Weighted Norm Estimates for Gradients of Half-space Extensions

RICHARD L. WHEEDEN & J. MICHAEL WILSON

ABSTRACT. Let $u(x, y) = f * \varphi_y(x)$, with f in $\bigcup_{1 \leq p < \infty} L^p(\mathbb{R}^d)$ and φ a smooth convolution kernel with decay at infinity. We prove sufficient conditions on positive measures μ and non-negative weights v which ensure that

$$(\alpha) \qquad \left(\int_{\mathbb{R}_+^{d+1}} |\nabla u|^q d\mu \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

holds for all suitable functions f . We prove necessary conditions in the case where φ = the Poisson kernel. We also consider a discrete (i.e., martingale) analogue to (α) , and we prove sufficient (and necessary) conditions in that case as well.

1. Introduction. In this paper we study the following norm inequality for the Poisson integral $u(x, y)$ ($x \in \mathbb{R}^d$, $y > 0$) of a function $f(x)$:

$$(1.1) \qquad \left(\int_{\mathbb{R}_+^{d+1}} |\nabla u(x, y)|^q d\mu(x, y) \right)^{1/q} \leq C \left(\int_{\mathbb{R}^d} |f(x)|^p v(x) dx \right)^{1/p}.$$

Here, $1 < p \leq q < \infty$, ∇ denotes the usual gradient operator in both the x and y variables (i.e., $\nabla = (\partial/\partial x_1, \dots, \partial/\partial x_d, \partial/\partial y)$ if $x = (x_1, \dots, x_d)$), μ is a non-negative Borel measure on $\mathbb{R}_+^{d+1} = \mathbb{R}^d \times (0, \infty)$, and v is a non-negative, locally integrable function on $\mathbb{R}^d$. The problem is to determine the conditions on μ and v which ensure that (1.1) holds for all f , with a constant C that is independent of f .

In case $v(x) \equiv 1$, this problem has been carefully studied in [L1], [L2], [Sh1], [Sh2], [Ve] and [Vi], where a complete characterization of the measure μ is given for all p, q satisfying $0 < p, q < \infty$, with the L^p norm of f replaced by the Hardy

H^p norm in case $0 < p \leq 1$. In this paper, we consider only the weighted L^p problem for $1 < p \leq q < \infty$, although we shall study the weighted H^p problem in a sequel to this paper.

We will also consider some variants of (1.1), including similar inequalities with $\nabla u(x, y)$ replaced by $y^{-1}(f * \varphi_y)(x)$, where φ is a smooth function which decays at infinity and satisfies

$$(1.2) \quad \int_{\mathbb{R}^d} \varphi(x) dx = 0,$$

and where φ_y is the standard dilation of φ defined by $\varphi_y(x) = y^{-d}\varphi(x/y)$, $x \in \mathbb{R}^d$, $y > 0$. The derivatives $\partial u/\partial x_i$ and $\partial u/\partial y$ are included in this situation. In fact, if we write $u(x, y) = P_y * f(x)$, where

$$P_y(x) = c_d \frac{y}{(y^2 + |x|^2)^{(d+1)/2}} = y^{-d} P\left(\frac{x}{y}\right)$$

is the Poisson kernel for $\mathbb{R}_+^{d+1}$ and $P(x) = c_d(1 + |x|^2)^{-(d+1)/2}$, then we obtain by direct computation that

$$\frac{\partial u}{\partial x_i}(x, y) = y^{-1}(f * \varphi_i)_y(x), \quad i = 1, \dots, d,$$

and

$$\frac{\partial u}{\partial y}(x, y) = y^{-1}(f * \varphi_0)_y(x),$$

with

$$\varphi_i(x) = \frac{\partial P}{\partial x_i}(x), \quad i = 1, \dots, d \quad \text{and} \quad \varphi_0(x) = -dP(x) - \sum_{i=1}^d x_i \frac{\partial P}{\partial x_i}(x).$$

Thus,

$$\int_{\mathbb{R}^d} \varphi_i(x) dx = 0 \quad i = 1, \dots, d,$$

and, by integration by parts,

$$\int_{\mathbb{R}^d} \varphi_0(x) dx = 0.$$

The rate of decay of φ plays an important role in our results. We note for example that the kernels φ_0 and φ_i defined above have different rates of decay at infinity. In fact, since

$$(1.3a) \quad \varphi_i(x) = -c_d(d+1) \frac{x_i}{(1+|x|^2)^{(d+3)/2}} \quad i = 1, \dots, d,$$

$$(1.3b) \quad \varphi_0(x) = c_d \frac{|x|^2 - d}{(1+|x|^2)^{(d+3)/2}},$$

it follows that φ_i decays like $|x|^{-d-2}$, while φ_0 decays at the slower rate $|x|^{-d-1}$. This difference will be reflected in our results. We shall also consider separately the simpler case when φ has compact support, as well as analogues of (1.1) for Haar-function and wavelet expansions.

For the moment, however, we consider only the estimate (1.1) for the full gradient of the Poisson integral. For that case, we will prove the following result when $p = q \geq 2$. In stating the result, and throughout the paper, if Q is a cube in $\mathbb{R}^d$ with center x_Q and edglength $\ell(Q)$, we use the notation $T(Q)$ for the set in $\mathbb{R}_+^{d+1}$ defined by

$$T(Q) = \left\{ (x, y) : x \in Q, \frac{\ell(Q)}{2} < y \leq \ell(Q) \right\}.$$

Thus, $T(Q)$ is the top half of the usual Carleson “box” $\hat{Q}$, defined by

$$\hat{Q} = \{(x, y) : x \in Q, 0 < y \leq \ell(Q)\}.$$

We denote the μ -measure of $T(Q)$ by $\mu(T(Q))$.

Theorem A. *Let v be a non-negative weight in $L_{\text{loc}}^1(\mathbb{R}^d)$ and μ be a positive measure on $\mathbb{R}_+^{d+1}$. Let $2 \leq p < \infty$, $1/p + 1/p' = 1$, and $\sigma = v^{1-p'}$.*

(i) *If the norm inequality,*

$$(1.4) \quad \left(\int_{\mathbb{R}_+^{d+1}} |\nabla u(x, y)|^p d\mu(x, y) \right)^{1/p} \leq \left(\int_{\mathbb{R}^d} |f(x)|^p v(x) dx \right)^{1/p},$$

holds for all f , where $u(x, y)$ is the Poisson integral of f , then there is a constant C depending only on p , σ , and d such that

$$(1.5) \quad \mu(T(Q))^{1/p} \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'}} dx \right)^{1/p'} \leq C$$

for all cubes Q in $\mathbb{R}^d$.

(ii) *Conversely, if σ is a weight in the Muckenhoupt class A_∞ and if*

$$(1.6) \quad \mu(T(Q))^{1/p} \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{d+1}} dx \right)^{1/p'} \leq C \ell(Q)^{(d+1)/p}$$

for all cubes Q , for a suitable constant C that depends only on p , d , and σ , then (1.4) holds for all f .

We will also derive a similar result in the case $q > p$. Moreover, we will derive a sufficiency result (i.e., an analogue of part (ii) of Theorem A) in the case when σ is not an A_∞ weight. There will always be, however, a gap between the conditions for sufficiency and those for necessity. In order to prove that (1.5) is necessary for (1.4), we need the presence of the full gradient in (1.4). On the other hand, for part (ii), we shall see that the analogue of (1.4) for just the x -derivatives $\partial u / \partial x_i$, excluding $\partial u / \partial y$, is true if we replace (1.6) by the weaker condition

$$(1.7) \quad \mu(T(Q))^{1/p} \left(\int_{\mathbb{R}^d} \frac{\log(e + |x - x_Q|/\ell(Q)) \sigma(x)}{(\ell(Q) + |x - x_Q|)^{d+2}} dx \right)^{1/p'} \\ \leq C \ell(Q)^{d+1-(d+2)/p'},$$

still assuming that $\sigma \in A_\infty$. The exponent $d+2$ in the denominator of the second factor on the left in (1.7) arises from the decay rate of the kernels φ_i , $i = 1, \dots, d$. More generally, we shall consider kernels φ with other rates of decay, and this will affect the form of the second factor in (1.7).

The proof of part (ii) in Theorem A is partly based on square function estimates of Littlewood-Paley type, and these estimates lead to restrictions on μ and σ . We say that a function f is in standard form if

$$(1.8) \quad f(x) = \sum_Q \lambda_Q b_Q(x),$$

where the sum is extended over a finite number (independent of x) of dyadic cubes Q , the λ_Q are constants, and

$$(1.9a) \quad b_Q(x) \text{ is continuously differentiable;}$$

$$(1.9b) \quad b_Q(x) \text{ is supported in } \tilde{Q}, \text{ the concentric triple of } Q;$$

$$(1.9c) \quad \int b_Q(x) dx = 0;$$

$$(1.9d) \quad |\nabla b_Q(x)| \leq \ell(Q)^{-1} |Q|^{-1/2}.$$

Such a function $b_Q(x)$ will be called an *adapted function* for Q .

For a function f in standard form (1.8), we will consider the square function $\tilde{S}(f)$, defined by (see [W3])

$$\tilde{S}(f)(x) = \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \chi_{\tilde{Q}}(x) \right)^{1/2},$$

where, as before, $\tilde{Q}$ denotes the cube with the same center as Q and with edge-length $3\ell(Q)$.

If σ and w are a pair of non-negative functions in $L^1_{\text{loc}}(\mathbb{R}^d)$, and if $1 < r < \infty$, we say that (σ, w) is an r -pair if $\sigma \leq w$ almost everywhere and if

$$\int_{\mathbb{R}^d} |f|^r \sigma dx \leq \int_{\mathbb{R}^d} [\tilde{S}(f)]^r w dx$$

for all functions f in standard form. For example, if σ is an A_∞ weight, then (by, e.g., [W3]) the pair $(\sigma, C\sigma)$ is an r -pair for any r , if C is a sufficiently large constant depending only on d , σ , and r . Also, by [W3], if k is an integer that is larger than $r/2$, then $(\sigma, CM^k\sigma)$ is an r -pair for any σ , where M^k denotes the k -fold iteration of the Hardy-Littlewood maximal operator M , and C is a constant which only depends on k , d , and r . We shall prove a stronger result in the case $r \geq 2$ which will be useful in combination with our results concerning (1.1) and its variants. In order to state this result, first let $\alpha \geq 0$ and define

$$\mathcal{W}_\alpha \sigma(x) = \sup_{Q: x \in Q} \frac{1}{|Q|} \int_Q \sigma(t) \log^\alpha \left(e + \frac{\sigma(t)}{\sigma_Q} \right) dt,$$

where, as usual,

$$\sigma_Q = \frac{1}{|Q|} \int_Q \sigma(t) dt,$$

the average of σ over Q , and Q is any cube, dyadic or otherwise. In the case when α is an integer, $\mathcal{W}_\alpha$ has been studied in [GI]. Clearly, $\mathcal{W}_0 = M$, the Hardy-Littlewood maximal operator, and also, by using results in [St], $\mathcal{W}_k \leq C_{k,d} M^{k+1}$ for all non-negative integers k .

We will prove the following result in Section 3.

Theorem B. *If $r \geq 2$ and $\alpha + 1 > r/2$, then $(\sigma, C_{\alpha,d} \mathcal{W}_\alpha \sigma)$ is an r -pair.*

We will use the notion of r -pairs to derive analogues of Theorem A in case σ is not an A_∞ weight. For example, we prove the following “real-variable” (non-Poisson kernel) result in Section 3.

Theorem C. *Let $1 < p \leq q < \infty$ and $q \geq 2$. Let $\varphi \in C^\infty(\mathbb{R}^d)$ satisfy $\int \varphi = 0$, and suppose that, for some non-negative integer m , $|\varphi(x)| \leq (1 + |x|)^{-d-2-m}$ and $|\nabla \varphi(x)| \leq (1 + |x|)^{-d-3-m}$ for all $x \in \mathbb{R}^d$. In order for*

$$(1.10) \quad \left(\int_{\mathbb{R}^{d+1}_+} |y^{-1} \varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , it is sufficient that there exist a p' -pair (σ, w) such that

$$(1.11) \quad \mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2+m)p'/q'}} dx \right)^{1/p'} \\ \leq C \ell(Q)^{d+1-(d+2+m)/q'}$$

holds for all cubes Q . The constant C in (1.11) depends on p, q, d , and m .

By choosing $m = 0$, Theorem C includes the x -components of the gradient of the Poisson kernel, but not the y -component, because the kernels $\varphi_i, i = 1, \dots, d$, defined earlier, satisfy the decay condition in the hypothesis of Theorem C, but φ_0 does not. For the y -derivative, which we study separately in Section 4, we require extra arguments and obtain several different results. One of these has already been stated in part (ii) of Theorem A. Another is given in Corollary 7 in Section 4, and is proved by using a result due to C. Pérez [P] for Calderón-Zygmund singular integral operators. In the special case when $q = p \geq 2$ and v satisfies the A_p condition

$$\left(\frac{1}{|Q|} \int_Q v dx \right)^{1/p} \left(\frac{1}{|Q|} \int_Q \sigma dx \right)^{1/p'} \leq C$$

for all cubes Q , this corollary may be stated as follows.

Theorem D. Let μ be a positive measure on $\mathbb{R}_+^{d+1}$, $2 \leq p < \infty$, $v \in A_p$, and $\sigma = v^{1-p'}$. Then

$$\left(\int_{\mathbb{R}_+^{d+1}} \left| \frac{\partial u}{\partial y} \right|^p d\mu(x, y) \right)^{1/p} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

for all functions f , where u is the Poisson integral of f , if

$$\mu(T(Q))^{1/p} \left(\int_{\mathbb{R}^d} \frac{\log(e + |x - x_Q|/\ell(Q)) \sigma(x)}{(\ell(Q) + |x - x_Q|)^{d+2}} dx \right)^{1/p'} \leq C \ell(Q)^{d+1-(d+2)/p'}$$

for all cubes Q . The constant C depends on p, d , and v .

By Theorem C, this result also holds with $\partial u/\partial y$ replaced by the full gradient ∇u .

Our main results and their proofs are organized as follows. We begin in Section 2 with analogous results for Haar function expansions and for the case when φ has compact support, including some results for wavelets. The proofs

in these situations are less complicated than in the non-compact support cases, but still contain many of the ideas used in the harder cases. The conditions imposed on μ and σ in Section 2 are more local in nature than those we have mentioned so far; in particular, in these conditions, the integral involving σ can be replaced by one which is extended only over $\tilde{Q}$. Theorem C, which deals with the case when φ is not compactly supported, but decays at infinity at least as fast as $|x|^{-d-2}$, is treated in Section 3. Theorem C includes the x -derivatives of the Poisson integral, but not the y -derivative. Theorem B is also proved in Section 3. Several results concerning the y -derivative of the Poisson integral are proved in Section 4, including Theorems A(ii) and D. In Section 5 we prove the necessity result in Theorem A(i).

The cancellation property of φ (i.e., the fact that $\int \varphi = 0$) plays an important role in the proofs of all the results discussed above. There is, however, an alternate approach to the sufficiency results which ignores this cancellation. In this approach, one simply majorizes the quantity $|y^{-1}\varphi_y * f(x)|$ which appears on the left in (1.10) by $y^{-1}|\varphi_y| * |f|(x)$, and then applies known results for operators with non-negative kernels, the non-negative kernel in this case being $|\varphi_y(x)|$ multiplied by an appropriate power of y . For example, in case φ satisfies the decay estimate

$$|\varphi(x)| \leq (1 + |x|)^{-M}$$

for some $M > 0$, and all x , it is convenient to use the inequality

$$|y^{-1}\varphi_y * f(x)| \leq y^{M-d-1} T f(x, y),$$

where $T f(x, y)$ is defined by

$$T f(x, y) = \int_{\mathbb{R}^d} |f(t)| \frac{1}{(y + |x - t|)^M} dt.$$

We then incorporate the factor y^{M-d-1} , raised to the power q when we consider an L^q norm, as part of the measure on $\mathbb{R}_+^{d+1}$, and apply the Carleson-type results in [SWZ] (and earlier in [SW] for the Poisson kernel) to estimate the resulting norm of $T f(x, y)$.

This process yields conclusions like (1.10) under conditions on μ and σ which are generally different from those that have been imposed so far. The new conditions involve an integral of μ extended over a set which is larger than $T(Q)$, such as $\tilde{Q}$ or all of $\mathbb{R}_+^{d+1}$, but in the integral for σ , they involve a smaller integrand, after normalization. Also, the new conditions do not require the restriction that $q \geq 2$. Sometimes the new conditions are weaker than those already considered, but there are also cases, most notably when $p = q$, when the opposite is true.

For example, in the case when $p < q$, we obtain the following result in this way.

Theorem E. Let $1 < p < q < \infty$, and let φ satisfy $|\varphi(x)| \leq (1 + |x|)^{-M}$ for some $M > 0$ and all $x \in \mathbb{R}^d$. Then the inequality

$$\left(\int_{\mathbb{R}_+^{d+1}} |y^{-1} \varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

holds for all f if

$$(1.12) \quad \left(\int_{\mathbb{R}_+^{d+1}} \frac{y^{(M-d-1)q}}{(\ell(Q) + y + |x - x_Q|)^{Mq}} d\mu(x, y) \right)^{1/q} \cdot \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{Mp'}} dx \right)^{1/p'} \leq C \ell(Q)^{-M}$$

for all cubes Q . The constant C depends on p, q, d , and M .

The x -derivatives $\partial u / \partial x_i$ of the Poisson kernel correspond to choosing $M = d + 2$, and $\partial u / \partial y$ corresponds to $M = d + 1$.

In order to see how (1.12) is related to our earlier conditions, we first note that arguments in [SW] and [SWZ] show that (1.12) is equivalent to the following two weaker conditions together when $q > p$:

$$(1.12' a) \quad \left(\int_{\hat{Q}} y^{(M-d-1)q} d\mu(x, y) \right)^{1/q} \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{Mp'}} dx \right)^{1/p'} \leq C,$$

$$(1.12' b) \quad \left(\int_{\mathbb{R}_+^{d+1}} \frac{y^{(M-d-1)q}}{(\ell(Q) + y + |x - x_Q|)^{Mq}} d\mu(x, y) \right)^{1/q} \sigma(Q)^{1/p'} \leq C.$$

If σ also satisfies the reverse doubling condition:

$$(RD) \quad \sigma(2Q) \geq \alpha \sigma(Q)$$

for some $\alpha > 1$, for all cubes Q , where $2Q$ denotes the cube with center x_Q and edglength $2\ell(Q)$ (e.g., if σ is an A_∞ weight), then as in ([SW], p. 658), the condition (1.12'b) is the same as

$$\left(\int_{\hat{Q}} y^{(M-d-1)q} d\mu(x, y) \right)^{1/q} \sigma(Q)^{1/p'} \leq C \ell(Q)^M,$$

which clearly follows from (1.12'a). Thus, for $q > p$ and σ satisfying (RD), (1.12) is equivalent to

$$(1.12'') \quad \left(\int_{\hat{Q}} y^{(M-d-1)q} d\mu(x, y) \right)^{1/q} \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{Mp'}} dx \right)^{1/p'} \leq C.$$

Even though the integral in the first factor of (1.12'') is extended over all of $\hat{Q}$, a simple addition argument shows that (1.12'') is weaker than the conditions we obtain in Theorems 3 and 6 as analogues of Theorems C and A(ii) for $q > p$. Of course, in Theorems 3 and 6, we do not assume that σ satisfies (RD).

It is also possible to simplify (1.12) if $y^{(M-d-1)q}d\mu(x, y)$ satisfies a sort of (RD) condition in the half-space, but we have to chosen to emphasize the case when σ satisfies (RD) for purposes of comparison with our earlier results.

The technique indicated above for obtaining results by ignoring cancellation does not work as well when $p = q$. For example, in the basic case when $u(x, y)$ is the Poisson integral of f , since $|(\varphi_i)_y| \leq c_d P_y(x)$ for all $i = 0, 1, \dots, d$, we have $|\nabla u(x, y)| \leq c_d y^{-1} |f| * P_y(x)$, and consequently,

$$\left(\int_{\mathbb{R}_+^{d+1}} |\nabla u(x, y)|^p d\mu(x, y) \right)^{1/p} \leq c_d \left(\int_{\mathbb{R}_+^{d+1}} (y^{-1} |f| * P_y)(x)^p d\mu(x, y) \right)^{1/p}.$$

We can now apply the results of [SW] to bound the integral on the right by a weighted norm of f . In the simple case when $v = 1$ and $1 < p < \infty$, this requires assuming that μ satisfy the condition

$$(1.13) \quad \int_{\hat{Q}} y^{-p} d\mu(x, y) \leq C|Q|,$$

whereas Theorem A(ii) shows that the weaker condition,

$$(1.14) \quad \ell(Q)^{-p} \int_{T(Q)} d\mu(x, y) \leq C|Q|,$$

is enough when $p \geq 2$. Clearly, $d\mu = y^{p-1} dx dy$ satisfies (1.14) but not (1.13).

For any σ in A_∞ and $1 < p < \infty$, the condition required for (1.1) by using this approach is

$$(1.15) \quad \left(\int_{\hat{Q}} y^{-p} d\mu(x, y) \right)^{1/p} \sigma(Q)^{1/p'} \leq C|Q|$$

(see [SW], p. 645). For general σ , it is also possible to formulate an appropriate analogue based on Theorem 2 in [SW].

2. When the kernels have compact support. We work on the unit interval $[0, 1]$. For each dyadic interval $I \subset [0, 1]$, we define the corresponding Haar function as follows. Let $I = I_1 \cup I_2$, where I_1 and I_2 are respectively the right and left halves of I . Then $h_{(I)}(x) = |I|^{-1/2}(\chi_{I_1} - \chi_{I_2})$. It is easy to see

that the family $\{h_{(I)}\}$ forms an orthonormal basis for $L_0^2[0,1] = \{f \in L^2[0,1] : \int f = 0\}$. If f is integrable and supported on $[0,1]$, we define $\lambda_I(f) = \langle f, h_{(I)} \rangle$. It is important to notice that $\lambda_I(f)$ is obtained by a “discrete” convolution of f with a function that has integral zero; this is what makes it analogous to ∇u .

Let $v \geq 0$ be an integrable function supported on $[0,1]$, and let $\{\mu_I\}$ be a sequence of non-negative numbers indexed on the family $\mathcal{D}$ of dyadic subintervals of $[0,1]$. Our first task is to find necessary and (strong) sufficient conditions on v and $\{\mu_I\}$ so that the inequality

$$(*) \quad \left(\sum_I |\lambda_I(f)|^q \mu_I \right)^{1/q} \leq \left(\int |f|^p v dx \right)^{1/p}$$

holds for all f .

Inequality $(*)$ has an easy necessary condition. Let $\sigma = v^{1-p'}$, where p' is the dual exponent to p . For fixed I , let $f = \sigma \chi_{I^*}$, where I^* is the right or left half of I . A quick computation shows that $(*)$ implies

$$(2.1) \quad \frac{\mu_I^{1/q} \cdot \sigma(I)^{1/p'}}{|I|^{1/2}} \leq C.$$

(If $\sigma \notin L^1[0,1]$, an approximation argument shows that $(*)$ still requires (2.1) to hold.) Our first theorem asserts that, if $1 < p \leq q < \infty$ and $q \geq 2$, inequality (1) is almost sufficient.

Definition. For η a positive number, I a dyadic interval, and σ as above, we define

$$\sigma(I, \eta) = \int_I \sigma(x) \log^\eta \left(e + \frac{\sigma(x)}{\sigma_I} \right) dx,$$

where $\sigma_I = |I|^{-1} \int_I \sigma$.

Theorem 1. If $1 < p \leq q < \infty$, $q \geq 2$, and $\tau > p'/2$, then $(*)$ holds if

$$(2.2) \quad \frac{\mu_I^{1/q} \cdot \sigma(I, \tau)^{1/p'}}{|I|^{1/2}} \leq C,$$

for all intervals I , where the constant C depends on p , q , and τ .

Proof. The proof has two main steps:

- (a) rephrasing (*) in a dual form, which makes our problem amenable to the machinery of weighted Littlewood-Paley theory;
- (b) applying the machinery. The first step is short and the second step is long.

For $g : \mathcal{D} \mapsto \mathbb{R}$, a *finitely supported* function, define $Tg(x) = \sum_I g(I)\mu_I \cdot h_{(I)}(x)$. By duality, inequality (*) holds if and only if

$$(2.3) \quad \left(\int_0^1 |Tg|^{p'} \sigma dx \right)^{1/p'} \leq \left(\sum_I |g(I)|^{q'} \mu_I \right)^{1/q'}$$

for all such g . So much for step one.

Now let us set $\gamma_I = g(I)\mu_I$. We need to show that (2.2) implies

$$\left(\int \left| \sum_I \gamma_I h_{(I)} \right|^{p'} \sigma dx \right)^{1/p'} \leq C \left(\sum_I |\gamma_I|^{q'} \mu_I^{1-q'} \right)^{1/q'}$$

for all finite sequences $\{\gamma_I\}$. This is a Littlewood-Paley inequality. Let us justify this assertion by considering the special case $p' = q' = 2$, and with the additional hypothesis that σ belongs to the Muckenhoupt class A_∞ . (Considering this simple case should make the proof in the general case easier to digest.) Recall the definition of the *dyadic square function* $S(f)$:

$$S(f)(x) = \left(\sum_I \frac{|\lambda_I(f)|^2}{|I|} \chi_I(x) \right)^{1/2},$$

and observe that $\lambda_I(Tg) = \gamma_I$. For this special case, it is well-known that

$$\begin{aligned} \int \left| \sum_I \gamma_I h_{(I)} \right|^2 \sigma dx &\leq C \int S^2(Tg) \sigma dx \\ &= C \sum_I |\gamma_I|^2 \sigma_I, \end{aligned}$$

where the constant C depends on σ . If, in addition, (2.2) (or even (2.1)) holds, then the last line is less than or equal to

$$\sum_I |\gamma_I|^2 \mu_I^{-1},$$

and this is (2.3).

If $\sigma \notin A_\infty$ then we can use a result from [W4] (p. 424):

Result 1. If $\tau > 1$ then there is a constant C_τ such that

$$\int \left| \sum_I \gamma_I h_{(I)} \right|^2 \sigma dx \leq C_\tau \sum_I \frac{|\gamma_I|^2}{|I|} \sigma(I, \tau),$$

for all finite sequences $\{\gamma_I\}$.

The result, when combined with (2.2), yields (2.3).

The preceding argument embodies the *main idea* of the proofs for the Haar function, compactly supported, and general continuous cases.

Now we consider the general case $1 < p \leq q < \infty$, with $q \geq 2$. This naturally divides into two subcases, $p \leq 2$ and $p > 2$. Both of these use the following result, which is essentially contained in [W3] (Lemma 2.3):

Result 2. Let r , A , and η be positive numbers, and suppose that $\mathcal{F} \subset \mathcal{D}$ is a finite family of dyadic intervals such that $\sigma(I, \eta) \leq A\sigma(I)$ for all $I \in \mathcal{F}$. If $f = \sum_{I \in \mathcal{F}} \lambda_I h_{(I)}$ then

$$\int |f|^r \sigma dx \leq CA^{r/2\eta} \int S^r(f) \sigma dx,$$

where C is a constant independent of A .

When we use Result 2, r will be p' . When $p' \geq 2$ we will take $\eta = 1$. When p' is less than 2, η will equal τ (which is larger than $p'/2$).

▷ SUBCASE 1: $2 < p < \infty$. We let $\Phi(x) = \sum_I \gamma_I h_{(I)}$; remember that $\gamma_I \neq 0$ for only finitely many I . Let $p'/2 < \tau$. Set $\eta = \tau$, and for each non-negative integer k , let $\mathcal{F}_k$ be those dyadic intervals I such that $2^k \leq \sigma(I, \eta)/\sigma(I) < 2^{k+1}$ (where we take the ratio to be 1 if $\sigma(I) = 0$). Let $\Phi_k = \sum_{I \in \mathcal{F}_k} \gamma_I h_{(I)}$. Let ε be a small positive number, to be chosen later. By Hölder's inequality and Result 2:

$$\begin{aligned} \int |\Phi|^{p'} \sigma dx &\leq C_\varepsilon \sum_k 2^{\varepsilon k} \int |\Phi_k|^{p'} \sigma dx \\ &\leq C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2\tau} \int S^{p'}(\Phi_k) \sigma dx \\ &\leq C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2\tau} \int \sum_{I \in \mathcal{F}_k} \frac{|\gamma_I|^{p'}}{|I|^{p'/2}} \chi_I \sigma dx \\ &\leq C \sum_k 2^{k(-1+\varepsilon+p'/2\tau)} \sum_{I \in \mathcal{F}_k} \frac{|\gamma_I|^{p'}}{|I|^{p'/2}} \cdot \left(\frac{\sigma(I, \tau)}{\sigma(I)} \right) \cdot \sigma(I), \end{aligned}$$

where the third inequality comes from the fact that $p'/2 < 1$ and the fourth comes from the definition of $\mathcal{F}_k$. Since $\tau > p'/2$, we have $-1 + \varepsilon + p'/2\tau < 0$ for small ε . Therefore, the last line in our chain of inequalities is less than or equal to

$$C \sum_I \frac{|\gamma_I|^{p'}}{|I|^{p'/2}} \sigma(I, \tau).$$

The hypothesis of the theorem now implies that the last line is bounded by

$$C \sum_I |\gamma_I|^{p'} \cdot \mu_I^{-p'/q} \leq C \left(\sum_I |\gamma_I|^{q'} \mu_I^{1-q'} \right)^{p'/q'},$$

since $p' \geq q'$. Subcase 1 is proved.

▷ SUBCASE 2: $1 < p \leq 2$. With $\eta = 1$ now, we define Φ and Φ_k as before. Reasoning as before, we get, for any $\varepsilon > 0$,

$$(2.4) \quad \int |\Phi|^{p'} \sigma dx \leq C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2} \int S^{p'}(\Phi_k) \sigma dx.$$

For each k , let $\beta_k(x)$ be a function such that $\|\beta_k\|_{L^{(p'/2)'}(\sigma dx)} = 1$ and

$$\int S^{p'}(\Phi_k) \sigma dx = \left(\int S^2(\Phi_k) \beta_k \sigma dx \right)^{p'/2}.$$

The preceding integral (the one inside the parentheses) can be easily estimated via Hölder's Inequality:

$$\begin{aligned} \int S^2(\Phi_k) \beta_k \sigma dx &= \sum_{I \in \mathcal{F}_k} \frac{|\gamma_I|^2}{|I|} \int_I \beta_k \sigma dx \\ &\leq \sum_{I \in \mathcal{F}_k} \frac{|\gamma_I|^2}{|I|} \cdot \sigma(I)^{2/p'}. \end{aligned}$$

Thus, the left-hand side of (2.4) is less than or equal to

$$\begin{aligned} &C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2} \left(\sum_{I \in \mathcal{F}_k} \frac{|\gamma_I|^2}{|I|} \cdot \sigma(I)^{2/p'} \right)^{p'/2} \\ &\leq C \left(\sum_I \frac{|\gamma_I|^2}{|I|} (\sigma(I, 1)/\sigma(I))^{1+2\varepsilon/p'} \cdot \sigma(I)^{2/p'} \right)^{p'/2}. \end{aligned}$$

Because of Hölder's Inequality, $\sigma(I, 1)/\sigma(I) \leq [\sigma(I, p'/2 + \varepsilon)/\sigma(I)]^{1/(p'/2 + \varepsilon)}$. It is a remarkable and deep fact (RDF) that $(1 + 2\varepsilon/p')/(p'/2 + \varepsilon) = 2/p'$. Thanks to the RDF, the last line in our chain is less than or equal to

$$(2.5) \quad C \left(\sum_I \frac{|\gamma_I|^2}{|I|} \cdot \sigma(I, p'/2 + \varepsilon)^{2/p'} \right)^{p'/2}.$$

Choose ε so that $p'/2 + \varepsilon < \tau$. Since $\sigma(I, \eta)$ is an increasing function of η , the hypothesis of the theorem (2.2) implies that (2.5) is less than or equal to

$$C \left(\sum_I |\gamma_I|^2 \mu_I^{-2/q} \right)^{p'/2} \leq C \left(\sum_I |\gamma_I|^{q'} \mu_I^{1-q'} \right)^{p'/q'},$$

because $q' \leq 2$. This finishes the proof of Subcase 2 and the theorem. $\square$

The next theorem we shall prove gives a sufficient condition for the inequality

$$(2.6) \quad \left(\int_{\mathbb{R}^{d+1}_+} |\nabla u|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , where $u(x, y) = \varphi_y * f(x)$ and φ is a *smooth, compactly supported kernel*. By ∇u we mean the full gradient:

$$\nabla u = \left(\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_d} \right).$$

Let us take a look at one of the components of ∇u . If it is $\partial u / \partial x_i$, it is clearly equal to $y^{-1} \psi_y * f(x)$, where ψ is a smooth, compactly supported function that satisfies $\int \psi = 0$. If it is $\partial u / \partial y$, it is equal to $y^{-1} \psi_y * f(x)$, where $\psi(x) = -d\varphi(x) - \vec{x} \cdot \nabla_x \varphi(x)$. This function is smooth and compactly supported, and integration by parts shows that it has integral 0, too.

The components of ∇u are analogous to "continuous" Haar coefficients. The proof we shall give for this compactly supported case will not use the preceding Haar function theorem, but it will use results from weighted Littlewood-Paley theory (in d dimensions) which had their origin in the study of dyadic martingales (i.e., Haar expansions). In order to state our theorem, we must generalize our previous definition to d dimensions.

Definition. For $Q \subset \mathbb{R}^d$ a cube (not necessarily dyadic), σ a weight defined on $\mathbb{R}^d$, and η a positive number, set

$$\sigma(Q, \eta) = \int_Q \sigma(x) \log^\eta(e + \sigma(x)/\sigma_Q) dx,$$

where $\sigma_Q = |Q|^{-1} \int_Q \sigma$, σ 's average over Q .

Theorem 2. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ satisfy $\text{supp } \psi \subset \{|x| \leq 1\}$ and $\int \psi = 0$. Let $1 < p \leq q < \infty$ and $q \geq 2$. Let $\tau > p'/2$. In order that

$$(2.7) \quad \left(\int_{\mathbb{R}_+^{d+1}} |y^{-1} \psi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

should hold for all f , it is sufficient that

$$(2.8) \quad \frac{\mu(T(Q))^{1/q} \cdot \sigma(\tilde{Q}, \tau)^{1/p'}}{|\hat{Q}|} \leq C,$$

for all cubes Q , where the constant C depends on ψ , p , τ , and q .

Proof. We shall first re-cast (2.7) in a dual form. Let $g : \mathbb{R}_+^{d+1} \mapsto \mathbb{R}$ be a bounded, Borel measurable, and compactly supported function; in particular, we assume that $g(x, y) = 0$ for small y , independently of x . We define

$$Tg(x) = \int_{\mathbb{R}_+^{d+1}} g(t, y) \cdot y^{-1} \psi_y(t - x) d\mu(t, y).$$

Inequality (2.7) holds if

$$(2.9) \quad \left(\int_{\mathbb{R}^d} |Tg|^{p'} \sigma dx \right)^{1/p'} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(t, y) \right)^{1/q'}$$

for all such g .

Now we apply Littlewood-Paley theory. Note that $\mathbb{R}_+^{d+1} = \bigcup T(Q)$, and the $T(Q)$'s are pairwise disjoint. For each dyadic cube $Q \subset \mathbb{R}^d$, define

$$\beta_{(Q)}(x) = \int_{T(Q)} g(t, y) \cdot y^{-1} \psi_y(t - x) d\mu(t, y).$$

The function $\beta_{(Q)}$

- (a) has some smoothness,
- (b) satisfies $\int \beta_{(Q)} = 0$, and
- (c) is supported on $\tilde{Q}$, the triple of Q .

Indeed, we can write $\beta_{(Q)} = \lambda_Q \cdot b_{(Q)}(x)$, where $b_{(Q)}$ has (b) and (c) and is normalized so that

(a') $\|\nabla b_{(Q)}\|_\infty \leq \ell(Q)^{-1}|Q|^{-1/2}$ holds;
and λ_Q is a real number which satisfies

$$(2.10) \quad |\lambda_Q| \leq C \int_{T(Q)} |g(t, y)| \frac{d\mu(t, y)}{y^{(d+2)/2}} \\ \leq C \left(\int_{T(Q)} |g(t, y)|^{q'} d\mu(t, y) \right)^{1/q'} \cdot \mu(T(Q))^{1/q} \cdot \ell(Q)^{-1/2} |\hat{Q}|^{-1/2}.$$

(Note that the constant C depends on ψ .) With these definitions, and assuming (2.10), we write

$$(2.11) \quad Tg(x) = \sum_Q \beta_{(Q)}(x) \\ = \sum_Q \lambda_Q \cdot b_{(Q)}(x).$$

The right-hand side of (2.11) is a finite sum, independent of x , of constants times $b_{(Q)}(x)$'s, where the $b_{(Q)}$'s are supported on the triples of dyadic cubes and satisfy (a'), (b), and (c). Thus, Tg is a function in standard form (see (1.8)-(1.9d)).

The following result ([W3], Lemma 2.3) is the analogue of our Result 2 for the Haar function case.

Result 3. *Let r , A , and η be positive numbers, and let $\mathcal{F}$ be a finite family of triples of dyadic cubes Q such that $\sigma(\tilde{Q}, \eta)/\sigma(\tilde{Q}) \leq A$ for all $\tilde{Q} \in \mathcal{F}$. If $h = \sum_{\tilde{Q} \in \mathcal{F}} \lambda_Q b_{(Q)}(x)$ is in standard form, then*

$$\int |h|^r \sigma dx \leq C A^{r/2\eta} \int \tilde{S}^r(h)(x) \sigma dx,$$

where $\tilde{S}(h)(x) = (\sum_{\tilde{Q} \in \mathcal{F}} |\lambda_Q|^2 / |Q| \cdot \chi_{\tilde{Q}}(x))^{1/2}$, and the constant C depends on r , η , and d .

Remark. Result 3 follows fairly easily from the corresponding Haar function (martingale) inequality and this (essentially combinatorial) fact: *The collection $\mathcal{G}$ of all the triples of dyadic cubes can be decomposed into 3^d disjoint subfamilies $\mathcal{G}_k$, and the cubes in each of these subfamilies have the same inclusion/exclusion properties as the dyadic cubes; i.e., if Q and Q' belong to the same $\mathcal{G}_k$, then they are either disjoint, or one is contained in the other; in addition, if $Q \subset Q'$ and $Q \neq Q'$, then $\ell(Q) \leq (\frac{1}{2})\ell(Q')$. (See [W3], Lemmas 2.1 and 2.2.)*

As in the Haar function case, we will take $r = p'$, with $\eta = \tau$ when $p' < 2$ and $\eta = 1$ otherwise. As in the Haar case, these correspond to two subcases.

▷ SUBCASE 1': $2 < p < \infty$. Setting $\eta = \tau$, for each non-negative integer k , we let $\mathcal{F}_k$ be the family of cubes $\tilde{Q}$ which are triples of dyadic cubes and for which $2^k \leq \sigma(\tilde{Q}, \eta)/\sigma(\tilde{Q}) < 2^{k+1}$. Set $h_k(x) = \sum_{\tilde{Q} \in \mathcal{F}_k} \lambda_Q \cdot b_{(Q)}(x)$. Note that $Tg = \sum_k h_k$. Let $0 < \varepsilon < 1 - p'/2\tau$. Arguing as in the Haar function case:

$$\begin{aligned} \int |Tg|^{p'} \sigma dx &\leq C_\varepsilon \sum_k 2^{\varepsilon k} \int |h_k|^{p'} \sigma dx \\ &\leq C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2\tau} \int \tilde{S}^{p'}(h_k) \sigma dx \\ &\leq C \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2\tau} \sum_{Q \in \mathcal{F}_k} \frac{|\lambda_Q|^{p'}}{|Q|^{p'/2}} \cdot \sigma(\tilde{Q}), \end{aligned}$$

where the last line follows from $p'/2 < 1$. Arguing as in the Haar function case (and using the fact that $\varepsilon < 1 - p'/2\tau$, this last line is seen to be less than or equal to

$$C \sum_Q \frac{|\lambda_Q|^{p'}}{|Q|^{p'/2}} \sigma(\tilde{Q}, \tau).$$

The hypothesis (2.8) implies that this is less than or equal to

$$C \sum_Q \frac{|\lambda_Q|^{p'}}{|Q|^{p'/2}} \mu(T(Q))^{-p'/q} |\hat{Q}|^{p'}.$$

Recalling that $|\hat{Q}| = \ell(Q)^{d+1}$, $|Q| = \ell(Q)^d$, and our bound (2.10) on λ_Q , we see that this quantity is less than or equal to

$$C \sum_Q \left(\int_{T(Q)} |g|^{q'} d\mu(x, y) \right)^{p'/q'} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(x, y) \right)^{p'/q'},$$

because $p'/q' \geq 1$. This proves Subcase 1'.

▷ SUBCASE 2': $1 < p \leq 2$. Take $\eta = 1$ and define $\mathcal{F}_k$ and h_k as before. Let $0 < \varepsilon < \tau - p'/2$. Arguing as in Subcase 2:

$$\int |Tg|^{p'} \sigma dx \leq C_\varepsilon \sum_k 2^{\varepsilon k} \cdot 2^{kp'/2} \int \tilde{S}^{p'}(h_k) \sigma dx.$$

If we choose optimal functions β_k , with $\|\beta_k\|_{L^{(p'/2)'}(\sigma)} = 1$, exactly as we did for Subcase 2, we see that the preceding line is bounded by

$$(2.12) \quad C \sum_k \left(\sum_{\tilde{Q} \in \mathcal{F}_k} \frac{|\lambda_Q|^2}{|Q|} \cdot \left(\frac{\sigma(\tilde{Q}, 1)}{\sigma(\tilde{Q})} \right)^{(2/p')(p'/2+\varepsilon)} \cdot \sigma(\tilde{Q})^{2/p'} \right)^{p'/2} \\ \leq C \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \cdot \left(\frac{\sigma(\tilde{Q}, 1)}{\sigma(\tilde{Q})} \right)^{(2/p')(p'/2+\varepsilon)} \cdot \sigma(\tilde{Q})^{2/p'} \right)^{p'/2},$$

since $p'/2 \geq 1$. By Hölder's inequality,

$$\frac{\sigma(\tilde{Q}, 1)}{\sigma(\tilde{Q})} \leq \left[\frac{\sigma(\tilde{Q}, p'/2 + \varepsilon)}{\sigma(\tilde{Q})} \right]^{1/(p'/2 + \varepsilon)}.$$

But, since $p'/2 + \varepsilon < \tau$, the right-hand side of (2.12) is bounded by

$$C \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \cdot \sigma(\tilde{Q}, \tau)^{2/p'} \right)^{p'/2}.$$

The hypothesis of the theorem implies that this is less than or equal to

$$C \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \mu(T(Q))^{-2/q} \cdot |\hat{Q}|^2 \right)^{p'/2}.$$

Because of our estimate on $|\lambda_Q|$ and because $q' \leq 2$, this is bounded by

$$C \left(\sum_Q \left[\int_{T(Q)} |g|^{q'} d\mu(x, y) \right]^{2/q'} \right)^{p'/2} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(x, y) \right)^{p'/q'}.$$

Theorem 2 is proved. □

Addendum. The methods and results just given have immediate consequences for wavelet decompositions. Let's suppose that we're working in $\mathbb{R}^1$. Let ψ be a C^1 function supported in $[-1, 2]$ and satisfying $\int \psi = 0$. If $I \subset \mathbb{R}^1$ is a dyadic interval with length 2^k and left endpoint x^* , we set $\psi_{(I)}(x) = 2^{-k/2} \psi(2^{-k}(x - x^*))$. If ψ is chosen cleverly, the family $\mathcal{G} = \{\psi_{(I)}\}$ will be an orthonormal basis for $L^2(\mathbb{R})$. (However, the result we are about to state will not depend on such a clever choice.)

For $f \in L^2(\mathbb{R})$, set $\lambda_I = \langle f, \psi_{(I)} \rangle$. This coefficient measures, roughly, how much of f has frequency $\sim |I|^{-1}$ near the interval I . If we are interested in how a weighted average of these coefficients is affected by the raw size of f ($= |f|$), we will want to consider inequalities like

$$(2.13) \quad \left(\sum_I |\lambda_I|^q \mu_I \right)^{1/q} \leq \left(\int_{\mathbb{R}} |f(x)|^p v(x) dx \right)^{1/p},$$

and know for what non-negative $v \in L^1_{\text{loc}}(\mathbb{R}^d)$ and sequences of weights $\{\mu_I\}_I$ they hold. But these are precisely what we have been handling in this section. If $\mathcal{F} \subset \mathcal{G}$ is finite, then a function $g = \sum_{I \in \mathcal{F}} \gamma_I \psi_{(I)}$ is in standard form. The methods of this section immediately imply:

Wavelet Theorem. *Let $\mathcal{G}$ be as above. Let $\{\mu_I\}_I$ be a sequence of non-negative numbers and let v be a non-negative function in $L^1_{\text{loc}}(\mathbb{R}^d)$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. If $1 < p \leq q < \infty$, $q \geq 2$, and $\tau > p'/2$, then (2.13) will hold for all $f \in L^2(\mathbb{R})$ if*

$$\frac{\mu_I^{1/q} \cdot \sigma(\tilde{I}, \tau)^{1/p'}}{|I|^{1/2}} \leq C,$$

for all dyadic intervals I , where the constant C depends on p, q, τ , and ψ .

We leave the details to the interested reader.

3. When the kernels do not have compact support. In this section we prove a result which is in some sense intermediate between the “compactly supported” Theorem 2 and the “Poisson kernel” Theorem 5 of the next section. In particular, our theorem applies to the x -derivatives of the Poisson kernel. We believe that it is also of independent interest: it includes the case in which $u(x, y) = f * \Phi_y(x)$ and Φ is a Schwartz function.

We prove our result in two stages. Theorem 3 is a general statement involving special pairs of weights. Its proof is essentially an application of Schur’s test for positive operators, plus interpolation. It does not make use of any results from weighted Littlewood-Paley theory. Theorem 4 asserts that certain tractable pairs of weights do indeed satisfy the hypothesis of Theorem 3. The actual proof of Theorem 4 will not involve Littlewood-Paley estimates, but it will use a key result from [W3].

Definition. Let σ and w be a pair of non-negative functions in $L^1_{\text{loc}}(\mathbb{R}^d)$, and let $1 < r < \infty$. We say that (σ, w) is an r -pair if $\sigma \leq w$ almost everywhere and if

$$\int |h|^r \sigma dx \leq \int \tilde{S}^r(h) w dx$$

for all functions h in standard form.

If σ belongs to the Muckenhoupt A_∞ class, then $(\sigma, C\sigma)$ is an r -pair for a sufficiently large constant C (depending on σ and r).

Theorem 3. *Let $1 < p \leq q < \infty$ and $q \geq 2$. Let $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ have $\int \varphi = 0$ and, for some non-negative integer m , satisfy $|\varphi(x)| \leq (1 + |x|)^{-d-2-m}$ and $|\nabla \varphi(x)| \leq (1 + |x|)^{-d-3-m}$ for all $x \in \mathbb{R}^d$. In order for*

$$\left(\int_{\mathbb{R}_+^{d+1}} |y^{-1} \varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , it is sufficient that there exist a p' -pair (σ, w) such that

$$(3.1) \quad \mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2+m)p'/q'}} dx \right)^{1/p'} \\ \leq C \ell(Q)^{d+1-(d+2+m)/q'}$$

hold for all cubes Q . The constant C in (3.1) depends on p, q, d , and m .

Proof. We require the following lemma from [U]:

Lemma 3.1. *Let m be a non-negative integer. Let $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ satisfy*

$$|\varphi(x)| + |\nabla \varphi(x)| \leq (1 + |x|)^{-d-2-m}$$

for all x . Moreover, suppose that $\int \varphi(x) P(x) dx = 0$ for all polynomials P of degree $\leq m+1$. Then there exists a sequence of $\mathcal{C}^\infty$ bounded functions $\{v_i\}_0^\infty$ such that

$$\varphi(x) = \sum_i 2^{-i(m+2)} (v_i)_{2^i}(x), \quad \text{supp } v_i \subset \{|x| \leq 1\}, \quad \|v_i\|_\infty \leq C, \quad \text{and}$$

$$\int v_i(x) P(x) dx = 0,$$

for all polynomials P of degree $\leq m+1$. In addition, there is a positive c so that, for each $i > 0$, $(v_i)_{2^i}(x) = 0$ for $|x| \leq c2^i$.

Remark. Uchiyama actually proves this lemma under the stronger hypothesis that $0 \notin \text{supp } \hat{\varphi}$, but his argument only requires that φ have zero moments to sufficiently high order. The ‘in addition’ is not in the original statement of the lemma, but follows from the construction in its proof. Ditto the smoothness.

We wish to apply Uchiyama's lemma to *our* kernel φ . Now, our φ has enough decay, but it need not have the vanishing moments. Let $\varphi_1 \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ have support contained in $\{|x| \leq 1\}$ and be such that $\int \varphi_1(x)P(x)dx = \int \varphi(x)P(x)dx$ for all polynomials of degree $\leq m+1$. Note that the integrals on the right are defined, because of φ 's decay, and that $\int \varphi_1 = 0$. Write $\varphi = \varphi_1 + (\varphi - \varphi_1) = \varphi_1 + \varphi_2$. We want a sufficient condition for the inequality

$$(3.2) \quad \left(\int_{\mathbb{R}_+^{d+1}} |y^{-1}\varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f . For this, it is enough to have a condition that ensures that (3.2) holds when φ is replaced by φ_1 or φ_2 . The first function, φ_1 , has only one moment of cancellation, but is compactly supported. The second function, φ_2 , is not compactly supported, but has lots of cancellation. The cancellation makes it tractable—via the Uchiyama lemma.

Let us quickly dispose of φ_1 . Define

$$Tg(x) = \int_{\mathbb{R}_+^{d+1}} g(t, y) \cdot y^{-1}(\varphi_1)_y(t - x) d\mu(t, y)$$

for all compactly supported $g : \mathbb{R}_+^{d+1} \mapsto \mathbb{R}$ as we did in the previous section. Inequality (3.2) will hold for φ_1 if

$$\left(\int_{\mathbb{R}^d} |Tg|^{p'} \sigma dx \right)^{1/p'} \leq \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(x, y) \right)^{1/q'}$$

holds for all such g . Therefore, under the hypothesis that (σ, w) is a p' -pair, inequality (3.2) will hold for φ_1 if

$$\left(\int_{\mathbb{R}^d} \left(\sum_{x \in \tilde{Q}} \frac{|\lambda_Q|^2}{|Q|} \right)^{p'/2} w dx \right)^{1/p'} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(x, y) \right)^{1/q'}$$

for all appropriate g , where

$$\lambda_Q = \ell(Q)^{-1}|Q|^{-1/2} \cdot \left(\int_{T(Q)} |g|^{q'} d\mu(x, y) \right)^{1/q'} \mu(T(Q))^{1/q}$$

for all dyadic cubes Q . Let us take $p' \geq 2$ and let ρ be the conjugate exponent to $p'/2$. There exists a function $h \in L^\rho(w dx)$ such that $\|h\|_{L^\rho(w dx)} = 1$ and

$$\begin{aligned} \left(\int_{\mathbb{R}^d} \left(\sum_{x \in \tilde{Q}} \frac{|\lambda_Q|^2}{|Q|} \right)^{p'/2} w dx \right)^{1/p'} &= \left(\int_{\mathbb{R}^d} \left(\sum_{x \in \tilde{Q}} \frac{|\lambda_Q|^2}{|Q|} \right) h w dx \right)^{1/2} \\ &= \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \int_{\tilde{Q}} h w dx \right)^{1/2} \\ &\leq \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} w(\tilde{Q})^{2/p'} \right)^{1/2}, \end{aligned}$$

where the last line follows from Hölder's inequality and the fact that $\|h\|_{L^\rho(w dx)} = 1$. This is precisely the kind of sum we obtained in the preceding section, when we considered the “compactly supported” case. The simple-minded argument (essentially, term-by-term comparison) used there shows that (3.2) will hold, with φ_1 in place of φ , if

$$(3.3) \quad \frac{\mu(T(Q))^{1/q} \cdot w(\tilde{Q})^{1/p'}}{|\hat{Q}|} \leq C$$

for all cubes Q . But it is trivial that (3.1) implies (3.3). This takes care of φ_1 when $p' \geq 2$. The case when $1 < p' < 2$ is similar (and easier), and we leave it to the interested reader.

For simplicity of notation, let us set $\varphi = \varphi_2$, so that now we can assume that φ satisfies the hypotheses of the Uchiyama lemma and, in addition, that $\nabla \varphi$ decays like $(1 + |x|)^{-d-m-3}$.

Let $g : \mathbb{R}_+^{d+1} \mapsto \mathbb{R}$ be as in Section 2, and define

$$Tg(x) = \int_{\mathbb{R}_+^{d+1}} g(t, y) \cdot y^{-1} \varphi_y(t - x) d\mu(t, y)$$

for $x \in \mathbb{R}^d$. By duality, inequality (3.2) will hold if

$$(3.4) \quad \left(\int_{\mathbb{R}^d} |Tg|^{p'} \sigma dx \right)^{1/p'} \leq \left(\int_{\mathbb{R}_+^{d+1}} |g|^{q'} d\mu(x, y) \right)^{1/q'}$$

is true for all such g . It is useful to point out here that inequality (3.1) implies that the left-hand side of (3.4) is finite for all such g . This is because

Tg is a bounded (and, indeed, smooth) function that decays at infinity like $(1 + |x|)^{-d-m-2}$. Inequality (3.1) implies that such a function belongs to $L^{p'}(\sigma)$. (Recall that $\sigma \leq w$ a.e.) More is true: the function Tg can be nicely approximated by functions h in standard form. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ be real and radial, satisfy $\int \psi = 0$, have support contained in $\{|x| \leq 1\}$, and be normalized so that

$$\int_0^\infty \left| \hat{\psi}(t\xi) \right|^2 \frac{dt}{t} = 1$$

for all $\xi \in \mathbb{R}^d \setminus \{0\}$. If $f \in L^2(\mathbb{R}^d)$ (as Tg certainly is), then

$$(3.5) \quad f(x) = \lim_{R \rightarrow \infty} \int_{\substack{|t| \leq R \\ R^{-1} < y < R}} [f * \psi_y(t)] \cdot \psi_y(x - t) \frac{dt dy}{y}$$

in L^2 . (This is the famous *Calderón-Torchinsky reproducing formula*.) Thus, there is a sequence $R_j \rightarrow \infty$ such that the right-hand side of (3.5) goes to f almost everywhere. For fixed R , the integral on the right side of (3.5) can be expressed as a function in standard form, accomplished by cutting up $\mathbb{R}_+^{d+1}$ into $T(Q)$'s. If we do that, then the corresponding square function $\tilde{S}(f)$ is seen to be less than or equal to $S_c(f) = (\sum_{x \in \tilde{Q}} |\lambda_Q|^2 / |Q|)^{1/2}$, where

$$|\lambda_Q| \leq C \left(\int_{T(Q)} |f * \psi_y(t)|^2 \frac{dt dy}{y} \right)^{1/2},$$

and, most important, all of these estimates hold *independently* of R . If (σ, w) is an r -pair, then Fatou's Lemma now implies that

$$\int |f|^r \sigma dx \leq C \int S_c^r(f) w dx$$

for all such f , and, in particular, for Tg .

Thus, the key to proving Theorem 3 is to obtain good estimates for $Tg * \psi_y(t)$, where ψ is as in the preceding paragraph. Henceforth, the function ψ , having the properties listed above, will be assumed to be *fixed*.

In order to estimate $Tg * \psi_y(t)$, we must first estimate $\varphi_\eta * \psi_y(x)$, for arbitrary positive numbers η and y . Since this result (actually a collection of results) is rather technical, we shall approach it gradually. First we need an easy lemma.

Lemma 3.2. *Let $\tilde{\varphi} \in C_0^\infty(\mathbb{R}^d)$ have support contained in $\{|x| \leq 1\}$ and satisfy $\int \tilde{\varphi}(x)P(x)dx = 0$ for all polynomials of degree $\leq m+1$. If $y \neq 0$ then $|\tilde{\varphi} * \psi_y(x)| \leq Cy^{-d-m-2}$, where C depends on φ and ψ but not on x .*

Proof. We wish to estimate

$$\int_{|t| \leq 1} \tilde{\varphi}(t) y^{-d} \psi\left(\frac{x-t}{y}\right) dt.$$

Because of the moment condition on $\tilde{\varphi}$, it equals

$$\int_{|t| \leq 1} \tilde{\varphi}(t) y^{-d} \left[\psi\left(\frac{x-t}{y}\right) - P\left(\frac{-t}{y}\right) \right] dt,$$

where P is ψ 's Taylor polynomial of degree $m+1$, expanded around x/y . The quantity in the square brackets has absolute value less than or equal to $C|t/y|^{m+2} \leq Cy^{-m-2}$, where C depends on ψ but is independent of x . This proves the lemma. $\square$

Before continuing, we should recall the precise properties of ψ and φ that will be relevant in the following estimates. The function ψ belongs to $\mathcal{C}_0^\infty(\mathbb{R}^d)$, is real and radial, satisfies $\int \psi = 0$, and has support contained in $\{|x| \leq 1\}$. The function φ belongs to $\mathcal{C}^\infty(\mathbb{R}^d)$. For some non-negative integer m , φ satisfies $|\varphi(x)| \leq (1+|x|)^{-d-m-2}$ and $|\nabla \varphi(x)| \leq (1+|x|)^{-d-m-3}$ for all $x \in \mathbb{R}^d$. In addition, $\int \varphi(x) P(x) dx = 0$ for all polynomials of degree $\leq m+1$. We have applied the Uchiyama lemma to write φ as a rapidly converging sum of smooth, compactly supported functions, and these functions also annihilate all polynomials of degree $\leq m+1$.

▷ EASY CASE: $\eta = 1$, $y \leq 1$. Because of φ 's smoothness and decay, and ψ 's cancellation,

$$\begin{aligned} |\varphi * \psi_y(x)| &= \left| \int_{|t| \leq y} \varphi(x-t) \psi_y(t) dt \right| \\ &= \left| \int_{|t| \leq y} [\varphi(x-t) - \varphi(x)] \psi_y(t) dt \right| \\ &\leq C \frac{y}{(1+|x|)^{d+m+3}}. \end{aligned}$$

▷ LESS EASY CASE: $\eta = 1$, $y \geq 1$, AND $|x| \geq 5y$. Note that now $|t| \leq |x|/2$. We get the same estimate by considering a similar integral.

▷ TRICKY CASE: $\eta = 1$, $y \geq 1$, $|x| \leq 5y$. Write:

$$(3.6) \quad \psi_y * \varphi(x) = \sum_0^\infty 2^{-i(m+2)} \psi_y * (v_i)_{2^i}(x),$$

where the v_i are as in the Uchiyama lemma. The function ψ_y is supported in $\{|x| \leq y\}$ and each $(v_i)_{2^i}$ is supported in $\{|x| \sim 2^i\}$ when $i > 0$.

Thus, since $|x| \leq 5y$, the convolutions on the right-hand side of (3.6) will be zero unless $2^i \leq Cy$. Now Lemma 3.2 implies that, for such i ,

$$\begin{aligned} |\psi_y * (v_i)_{2^i}(x)| &= 2^{-id} |\psi_{y2^{-i}} * v_i(2^{-i}x)| \\ &\leq C2^{-id} (y2^{-i})^{-d-m-2} \\ &= C2^{-id} (2^i/y)^{d+m+2}, \end{aligned}$$

where we used our Lemma 3.2 to get the second inequality. If we multiply this by $2^{-i(m+2)}$ and sum for $2^i \leq Cy$, the result is less than or equal to $C \log(e+y)/y^{d+m+2}$.

▷ GENERAL CASES:

1. If $\eta \geq y$, the easy case plus a change of variable yields

$$\begin{aligned} |\psi_y * \varphi_\eta(x)| &= \left| \eta^{-d} \left(\psi_{y/\eta} * \varphi \left(\frac{x}{\eta} \right) \right) \right| \\ &\leq \frac{C\eta^{-d}(y/\eta)}{(1+|x/\eta|)^{d+m+3}} \\ &= \frac{Cy\eta^{m+2}}{(\eta+|x|)^{d+m+3}}. \end{aligned}$$

2. If $\eta \leq y$ and $|x| \geq 5y$, we get the same estimate.

3. If $\eta \leq y$ and $|x| \leq 5y$, the tricky case plus a change of variable yields

$$\begin{aligned} |\psi_y * \varphi_\eta(x)| &= \left| \eta^{-d} \left(\psi_{y/\eta} * \varphi \left(\frac{x}{\eta} \right) \right) \right| \\ &\leq C\eta^{-d} \log \left(\frac{e+y}{\eta} \right) \left(\frac{\eta}{y} \right)^{d+m+2} \\ &= \frac{C\eta^{m+2} \log(e+y/\eta)}{y^{d+m+2}}. \end{aligned}$$

Before continuing, we observe that these inequalities are in a sense stable: if we replace $\eta \leq y$ by $\eta \leq Cy$, for instance, the forms of the estimates are unaffected.

Now we can estimate $\psi_y * Tg(x)$. Let us fix a dyadic cube Q' and suppose that $(x, y) \in T(Q')$. Using the preceding estimates and the fact that ψ is even,

$$\begin{aligned}
 |\psi_y * Tg(x)| &\leq \int_{\mathbb{R}_+^{d+1}} |g(t, \eta)| \eta^{-1} |\psi_y * \varphi_\eta(t - x)| d\mu(t, \eta) \\
 &= \sum_Q \int_{T(Q)} |g(t, \eta)| \eta^{-1} |\psi_y * \varphi_\eta(t - x)| d\mu(t, \eta) \\
 &= \sum_{Q \subset \tilde{Q}'} \int_{T(Q)} |g(t, \eta)| \eta^{-1} |\psi_y * \varphi_\eta(t - x)| d\mu(t, \eta) \\
 &\quad + \sum_{Q \not\subset \tilde{Q}'} \int_{T(Q)} |g(t, \eta)| \eta^{-1} |\psi_y * \varphi_\eta(t - x)| d\mu(t, \eta) \\
 &\leq C \sum_{Q \subset \tilde{Q}'} \left[\int_{T(Q)} |g(t, \eta)| d\mu(t, \eta) \right] \cdot \left(\frac{\ell(Q)^{m+1}}{\ell(Q')^{d+m+2}} \right) \cdot \log \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \\
 &\quad + C \sum_{Q \not\subset \tilde{Q}'} \left[\int_{T(Q)} |g(t, \eta)| d\mu(t, \eta) \right] \cdot \frac{\ell(Q') \ell(Q)^{m+1}}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+m+3}} \\
 &= (I) + (II),
 \end{aligned}$$

where, as usual, x_Q denotes the center of Q . We have obtained the summands in (I) using our estimate for the general tricky case above. We have used the easy and less easy cases to obtain (II). In the sums, $\ell(Q)$ plays the role of η and $\ell(Q')$ plays the role of y . The reason that we have $\ell(Q)^{m+1}$, instead of $\ell(Q)^{m+2}$, is that one power of η gets cancelled by an η^{-1} in the integral that defines Tg .

We can now express our problem in terms of sequence spaces. For dyadic cubes Q and Q' , set:

$$\begin{aligned}
 \Lambda(Q) &= \left(\int_{T(Q)} |\psi_y * Tg(x)|^2 \frac{dx dy}{y} \right)^{1/2}, \\
 G(Q) &= \int_{T(Q)} |g(t, \eta)| d\mu(t, \eta), \\
 a_1(Q', Q) &= \left(\frac{\ell(Q)^{m+1}}{\ell(Q')^{d+m+2}} \right) \log \left(e + \frac{\ell(Q')}{\ell(Q)} \right),
 \end{aligned}$$

$$\begin{aligned}
 a_2(Q', Q) &= \frac{\ell(Q')\ell(Q)^{m+1}}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+m+3}}, \\
 a(Q', Q) &= \begin{cases} a_1(Q', Q) & \text{if } Q \subset \tilde{Q}', \\ a_2(Q', Q) & \text{if } Q \not\subset \tilde{Q}', \end{cases} \\
 \rho(Q) &= \mu(T(Q))^{1-q'}.
 \end{aligned}$$

We have proved that

$$\begin{aligned}
 \Lambda(Q') &\leq C[(I) + (II)]|Q'|^{1/2} \\
 &\leq C|Q'|^{1/2} \left[\sum_{Q \subset \tilde{Q}'} a_1(Q', Q)G(Q) + \sum_{Q \not\subset \tilde{Q}'} a_2(Q', Q)G(Q) \right].
 \end{aligned}$$

The square function $\tilde{S}(Tg)$ is dominated by $(\sum_{\tilde{Q} \ni x} |\Lambda(Q)|^2/|Q|)^{1/2}$. And, by Hölder's inequality,

$$(3.7) \quad \left(\sum_Q G(Q)^{q'} \rho(Q) \right)^{1/q'} \leq \left(\int_{\mathbb{R}^{d+1}_+} |g|^{q'} d\mu(x, y) \right)^{1/q'}.$$

Our goal, given the assumption that (σ, w) is a p' -pair, is to show that hypothesis (3.1) implies that

$$\left(\int_{\mathbb{R}^d} [\tilde{S}(Tg)]^{p'} w dx \right)^{1/p'}$$

is bounded by a constant times the left-hand side of (3.7). As in the preceding section, we need to consider two subcases. For each of these, we will have to show that the appropriate bound holds for both $a_1(Q', Q)$ and $a_2(Q', Q)$.

▷ SUBCASE 1: $1 < p \leq 2$. The goal is to show that

$$\begin{aligned}
 (3.8) \quad &\left(\int \left[\sum_{x \in \tilde{Q}'} \left(\sum_Q a(Q', Q)G(Q) \right)^2 \right]^{p'/2} w dx \right)^{1/p'} \\
 &\leq C \left(\sum_Q G(Q)^{q'} \rho(Q) \right)^{1/q'}.
 \end{aligned}$$

To this end, we let $h \in L^{(p'/2)'}(w dx)$ be non-negative, satisfy $\|h\|_{L^{(p'/2)'}(w dx)} \leq 1$ and be chosen so that

$$\begin{aligned} & \left(\int \left[\sum_{x \in \tilde{Q}'} \left(\sum_Q a(Q', Q) G(Q) \right)^2 \right] h w dx \right)^{1/2} \\ &= \left(\int \left[\sum_{x \in \tilde{Q}'} \left(\sum_Q a(Q', Q) G(Q) \right)^2 \right]^{p'/2} w dx \right)^{1/p'}. \end{aligned}$$

Having done this, our problem is to show that

$$(3.9) \quad \left(\int \left[\sum_{x \in \tilde{Q}'} \left(\sum_Q a(Q', Q) G(Q) \right)^2 \right] h w dx \right)^{1/2} \leq C \left(\sum_Q G(Q)^{q'} \rho(Q) \right)^{1/q'}$$

for all finite positive sequences $\{G(Q)\}$ and all h as above. Define $\nu(Q') = \int_{\tilde{Q}'} h w dx$. Then the left-hand side of (3.9) is

$$\left(\sum_{Q'} \left(\sum_Q a(Q', Q) G(Q) \right)^2 \nu(Q') \right)^{1/2},$$

and thus our problem has been reduced to one about sequence spaces. We need to show that the positive “kernel” $a(Q', Q)$ maps boundedly from $\ell^{q'}(\rho(Q))$ into $\ell^2(\nu(Q'))$. We will do this via the Riesz-Thorin Interpolation Theorem, and in order to facilitate this we will redefine our problem as follows. Set $G(Q) = Y(Q) \left| \hat{Q} \right|$, and let $\{Y(Q)\}$ be the sequence we operate on. Then we must replace $\rho(Q)$ with $\left| \hat{Q} \right|^{q'} \rho(Q)$ and $a(Q', Q)$ with $A(Q', Q) = \left| \hat{Q} \right| a(Q', Q)$. We will show that the kernel $A(Q', Q)$ maps boundedly between the two ℓ^∞ spaces, and also between $\ell^1(\left| \hat{Q} \right|^{q'} \rho(Q))$ and $\ell^{2/q'}(\nu(Q'))$. (Recall that $q' \leq 2$.) The desired result then follows by interpolation.

$\ell^\infty \mapsto \ell^\infty$. We show the appropriate bounds for $A_1 = a_1 \cdot \left| \hat{Q} \right|$ and $A_2 = a_2 \cdot \left| \hat{Q} \right|$. The A_1 term is easier so we'll do it first.

The ℓ^∞ boundedness is equivalent to having $\sum_{Q: Q \subset \tilde{Q}'} a_1(Q', Q) \cdot \left| \hat{Q} \right| \leq C$ for all Q' . But this is trivial, since

$$\begin{aligned} \sum_{\substack{Q \\ Q \subset \tilde{Q}'}} a_1(Q', Q) \cdot \left| \hat{Q} \right| &\leq C \sum_{\substack{Q \\ Q \subset \tilde{Q}'}} \left(\frac{\ell(Q)}{\ell(Q')} \right)^{d+m+2} \log \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \\ &\leq C_\delta \sum_{\substack{Q \\ Q \subset \tilde{Q}'}} \left[\frac{|Q|}{|Q'|} \right]^{1+\delta} \end{aligned}$$

for some $\delta > 0$, because $d + m + 2 > d$. We claim that this last sum has a bound independent of Q' . *Proof of claim:* If $Q \subset \tilde{Q}'$, then $|Q|/|Q'| = 2^{kd}$ for some integer $k \leq 2$. For each such k , there are at most $C_d 2^{-kd}$ such cubes. Thus, their contribution to the total is at most $C_d 2^{\delta kd}$, which we may sum from $k = 2$ to $-\infty$. (Note: It was in order to make this part go through that we defined $Y(Q)$ the way we did.)

For A_2 , we need to show that $\sum_{Q: Q \not\subset \tilde{Q}'} a_2(Q', Q) \cdot |\hat{Q}| \leq C$. We divide this sum into two pieces: $\ell(Q) \leq \ell(Q')$ and $\ell(Q) > \ell(Q')$. Take the first piece and consider those Q 's such that $\ell(Q) = 2^{-k}\ell(Q')$, with $k \geq 0$ temporarily fixed. We get:

$$\begin{aligned} \sum_{\substack{Q: Q \not\subset \tilde{Q}' \\ \ell(Q) = 2^{-k}\ell(Q')}} a_2(Q', Q) \cdot |\hat{Q}| &\leq C \int_{x \notin \tilde{Q}'} \frac{\ell(Q')(2^{-k}\ell(Q'))^{m+2}}{(2^{-k}\ell(Q') + |x - x_{Q'}|)^{d+m+3}} dx \\ &\leq C \ell(Q')(2^{-k}\ell(Q'))^{m+2} \ell(Q')^{-m-3} \\ &\leq C 2^{-k(m+2)}, \end{aligned}$$

and these sum (for $k \geq 0$) to something $\leq C$.

Now take $\ell(Q) = 2^k\ell(Q')$. We get

$$\begin{aligned} \sum_{\substack{Q: Q \not\subset \tilde{Q}' \\ \ell(Q) = 2^k\ell(Q')}} a_2(Q', Q) \cdot |\hat{Q}| &\leq C \int_{\mathbb{R}^d} \frac{\ell(Q')(2^k\ell(Q'))^{m+2}}{(2^k\ell(Q') + |x - x_{Q'}|)^{d+m+3}} dx \\ &\leq C 2^{-k}, \end{aligned}$$

and these also sum for $k \geq 0$. The $\ell^\infty \mapsto \ell^\infty$ bound is established.

By Minkowski's inequality,

$$\left(\sum_{Q'} \left[\sum_Q A(Q', Q) Y(Q) \right]^{2/q'} \nu(Q') \right)^{q'/2} \leq \sum_Q \left[\sum_{Q'} A(Q', Q)^{2/q'} \nu(Q') \right]^{q'/2} Y(Q).$$

Therefore, in order to get the $\ell^1 \mapsto \ell^{2/q'}$ bound, it is enough to show that

$$(3.10) \quad \left(\sum_{Q'} A(Q', Q)^{2/q'} \nu(Q') \right)^{q'/2} \leq C |\hat{Q}|^{q'} \rho(Q)$$

for all cubes Q . As before, we will bound the sums corresponding to $A_1(Q', Q)$ and $A_2(Q', Q)$ separately.

▷ THE FAIRLY EASY TERM: $A_1(Q', Q)$. We first recall an elementary fact about infinite series: *If α and β are positive numbers then, for any positive integer j , $\sum_{k \geq j} k^\alpha e^{-k\beta} \leq C_{\alpha, \beta} j^\alpha e^{-j\beta}$; i.e., the series is bounded by the first (and essentially the largest) term. Write:*

$$(3.11) \quad \sum_{Q'} A_1(Q', Q)^{2/q'} \nu(Q') \\ \leq C \sum_{\substack{Q' \\ Q \subset \tilde{Q}'}} \left(\frac{\ell(Q)}{\ell(Q')} \right)^{(d+m+2)2/q'} \log^{2/q'} \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \int_{\tilde{Q}'} h w \, dx,$$

which is less than or equal to

$$(3.12) \quad C \int_{\mathbb{R}^d} \left[\sum_{\substack{Q' \\ Q \subset \tilde{Q}'}} \left(\frac{\ell(Q)}{\ell(Q')} \right)^{(d+m+2)2/q'} \log^{2/q'} \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \chi_{\tilde{Q}'}(x) \right] h(x) w(x) \, dx.$$

As a function of x , the summands in the integral (3.12) will be zero unless $|x - x_Q| \leq c\ell(Q')$. No x belongs to more than $C(d)$ of the cubes $\tilde{Q}'$ that have the same sidelength, and these sidelengths increase geometrically (they equal $3 \cdot 2^k$, for $k \in \mathbb{Z}$). Thus, the infinite series in the square brackets in (3.12) is bounded by

$$C \sum_{\substack{Q': \ell(Q') \geq c|x - x_Q| + c\ell(Q) \\ Q \subset \tilde{Q}'}} \log^{2/q'} \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \cdot \left(\frac{\ell(Q)}{\ell(Q')} \right)^{(d+m+2)2/q'} \\ \leq C \log^{2/q'} \left(e + \frac{|x - x_Q|}{\ell(Q)} \right) \cdot \ell(Q)^{(d+m+2)2/q'} (\ell(Q) + |x - x_Q|)^{(-d-m-2)2/q'}.$$

Therefore, the integral (3.12) is bounded by

$$(3.13) \quad C \ell(Q)^{(d+m+2)2/q'} \int_{\mathbb{R}^d} \frac{\log^{2/q'}(e + |x - x_Q|/\ell(Q)) h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)2/q'}} w(x) \, dx.$$

Since $\|h\|_{L^{(p'/2)'}(w \, dx)} \leq 1$, the last line is less than or equal to

$$C \ell(Q)^{(d+m+2)2/q'} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)p'/q'}} \, dx \right)^{2/p'}.$$

The hypothesis (3.1) of the theorem implies that the last line is less than or equal to

$$C \mu(T(Q))^{-2/q} \ell(Q)^{(d+m+2)2/q'} \ell(Q)^{2(d+1-(d+2+m)/q')} \leq C \rho(Q)^{2/q'} \left| \hat{Q} \right|^2.$$

The last line, when raised to the power $q'/2$, is less than or equal to $C \rho(Q) \left| \hat{Q} \right|^{q'}$, and this is what we wanted to show.

▷ THE SOMEWHAT LESS EASY TERM: $A_2(Q', Q)$. We begin as before:

$$\begin{aligned} \sum_{Q'} A_2(Q', Q)^{2/q'} \nu(Q') &\leq C \sum_{Q': Q \not\subset \tilde{Q}'} \left(\frac{\ell(Q') \ell(Q)^{d+m+2}}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+m+3}} \right)^{2/q'} \nu(Q') \\ &= (i) + (ii), \end{aligned}$$

where (i) is the sum over Q' 's such that $\ell(Q') \leq \ell(Q)$ and (ii) is the sum over Q' 's such that $\ell(Q') > \ell(Q)$.

Handling (i): Fix $k \geq 0$ and look at the sum for Q' 's with $\ell(Q') = 2^{-k} \ell(Q)$. The sum is dominated by

$$\begin{aligned} C 2^{-2k/q'} \ell(Q)^{(d+m+3)2/q'} &\int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+3)2/q'}} w(x) dx \\ &\leq C 2^{-2k/q'} \ell(Q)^{(d+m+2)2/q'} \int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)2/q'}} w(x) dx \\ &\leq C 2^{-2k/q'} \ell(Q)^{(d+m+2)2/q'} \int_{\mathbb{R}^d} \frac{\log^{2/q'}(e + |x - x_Q|/\ell(Q)) h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)2/q'}} w(x) dx, \end{aligned}$$

which, except for the convenient factor $2^{-2k/q'}$, is (3.13). If we sum over k , the result is bounded by (3.13), which we estimated above.

Handling (ii): Momentarily fix $k > 0$; we sum over the Q' 's such that $\ell(Q') = 2^k \ell(Q)$. Observe that, in this case, $Q \not\subset \tilde{Q}'$ forces $|x_Q - x_{Q'}| \geq C \ell(Q')$. Our sum is dominated by

$$2^{2k/q'} \ell(Q)^{(d+m+3)2/q'} \int_{|x - x_Q| > c 2^k \ell(Q)} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+3)2/q'}} w(x) dx.$$

If we now sum over k , and use the fact that $\sum_{k>0} 2^{2k/q'} \chi_{\{|x| \geq 2^k\}}(x) \leq C(1 + |x|)^{2/q'}$, we see that (ii) is dominated by

$$C \ell(Q)^{(d+m+2)2/q'} \int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)2/q'}} w(x) dx,$$

and we've seen this before. Subcase 1 is proved.

▷ SUBCASE 2: $2 < p < \infty$. We keep the notation from Subcase 1. Arguing as in the compactly supported case, it is enough to show that

$$\left(\int_{\mathbb{R}^d} \sum_{x \in \tilde{Q}} \frac{|\Lambda(Q)|^{p'}}{|Q|^{p'/2}} w dx \right)^{1/p'} \leq C \left(\sum_Q G(Q)^{q'} \rho(Q) \right)^{1/q'},$$

for all finite non-negative sequences $\{G(Q)\}$. Using our bound on $\Lambda(Q)$ and the conventions of Subcase 1, we see that this amounts to showing:

$$\left(\sum_{Q'} \left[\sum_Q A(Q', Q) Y(Q) \right]^{p'} w(\tilde{Q}') \right)^{1/p'} \leq C \left(\sum_Q Y(Q)^{q'} \left| \hat{Q} \right|^{q'} \rho(Q) \right)^{1/q'},$$

for all finite non-negative sequences $\{Y(Q)\}$. As in Subcase 1, we establish this by an interpolation. One of our endpoint bounds will be $\ell^\infty \mapsto \ell^\infty$, just like before. The second will be $\ell^1 \left(\left| \hat{Q} \right|^{q'} \rho(Q) \right) \mapsto \ell^{p'/q'}(w(\tilde{Q}'))$.

$\ell^\infty \mapsto \ell^\infty$. Done in Subcase 1—no change here!

$\ell^1 \mapsto \ell^{p'/q'}$. By Minkowski's inequality, it is enough to show that

$$\left(\sum_{Q'} A(Q', Q)^{p'/q'} w(\tilde{Q}') \right)^{q'/p'} \leq C \left| \hat{Q} \right|^{q'} \rho(Q)$$

holds for all Q . We treat the A_1 and A_2 terms separately.

• A_1 :

$$\begin{aligned} & \sum_{Q'} A_1(Q', Q)^{p'/q'} w(\tilde{Q}') \\ & \leq C \sum_{\substack{Q' \\ Q \subset \tilde{Q}'}} \left(\frac{\ell(Q)}{\ell(Q')} \right)^{(d+m+2)p'/q'} \log^{p'/q'} \left(\frac{e + \ell(Q')}{\ell(Q)} \right) w(\tilde{Q}') \\ & \leq C \ell(Q)^{(d+m+2)p'/q'} \int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)p'/q'}} dx, \\ & \leq C \ell(Q)^{(d+m+2)p'/q'} \mu(T(Q))^{-p'/q} \ell(Q)^{p'[d+1] - (d+2+m)p'/q'}, \end{aligned}$$

where the last line follows from (3.1) and the next-to-the-last line follows by the same sort of argument used to get from (3.11) to (3.13). This simplifies to

$$\leq C \left| \hat{Q} \right|^{p'} \rho(Q)^{p'/q'}.$$

Now raising both sides to the power q'/p' yields the result.

• A_2 : This takes a little more doing. We split our sum into the parts corresponding to $\ell(Q') \leq \ell(Q)$ and $\ell(Q') > \ell(Q)$. Take the first part first, fix $k \geq 0$ and sum over those Q' such that $\ell(Q') = 2^{-k}\ell(Q)$. The corresponding sum is less than or equal to

$$\begin{aligned} & C 2^{-kp'/q'} \ell(Q)^{(d+m+3)p'/q'} \int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+3)p'/q'}} dx \\ & \leq C 2^{-kp'/q'} \ell(Q)^{(d+m+2)p'/q'} \int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)p'/q'}} dx. \end{aligned}$$

Except for the convenient convergence factor $2^{-kp'/q'}$, this is just what we got when we handled A_1 . So the corresponding sum has the right bound.

Now write $\ell(Q') = 2^k \ell(Q)$. We are summing over Q' 's which are larger than Q but whose triples do not contain Q , and this implies that $|x_Q - x_{Q'}| \geq c\ell(Q')$. The sum corresponding to k is less than or equal to

$$C2^{kp'/q'} \ell(Q)^{(d+m+3)p'/q'} \int_{|x_Q - x| \geq c2^k \ell(Q)} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+3)p'/q'}} dx.$$

If we now sum this over all $k > 0$ and argue as in Subcase 1, the result is less than or equal to

$$C\ell(Q)^{(d+m+2)p'/q'} \int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+m+2)p'/q'}} dx,$$

which is exactly what we want. Theorem 3 is proved. $\square$

Theorem 3 states its sufficient condition in terms of p' -pairs (σ, w) , but it says nothing about what those pairs are. The following definition and theorem partially fill this gap.

Definition. For σ a weight and α a non-negative number,

$$\mathcal{W}_\alpha \sigma(x) \equiv \sup_{\substack{Q \\ x \in Q}} \frac{1}{|Q|} \int_Q \sigma(t) \log^\alpha \left(e + \frac{\sigma(t)}{\sigma_Q} \right) dt,$$

where σ_Q , as usual, denotes the average of σ over Q .

Clearly, $\mathcal{W}_0 \sigma = M\sigma$, where M is the Hardy-Littlewood maximal operator. It is a consequence of [St, p. 23] that $\mathcal{W}_k \sigma \leq C_{k,d} M^{k+1} \sigma$, for all non-negative integers k , where M^{k+1} denotes a $(k+1)$ -fold application of M .

Theorems 2.5 and 2.7 from [W3] state that, if $k > r/2$, then $(\sigma, C_{k,d} M^k \sigma)$ is an r -pair. The following theorem provides a substantial improvement of this result for large values of r .

Theorem 4. Let $r \geq 2$. If $\alpha + 1 > r/2$ then $(\sigma, C_{\alpha,d} \mathcal{W}_\alpha \sigma)$ is an r -pair.

Proof. Let $\eta > r/2$. Theorems 2.5 and 2.7 from [W3] state that (σ, w) is an r -pair if

$$\int_Q \sigma(x) \log^\eta \left(e + \frac{\sigma(x)}{\sigma_Q} \right) dx \leq C_{\eta,d} \int_Q w(x) dx$$

for all cubes Q . Thus it will be enough to show that if $\alpha > 0$ and Q is any cube, then

$$(3.14) \quad \int_Q \sigma(x) \log^{\alpha+1} \left(e + \frac{\sigma(x)}{\sigma_Q} \right) dx \leq C_{\alpha,d} \int_Q \mathcal{W}_\alpha \sigma(x) dx.$$

We can take Q to be $Q_0 = \{x \in \mathbb{R}^d : \max |x_i| \leq 1\}$, the “unit cube.” Let $A = A(d)$ be a large constant, to be chosen presently. For $k = 1, 2, \dots$, let $\{Q_j^k\}_j \subset Q_0$ be the maximal dyadic subcubes such that $\sigma_{Q_j^k} > A^k \sigma_{Q_0}$, and set $Q_0^0 = Q_0$. We can (and shall) take A to be so large that the Q_j^k ’s satisfy the *Carleson packing condition*:

$$\sum_{Q_j^k \subset Q} |Q_j^k| \leq 2|Q|$$

for all dyadic cubes Q . To see that this is possible, recall that, because of the Q_j^k ’s maximality, we have $A^k < \sigma_{Q_j^k} \leq 2^d A^k$ for all k and j . The maximality also implies that, for fixed k , the collection $\{Q_j^k\}$ is pairwise disjoint. Now it is easy to show that, for any k and j :

$$(3.15) \quad \sum_{j': Q_{j'}^{k+1} \subset Q_j^k} |Q_{j'}^{k+1}| \leq \frac{2^d}{A} |Q_j^k|.$$

Take $A > 2 \cdot 2^d$. Note that, with this choice of A , each Q_ℓ^{k+1} is strictly contained in some Q_i^k . Inequality (3.15) implies that, for all Q_j^k and all $\ell \geq 0$,

$$\sum_{j': Q_{j'}^{k+\ell} \subset Q_j^k} |Q_{j'}^{k+\ell}| \leq 2^{-\ell} |Q_j^k|,$$

which, after summing on ℓ , implies our packing condition for the Q_j^k ’s. The inequality for general Q follows by applying the condition to the maximal $Q_j^k \subset Q$.

The packing condition has the following well-known consequence, which we will use below: *If $\{F(Q)\}_Q$ is a function mapping from the collection of dyadic cubes to the non-negative reals and $\{Q_j^k\}_{k,j}$ satisfies the Carleson packing condition, then $\sum_{k,j} F(Q_j^k) |Q_j^k| \leq C \int_{Q_0} \sup_{x \in Q} F(Q) dx$, where the constant C does not depend on F .* Proof of the well-known consequence: We compare the distribution functions of $F(Q_j^k)$ and $\sup_{x \in Q} F(Q)$ in their respective measure spaces. If $F(Q_j^k) > \lambda$, then Q_j^k is contained in some *maximal* Q'_i such that $F(Q'_i) > \lambda$. Because of the packing condition,

$$\begin{aligned} \sum_{Q_j^k: F(Q_j^k) > \lambda} |Q_j^k| &\leq \sum_i \sum_{Q_j^k \subset Q'_i} |Q_j^k| \\ &\leq 2 \sum_i |Q'_i| \\ &= 2 \left| \left\{ x : \sup_{Q: x \in Q} F(Q) > \lambda \right\} \right|. \end{aligned}$$

The integral inequality follows from standard measure theory.

For the rest of the proof of Theorem 4, all cubes will be assumed to be dyadic subcubes of Q_0 .

Fact 1. For any $\beta > 0$,

$$\int_{Q_0} \sigma(x) \log^\beta \left(e + \frac{\sigma(x)}{\sigma_{Q_0}} \right) dx \leq C_{\beta,d} \int_{Q_0} \sigma(x) \cdot \left(\sum_{k,j} \chi_{Q_j^k}(x) \right)^\beta dx.$$

Fact 2. If $\beta > 0$ and $Q \subset Q_0$,

$$\int_Q \sigma(x) \cdot \left(\sum_{Q_j^k \subset Q} \chi_{Q_j^k}(x) \right)^\beta dx \leq C_{\beta,d} \int_Q \sigma(x) \log^\beta \left(e + \frac{\sigma(x)}{\sigma_Q} \right) dx.$$

Let us temporarily assume Facts 1 and 2 and use them to finish the proof of the theorem.

$$\begin{aligned} & \int_{Q_0} \sigma(x) \log^{\alpha+1} \left(e + \frac{\sigma(x)}{\sigma_{Q_0}} \right) dx && \text{(by Fact 1)} \\ & \leq C \int_{Q_0} \sigma(x) \cdot \left(\sum_{k,j} \chi_{Q_j^k}(x) \right)^{\alpha+1} dx \\ & \leq C \int_{Q_0} \sigma(x) \cdot \left(\sum_{\ell=0}^{\infty} \left[\left(\sum_{\substack{k,j \\ k \geq \ell}} \chi_{Q_j^k}(x) \right)^{\alpha+1} - \left(\sum_{\substack{k,j \\ k \geq \ell+1}} \chi_{Q_j^k}(x) \right)^{\alpha+1} \right] \right) dx. \end{aligned}$$

All we have done is to re-write the big sum in the integral as a telescoping series. Calculus tells us that, if $t \geq 0$, then $(t+1)^{\alpha+1} - t^{\alpha+1} \leq C_\alpha(t+1)^\alpha$. Recall that, for fixed k , the cubes Q_j^k are pairwise disjoint. If $k > 0$, then each Q_j^k is contained in exactly one $Q_{j'}^{k-1}$. Thus, for each x and ℓ , $(\sum_{k,j:k \geq \ell} \chi_{Q_j^k}(x))$ and $(\sum_{k,j:k \geq \ell+1} \chi_{Q_j^k}(x))$ differ by at most one term—namely, $\chi_{Q_j^\ell}(x)$ for some j . Therefore, our calculus fact tells us that

$$\left[\left(\sum_{\substack{k,j \\ k \geq \ell}} \chi_{Q_j^k}(x) \right)^{\alpha+1} - \left(\sum_{\substack{k,j \\ k \geq \ell+1}} \chi_{Q_j^k}(x) \right)^{\alpha+1} \right] \leq C_\alpha \left(\sum_{\substack{k,j \\ k \geq \ell}} \chi_{Q_j^k}(x) \right)^\alpha.$$

Thus, our integral is bounded by

$$\begin{aligned}
& C \int_{Q_0} \sigma(x) \cdot \left(\sum_{\ell, j'} \chi_{Q_{j'}^\ell} \cdot \left[\sum_{Q_j^k \subset Q_{j'}^\ell} \chi_{Q_j^k}(x) \right]^\alpha \right) \\
& \leq C \sum_{\ell, j'} \int_{Q_{j'}^\ell} \sigma(x) \cdot \left[\sum_{Q_j^k \subset Q_{j'}^\ell} \chi_{Q_j^k}(x) \right]^\alpha dx \\
& \leq C \sum_{\ell, j'} \int_{Q_{j'}^\ell} \sigma(x) \log^\alpha \left(e + \frac{\sigma(x)}{\sigma_{Q_{j'}^\ell}} \right) dx, \\
& \leq C \int_{Q_0} \mathcal{W}_\alpha \sigma(x) dx,
\end{aligned}$$

where the next-to-last inequality is because of Fact 2 and the last inequality comes from the Carleson packing condition.

Proof of Fact 1. Notice simply that, for almost every x , $\sigma(x) > A^{k'} \sigma_{Q_0}$ implies that $x \in \bigcup_j Q_j^{k'}$, which means that $\sum_{k,j} \chi_{Q_j^k}(x) > ck'$. Thus, for almost every x , $\log(e + \sigma(x)/\sigma_{Q_0}) > c_1 k'$ implies $\sum \chi_{Q_j^k}(x) > c_2 k'$. Thus, Fact 1 follows from a pointwise inequality.

Proof of Fact 2. Unlike Fact 1, this does *not* follow from a pointwise bound. Indeed, the corresponding bound is likely to fail at many points. However, notice that it will fail at the points where $\sigma(x)$ is too small, and this set (by definition) has small σdx measure. That's the idea of the proof.

Now, fix $Q \subset Q_0$. Let $f(x) = \sum_{Q_j^k \subset Q} \chi_{Q_j^k}(x)$ and $g(x) = \log(e + \sigma(x)/\sigma_Q)$. We wish to show that

$$\int_Q (f(x))^\beta \sigma dx \leq C_{\beta,d} \int_Q (g(x))^\beta \sigma dx.$$

By standard facts from measure theory, it is enough to show that, for some positive constants c_1 and c_2 , and all $\lambda > 10$,

$$\sigma(\{x \in Q : f(x) > \lambda\}) \leq c_1 \sigma(\{x \in Q : g(x) > c_2 \lambda\}).$$

Now, if $f(x) > \lambda > 10$, then x lies in some $Q_j^k \subset Q$ such that $\sigma_{Q_j^k} > A^{\lambda/2} \sigma_Q$, and we can assume that Q_j^k is maximal for this property. Call these maximal cubes $\{Q'_i\}_i$. By definition, we have that $A^{\lambda/2} \sigma_Q < \sigma_{Q'_i}$ for each i . It is easy to see that, for any cube Q' ,

$$\sigma(\{x \in Q' : \sigma(x) > .5\sigma_{Q'}\}) \geq .5\sigma(Q').$$

Since $\sigma_{Q'_i} > A^{\lambda/2} \sigma_Q$, the definition of $g(x)$ now implies that

$$\sigma(\{x \in Q'_i : g(x) > c_2 \lambda\}) \geq .5 \sigma(Q'_i).$$

Since $\{x \in Q : f(x) > \lambda\} \subset \bigcup_i Q'_i$ and the Q'_i are disjoint, summing on i now yields Fact 2.

Theorem 4 is proved. $\square$

We can combine Theorems 3 and 4 to get the following

Corollary 4.1. *Let φ satisfy the hypotheses of Theorem 3.*

(a) *If $2 < p \leq q < \infty$, then*

$$\left(\int_{\mathbb{R}_+^{d+1}} |y^{-1} \varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

will hold for all f if

$$\begin{aligned} & \mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{M\sigma(x) \cdot \log^{p'/q'}(e + |x - x_Q|/\ell(Q))}{(\ell(Q) + |x - x_Q|)^{(d+2+m)p'/q'}} dx \right)^{1/p'} \\ & \leq C \ell(Q)^{d+1-(d+2+m)/q'} \end{aligned}$$

holds for all cubes Q , where M denotes the Hardy-Littlewood maximal operator, and the constant C depends on m, d, p, q , and φ .

(b) *If $1 < p \leq 2 \leq q < \infty$ and $\alpha > p'/2 - 1$, then*

$$\left(\int_{\mathbb{R}_+^{d+1}} |y^{-1} \varphi_y * f(x)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

will hold for all f if

$$\begin{aligned} & \mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\mathcal{W}_\alpha \sigma(x) \cdot \log^{p'/q'}(e + |x - x_Q|/\ell(Q))}{(\ell(Q) + |x - x_Q|)^{(d+2+m)p'/q'}} dx \right)^{1/p'} \\ & \leq C \ell(Q)^{d+1-(d+2+m)/q'} \end{aligned}$$

holds for all cubes Q . The constant C depends on m, d, p, q, φ , and α .

Remark. The converse inequality to (3.14) is easy. Let us prove it in the case where $Q = Q_0$, the “unit cube,” and $\sigma_{Q_0} = 1$. For any $Q \subset Q_0$, not necessarily dyadic,

$$\begin{aligned} & \frac{1}{|Q|} \int_Q \sigma(t) \log^\alpha \left(e + \frac{\sigma(t)}{\sigma_Q} \right) dt \\ & \leq C_\alpha \left[\frac{1}{|Q|} \int_Q \sigma(t) \log^\alpha(e + \sigma(t)) dt + \frac{1}{|Q|} \int_Q \sigma(t) \log^\alpha \left(e + \frac{1}{\sigma_Q} \right) dt \right]. \end{aligned}$$

Take the supremum over all $Q \ni x$ and integrate over $x \in Q_0$. The integral of the first term on the right, by the $L \log L$ result for the ordinary Hardy-Littlewood maximal function, is controlled by

$$\begin{aligned} & C \int_{Q_0} \sigma(t) \log^\alpha(e + \sigma(t)) \cdot \log[2 + \sigma(t) \log^\alpha(e + \sigma(t))] dt \\ & \leq C \int_{Q_0} \sigma(t) \log^{\alpha+1} \left(e + \frac{\sigma(t)}{\sigma_{Q_0}} \right) dt. \end{aligned}$$

(Here we are also using the normalization $\sigma_{Q_0} = 1$.) For the second term, we have two cases. Either $\sigma_Q > 1$, and it is dominated by the Hardy-Littlewood maximal function, or $\sigma_Q \leq 1$, and it is bounded by an absolute constant, since $s \log^\alpha(e + 1/s)$ is bounded for $0 < s \leq 1$. Thus, the integral of the second term is no bigger than

$$C \int_{Q_0} \sigma(t) \log \left(e + \frac{\sigma(t)}{\sigma_{Q_0}} \right) dt.$$

4. The Poisson kernel. The problem of the Poisson kernel naturally divides into two subcases, corresponding to the two parts of ∇u : the x -components $\partial u / \partial x_i$ ($i = 1, \dots, d$), and the y -component $\partial u / \partial y$. It turns out that our result for the x -derivatives is a straightforward consequence of the theorems in Section 2. The corresponding result for the y -derivative (and thus for the full gradient) requires some additional trickery.

If

$$\int_{\mathbb{R}^d} \frac{|f(x)|}{(1 + |x|)^{d+1}} dx < \infty,$$

then $u(x, y) = P_y * f(x)$ is defined and, for $i = 1, \dots, d$, $\partial u / \partial x_i$ is equal to $y^{-1} f * (\varphi_i)_y(x)$, where φ_i satisfies

$$\begin{aligned} |\varphi_i(x)| & \leq C_d (1 + |x|)^{-d-2}, \\ |\nabla \varphi_i(x)| & \leq C_d (1 + |x|)^{-d-3}, \\ \int_{\mathbb{R}^d} \varphi_i(x) dx & = 0. \end{aligned}$$

Thus, Theorem 3 from the previous section immediately implies:

Theorem 5. *Let v be a non-negative weight in $L^1_{\text{loc}}(\mathbb{R}^d)$ and μ a positive measure on $\mathbb{R}^{d+1}_+$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. If $1 < p \leq q < \infty$ and $q \geq 2$, then, in order for*

$$\left(\int_{\mathbb{R}^{d+1}_+} \left| \frac{\partial u}{\partial x_i} \right|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , it is sufficient that there exist a p' -pair (σ, w) such that

$$\mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q))w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2)p'/q'}} dx \right)^{1/p'} \leq C\ell(Q)^{d+1-(d+2)/q'}$$

holds for all cubes Q . The constant C depends on p , q , and d .

We would like to say the same thing for the y -derivative. Unfortunately, the corresponding kernel for $\partial u/\partial y$ —which we shall call φ_0 —while it has cancellation, decays no faster than $(1 + |x|)^{-d-1}$. This fact will make our sufficient condition for the y -derivative, and thus for the full gradient, significantly stronger (= worse) than that for the x -derivatives. It also makes it impossible (or at least unwise) for us to follow the procedure that worked so well with the x -derivatives. This is because the Uchiyama lemma (Lemma 3.1) requires that we have decay of order at least $(1 + |x|)^{-d-2}$ in our kernel.

We are going to get some extra decay—but not enough, unfortunately, to use Uchiyama's result—by exploiting harmonicity. Let ψ be the convolution kernel for $\partial^2 u/\partial y^2$, i.e., such that

$$\frac{\partial^2 u}{\partial y^2}(x, y) = y^{-2}\psi_y * f(x)$$

when $u = P_y * f$. By the semigroup property,

$$y^{-1}(\varphi_0)_y * \eta^{-1}(\varphi_0)_\eta = C \cdot (y + \eta)^{-2}\psi_{y+\eta}.$$

This identity can also be verified by Fourier transforms, and recalling that $\hat{P}_y(\xi) = e^{-y|\xi|}$. By a computation, or simply observing that $u_{yy} = -\sum_i u_{x_i x_i}$, we see that

$$\begin{aligned} |\psi(x)| &\leq C(1 + |x|)^{-d-3}, \\ |\nabla_x \psi| &\leq C(1 + |x|)^{-d-4}. \end{aligned}$$

Let $\rho \in C_0^\infty(\mathbb{R}^d)$ be real, radial, satisfy $\int \rho = 0$, and have support contained in $\{|x| \leq 1\}$. Furthermore, let ρ be normalized so that

$$\int_0^\infty \hat{\varphi}_0(y\xi) \hat{\rho}(y\xi) \frac{dy}{y} \equiv 1$$

for $\xi \neq 0$. If $u = P_y * f$ and $f \in L^2(\mathbb{R}^d)$, then, by Fourier inversion,

$$\begin{aligned} (4.1) \quad f(x) &= \int_{\mathbb{R}_+^{d+1}} \frac{\partial u}{\partial y}(t, y) \rho_y(x-t) dt dy \\ &= \int_{\mathbb{R}_+^{d+1}} [f * (\varphi_0)_y(t)] \rho_y(x-t) \frac{dt dy}{y} \end{aligned}$$

in L^2 . Because ρ has compact support, this integral can be cut up into a sum of constants times smooth functions that are supported on the triples of dyadic cubes. In other words, f is a limit of functions f_n in *standard form*, which we considered in the previous section, and f 's size can be controlled by the same sort of square function we used there; i.e., a square function of the form

$$S_c(f)(x) = \left(\sum_{x \in \tilde{Q}} \frac{|\lambda_Q|^2}{|Q|} \right)^{1/2},$$

where

$$|\lambda_Q| \leq C \left(\int_{T(Q)} |f * (\varphi_0)_y(t)|^2 \frac{dt dy}{y} \right)^{1/2}.$$

We can now state and prove our theorem for the full gradient.

Theorem 6. *Let v be a non-negative weight in $L_{\text{loc}}^1(\mathbb{R}^d)$ and μ a positive measure on $\mathbb{R}_+^{d+1}$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. If $1 < p \leq q < \infty$ and $q \geq 2$, then, in order for*

$$\left(\int_{\mathbb{R}_+^{d+1}} |\nabla u(x, y)|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , with $u = P_y * f$, it is sufficient that there exist a p' -pair (σ, w) such that

$$(4.2) \quad \mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'/q'}} dx \right)^{1/p'} \leq C \ell(Q)^{(d+1)/q}$$

holds for all cubes Q . The constant C depends on p , q , and d .

Proof. By Theorem 5, we only need to estimate the u_y component of the gradient. If $g : \mathbb{R}_+^{d+1} \mapsto \mathbb{R}$ is bounded and has compact support, define

$$T_0 g(x) = \int_{\mathbb{R}_+^{d+1}} g(t, \eta) \cdot \eta^{-1} (\varphi_0)_\eta(t - x) d\mu(t, \eta).$$

We are going to apply the observations from the preceding paragraph to the function $f = T_0 g$. Let $(x, y) \in T(Q')$, where $Q' \subset \mathbb{R}^d$ is a dyadic cube. We need to estimate $(\varphi_0)_y * T_0 g(x)$. As in the proof of Theorem 3, this must begin with estimates for $(\varphi_0)_y * (\varphi_0)_\eta$. Fortunately, we know precisely what this is:

$$\begin{aligned} |(\varphi_0)_y * (\varphi_0)_\eta(x)| &= |y\eta(y + \eta)^{-2} \psi_{y+\eta}(x)| \\ &\leq C \frac{y\eta}{(y + \eta)^2} (y + \eta)^{-d} \frac{1}{(1 + |x|/(y + \eta))^{d+3}} \\ &\leq \frac{Cy\eta(y + \eta)}{(y + \eta + |x|)^{d+3}}. \end{aligned}$$

In estimating the last term, we separately consider the two cases $y < \eta$ and $y \geq \eta$. In the first case our bound is

$$(4.3) \quad \frac{Cy\eta^2}{(\eta + |x|)^{d+3}}$$

and in the second it is

$$(4.4) \quad \frac{C\eta y^2}{(y + |x|)^{d+3}}.$$

Now, taking $(x, y) \in T(Q')$, we may estimate $(\varphi_0)_y * T_0 g(x)$ as we did in the preceding section:

$$\begin{aligned} |(\varphi_0)_y * T_0 g(x)| &\leq \int_{\mathbb{R}_+^{d+1}} |g(t, \eta)| |(\varphi_0)_y * \eta^{-1} (\varphi_0)_\eta(t - x)| d\mu(t, \eta) \\ &= \sum_{Q: \ell(Q) \leq \ell(Q')} \int_{T(Q)} |g(t, \eta)| |(\varphi_0)_y * \eta^{-1} (\varphi_0)_\eta(x - t)| d\mu(t, \eta) \\ &\quad + \sum_{Q: \ell(Q) > \ell(Q')} \int_{T(Q)} |g(t, \eta)| |(\varphi_0)_y * \eta^{-1} (\varphi_0)_\eta(x - t)| d\mu(t, \eta) \\ &\leq C \sum_{Q: \ell(Q) \leq \ell(Q')} \left(\int_{T(Q)} |g(t, \eta)| d\mu(t, \eta) \right) \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} \right) \\ &\quad + C \sum_{Q: \ell(Q) > \ell(Q')} \left(\int_{T(Q)} |g(t, \eta)| d\mu(t, \eta) \right) \left(\frac{\ell(Q)\ell(Q')}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+3}} \right), \end{aligned}$$

where, as usual, x_Q denotes the center of the cube Q .

We write the two sums in the last line as one sum, to wit:

$$C \sum_Q a^*(Q', Q) G^*(Q),$$

where

$$a^*(Q', Q) = \begin{cases} \frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} & \text{if } \ell(Q) \leq \ell(Q'), \\ \frac{\ell(Q)\ell(Q')}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+3}} & \text{if } \ell(Q) > \ell(Q'), \end{cases}$$

and

$$G^*(Q) = \int_{T(Q)} |g(t, \eta)| d\mu(t, \eta).$$

These are analogous to expressions we obtained in the previous section. (Indeed, the $G^*(Q)$ is identical to the term we called $G(Q)$.) We need to show the same sequence-space estimates we proved there. Fortunately, half of our work has already been done. Notice that, if (4.2) holds, then inequality (3.1), the hypothesis of Theorem 3, also holds, for $m = 0$. We now claim that if $\ell(Q) > \ell(Q')$ and we take $m = 0$, then $a^*(Q', Q) \leq Ca(Q', Q)$. The proof of this fact is simple. If $\ell(Q) \geq 4\ell(Q')$, then $Q \not\subset \tilde{Q}'$ and $a^*(Q', Q) = a_2(Q', Q) = a(Q', Q)$. If $\ell(Q') < \ell(Q) < 4\ell(Q')$ and $Q \subset \tilde{Q}'$, then

$$\begin{aligned} a^*(Q', Q) &= \frac{\ell(Q)\ell(Q')}{(\ell(Q) + |x_Q - x_{Q'}|)^{d+3}} \\ &\leq C\ell(Q)^{-d-1} \\ &\leq C \left(\frac{\ell(Q)}{\ell(Q')^{d+2}} \right) \log \left(e + \frac{\ell(Q')}{\ell(Q)} \right) \\ &= Ca_1(Q', Q) \\ &= Ca(Q', Q). \end{aligned}$$

If $\ell(Q') < \ell(Q) < 4\ell(Q')$ and $Q \not\subset \tilde{Q}'$ then $a^*(Q', Q) = a_2(Q', Q) = a(Q', Q)$ as before. Therefore, for example, the sequence-space estimate analogous to inequality (3.10) (which gave the $\ell^1 \mapsto \ell^{2/q'}$ bound) will automatically hold for the “kernel” $A^*(Q', Q) = |\hat{Q}| a^*(Q', Q)$, as long as we sum only over cubes Q' such that $\ell(Q) > \ell(Q')$; and a similar observation applies to the other sequence-space bounds. Therefore, in proving Theorem 6, we only need to worry about

pairs of cubes (Q', Q) such that $\ell(Q) \leq \ell(Q')$. Having said this, we shall now redefine $a^*(Q', Q)$:

$$a^*(Q', Q) = \begin{cases} \frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} & \text{if } \ell(Q) \leq \ell(Q'), \\ 0 & \text{if } \ell(Q) > \ell(Q'). \end{cases}$$

We have just thrown out the pairs (Q', Q) that we do not need to consider.

We consider two subcases: $1 < p \leq 2$ and $p > 2$. As before, we shall consider the first subcase first. Let $h \in L^{(\overline{p}/2)'}(w dx)$ satisfy $\|h\|_{L^{(p'/2)'}(w dx)} \leq 1$, and set $\nu(Q') = \int_{\tilde{Q}'} h w dx$. Define $A^*(Q', Q) = |\hat{Q}| a^*(Q', Q)$ and let $G^*(Q) = |\hat{Q}| Y^*(Q)$. Let $\rho(Q) = \mu(T(Q))^{1-q'}$, just as in Section 2. We need to show that, under hypothesis (4.2),

$$(4.5) \quad \sum_Q A^*(Q', Q) \leq C$$

for all cubes Q' , and

$$(4.6) \quad \left(\sum_{Q'} A^*(Q', Q)^{2/q'} \nu(Q') \right)^{q'/2} \leq C |\hat{Q}|^{q'} \rho(Q)$$

for all Q . The first inequality implies the $\ell^\infty \mapsto \ell^\infty$ bound and the second gives the desired $\ell^1 \mapsto \ell^{2/q'}$ estimate.

Proof of (4.5). Remember: we are summing over cubes Q such that $\ell(Q) \leq \ell(Q')$. If $Q \subset \tilde{Q}'$ then $\ell(Q') + |x_Q - x_{Q'}| \sim \ell(Q')$ and $a^*(Q', Q) \sim |\hat{Q}'|^{-1}$. Therefore,

$$\begin{aligned} \sum_{Q: Q \subset \tilde{Q}'} A^*(Q', Q) &= \sum_{Q: Q \subset \tilde{Q}'} a^*(Q', Q) |\hat{Q}| \\ &\leq C |\hat{Q}'|^{-1} \sum_{Q \subset \tilde{Q}'} |\hat{Q}| \\ &\leq C, \end{aligned}$$

as before. Let k be a non-negative integer. If $|x_Q - x_{Q'}| \sim 2^k \ell(Q')$, then $a^*(Q', Q) \sim 2^{-k(d+3)} |\hat{Q}'|^{-1}$. Therefore,

$$(4.7) \quad \sum_{\substack{Q: \ell(Q) \leq \ell(Q') \\ |x_Q - x_{Q'}| \sim 2^k \ell(Q')}} A^*(Q', Q) = \sum_{\substack{Q: \ell(Q) \leq \ell(Q') \\ |x_Q - x_{Q'}| \sim 2^k \ell(Q')}} a^*(Q', Q) |\hat{Q}| \leq$$

$$\begin{aligned}
&\leq C 2^{-k(d+3)} \left| \hat{Q}' \right|^{-1} \sum_{\substack{Q: \ell(Q) \leq \ell(Q') \\ |x_Q - x_{Q'}| \sim 2^k \ell(Q')}} \left| \hat{Q} \right| \\
&\leq C 2^{-k(d+3)} \left| \hat{Q}' \right|^{-1} \cdot 2^{kd} \left| \hat{Q}' \right| \\
&\leq C 2^{-3k}.
\end{aligned}$$

(The sum of the measures of the $\hat{Q}$'s in (4.7) is less than or equal to $C|2^k Q' \times (0, \ell(Q'))|$, where, as usual, $2^k Q'$ means the cube concentric with Q' and with diameter 2^k times as large.) Summing over k now yields (4.5). (*Remark*: Notice that we got (4.5) “for free,” without using our hypothesis (4.2). This also happened in the proof of Theorem 3.)

Proof of (4.6). We need to show that, if (4.2) holds, then

$$(4.8) \quad \left[\sum_{Q': \ell(Q') \geq \ell(Q)} \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} \left| \hat{Q} \right| \right)^{2/q'} \nu(Q') \right]^{q'/2} \leq C \left| \hat{Q} \right|^{q'} \rho(Q).$$

We have handled objects like this in the previous section (see inequality (3.10) there and *passim*). We have two sorts of cubes Q' to consider here: those whose triples contain Q , and the others. For fixed k , a non-negative integer, there are at most $C(d)$ cubes Q' such that $Q \subset \hat{Q}'$ and $\ell(Q') = 2^k \ell(Q)$. Thus, by the sort of argument we used for (3.11)–(3.13), all of these cubes (meaning, for all k) contribute no more than

$$\left[C \left| \hat{Q} \right|^{2/q'} \int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)2/q'}} w(x) dx \right]^{q'/2}$$

to the left-hand side of (4.8). Hölder's inequality and the fact that $\|h\|_{L^{(p'/2)'}(w dx)} \leq 1$ imply that this last expression is less than or equal to

$$C \left| \hat{Q} \right| \left(\int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'/q'}} dx \right)^{q'/p'},$$

which, by hypothesis (4.2), is bounded by

$$C \left| \hat{Q} \right|^{1+q'/q} / \mu(T(Q))^{q'/q} = C \left| \hat{Q} \right|^{q'} \rho(Q).$$

Now we consider the cubes Q' such that $\ell(Q') \geq \ell(Q)$ and $Q \not\subset \tilde{Q}'$. Fix k , and let $\mathcal{F}_k$ be those cubes Q' such that $\ell(Q') = 2^k \ell(Q)$ and $Q \not\subset \tilde{Q}'$. Then:

$$\begin{aligned} & \sum_{Q' \in \mathcal{F}_k} \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} |\hat{Q}| \right)^{2/q'} \nu(Q') \\ & \leq C 2^{4k/q'} \ell(Q)^{(d+3)2/q'} \left(\int_{|x - x_Q| > c 2^k \ell(Q)} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+3)2/q'}} w(x) dx \right). \end{aligned}$$

If we sum this over k (as we did in Section 3) and (just as in Section 3) use the fact that $\sum_{k>0} 2^{4k/q'} \chi_{\{|x|>2^k\}} \leq C(1+|x|)^{4/q'}$, we see that the sum is bounded by

$$\begin{aligned} & C \ell(Q)^{(d+1)2/q'} \int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)2/q'}} w(x) dx \\ & = C |\hat{Q}|^{2/q'} \int_{\mathbb{R}^d} \frac{h(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)2/q'}} w(x) dx, \end{aligned}$$

which is exactly what we got a few lines ago, and therefore has the right bound. The case $1 < p \leq 2$ is proved.

Let us now take $p > 2$. Just as in the proof of Theorem 3, we shall get two endpoint estimates and then interpolate. The $\ell^\infty \mapsto \ell^\infty$ is a repeat of the one we did for $1 < p \leq 2$, and we don't need to do it again. The $\ell^1 \mapsto \ell^{p'/q'}$ bound will follow when we show that, with hypothesis (4.2), we have

$$\left(\sum_{Q'} A^*(Q', Q)^{p'/q'} w(\tilde{Q}') \right)^{q'/p'} \leq C |\hat{Q}|^{q'} \rho(Q)$$

for all cubes Q . (Recall that we are summing only over cubes Q' for which $\ell(Q') \geq \ell(Q)$.)

Let us begin:

$$\begin{aligned} \sum_{Q'} A^*(Q', Q)^{p'/q'} w(\tilde{Q}') & = \sum_{Q'} \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} |\hat{Q}| \right)^{p'/q'} w(\tilde{Q}') \\ & = \sum_{Q': Q \subset \tilde{Q}'} \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} |\hat{Q}| \right)^{p'/q'} w(\tilde{Q}') \\ & \quad + \sum_{Q': Q \not\subset \tilde{Q}'} \left(\frac{\ell(Q')^2}{(\ell(Q') + |x_Q - x_{Q'}|)^{d+3}} |\hat{Q}| \right)^{p'/q'} w(\tilde{Q}') \\ & = \text{(I)} + \text{(II)}. \end{aligned}$$

The treatment of (I) follows the same track as the argument in inequalities (3.11) through (3.13), and leads to the bound:

$$(I) \leq C \left| \hat{Q} \right|^{p'/q'} \int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'/q'}} dx,$$

which is what we want.

We handle (II) by first summing over the Q' such that $\ell(Q') = 2^k \ell(Q)$, where k is a non-negative integer. The sum over these Q' is bounded by

$$C 2^{2kp'/q'} \ell(Q)^{(d+3)p'/q'} \int_{|x-x_Q| > c2^k \ell(Q)} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+3)p'/q'}} dx.$$

When we sum this over $k > 0$, the result is less than or equal to

$$C \ell(Q)^{(d+1)p'/q'} \int_{\mathbb{R}^d} \frac{w(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'/q'}} dx,$$

which is what we want. This concludes the proof of Theorem 6. $\square$

By an adroit use of singular integral operators, we can improve our sufficient condition in the case when $\sigma \in A_{p'}$. Let $g : \mathbb{R}_+^{d+1} \mapsto \mathbb{R}$ be bounded and have compact support. Recall that we earlier defined φ_i ($1 \leq i \leq d$) to be the convolution kernel that “generates” $\partial u / \partial x_i$; i.e., such that $(\partial u / \partial x_i)(x, y) = y^{-1}(\varphi_i)_y * f(x)$ when $u = P_y * f$. For $i = 1, \dots, d$, set

$$T_i g(x) = \int_{\mathbb{R}_+^{d+1}} g(t, y) \cdot y^{-1}(\varphi_i)_y(t - x) d\mu(t, y).$$

If R_i is the i^{th} Riesz transform, then, by Fourier transforms,

$$\begin{aligned} T_0 g(x) &= - \int_{\mathbb{R}_+^{d+1}} g(t, y) \cdot y^{-1} \left[\sum_{i=1}^d R_i(\varphi_i)_y(t - x) \right] d\mu(t, y) \\ &= - \sum_{i=1}^d R_i T_i g(x). \end{aligned}$$

Therefore, our norm inequality will hold for the full gradient ∇u if

$$\sum_i \left(\int_{\mathbb{R}^d} (|T_i g|^{p'} + |R_i T_i g|^{p'}) \sigma dx \right)^{1/p'} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g(t, y)|^{q'} d\mu(t, y) \right)^{1/q'}$$

holds for all such g . However, if $\sigma \in A_{p'}$, then each R_i maps boundedly from $L^{p'}(\sigma dx) \mapsto L^{p'}(\sigma dx)$, and so the last inequality will follow from

$$\sum_i \left(\int_{\mathbb{R}^d} |T_i g|^{p'} \sigma dx \right)^{1/p'} \leq C \left(\int_{\mathbb{R}_+^{d+1}} |g(t, y)|^{q'} d\mu(t, y) \right)^{1/q'}.$$

If $\sigma \in A_{p'}$, then $\sigma \in A_\infty$ and $(\sigma, C\sigma)$ is a p' -pair. When we put these remarks together, we get:

Corollary 7. *Let v be a non-negative weight in $L^1_{\text{loc}}(\mathbb{R}^d)$ and μ be a positive measure on $\mathbb{R}^{d+1}_+$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. Suppose that $\sigma \in A_{p'}$ (which is equivalent to having $v \in A_p$). If $1 < p \leq q < \infty$ and $q \geq 2$, then, in order for*

$$\left(\int_{\mathbb{R}^{d+1}_+} |\nabla u|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , where $u = P_y * f$, it is sufficient that

$$\mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q))w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2)p'/q'}} dx \right)^{1/p'} \leq C \ell(Q)^{d+1-(d+2)/q'}$$

holds for all cubes Q . The constant C depends on p, q, d , and the $A_{p'}$ constant of σ .

Recent work of Carlos Pérez [P] allows us to partially salvage this result even when $\sigma \notin A_{p'}$. He has shown that, for any Calderón-Zygmund operator K , $1 < r < \infty$, and weight σ ,

$$\int_{\mathbb{R}^d} |Kf|^r \sigma dx \leq C \int_{\mathbb{R}^d} |f|^r M^{[r+1]} \sigma dx,$$

where M^k means a k -fold application of the Hardy-Littlewood maximal operator and $[\cdot]$ is the greatest integer function. Since $M^k \sigma \geq \sigma$ almost everywhere, the full gradient inequality will hold, therefore, if we have

$$\sum_i \left(\int_{\mathbb{R}^d} |T_i g|^{p'} M^{[p'+1]} \sigma dx \right)^{1/p'} \leq C \left(\int_{\mathbb{R}^{d+1}_+} |g(t, y)|^{q'} d\mu(t, y) \right)^{1/q'}.$$

These facts immediately yield a second corollary to Theorem 5:

Corollary 8. *Let v be a non-negative weight in $L^1_{\text{loc}}(\mathbb{R}^d)$ and μ be a positive measure on $\mathbb{R}^{d+1}_+$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. If $1 < p \leq q < \infty$ and $q \geq 2$, then, in order for*

$$\left(\int_{\mathbb{R}^{d+1}_+} |\nabla u|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all f , where $u = P_y * f$, it is sufficient that there exists a p' -pair $(M^{[p'+1]} \sigma, w)$ such that

$$\mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q))w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2)p'/q'}} dx \right)^{1/p'} \leq C \ell(Q)^{d+1-(d+2)/q'}$$

holds for all cubes Q . The constant C depends on p, q , and d .

5. Necessary conditions. Because of the specific forms of the Poisson kernel and its derivatives, we can readily obtain a necessary condition for the weighted norm inequality

$$\left(\int_{\mathbb{R}_+^{d+1}} |\nabla u|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

to hold for all $f \in L^p(v)$. The idea is to consider separate cases which allow one to reduce the problem to non-negative functions and non-negative kernels. Unlike our sufficiency proofs, the necessity argument requires no special hypotheses on q .

Theorem 8. *Let v be a non-negative weight in $L^1_{\text{loc}}(\mathbb{R}^d)$ and let μ be a positive measure on $\mathbb{R}_+^{d+1}$. For $1 < p < \infty$, set $\sigma = v^{1-p'}$. If $0 < q < \infty$ and*

$$(5.1) \quad \left(\int_{\mathbb{R}_+^{d+1}} |\nabla u|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p}$$

holds for all $f \in L^p(v)$, then there is a constant C , depending only on p, q , and d , such that

$$\mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\sigma(x)}{(\ell(Q) + |x - x_Q|)^{(d+1)p'}} dx \right)^{1/p'} \leq C$$

for all cubes Q .

Proof. Take $x_Q = 0$. Recall that we have denoted the convolution kernel for $\partial u / \partial y$ by φ_0 , by which we mean that, if $u = P_y * f$, then $(\partial u / \partial y)(x, y) = y^{-1}(\varphi_0)_y * f(x)$. Now, by (1.3b), there is an $\varepsilon_0 > 0$, which we may take to be less than 1, such that $|(\varphi_0)_y(t)| > C_d$ if $|t| < \varepsilon_0$ and $\frac{1}{2} \leq y \leq 1$. Note that, since φ_0 is continuous, it will not change sign for this range of y and t . Fix this ε_0 . By considering how φ_0 behaves under dilations, it is easy to see that, for any $\lambda > 0$, if $\lambda/2 \leq y \leq \lambda$ and $|t| < \varepsilon_0 \lambda$, then $|(\varphi_0)_y(t)| > C_d \lambda^{-d}$. We are going to make use of this fact. Set $f = \sigma \cdot \chi_{\{|x| \leq \varepsilon_0 \ell(Q)/10\}}$ and put $u(x, y) = P_y * f(x)$. If $|x| < \varepsilon_0 \ell(Q)/10$ and $\ell(Q)/2 \leq y \leq \ell(Q)$, then, because of φ_0 's constant-sign property, mentioned above, $|(\partial u / \partial y)(x, y)| > C_d \cdot \left(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \sigma dt \right) \cdot \ell(Q)^{-d-1}$. Thus, if (5.1) holds, we have

$$\begin{aligned} & \left(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \sigma dt \right) \cdot \ell(Q)^{-d-1} \cdot \mu \left(\left\{ |x| < \varepsilon_0 \frac{\ell(Q)}{10} \right\} \times \left[\frac{\ell(Q)}{2}, \ell(Q) \right] \right)^{1/q} \\ & \leq C \left(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \sigma^p v dt \right)^{1/p}. \end{aligned}$$

But $\sigma^p v = \sigma$, which makes the two integrals in the preceding line equal. If we divide both sides by $(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \sigma dt)^{1/p}$, we get

$$\begin{aligned} & \mu \left(\left\{ |x| < \varepsilon_0 \frac{\ell(Q)}{10} \right\} \times \left[\frac{\ell(Q)}{2}, \ell(Q) \right] \right)^{1/q} \\ & \cdot \ell(Q)^{-d-1} \cdot \left(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \sigma dt \right)^{1/p'} \leq C_d; \end{aligned}$$

i.e.,

$$\begin{aligned} & \mu \left(\left\{ |x| < \varepsilon_0 \frac{\ell(Q)}{10} \right\} \times \left[\frac{\ell(Q)}{2}, \ell(Q) \right] \right)^{1/q} \\ & \cdot \left(\int_{\{|t| \leq \varepsilon_0 \ell(Q)/10\}} \frac{\sigma}{(\ell(Q) + |t|)^{(d+1)p'}} dt \right)^{1/p'} \leq C. \end{aligned}$$

There is an $\eta_0 > 1$ such that if $\frac{1}{2} \leq y \leq 1$ and $|t| > \eta_0$ then $|(\varphi_0)_y(t)| > C_d \lambda (1 + |t|)^{-d-1}$. In addition, because of continuity, $(\varphi_0)_y(t)$ does not change sign on $\{|t| > \eta_0\}$. Fix this η_0 . Because of dilation invariance, if $\lambda > 0$, $|t| > \eta_0 \lambda$, and $\lambda/2 \leq y \leq \lambda$, then $|(\varphi_0)_y(t)| \geq C_d \lambda (\lambda + |t|)^{-d-1}$. Let $f = \sigma \cdot \chi_{\{|t| > 2\eta_0 \ell(Q)\}} \cdot (\ell(Q) + |t|)^{-(d-1)(p'-1)}$. If $|x| \leq \ell(Q)/10$ and $\ell(Q)/2 \leq y \leq \ell(Q)$ then

$$|y^{-1}(\varphi_0)_y * f(x)| \geq C_d \int_{\{|t| > 2\eta_0 \ell(Q)\}} \frac{\sigma(t)}{(\ell(Q) + |t|)^{(d+1)p'}} dt.$$

If (5.1) holds, then

$$\begin{aligned} & \mu \left(\left\{ |x| < \frac{\ell(Q)}{10} \right\} \times \left[\frac{\ell(Q)}{2}, \ell(Q) \right] \right)^{1/q} \cdot \left(\int_{\{|t| > 2\eta_0 \ell(Q)\}} \frac{\sigma(t)}{(\ell(Q) + |t|)^{(d+1)p'}} dt \right) \\ & \leq C_d \left(\int_{\mathbb{R}^d} |f|^p v dx \right)^{1/p} \\ & = C_d \left(\int_{\{|t| > 2\eta_0 \ell(Q)\}} \frac{\sigma(t)}{(\ell(Q) + |t|)^{(d+1)p'}} dt \right)^{1/p}, \end{aligned}$$

where the last line follows from the equations $\sigma^p v = \sigma$ and $(p' - 1)p = p'$. Now dividing both sides by

$$\left(\int_{\{|t| > 2\eta_0 \ell(Q)\}} \frac{\sigma(t)}{(\ell(Q) + |t|)^{(d+1)p'}} dt \right)^{1/p}$$

yields

$$\mu\left(\left\{|x| < \frac{\ell(Q)}{10}\right\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q} \cdot \left(\int_{\{|t| > 2\eta_0 \ell(Q)\}} \frac{\sigma(t)}{(\ell(Q) + |t|)^{(d+1)p'}} dt\right)^{1/p'} \leq C.$$

We have proved two integral estimates involving σ . One treats points close to 0 and the other concerns points that are “far” from 0. The key to both of our estimates was the fact that the kernel for the y -derivative does not become too small on $\mathbb{R}^d$ when x is very small or very large. Unfortunately, it does become small when $|x| \sim 1$. We are going to use the x_i -derivatives of u ($1 \leq i \leq d$) and simple facts about their corresponding convolution kernels, φ_i , to fill the gap between our estimates for large and small x .

Let φ_1 be the convolution kernel for $\partial u / \partial x_1$; i.e., such that $(\partial u / \partial x_1)(x, y) = y^{-1}(\varphi_1)_y * f(x)$ when $u = P_y * f(x)$. For every $0 < M_1 < 1$ and $1 < M_2 < \infty$, by (1.3b), there is a positive m such that if

$$(5.2) \quad M_1 < \frac{|x_1|}{|x|},$$

$$(5.3) \quad M_2^{-1} < |x| < M_2,$$

$$(5.4) \quad \frac{1}{2} \leq y \leq 1,$$

then $|(\varphi_1)_y(x)| > m$. For such M_1 , M_2 , and m , there is a positive number δ —which we may take to be less than $\varepsilon_0/100$ —such that if $|t| \leq \delta$ and (5.2), (5.3), and (5.4) hold, then $|(\varphi_1)_y(x-t)| > m/2$. It is important to observe that if (5.2), (5.3), and (5.4) hold and x_1 does not change sign, then $(\varphi_1)_y(x-t)$ will not change sign either (this is a consequence of φ_1 's special form). Obviously, analogous estimates will hold for the other kernels φ_i , $i = 2, \dots, d$ as well, when we replace x_1 with x_i . It is important to observe that these estimates will hold for the same numbers M_1 , M_2 , etc.

Let us now take $M_1 < d^{-1/2}$, $M_2 > \max(10\eta_0, 100\varepsilon_0^{-1})$, and fix them. Let m and δ be as specified in the preceding paragraph. Let

$$S_1^+ = \left\{x \in \mathbb{R}^d : M_1 < \frac{|x_1|}{|x|}, M_2^{-1}\ell(Q) < |x| < M_2\ell(Q), x_1 > 0\right\},$$

$$S_1^- = \left\{x \in \mathbb{R}^d : M_1 < \frac{|x_1|}{|x|}, M_2^{-1}\ell(Q) < |x| < M_2\ell(Q), x_1 < 0\right\}.$$

Let $f = \sigma(x) \cdot \chi_{S_1^+}$. By the preceding and by dilation invariance, if $\ell(Q)/2 \leq y \leq \ell(Q)$ and $|t| \leq \delta\ell(Q)$, we must have

$$|(\varphi_1)_y * f(t)| \geq C_d \cdot \left(\frac{m}{2}\right) \ell(Q)^{-d-1} \left(\int_{S_1^+} \sigma dx\right).$$

(Here we are using the fact that if $|t| \leq \delta \ell(Q)$ and $\ell(Q)/2 \leq y \leq \ell(Q)$, then $(\varphi_1)_y(t-x)$ does not change sign for $x \in S_1^+$.) Now, following the sort of argument we have already seen, we obtain, under the assumption that (5.1) holds,

$$\mu\left(\{|x| \leq \delta \ell(Q)\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q} \cdot \left(\int_{S_1^+} \sigma dt\right)^{1/p'} \leq C \ell(Q)^{d+1}.$$

Clearly, we will obtain the same estimate for S_1^- and for the other corresponding sets S_i^+ and S_i^- . Calling the union of these $2d$ sets S , we obtain:

$$\mu\left(\{|x| \leq \delta \ell(Q)\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q} \cdot \left(\int_S \sigma dt\right)^{1/p'} \leq C \ell(Q)^{d+1}.$$

But $\{x : (\varepsilon_0/10)\ell(Q) \leq |x| \leq 10\eta_0\ell(Q)\} \subset S$ —because $M_1 < d^{-1/2}$ and $\max(|x_i|) \geq d^{-1/2}|x|$. Therefore,

$$\begin{aligned} & \mu\left(\{|x| \leq \delta \ell(Q)\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q} \\ & \cdot \left(\int_{\{t: (\varepsilon_0/10)\ell(Q) \leq |t| \leq 10\eta_0\ell(Q)\}} \frac{\sigma}{(\ell(Q) + |t|)^{(d+1)p'}} dt\right)^{1/p'} \leq C. \end{aligned}$$

When we combine this with our two earlier estimates (for “big” and “little” t), and recall that $0 < \delta < \varepsilon_0/100 < 1/10$, we get:

$$\mu\left(\{|x| \leq \delta \ell(Q)\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q} \cdot \left(\int_{\mathbb{R}^d} \frac{\sigma}{(\ell(Q) + |t|)^{(d+1)p'}} dt\right)^{1/p'} \leq C.$$

We are almost done. We have proved that (5.1) implies an inequality in which the integral (‘ σ ’) term is as we claimed, but the ‘ μ ’ term only covers the set $\{x : |x| \leq \delta \ell(Q)\} \times [\ell(Q)/2, \ell(Q)]$, instead of all of $T(Q) = Q \times [\ell(Q)/2, \ell(Q)]$. (Recall that $x_Q = 0$.) We will remedy this by “moving our estimate around,” to cover all of Q .

If we repeat the preceding arguments with some x' in place of 0, and if $|x'| \leq \sqrt{d}\ell(Q)$, then we get

$$C \geq \mu\left(\{x : |x - x'| \leq \delta \ell(Q)\} \times \left[\frac{\ell(Q)}{2}, \ell(Q)\right]\right)^{1/q}$$

$$\begin{aligned} & \cdot \left(\int_{\mathbb{R}^d} \frac{\sigma}{(\ell(Q) + |x' - t|)^{(d+1)p'}} dt \right)^{1/p'} \\ & \geq C \left(\mu \left(\{x : |x - x'| \leq \delta \ell(Q)\} \times \left(\frac{\ell(Q)}{2}, \ell(Q) \right) \right) \right)^{1/q} \\ & \cdot \left(\int_{\mathbb{R}^d} \frac{\sigma}{(\ell(Q) + |t|)^{(d+1)p'}} dt \right)^{1/p'}, \end{aligned}$$

because the σ -integrands for θ and x' are comparable as long as $|x'| \leq \sqrt{d}\ell(Q)$. There are $C(d)$ points $x^{(1)}, \dots, x^{(C(d))}$ such that each $|x^{(i)}| \leq \sqrt{d}\ell(Q)$ and $Q \subset \cup_i \{t : |t - x^{(i)}| < \delta \ell(Q)\}$. When we add the corresponding estimates for these points, we get

$$\mu(T(Q))^{1/q} \left(\int_{\mathbb{R}^d} \frac{\sigma}{(\ell(Q) + |x|)^{(d+1)p'}} dx \right)^{1/p'} \leq C.$$

Theorem 8 is proved. $\square$

REFERENCES

- [GI] L. GRECO & T. IWANIEC, *New inequalities for the Jacobian*, 1993, (Preprint).
- [L1] D. LUECKING, *Forward and reverse Carleson inequalities for functions in the Bergman spaces and their derivatives*, Amer. J. Math. **107** (1985), 85-111.
- [L2] ———, *Embedding derivatives of Hardy spaces into Lebesgue spaces*, Proc. London Math. Soc. **3** (1991), 595-619.
- [P] C. PÉREZ, *Weighted norm inequalities for singular integral operators*, J. London Math. Soc. **49** (1994), 296-308.
- [Sh1] N. A. SHIROKOV, *Some generalizations of the Littlewood-Paley Theorem*, J. Soviet Math. **8** (1977), 119-129.
- [Sh2] ———, *Some embedding theorems for spaces of harmonic functions*, J. Soviet Math. **14** (1980), 1173-1176.
- [SW] E. T. SAWYER & R. L. WHEEDEN, *Carleson conditions for the Poisson integral*, Indiana Univ. Math. J. **40** (1991), 639-676.
- [SWZ] E. T. SAWYER, R. L. WHEEDEN & S. ZHAO, *Weighted norm inequalities for operators of potential type and fractional maximal functions*, Potential Analysis (to appear).
- [St] E. M. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
- [U] A. UCHIYAMA, *Extension of the Hardy-Littlewood-Fefferman-Stein inequality*, Pacific J. Math. **120** (1985), 229-255.
- [Ve] I. E. VERBITSKY, *Imbedding theorems for the spaces of analytic functions with mixed norms*, Acad. Sci. Kishinev, Moldova, 1987, (Preprint).
- [Vi] I. V. VIDENSKII, *On an analogue of Carleson measures*, Soviet Math. Dokl. **37** (1988), 186-190.

- [W1] J. M. WILSON, *A sharp inequality for the square function*, Duke Math. J. **55** (1987), 879-887.
- [W2] ———, *L^p weighted norm inequalities for the continuous square function* $0 < p < 2$, Illinois J. Math. **33** (1989), 361-366.
- [W3] ———, *Weighted norm inequalities for the continuous square function*, Trans. Amer. Math. Soc. **314** (1989), 661-692.
- [W4] ———, *Some two-parameter square function inequalities*, Indiana Univ. Math. J. **40** (1991), 419-442.

The first author was partially supported by NSF grant DMS 93-02991. The second author was partially supported by NSF grant DMS 94-01498.

RICHARD L. WHEEDEN
Department of Mathematics
Rutgers University
New Brunswick, New Jersey 08903

J. MICHAEL WILSON
Department of Mathematics
University of Vermont
Burlington, Vermont 05405

Submitted: January 9, 1995, revised: September 18, 1995.