivalent to dyadic C_p (since C_p weights are not necessarily doubling, it was also important that we did not need $w \in \mathbf{Db}^d$ to manipulate $S_1(Q,w)$).

Before proving Theorem 2.2, we will state and prove a couple of lemmas which are needed, but we first need to introduce some notation. We define functions ϕ_r for every real number r as follows:

$$\phi_r(x) = \frac{x^r}{r(r-1)} \quad \text{for } r \ne 0,1;$$

$$\phi_1(x) = x\log(2+x);$$

$$\phi_0(x) = \begin{cases} -\log(x) & \text{for } 0 \le x \le 1, \\ -3/2 + 2/x - 1/2x^2, & \text{for } 1 < x. \end{cases}$$

These functions are defined for all $x \ge 0$ ($\phi_r(0) = \infty$ if $r \le 0$). A little calculation shows that each of these functions is convex (i.e., $\phi_r'' > 0$ for every r). The definition of ϕ_0 seems rather strange, but it is chosen so as to be a C^2 function which is bounded below. Note also that $\phi_0(x) \ge -\log(x)$ for all $x > 0$.

LEMMA 2.5. *Suppose $a_i \ge 0$ for $1 \le i \le N$ and $\bar{a} = (\sum_{i=1}^N a_i)/N > 0$. Then:*

(i) *For all $r \ne 0,1$,*

$$\sum_{i=1}^N (a_i - \bar{a})^2 \bar{a}^{r-2} \le C_N^{|r|+1} \sum_{i=1}^N (\phi_r(a_i) - \phi_r(\bar{a})).$$

(ii) *If, in addition, $\epsilon < a_i/a_j$ for any pair a_i, a_j, then for all $r \in \mathbf{R}$,*

$$\sum_{i=1}^N (a_i - \bar{a})^2 \bar{a}^{r-2} \ge C_{N,\epsilon}^{|r|+1} \sum_{i=1}^N (\phi_r(a_i) - \phi_r(\bar{a})).$$

The condition "$\epsilon < a_i/a_j$" is needed only for $r < 1$.

The constants C_N and $C_{N,\epsilon}$ depend only on their subscripted variables.

REMARK 2.6. Actually, it is easy to show that, for $r > 0$, $r \ne 1$, the condition on a_i/a_j in (ii) can be eliminated, as long as we do not require the dependence on r to be quite as good near 0 and 1. To see this, note that we can assume, by normalization, that $a_1 = 1 \geq a_2 \geq \cdots \geq a_N$. If $a_N < 1/2N$, then

$$\sum_{i=1}^{N}(a_i-\bar{a})^2\bar{a}^{r-2} \geq \left(\frac{1}{2N}\right)^2\left(\frac{1}{N}\right)^{|r-2|} \geq |r(r-1)|C_N^{|r|+1}\sum_{i=1}^{N}(\phi_r(a_i)-\phi_r(\bar{a})).$$

We postpone the proof of this lemma until after the proof of Theorem 2.2. However, we will now prove it in the case $N=2$ by a more intuitive method than used to prove the general case, which will show why the lemma "should" be true. To this end, let us restate the lemma in the case $N=2$, using notation more suitable to that case. For any function f, we define $D_f(a,b) = f(a)+f(b)-2f((a+b)/2)$. For ease of notation, if r is a real number we also define D_r to be D_{ϕ_r}, where ϕ_r is as defined above.

LEMMA 2.7. *Suppose $a, b > 0$. Then:*

(i) *For all $r \neq 0, 1$,*

$$(a-b)^2(a+b)^{r-2} \leq C^{|r|+1}D_r(a,b).$$

(ii) *If $\epsilon < b/a < 1/\epsilon$ then, for all $r \in \mathbf{R}$,*

$$(a-b)^2(a+b)^{r-2} \geq C_\epsilon^{|r|+1}D_r(a,b).$$

The size condition on b/a is needed only for $r<1$.

The proof of Lemma 2.7 involves the following elementary lemma, which follows immediately from the fundamental theorem of calculus.

LEMMA 2.8. *If $\phi \in C^2(0,\infty)$, and if $0<b<a$, then*

$$D_\phi(a,b) = \int_b^{(a+b)/2}\int_0^{(a-b)/2}\phi''(x+y)\,dy\,dx.$$

Proof of Lemma 2.7. Without loss of generality, we may assume that $b<a$. We first prove (i). By the previous lemma,

$$D_r(a,b) = \int_R (x+y)^{r-2}\,dx\,dy \quad \text{for any } r \neq 0, 1,$$

where $R = [b, (a+b)/2] \times [0, (a-b)/2]$. But now, if

$$(x,y) \in R' \equiv \left[b, \frac{a+b}{2}\right] \times \left[\frac{a-b}{4}, \frac{a-b}{2}\right],$$

then $x+y \sim a+b$ and so

$$D_\phi(a,b) \geq \int_{R'}\phi''(x+y)\,dx\,dy \sim \int_{R'}(a+b)^{r-2}\,dx\,dy \sim (a-b)^2(a+b)^{r-2}.$$

As for (ii), a little calculation shows that our size assumption on a/b implies $\phi''(x+y) \leq C(a+b)^{r-2}$ for all $(x,y) \in R$, and so

$$D_\phi(a,b) = \int_R \phi''(x+y)\,dx\,dy \leq \int_R C(a+b)^{r-2}\,dx\,dy \leq C(a-b)^2(a+b)^{r-2}.$$

It is also clear that the size condition is unnecessary when $r>2$. If $0<r<2$, then

$$\int_0^a \int_0^a |x+y|^{r-2}\, dx\, dy \le \frac{C}{r} a^r,$$

from which it follows that the size condition is unnecessary for $1 \le r < 2$ (or even for $0<\delta<r<2$). □

The second lemma used in the proof of Theorem 2.2 follows easily from Fatou's lemma but, before we state it, we need to introduce some additional notation. The set of all dyadic cubes of side-length 2^m will be denoted as $\mathfrak{D}_m(\mathbf{R}^n)$. If Q is a dyadic cube of side-length at least 2^m, then $\mathfrak{D}_m(Q)$ denotes all dyadic subcubes of Q of side-length 2^m. If $f\colon \mathbf{R}^n \to \mathbf{R}$ is any $L^{1,\infty}$ function, we define $f_m(x)$ to be the average value of f on $Q_{m,x}$, where $Q_{m,x}$ is the cube for which $x \in Q_{m,x} \in \mathfrak{D}_m(\mathbf{R}^n)$.

Lemma 2.9. *Suppose* $|S|<\infty$, $r>0$, *and* $\phi\colon [0,\infty) \to (-r,\infty]$ *is continuous. Then, for any weight* w,

$$\int_S \phi \circ w \le \liminf_{m \to -\infty} \int_S \phi \circ w_m.$$

With these two lemmas in hand, we are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. Let us fix $Q_0 \in \mathfrak{D}$, with side-length 2^{m_0}, say. Suppose $w \in B_p^d$. Then, using Lemma 2.5, we get

$$\begin{aligned} S_p(Q_0,w) &\le K \sum_{Q \in \mathfrak{D}(Q_0)} (w_{\tilde{Q}}^p - w_Q^p)|Q| \\ &= K \sum_{m=1}^{\infty} (a_m - a_{m-1}) = K\big(\lim_{m\to\infty} a_m - w_{Q_0}^p |Q_0|\big), \end{aligned}$$

where

$$a_m = \sum_{Q \in \mathfrak{D}_{m_0-m}(Q_0)} w_Q^p |Q| = \int_{Q_0} w_{m_0-m}^p.$$

Letting $\sigma = w^p$ and using Hölder's inequality, we get $w_Q^p \le \sigma_Q$ and so $a_m \le \sigma(Q_0)$. It follows that $S_p(Q_0,w) \le K(\sigma(Q_0) - w_{Q_0}^p |Q_0|)$.

Now, if $w \in B_p^d$ with B_p^d norm K_w^*, then $\sigma_{Q_0} \le K_w^{*p} w_{Q_0}^p$ and so $S_p(Q,w) \le K w_{Q_0}^p$, where K is a constant of the required form. This proves half of (i) (in fact, half of (i′)). The proof of the corresponding half of (ii) is very similar, so we omit it.

To prove (iv), let $0<r<1$ and use Lemma 2.5 to get

$$\begin{aligned} S_r(Q_0,w) &\le K \sum_{Q \in \mathfrak{D}(Q_0)} (w_{\tilde{Q}}^r - w_Q^r)|Q| \\ &= K \sum_{m=1}^{\infty} (a_{m-1} - a_m), \quad \text{where } a_m = \int_{Q_0} w_{m_0-m}^r \\ &\le K a_0 = C w_{Q_0}^r |Q_0|. \end{aligned}$$

To prove half of (iii), it is convenient to first define a measure μ on $\mathfrak{D}(Q_0)$, defined by

$$\mu(\{Q\}) = \left(\frac{\Delta_Q w}{w_{\tilde{Q}}}\right)^2 |Q|.$$

Using this notation, $S_r(Q_0, w) = \int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}}^r \, d\mu$.

Now, if $w \in A_\infty^d$, then $w \in A_p^d$ for large enough $p < \infty$ and so, by Hölder's inequality, part (iv), and the proven half of (ii), we have

$$\begin{aligned} S_0(Q_0, w) &= \int_{\mathfrak{D}(Q_0)} 1 \, d\mu \\ &\leq \left(\int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}}^{\epsilon} \, d\mu\right)^{1/2} \left(\int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}}^{-\epsilon} \, d\mu\right)^{1/2} \\ &\leq K(w_{Q_0}^{\epsilon}|Q_0|)^{1/2}(w_{Q_0}^{-\epsilon}|Q_0|)^{1/2} \\ &= K|Q_0|, \end{aligned}$$

as long as $0 < \epsilon < \min(1, 1/(p-1))$.

Also, if $w \in A_\infty^d$ then $w \in B_p^d$ for some $p > 1$. Thus, by Hölder's inequality, part (iv), and the proven half of (i),

$$\begin{aligned} S_1(Q_0, w) &= \int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}} \, d\mu \\ &\leq \left(\int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}}^{1+\epsilon} \, d\mu\right)^{1/2} \left(\int_{\mathfrak{D}(Q_0)} w_{\tilde{Q}}^{1-\epsilon} \, d\mu\right)^{1/2} \\ &\leq K(w_{Q_0}^{1+\epsilon}|Q_0|)^{1/2}(w_{Q_0}^{1-\epsilon}|Q_0|)^{1/2} \\ &= K w_{Q_0}|Q_0|, \end{aligned}$$

as long as $0 < \epsilon < \min(1, p-1)$. This finishes the proof of half of (iii).

We are left with proving the other halves of (i)–(iii). Suppose that

$$S_p(Q_0, w) \leq C w_{Q_0}^p(Q_0, w)$$

for some $p > 1$. Then, by Lemma 2.5,

$$\begin{aligned} w_{Q_0}^p|Q_0| \geq C S_p(Q_0, w) &\geq \sum_{Q \in \mathfrak{D}(Q_0)} (w_Q^p - w_{\tilde{Q}}^p)|Q| \\ &= C\left(\left(\lim_{n \to -\infty} \int_{Q_0} w_n^p\right) - w_{Q_0}^p|Q_0|\right). \end{aligned}$$

Notice that, by Jensen's lemma and Lemma 2.9, the limit actually exists and equals $\sigma(Q_0)$, so we deduce that $\int_{Q_0} \sigma \leq K w_{Q_0}^p |Q_0|$. This proves the other half of (i′). The proofs of the other halves of (ii) and (iii) (in the case $r = 1$) are so similar that we omit them. (Note that, because of the "ϵ" condition in Lemma 2.5, we need to assume $w \in \mathbf{Db}^d$ for (ii).)

In the case $r = 0$, we can assume $|Q_0| = w(Q_0) = 1$ by normalization (and so $\phi_0(w_{Q_0}) = 0$). Now, by Lemma 2.5 and Lemma 2.9, we get

$$1 \geq C \lim_{m \to -\infty} \int_{Q_0} \phi_0(w_m) \geq C \int_{Q_0} \phi_0(w) \geq -C \int_{Q_0} \log w.$$

It follows that $\exp \int_{Q_0} \log w \geq e^{-1/C}$, which is equivalent to $w \in A_\infty^d$ (see [8, Thm. IV.2.15]). This finishes the proof of the theorem. □

Now that we have proved Theorem 2.2, we go back to Lemma 2.5 and give a proof of it for all N, as promised.

Proof of Lemma 2.5. We assume without loss of generality that $a_1 \geq a_2, \dots, a_N$. We fix $a_1 > 0$ and let $A = \{(a_2, \dots, a_N) : 0 \leq a_2, \dots, a_N \leq a_1\}$. We define f on A by

$$f(a_2, \dots, a_N) = \sum_{j=1}^{N} (a_j - \bar{a})^2 a_1^{r-2} - C \sum_{j=1}^{N} (\phi_r(a_j) - \phi_r(\bar{a})).$$

Proving (i) is equivalent to showing $f \leq 0$, as long as $C = C_N^{|r|+1}$ is large enough (because $a_1 \sim \bar{a}$). To prove this, suppose $a_i < \bar{a}$ for some i. Writing ∂_i for ∂_{a_i}, we get

$$\partial_i f(a_2, \dots, a_N) = -2(\bar{a} - a_i) a_1^{r-2} + C(\phi_r'(\bar{a}) - \phi_r'(a_i)). \tag{2}$$

Now, for $r \neq 0, 1$, $\phi_r''(x) = x^{r-2}$, which is monotonic and so

$$\partial_i f(a_2, \dots, a_N) \geq -2(\bar{a} - a_i) a_1^{r-2} + C(\bar{a} - a_i) \min(\bar{a}^{r-2}, a_i^{r-2}).$$

If $r \leq 2$ or $a_i \geq a_1/2N$, then we can conclude easily that $\partial_i f(a_2, \dots, a_N) \geq 0$ for large enough C. If, on the other hand, $r > 2$ and $a_i < a_1/2N$, then

$$(r-1)(\phi_r'(\bar{a}) - \phi_r'(a_i)) = \bar{a}^{r-1} - a_i^{r-1} \geq \tfrac{1}{2} \bar{a}^{r-1},$$

and so

$$\partial_i f(a_2, \dots, a_N) \geq -2\bar{a} a_1^{r-2} + \frac{C}{2(r-1)} \bar{a}^{r-1} \geq 0$$

for large enough $C = C_N^{|r|+1}$.

In either case, we have shown that $\partial_i f(a_2, \dots, a_N) \geq 0$ for any i for which $a_i < \bar{a}$. Thus, f can only achieve its maximum on its compact domain A when $a_i \geq \bar{a}$ for all i. This clearly implies that $a_i = a_1$ for all i and that $f(a_1, \dots, a_1) = 0$, giving us the required result.

To prove (ii), it suffices to show that $f(a_2, \dots, a_N) \geq 0$, as long as $a_i \geq \epsilon a_1$ and $C = C_N^{|r|+1}$ is small enough, and that the size restriction on a_i is unnecessary if $r \geq 1$. Without loss of generality, we may assume that $\epsilon \leq 1/2N$. Suppose again that $a_i < \bar{a}$ for some i. From (2), we get

$$\partial_i f(a_2, \dots, a_N) \leq -2(\bar{a} - a_i) a_1^{r-2} + C(\bar{a} - a_i) \max(\bar{a}^{r-2}, a_i^{r-2}),$$

because a little calculation shows that $\phi_r''(x) \leq 2x^{r-2}$ for any $r \in \mathbf{R}$. Thus, if $r \geq 2$ or $a_i \geq \epsilon a_1$, then $\partial_i f(a_2, \dots, a_N) \leq 0$ for small enough $C > 0$. This allows us to argue, as in (i), that $f(a_2, \dots, a_N) \geq 0$ when $r \geq 2$ or for any $r \in \mathbf{R}$ if $a_i \geq \epsilon a_1$ for all i.

We are left only with the case where $1 \le r < 2$ and $a_i \le \epsilon a_1$ for some i. We first assume $1 < r < 2$ and normalize so that $a_1 = 1$. In this case,

$$f(a_2, \ldots, a_N) \ge \frac{1}{4N^2} - C \sum_{j=1}^{N} (\phi_r(a_j) - \phi_r(\bar{a})).$$

Thus it suffices to show that, for large enough $C = C_N$, $\phi_r(a_j) - \phi_r(\bar{a}) \le C$ if $a_j > \bar{a}$ (since it is trivial if $a_j \le \bar{a}$). This follows from the fact that $1 - x^r \le C_N(r-1)$ whenever $1 < r < 2$ and $x \ge 1/N$.

In the case $r = 1$, we have

$$\phi_1''(x) = \frac{2}{(2+x)^2} + \frac{1}{2+x} \le \frac{2}{2+x}.$$

If $a_i < \bar{a}$, then it follows from (2) that

$$\partial_i f(a_2, \ldots, a_N) \le \frac{-2(\bar{a} - a_i)}{a_1} + \frac{C(\bar{a} - a_i)}{2 + a_i} \le 0$$

for small enough $C = C_N$, if a_1 is not very large ($a_1 \le 10N$, say), and so we get that $f(a_2, \ldots, a_N) \ge 0$ by the usual argument. If $a_1 > 10N$ (and so $\bar{a} > 10$), and if $a_i < \epsilon a_1$ for some i, then let $\Phi = \sum_{j=1}^{N} (\phi_1(a_j) - \phi_1(\bar{a}))$. We need to show that $\Phi \le C\bar{a}$. To see this, we write

$$\begin{aligned}\Phi &= \bar{a} \sum_{j=1}^{N} (\phi_1(a_j/\bar{a}) - \phi_1(1)) + \sum_{j=1}^{N} (\phi_1(a_j) - \bar{a}\phi_1(a_j/\bar{a}) - \phi_r(\bar{a}) + \bar{a}\phi_1(1)) \\ &\equiv \Phi_1 + \Phi_2.\end{aligned}$$

Now $\Phi_1 \le C\bar{a}$, since it covered by the case $a_1 \le 10N$.

As for Φ_2, we note that

$$\phi_1(a_j) - \bar{a}\phi_1(a_j/\bar{a}) = a_j \log\left(\frac{2 + a_j}{2 + a_j/\bar{a}}\right)$$

and that

$$\phi_1(\bar{a}) - \bar{a}\phi_1(1) = \bar{a} \log\left(\frac{2+\bar{a}}{2+1}\right) = \frac{1}{N} \sum_{j=1}^{N} a_j \log\left(\frac{2+\bar{a}}{3}\right),$$

and so

$$\Phi_2 = \sum_{j=1}^{N} a_j \log\left(\frac{3(2+a_j)}{(2+a_j/\bar{a})(2+\bar{a})}\right) \le C\bar{a}.$$

This completes the proof of the lemma. □

3. The Dyadic Square Function on $L^2(w)$

For any locally integrable function f, let us define the dyadic square function $S_d f$ (subscripting the "d", meaning "dyadic", is inconsistent with our previous notation, but convenient because we often wish to consider $S_d^2 f$, the square of $S_d f$) by the equation

$$S_d^2 f(x) = \sum_{x \in Q \in \mathfrak{D}} (f_Q - f_{\tilde{Q}})^2.$$

One can prove that S_d is bounded on $L^p(w)$ for $w \in A_p^d$ using a good-λ argument, similar to Coifman and C. Fefferman's [4] proof that a singular integral operator is bounded on $L^p(w)$. We shall get this result (Theorem 3.6) for $p=2$ in a completely different manner, using the results of the previous section (Rubio de Francia's extrapolation theorem [14] can then be used to show S_d is bounded on $L^p(w)$ for all $1<p<\infty$, and all $w \in A_p^d$).

Interestingly, the K_w-dependence of the operator norm which we get in Theorem 3.6 is actually the same dependence as we could get by good-λ methods, if we used Chang, Wilson, and Wolff's [2] sharp good-λ estimate for S_d.

LEMMA 3.1. *Let $Q_0 \in \mathfrak{D}$. Let μ be a positive measure on $\mathfrak{D}(Q_0)$ and let ν be a positive measure on $\mathbf{R}^n$. If, for all $Q \in \mathfrak{D}(Q_0)$, $\mu(\mathfrak{D}(Q)) \leq \nu(Q)$, then for all $f \in L^{1,\infty}$,*

$$\int_{\mathfrak{D}(Q_0)} |f_{\tilde{Q}}|^p \, d\mu(Q) \leq C \int_{Q_0} (M^d f)^p \, d\nu, \tag{3}$$

where M^d denotes the dyadic maximal function.

Proof: Without loss of generality we may assume that $0<\nu(Q_0)<\infty$, that $f \geq 0$, and that f is supported on Q_0 (since truncating f in this fashion decreases the right-hand side of (3) but leaves the left-hand side unchanged). Let $\{Q_j^k\}$ be the set of maximal dyadic cubes Q for which $(1/|Q|)\int_Q f \geq 2^k$. Also let $\alpha_Q = \mu(\{Q\})$. Then

$$\begin{aligned}\int_{\mathfrak{D}(Q_0)} f_{\tilde{Q}}^p \, d\mu(Q) &= \sum_{\mathfrak{D}(Q_0)} f_{\tilde{Q}}^p \alpha_Q \\ &\leq 2^p \sum_{k,j} 2^{kp} \sum_{Q \in \mathfrak{D}(Q_j^k)} \alpha_Q \\ &\leq 2^p \sum_{k,j} 2^{kp} \nu(Q_j^k) \leq C \int_{Q_0} (M^d f)^p \, d\nu. \qquad \square\end{aligned}$$

REMARK 3.2. Clearly, we could have replaced $f_{\tilde{Q}}$ by f_Q in the statement of the above lemma because $f_Q \leq 2^n f_{\tilde{Q}}$. However, the lemma is useful in its present slightly stronger form.

LEMMA 3.3. *If $w \in A_p^d$, then $\|M^d f\|_{L^p(w)}^p \leq C K_w^{p'} \|f\|_{L^p(w)}^p$. The power $K_w^{p'}$ is best possible.*

REMARK 3.4. The nondyadic version of this result is basically due to Muckenhoupt [12] in the 1-dimensional case, and Coifman and C. Fefferman [4] in the n-dimensional case, but they did not examine the dependence on K_w. This examination and another proof of this result can be found in [1]. Any of these proofs can be easily modified to handle the dyadic case.

COROLLARY 3.5. *If μ and ν are as above and if, in addition, $d\nu(x) = w(x)\,dx$ for some $w \in A_p^d$, then*

$$\int_{\mathfrak{D}(Q_0)} f_{\tilde{Q}}^p \, d\mu(Q) \leq C K_w^{p'} \int_{Q_0} f^p \, d\nu.$$

Proof. Just combine Lemma 3.1 and the dyadic version of Lemma 3.3. □

THEOREM 3.6. *If* $w \in A_2^d$, *then* $\int (S_d^2 f) w \le C K_w^3 \int f^2 w$.

Proof. First, note that

$$\int (S_d^2 f) w = \sum_{Q \in \mathfrak{D}} w(Q)(f_Q - f_{\tilde{Q}})^2 \le \sum_{Q \in \mathfrak{D}} w(\tilde{Q})(f_Q - f_{\tilde{Q}})^2 \equiv W,$$

so it suffices to control W. An examination of the proof of Lemma 2.5 indicates that, in the case $r=2$, the inequalities in (i) and (ii) are replaced by an equality. Using this equality, we get

$$\begin{aligned} W &= \sum_{Q \in \mathfrak{D}} w(\tilde{Q})(f_Q^2 - f_{\tilde{Q}}^2) \\ &= \sum_{Q \in \mathfrak{D}} (2^n w(Q) f_Q^2 - w(\tilde{Q}) f_{\tilde{Q}}^2) + \sum_{Q \in \mathfrak{D}} (w(\tilde{Q}) - 2^n w(Q)) f_Q^2 \\ &= W_1 + W_2. \end{aligned}$$

Now $W_1 = \sum_{m=-\infty}^{\infty} (a_m - a_{m+1})$, where

$$a_m = \sum_{Q \in \mathfrak{D}_m} 2^n w(Q) f_Q^2 = 2^n \int f_m^2 w.$$

Clearly,

$$a_m \le C \int (M^d f)^2 w \le C K_w^2 \int f^2 w,$$

by Lemma 3.3. Thus $W_1 \le C K_w^2 \int f^2 w$.

Next,

$$\begin{aligned} W_2 &= \sum_{Q \in \mathfrak{D}} (w(\tilde{Q}) - 2^n w(Q))(f_Q^2 - f_{\tilde{Q}}^2) \\ &\le \left(\sum_{Q \in \mathfrak{D}} \frac{(w(\tilde{Q}) - 2^n w(Q))^2}{w(\tilde{Q})} (f_Q + f_{\tilde{Q}})^2 \right)^{1/2} \left(\sum_{Q \in \mathfrak{D}} w(\tilde{Q})(f_Q - f_{\tilde{Q}})^2 \right)^{1/2} \\ &= W_3^{1/2} W^{1/2} \le (W_3 + W)/2. \end{aligned}$$

Thus, $W \le C K_w^2 \int f^2 w + \frac{1}{2} W_3 + \frac{1}{2} W$, so it suffices to show that

$$W_3 \le C K_w^3 \int f^2 w.$$

Since $f_Q \le C f_{\tilde{Q}}$ we have $W_3 \le C \sum_{Q \in \mathfrak{D}} \alpha_Q f_{\tilde{Q}}^2$, where

$$\alpha_Q = w_{\tilde{Q}} \left(\frac{\Delta_Q w}{w_{\tilde{Q}}} \right)^2 |Q|$$

and Δ_Q is as in the definition of $S_r(Q, w)$. Now $w \in B_{1+\delta}^d$, where $\delta \sim K_w^{-1}$ and $K_{w,1+\delta}^* \le 2$ (as is revealed by an examination of the constants in the proof of [4, Thm. IV]). Therefore, by Hölder's inequality and the estimates for the constants in Theorem 2.2,

$$\begin{aligned} \sum_{Q \in \mathfrak{D}(Q_0)} \alpha_Q &= S_1(Q_0, w) \le (S_{1-\delta}(Q_0, w) S_{1+\delta}(Q_0, w))^{1/2} \\ &\le \frac{C w(Q_0)}{\delta} \le C K_w w(Q_0). \end{aligned}$$

Thus, it follows from Corollary 3.5 that

$$\sum_{Q\in\mathfrak{D}} \alpha_Q f_Q^2 \le CK_w \cdot K_w^2 \int f^2 w.$$

4. Summation Conditions on Product Spaces

In this section we look for a generalization of Theorem 2.2 in the setting of so-called "product spaces". By a product space, we mean a space $\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}\times\cdots\times\mathbf{R}^{d_k}$, where we are concerned with classes of operators invariant under the k-parameter family of dilations

$$(4)\qquad (x_1,\ldots,x_k)\to(\delta_1 x_1,\ldots,\delta_k x_k).$$

The main result of this section is true for such general product spaces but, for ease of notation, we will confine our attention to the case $k=2$ (letting $d\equiv d_1+d_2$). Examples of operators invariant under such dilations on $\mathbf{R}\times\mathbf{R}$ are the strong maximal function

$$M_s f(x)=\sup_{x\in R}\frac{1}{|R|}\int_R f,$$

where we take the supremum over all rectangles with sides parallel to the axes, and the double Hilbert transform

$$H_2 f(x_1,x_2)=\iint_{\mathbf{R}^2} f(x_1-y_1,x_2-y_2)\,\frac{dy_1\,dy_2}{y_1 y_2}.$$

Some of the simpler results, such as the boundedness of M_s on $L^p(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$, which was first proved by Jessen, Marcinkiewicz, and Zygmund [10], can be proved simply by iterating the 1-parameter results, but many problems cannot be treated so simply.

Product versions of singular integrals were introduced by R. Fefferman [5] and shown to be bounded on L^p for $1<p<\infty$. R. Fefferman and Stein [7] later showed that such singular integral operators are bounded on $L^p(w,\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$ if $w\in A_p(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$. (This weight space is the class of all weights w on $\mathbf{R}^{d_1+d_2}$ for which all dilations of w of the form (4) are uniformly in $A_p(\mathbf{R}^d)$; equivalently, it is defined just like the ordinary A_p space, except we replace arbitrary cubes by arbitrary "rectangles", a "rectangle" meaning anything one can get from a cube by a dilation of the form (4).) This weight condition is in fact necessary and sufficient for the boundedness of both product singular integral operators and M_s on $L^p(w,\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$, exactly analogous to the 1-parameter case.

Definition. The space of dyadic rectangles $\mathfrak{D}(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$ is the set of all rectangles $R=Q_1\times Q_2$, where $Q_1\in\mathfrak{D}(\mathbf{R}^{d_1})$ and $Q_2\in\mathfrak{D}(\mathbf{R}^{d_2})$. We also define $\mathfrak{D}_{n,m}(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})=\mathfrak{D}_n(\mathbf{R}^{d_1})\times\mathfrak{D}_m(\mathbf{R}^{d_2})$, the set of all dyadic rectangles of size $2^n\times 2^m$.

Definition. The product weight spaces $B_p^d(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$ and $A_p^d(\mathbf{R}^{d_1}\times\mathbf{R}^{d_2})$ are defined in a similar fashion to the corresponding 1-parameter spaces, except that we replace the dyadic cubes Q with dyadic rectangles R.

Since this section will be concerned only with dyadic rectangles and product weights on $\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}$, we will, for ease of notation, drop references to $\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}$ and simply write $\mathfrak{D}$, B_p^d, and A_p^d. We do this also for the product dyadic square function which we will soon define.

For any $R=Q_1\times Q_2\in\mathfrak{D}$, we define $^*R=Q_1^*\times Q_2$, $R^*=Q_1\times Q_2^*$, and $^*R^*=Q_1^*\times Q_2^*$, where Q_i^* denotes the dyadic double of Q_i. For any subset A of $\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}$, we define $\mathfrak{D}(A)$ to be the set of dyadic rectangles R such that $^*R^*\subseteq A$. We define $\mathfrak{D}_{n,m}(A)$ to be the set of all dyadic rectangles in $\mathfrak{D}(A)$ of size $2^n\times 2^m$.

For any $f\in L^{1,\infty}$, we define the second-order difference

$$\Delta_R f=f_R-f_{^*R}-f_{R^*}+f_{^*R^*},$$

and the function $f_{m,n}(x)=f_{R_{m,n;x}}$ for $x\in R_{m,n;x}\in\mathfrak{D}_{n,m}$. For any weight w, we define the product space sum

$$S_r(A,w)=\sum_{R\in\mathfrak{D}(A)} w_R^r\left(\frac{\Delta_R w}{w_R}\right)^2|R|$$

for any $A\subseteq\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}$. (To be consistent with Section 2, we should replace w_R by $w_{^*R^*}$ in this definition, but this is unnecessary since we will only deal with weights $w\in A_\infty^d\subset\mathbf{Db}$ and so $w_R\sim w_{^*R^*}$.) Our principal interest in this sum lies in the case $A=R_0\in\mathfrak{D}$, but we need to consider more general sums to derive the following theorem, which is our 2-parameter analog of Theorem 2.2.

THEOREM 4.1.

(i) $w\in B_p^d\Rightarrow S_p(R,w)\le Kw_R^p|R|\ \forall R\in\mathfrak{D}$.
(ii) $w\in A_p^d\Rightarrow S_r(R,w)\le Kw_R^r|R|\ \forall R\in\mathfrak{D}$, *where* $r=-1/(p-1)$.
(iii) $w\in A_\infty^d\Rightarrow S_r(R,w)\le Kw_R^r|R|\ \forall R\in\mathfrak{D}$, *whenever* $0\le r\le 1$.

REMARK 4.2. We cannot get any satisfactory version of the opposite-direction implications in Theorem 2.2. In fact, if $w\colon\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}\to(0,\infty)$ is an arbitrary measurable function for which $w(x_1,x_2)$ depends on x_1 alone, then $S_r(R,w)=0$ for all $r\in\mathbf{R}$ and all $R\in\mathfrak{D}$. To address this problem, we could define a sum with first-order differences as well as the second-order difference. For example, we could define, for $R_0=Q_{1,0}\times Q_{2,0}\in\mathfrak{D}$,

$$S_r'(R_0,w)=S_r(R_0,w)+\sum_{Q_1\in\mathfrak{D}(Q_{1,0})} w_{Q_1\times Q_{2,0}}^r\left(\frac{\Delta^1_{Q_1\times Q_{2,0}}w}{w_{Q_1\times Q_{2,0}}}\right)^2|Q_1||Q_{2,0}|$$
$$+\sum_{Q_2\in\mathfrak{D}(Q_{2,0})} w_{Q_{1,0}\times Q_2}^r\left(\frac{\Delta^2_{Q_{1,0}\times Q_2}w}{w_{Q_{1,0}\times Q_2}}\right)^2|Q_{1,0}||Q_2|,$$

where Δ_R^1 and Δ_R^2 are first-order differences in each of the two directions (for example, $\Delta_R^1 w=w_R-w_{^*R}$). With this definition, we can get equivalence in our theorem (assuming $w\in\mathbf{Db}^d$, of course). However, the second-order sum is then almost superfluous, since we can leave it out and still get equivalence. This is because a weight w satisfies the product A_p^d (or B_p^d) condition if and only if both $w(\cdot, x_2)$ and $w(x_1, \cdot)$ are uniformly in 1-parameter A_p^d (resp., B_p^d) for almost all $x_1 \in \mathbf{R}^{d_1}$ and $x_2 \in \mathbf{R}^{d_2}$. The product space equivalence follows easily by an iterated application of Theorem 2.2 in both variables. Theorem 4.1, by contrast, does not follow easily from our previous results. To prove it, we need the following three lemmas.

LEMMA 4.3. *Let Ω be a bounded open set, let μ be a positive measure on $\mathfrak{D}(\Omega)$, and let ν be a positive measure on Ω. If $\mu(\mathfrak{D}(A)) \le \nu(A)$ for all open $A \subseteq \Omega$ then, for all $f \in L^{1,\infty}$,*

$$\int_{\mathfrak{D}(\Omega)} f_R^p \, d\mu(R) \le C \int_\Omega (M_s^d f)^p \, d\nu. \tag{5}$$

If, in addition, $d\nu = w(x)\,dx$ for some $w \in A_p^d$, then

$$\int_{\mathfrak{D}(\Omega)} |f_R|^p \, d\mu(R) \le C_w \int_\Omega |f|^p \, d\nu.$$

Proof. Without loss of generality, we may assume that $0 < \nu(\Omega) < \infty$, that $f \ge 0$, and that f is supported on Ω (since truncating f in this fashion decreases the right-hand side of (5) but leaves the left-hand side unchanged). We define $A_k = \{x \in \Omega : 2^{k-d} < M_s^d f\}$ and $\alpha_R = \mu(\{R\})$. If $2^k < f_R \le 2^{k+1}$ then $^*R^* \subseteq A_k$, and so

$$\begin{aligned}\int_{\mathfrak{D}(\Omega)} f_R^p \, d\mu(R) &= \sum_{\mathfrak{D}(\Omega)} f_R^p \alpha_R \\ &\le \sum_k 2^{(k+1)p} \sum_{R \in \mathfrak{D}(A_k)} \alpha_R \\ &\le \sum_k 2^{(k+1)p} \nu(A_k) \le C \int_\Omega (M_s^d f)^p \nu.\end{aligned}$$

The rest of the lemma follows by the boundedness of M_s^d on $L^p(w)$, if $w \in A_p^d$ (which in turn follows easily by an iteration of the corresponding 1-parameter result). □

We define the product square function by $S_d^2 f(x) = \sum_{x \in R \in \mathfrak{D}} (\Delta_R f)^2$. We now show that S_d is bounded on $L^2(w)$ for $w \in A_2^d$. The easy proof is by an iteration technique very similar to that used in [7].

LEMMA 4.4. *If $u \in A_2^d$, then $\int (S_d^2 f) u = K \int |f|^2 u$ for some constant K.*

Proof. First, we note that

$$\int (S_d^2 f) u = \sum_{R \in \mathfrak{D}} (\Delta_R f)^2 u(R).$$

By the 1-parameter theory (i.e. Theorem 3.6),

$$\int \sum_{Q_2 \in \mathfrak{D}(\mathbf{R}^{d_2})} (\Delta_{Q_2} f_{x_1})^2 u(Q_2; x_1) \, dx_1 \le K \int \int f^2(x_1, x_2) u(x_1, x_2) \, dx_1 \, dx_2$$

where $f_{x_1}(x_2)=f(x_1,x_2)$ and $u(Q_2;x_1)=\int_{Q_2}u(x_1,y_2)\,dy_2$. If we now apply the 1-parameter theory to the function $x_1\mapsto\Delta_{Q_2}f_{x_1}$ and the weight $x_1\mapsto u(Q_2;x_1)$, the lemma then follows because $\Delta_R=\Delta_{Q_1}\Delta_{Q_2}$. □

The next lemma is due in its original form to Peter Jones [11]. Jawerth [9] extended it to cover more general types of weight spaces, including the product spaces we are dealing with here.

LEMMA 4.5. *If $w\in A_p^d$, then $w=w_1w_2^{1-p}$ for some $w_1,w_2\in A_1^d$.*

Proof of Theorem 4.1. We will prove (i)–(iii) for a fixed rectangle R_0. First, for any open $\Omega\subseteq R_0$ we will show that, if $w\in A_\infty^d$, then

$$S_1(\Omega,w)\equiv\sum_{R\in\mathfrak{D}(\Omega)}w_R^{-1}(\Delta_Rw)^2|R|\le Kw(\Omega). \tag{6}$$

It suffices to show this under the additional assumption that w is bounded on Ω since, if u is a general A_∞^d weight, we would then get (6) for the bounded weights $u_{m,n}$ for all $m,n\in\mathbf{Z}$. This is easily seen to imply (6) for $w=u$ by letting $m,n\to-\infty$.

Since $w\in A_\infty^d$, $w=w_1w_2^{1-p}$ for some $w_1,w_2\in A_1^d$ and some $1<p<\infty$ (we may assume $p>2$). Letting $u=w_1^{-1}$ (and so $u\in A_2^d$), and applying Jensen's inequality and Lemma 4.4 (with $f=w\chi_\Omega$), we get

$$\sum_{R\in\mathfrak{D}(\Omega)}(w_1)_R^{-1}(\Delta_Rw)^2|R|\le\sum_{R\in\mathfrak{D}(\Omega)}(\Delta_Rf)^2u(R)$$
$$\le K\int_\Omega w^2u=K\int_\Omega w_1w_2^{2-2p}.$$

We now apply the A_1^d condition twice, Jensen's lemma, and Lemma 4.3 with $\mu(\{R\})=(w_1)_R^{-1}(\Delta_Rw)^2|R|$ and $d\nu=w_1w_2^{2-2p}\,dx$, to get

$$\sum_{R\in\mathfrak{D}(\Omega)}w_R^{-1}(\Delta_Rw)^2|R|\le\sum_{R\in\mathfrak{D}(\Omega)}(w_2)_R^{p-1}(w_1)_R^{-1}(\Delta_Rw)^2|R|$$
$$\le K\int_\Omega(M_s^dw_2)^{p-1}w_1w_2^{2-2p}$$
$$\le K\int_\Omega w_2^{p-1}w_1w_2^{2-2p}=Kw(\Omega),$$

which proves (6).

To prove (i), we know that $w\in A_q^d$ for some $1<q<\infty$. If $u=w^{(p-1)/q}$, then $w_R^{p-1}\le u_R^q$ (by Hölder's inequality if $p-1\ge q$, and because $w\in B_p^d$ otherwise). Thus

$$\sum_{R\in\mathfrak{D}(R_0)}w_R^{p-2}(\Delta_Rw)^2|R|\le\sum_{R\in\mathfrak{D}(R_0)}u_R^qw_R^{-1}(\Delta_Rw)^2|R|$$
$$\le K\int_{R_0}u^qw\quad\text{(by Lemma 4.3)}$$
$$=K\int_{R_0}w^p\le Kw_{R_0}^p|R_0|.$$

To prove (ii), we let $\sigma = w^{-p'/p}$ as usual, and now

$$\begin{aligned} S_r(R_0, w) &= \sum_{R \in \mathfrak{D}(R_0)} w_R^{-1-p'}(\Delta_R w)^2 |R| \\ &\leq \sum_{R \in \mathfrak{D}(R_0)} \sigma_R^p w_R^{-1}(\Delta_R w)^2 |R| \quad \text{(by Jensen's inequality)} \\ &\leq K \int_{R_0} \sigma^p w \quad \text{(by Lemma 4.3)} \\ &= K \int_{R_0} \sigma \leq K w_{R_0}^r |R_0|. \end{aligned}$$

Finally, part (iii) follows easily since if $w \in A_\infty^d$ then $w \in A_p^d$ for some p, so we simply interpolate (or use Hölder's inequality, as in Theorem 2.2) between (ii) and the case $r = 1$ of (iii) (which is implied by (6)). □

It is interesting to ask if we can prove Theorem 4.1 in a manner similar to the proof of Theorem 2.2. To do so, we would need an analog of Lemma 2.5 suited to the product setting. Lemma 4.6 below leads to a proof of (i) in the case $1 < p \leq 2$ (we only need the "evenly weighted" case, i.e. $t_i = 1/N$, $s_j = 1/M$ for all i, j). However, the lemma is false for values of r outside the range $1 < r \leq 2$, as random inspection will show, so it is not possible to give such a proof for all of Theorem 4.1. Another drawback to this method of proof is that it does not easily extend to the case of more than two parameters.

For Lemma 4.6, we have numbers $a_{i,j} > 0$ for $1 \leq i \leq N$, $1 \leq j \leq M$, and we assume in addition that $\epsilon < a_{i_1,j_1}/a_{i_2,j_2}$ for some $\epsilon > 0$ and all pairs a_{i_1,j_1} and a_{i_2,j_2}. We also have weights $t_i \geq 0$ for $1 \leq i \leq N$, not all of which are zero, and $s_j \geq 0$ for $1 \leq j \leq M$, not all zero. We define $\bar{a} = (\sum_{i,j} t_i s_j a_{i,j})/(\sum_{i,j} t_i s_j)$, where "$\sum_{i,j}$" stands for "$\sum_{i=1}^N \sum_{j=1}^M$". We also define the marginal weighted averages $\bar{a}_i = (\sum_j s_j a_{i,j})/(\sum_j s_j)$ and $\bar{a}_{,j} = (\sum_i t_i a_{i,j})/(\sum_i t_i)$.

Lemma 4.6. *Suppose $a_{i,j} > 0$, $t_i \geq 0$, and $s_j \geq 0$ are as described above. Then, for all $1 < r \leq 2$, we have*

$$\sum_{i,j} t_i s_j (a_{i,j} - \bar{a}_i - \bar{a}_{,j} + \bar{a})^2 \bar{a}^{r-2} \leq K \sum_{i,j} t_i s_j (a_{i,j}^r - \bar{a}_i^r - \bar{a}_{,j}^r + \bar{a}^r),$$

where K depends only on r, N, M, and ϵ.

Remark 4.7. There are two reasons for proving a weighted version of this lemma. First, it does not seem possible to prove a product version of the lemma using only the techniques of maximization, although it is possible to prove it in the case $N = M = 2$ by transforming the problem, using a method similar to the proof of Lemma 2.7, into a different inequality that can be more easily proven, using maximization techniques. This naturally suggests attempting to prove a weighted version of the lemma to allow us to get the general case by induction (we in fact do this in the proof of the lemma). The second reason is that, by considering partial derivatives with respect to the weighting variables, one gets a completely different and considerably easier proof than is otherwise possible.

REMARK 4.8. The above lemma cannot be interpreted in terms of convexity, unlike Lemma 2.5 (in the case $N=2$). It just happens to be true by "sheer luck". In fact, it is difficult to intuitively understand why it "should" be true for this range of r, but not for other r.

Proof of Lemma 4.6. We first look at the case $N=M=2$. We can assume that $t_1+t_2=1$ and $s_1+s_2=1$ by homogeneity. We can also normalize so that $\epsilon \leq a_{i,j} \leq 1$. We write $t=t_1$, $1-t=t_2$, $s=s_1$, $1-s=s_2$, $d=a_{1,1}-a_{1,2}-a_{2,1}+a_{2,2}$, and

$$f(s,t)=st(1-s)(1-t)d^2-K\sum_{i,j} t_i s_j(a_{i,j}^r-\bar{a}_i^r-\bar{a}_{,j}^r+\bar{a}^r).$$

Now, because of the restriction on the size of $a_{i_1,j_1}/a_{i_2,j_2}$, the lemma reduces to showing that $f(s,t)<0$ for large enough K, where K depends only on r and ϵ. We note that $f(0,t)=f(1,t)=f(s,0)=f(s,1)=0$ for all $s,t\in[0,1]$. Differentiating twice in both s and t yields $\partial_t^2\partial_s^2 f(s,t)=4d^2-K\partial_t^2\partial_s^2(\bar{a}^r)$, because the other terms are harmonic in either s or t. Now, if we define $b=(\bar{a}_{,1}-\bar{a}_{,2})(\bar{a}_1-\bar{a}_2)$, then a little calculation shows us that

$$\begin{aligned}\partial_t^2\partial_s^2(\bar{a}^r)&=r(r-1)\bar{a}^{r-4}(2d^2\bar{a}^2-4(2-r)db\bar{a}+(2-r)(3-r)b^2)\\&\geq c_r d^2\bar{a}^{r-2}\geq c_r d^2,\end{aligned}$$

where c_r depends only on r. If $r=2$, the first inequality is obvious. If $1<r<2$, it follows because the parenthesized expression can be written as a perfect square added to $c_r' d^2\bar{a}^2$, where c_r' depends only on r. To see this, note that this expression is a quadratic in $d\bar{a}$ and that

$$4(2)\{(2-r)(3-r)b^2\}-\{4(2-r)b\}^2=8(2-r)(r-1)b^2>0.$$

Thus, $\partial_t^2\partial_s^2 f(s,t)\leq 4d^2-K(c_r d^2)\leq 0$ for K sufficiently large, and so $\partial_s^2 f(s,t)$ is superharmonic in t. Since

$$\partial_s^2 f(s,0)=\partial_s^2 f(s,1)=0,$$

it follows that $\partial_s^2 f(s,t)\geq 0$, so that f is subharmonic in s. Since $f(0,t)=f(1,t)$, it follows that $f(s,t)\leq 0$, as required. This finishes the proof in the case $M=N=2$.

To prove the result inductively for all N and M, it suffices, by symmetry of N and M, to prove that it is true for $(N,M)=(N_0,M_0+1)$ whenever it is true for $(N,M)=(N_0,M_0)$. Let us assume the result for $(N,M)=(N_0,M_0)$ and suppose that we are given $a_{i,j}>0$, $t_i\geq 0$, and $s_j\geq 0$ for $1\leq i\leq N_0$, $1\leq j\leq M_0+1$. We can assume $\epsilon\leq a_{i,j}\leq 1$ and $\sum_{i=1}^{N_0} t_i=1$ by normalization. We can also assume $\sum_{j=1}^{M_0} s_j=1$ (this is only a problem if the sum is 0, in which case the inductive step is trivial). We define $\hat{a}=\sum_{i=1}^{N_0}\sum_{j=1}^{M_0} t_i s_j a_{i,j}$ and $\hat{a}_i=\sum_{j=1}^{M_0} s_j a_{i,j}$ ($\bar{a}$ and $\bar{a}_i$ will denote weighted averages for the larger set of numbers).

Now it is easy to show that, if $\{b_j\}_{j=1}^{M_0}$ is any set of numbers and $\hat{b}=\sum_{j=1}^{M_0} s_j b_j$, then for any $\lambda \in \mathbf{R}$,

$$\sum_{j=1}^{M_0} s_j(b_j-\lambda)^2=(\lambda-\hat{b})^2+\sum_{j=1}^{M_0} s_j(b_j-\hat{b})^2.$$

Using this equality, we get

$$\begin{aligned}(7)\quad \sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0+1}} t_i s_j(a_{i,j}-\bar{a}_i-\bar{a}_{,j}+\bar{a})^2 &= \sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0}} t_i s_j(a_{i,j}-\hat{a}_i-\bar{a}_{,j}+\hat{a})^2\\ &\quad+\sum_{1\le i\le N_0} t_i(\hat{a}_i-\hat{a}-\bar{a}_i+\bar{a})^2\\ &\quad+\sum_{1\le i\le N_0} t_i s_{M_0+1}(a_{i,M_0+1}-\bar{a}_i-\bar{a}_{,M_0+1}+\bar{a})^2.\end{aligned}$$

Also,

$$\begin{aligned}(8)\quad \sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0+1}} t_i s_j(a^r_{i,j}-\bar{a}^r_i-\bar{a}^r_{,j}+\bar{a}^r) &= \sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0}} t_i s_j(a^r_{i,j}-\hat{a}^r_i-\bar{a}^r_{,j}+\hat{a}^r)\\ &\quad+\sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0}} t_i s_j(\hat{a}^r_i-\hat{a}^r-\bar{a}^r_i+\bar{a}^r)\\ &\quad+\sum_{1\le i\le N_0} t_i s_{M_0+1}(a^r_{i,M_0+1}-\bar{a}^r_i-\bar{a}^r_{,M_0+1}+\bar{a}^r)\\ &= \sum_{\substack{1\le i\le N_0\\ 1\le j\le M_0}} t_i s_j(a^r_{i,j}-\hat{a}^r_i-\bar{a}^r_{,j}+\hat{a}^r)\\ &\quad+\sum_{1\le i\le N_0} t_i(\hat{a}^r_i-\hat{a}^r-\bar{a}^r_i+\bar{a}^r)\\ &\quad+\sum_{1\le i\le N_0} t_i s_{M_0+1}(a^r_{i,M_0+1}-\bar{a}^r_i-\bar{a}^r_{,M_0+1}+\bar{a}^r).\end{aligned}$$

The first term in (7) is a left-hand side of the inequality for $(N, M)=(N_0, M_0)$ and so, by the inductive hypothesis, is less than a constant times the first term in (8), which is the corresponding right-hand side. The sum of the last two terms in (7) is a left-hand side of the inequality for $(N, M)=(N_0, 2)$ and so, by the inductive hypothesis, is less than a constant times the sum of the last two terms in (8), which is the corresponding right-hand side. Thus the lemma is true by induction. □

References

[1] S. Buckley, *Harmonic analysis on weighted spaces,* Thesis, University of Chicago, 1990.

[2] S. Y. A. Chang, J. Wilson, and T. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators,* Comm. Math. Helv. 60 (1985), 217–246.

[3] S. Y. A. Chang and R. Fefferman, *Some recent developments in Fourier analysis and H^p-theory on product domains,* Bull. Amer. Math. Soc. (N.S.) 12 (1985), 1–43.

[4] R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals,* Studia Math. 51 (1974), 241–250.

[5] R. Fefferman, *Singular integrals on product domains,* Bull. Amer. Math. Soc. (N.S.) 4 (1981), 195–201.

[6] R. Fefferman, C. Kenig, and J. Pipher, *The theory of weights and the Dirichlet problem for elliptic equations,* Ann. of Math. (2) 134 (1991), 65–124.

[7] R. Fefferman and E. Stein, *Singular integrals on product spaces,* Adv. in Math. 45 (1982), 117–143.

[8] J. García-Cuerva and J. Rubio de Francia, *Weighted norm inequalities and related topics,* mathematics studies, North-Holland, 1985.

[9] B. Jawerth, *Weighted inequalities for maximal operators: linearization, localization and factorization,* Amer. J. Math. 108 (1986), 361–414.

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

[11] P. Jones, *Factorization of A_p weights,* Ann. of Math. (2) 111 (1980), 511–530.

[12] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function,* Trans. Amer. Math. Soc. 165 (1972), 207–226.

[13] ———, *Norm inequalities relating the Hilbert transform to the Hardy-Littlewood maximal function,* Functional analysis and approximation (Oberwolfach, 1980), pp. 219–231, Birkhäuser, Basel, 1981.

[14] J. Rubio de Francia, *Factorization theory and A_p weights,* Amer. J. Math. 106 (1984), 533–547.

Department of Mathematics
University of Michigan
Ann Arbor, MI 48109