WEIGHTED TWO-PARAMETER BERGMAN SPACE INEQUALITIES

J. Michael Wilson

Abstract

For f, a function defined on $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, take u to be its biharmonic extension into $\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$. In this paper we prove strong sufficient conditions on measures μ and weights v such that the inequality

$$(*) \quad \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |\nabla_1 \nabla_2 u|^q \, d\mu(x_1, x_2, y_1, y_2) \right)^{1/q}$$
$$\leq \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p v \, dx \right)^{1/p}$$

will hold for all f in a reasonable test class, for $1 < p \leq 2 \leq q < \infty$. Our result generalizes earlier work by R. L. Wheeden and the author on one-parameter harmonic extensions. We also obtain sufficient conditions for analogues of $(*)$ to hold when the entries of $\nabla_1 \nabla_2 u$ are replaced by more general convolutions.

1. Introduction

In an earlier paper [**WhWi**], Richard L. Wheeden and the author studied the following weighted norm inequality for the Poisson integral $u(x, y)$ ($x \in \mathbf{R}^d$, $y > 0$) of a function f:

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

In this inequality, ∇ denotes the full gradient in $\mathbf{R}_+^{d+1}$: $\nabla = (\partial/\partial x_1, \dots, \partial/\partial x_d, \partial/\partial y)$; $\mathbf{R}_+^{d+1}$ is the usual upper half space $\mathbf{R}^d \times (0, \infty)$; μ is a positive Borel measure defined on $\mathbf{R}_+^{d+1}$; and v is a non-negative function in $L_{\text{loc}}^1(\mathbf{R}^d)$. We studied this inequality *primarily* for p and q in the

range $1 < p \leq q < \infty$. For the case in which $q \geq 2$, we proved sufficient conditions on μ and v (depending on p, q, and d) for the inequality (1.1) to hold for all $f \in \cup_{1 \leq r < \infty} L^r(\mathbf{R}^d, dx)$.

The argument in [**WhWi**] began with the observation that (1.1) is a special case of a more general inequality. Let h be a smooth function with decay at infinity (precisely how much decay will be specified later), defined on $\mathbf{R}^d$. For $y > 0$, set $h_y(x) = y^{-d}h(x/y)$, the usual L^1-dilation. If we set $u(x, y) = f * h_y(x)$, then any component of $\nabla u(x, y)$ can be written as $f * (y^{-1}\phi_y)(x)$, where ϕ is smooth, has some decay, and in addition satisfies

$$(1.2) \qquad \int_{\mathbf{R}^d} \phi \, dx = 0.$$

This said, we may now shift our attention to an arbitrary smooth ϕ with decay (how *much*, again, to be specified presently), and which satisfies (1.2), and we may ask: What conditions on μ and v ensure that

$$(1.3) \quad \left(\int_{\mathbf{R}_+^{d+1}} |f * (y^{-1}\phi_y)(x)|^q \, d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbf{R}^d} |f|^p v \, dx \right)^{1/p}$$

holds for all f in our test class?

In this paper, we are concerned with *two-parameter generalizations* of (1.1) and (1.3), and especially the latter. What does "two-parameter" mean? Let $\mathbf{R}^d = \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$. For $i = 1, 2$, let ϕ_i be smooth functions with good decay, defined on $\mathbf{R}^{d_i}$, and which satisfy $\int_{\mathbf{R}^{d_i}} \phi_i \, dx_i = 0$. In our two-parameter problem, we look for sufficient conditions on measures μ, defined on $\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$, and non-negative weights $v \in L^1_{\text{loc}}(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$, which are sufficient for the inequality

$$(1.4) \quad \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |f * [(y_1^{-1}(\phi_1)_{y_1}) \cdot (y_2^{-1}(\phi_2)_{y_2})](x_1, x_2)|^q \, d\mu(x, y) \right)^{1/q}$$

$$\leq \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p v \, dx \right)^{1/p}$$

to hold for all f. (Here we are using '(x, y)' to stand for '(x_1, x_2, y_1, y_2).') When we write $(\phi_i)_{y_i}(x_i)$, we mean, of course, $y_i^{-d_i}\phi_i(x_i/y_i)$. In the case where the ϕ_i's are the kernels that "generate" the components of the Poisson kernel (in their respective upper half spaces!), such a result

would yield a sufficient condition for the inequality

$$(1.5) \quad \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |\nabla_1 \nabla_2 u|^q \, d\mu(x_1, x_2, y_1, y_2) \right)^{1/q}$$

$$\leq \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p v \, dx \right)^{1/p},$$

where u is f's biharmonic extension and ∇_i denotes the full gradient in the (x_i, y_i) variables. Thus, $\nabla_1 \nabla_2 u$ is a $(d_1 + 1) \times (d_2 + 1)$ matrix of functions, and $|\nabla_1 \nabla_2 u|$ can be taken to be the square root of the sum of the squares of its entries.

In this paper we prove sufficient conditions for inequality (1.4), valid for $1 < p \leq 2 \leq q < \infty$ and for a certain class of ϕ_i's. This class includes the kernels that generate the x-derivatives of the Poisson kernels, but not, alas, the y-derivatives. The reason for this troubling gap is that, while the convolution kernels for the x-derivatives of the d-dimensional Poisson kernel decay to order $(1 + |x|)^{-d-2}$, the corresponding y-derivative kernel only decays like $(1 + |x|)^{-d-1}$. Unfortunately, our general one-parameter result (Theorem 1.1 below) requires decay like $(1 + |x|)^{-d-2}$. In [**WhWi**], the authors treated the y-derivative by means a trick combining harmonicity and the Poisson kernel's semigroup property. The whole trick is given on [**WhWi**, pp. 955–959], but in a nutshell it's this. Our duality argument (which works so well with the x-derivatives) requires that we obtain good Littlewood-Paley estimates for a certain function $Tg(x)$, expressed as a weighted integral of $\partial P_y(x - t)/\partial y$ over $(t, y) \in \mathbf{R}_+^{d+1}$. For the x-derivatives, the corresponding integrals involved $\partial P_y(x - t)/\partial x_i$, and we got our Littlewood-Paley estimates by convolving with $\psi_\eta(\cdot)$, where ψ was a smooth, compactly-supported function with cancellation. The extra decay in the $\partial P_y(x - t)/\partial x_i$'s let us bound the resulting integrals nicely. *Lacking* that decay for the y-derivative, R. L. Wheeden and the author convolved Tg with $\partial P_y(\cdot)/\partial y$. By harmonicity and the semigroup property, the resulting integral could be expressed in terms of second partials in the x derivatives —for whose kernels we *do* have good bounds. In our final section we drag this additional argument in to obtain a sufficient (but not so good) condition for the y-derivatives in the bi-space setting as well.

Before stating our main theorem, we should state the one-parameter result from [**WhWi**] that motivated it. Even this earlier result is fairly technical, and the two-parameter result is, in our opinion, liable to be

completely indigestible to a reader who has not seen the one-parameter version first.

The one-parameter result.

As is traditional in this business, we begin with cubes $Q \subset \mathbf{R}^d$. We use $\ell(Q)$ to denote the sidelength of Q, and $|Q|$ is its Lebesgue measure. We denote the Euclidean center of Q by x_Q. By $\hat{Q}$ we mean the set

$$\hat{Q} = \{(x, y) \in \mathbf{R}^{d+1}_+ : x \in Q, \, 0 < y < \ell(Q)\},$$

the so-called "Carleson box" sitting above Q. We use $T(Q)$ to denote the "top half" of $\hat{Q}$:

$$T(Q) = \{(x, y) \in \mathbf{R}^{d+1}_+ : x \in Q, \, \ell(Q)/2 \leq y < \ell(Q)\}.$$

One more definition: If $\eta \geq 0$, $\sigma \in L^1_{\text{loc}}(\mathbf{R}^d)$ is a non-negative weight, and $Q \subset \mathbf{R}^d$ is a cube, we set

$$(1.6) \qquad \sigma^*(Q, \eta) \equiv \int_Q \sigma(x) \log^\eta(e + \sigma(x)/\sigma_Q) \, dx,$$

where $\sigma_Q = (1/|Q|) \int_Q \sigma$, σ's average over Q. Equation (1.6) defines an Orlicz-type norm that shows up in weighted Littlewood-Paley theory [**W1**], [**W2**], and whose properties underlie the results in [**WhWi**] as well as those of the present paper.

Theorem 1.1. *Let m be a non-negative integer. Let $\phi \in \mathcal{C}^\infty(\mathbf{R}^d)$ have $\int \phi = 0$. Let ϕ also satisfy $|\phi(x)| \leq (1 + |x|)^{-d-2-m}$ and $|\nabla \phi(x)| \leq (1 + |x|)^{-d-3-m}$ for all $x \in \mathbf{R}^d$. Let $v \in L^1_{\text{loc}}(\mathbf{R}^d)$ be a non-negative weight and let μ be a positive Borel measure on $\mathbf{R}^{d+1}_+$. Let $1 < p \leq 2 \leq q < \infty$. Set $\sigma = v^{1-p'}$, where p' is the dual exponent to p. Let $\eta > p'/2$. There is a positive constant $C = C(\eta, p, q, d, m)$ such that the following is true: If there exists a weight w satisfying*

$$(1.7) \qquad \sigma^*(Q, \eta) \leq \int_Q w$$

and

$$(1.8) \quad \mu(T(Q))^{1/q} \left(\int_{\mathbf{R}^d} \frac{\log^{p'/q'}(e + |x - x_Q|/\ell(Q)) w(x)}{(\ell(Q) + |x - x_Q|)^{(d+2+m)p'/q'}} \, dx \right)^{1/p'}$$

$$\leq C\ell(Q)^{d+1-(d+2+m)/q'}$$

for all cubes $Q \subset \mathbf{R}^d$, then (1.3) holds for all $f \in \cup_{1 \leq r < \infty} L^r(\mathbf{R}^d, dx)$.

Remark. The reader can see what we mean by indigestibility.

Remark. The theorem, as stated in [**WhWi**], actually gives a sufficient condition for the range $1 < p \leq q < \infty$, with $q \geq 2$. We have stated this limited form of the theorem to make it more closely resemble Theorem 1.3 below. The restriction in Theorem 1.3 comes about because our method of proof, in two parameters, requires $p' \geq 2$. This is related to another difference between Theorem 1.1 and the corresponding result in [**WhWi**]. The theorem in [**WhWi**] does not contain the hypothesis (1.7). Rather, it speaks of pairs of weights ('p'-pairs') (σ, w) for which w also satisfies (1.8). However, as is pointed out in [**WhWi**, p. 949] and in [**W1**], a pair that satisfies (1.7) *is* a p'-pair. Unfortunately, we have no good characterization of p'-pairs (for $p' \neq 2$) in the two-parameter setting. We express Theorem 1.1 in this fashion in order to make its statement look more like those of Theorem 1.3 and Theorem 5.3 (see below).

Remark. If σ belongs to the Muckenhoupt A_∞ class, then (1.7) holds for $w = c\sigma$, where c depends on η, d, and the A_∞ "box specs" of σ. In that case, (1.8) amounts to saying that μ and σ cannot put too much mass too near any cube Q. Since σ is big when v is small, this is a quantitative way of saying that v cannot be too small near points where μ is "large". Theorem 1.1 is a restatement of this fact for v's whose corresponding σ's are not in A_∞.

The two-parameter result.

We begin here with rectangles $R = Q_1 \times Q_2$, where the Q_i are cubes in $\mathbf{R}^{d_i}$. We use $|R|$ to mean the Lebesgue measure of R. We set $T(R) = T(Q_1) \times T(Q_2)$ and $\hat{R} = \hat{Q}_1 \times \hat{Q}_2$, where $T(Q_i)$ and $\hat{Q}_i$ are as defined above.

We will be using the next definition so often that it merits its own formal statement:

Definition 1.2. Let $\eta \geq 0$ be a number and let $\sigma \in L^1_{\mathrm{loc}}(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$ be a non-negative weight. If $R \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ is rectangle, we set

$$\sigma(R, \eta) = \int_R \sigma(x) \log^\eta(e + \sigma(x)/\sigma_R) \, dx,$$

where $\sigma_R = (1/|R|) \int_R \sigma$ denotes σ's average over R.

Remark. The only difference between Definition 1.2 and the one given earlier is that Definition 1.2 applies to rectangles.

Our main result, which we prove in Section 4, is:

Theorem 1.3. *Let m_1 and m_2 be non-negative integers. For $i = 1, 2$, let $\phi_i \in \mathcal{C}^\infty(\mathbf{R}^{d_i})$ have $\int \phi_i = 0$. Let the ϕ_i also satisfy $|\phi_i(x_i)| \leq (1 + |x_i|)^{-d_i - 2 - m_i}$ and $|\nabla\phi_i(x_i)| \leq (1 + |x_i|)^{-d_i - 3 - m_i}$ for all $x_i \in \mathbf{R}^{d_i}$. Let $v \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$ be a non-negative weight and let μ be a positive Borel measure on $\mathbf{R}^{d_1 + 1}_+ \times \mathbf{R}^{d_2 + 1}_+$. Let $1 < p \leq 2 \leq q < \infty$. Set $\sigma = v^{1 - p'}$, where p' is the dual exponent to p. Let $\eta > p'$ and let $\epsilon > 0$. There is a positive constant C,*

$$C = C(\eta, \epsilon, p, q, d_1, d_2, m_1, m_2),$$

such that the following is true: If there exists a weight w satisfying

$$(1.9) \qquad \sigma(R, \eta) \leq \int_R w$$

and

$$(1.10) \quad \mu(T(R))^{1/q} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left[\frac{1}{(\ell(Q_1) + |x_1 - x_{Q_1}|)^{(d_1 + 2 + m_1 - \epsilon)p'/q'}} \right. \right.$$

$$\times \frac{1}{(\ell(Q_2) + |x_2 - x_{Q_2}|)^{(d_2 + 2 + m_2 - \epsilon)p'/q'}}$$

$$\left. \left. \times w(x) \right] dx \right)^{1/p'}$$

$$\leq C\ell(Q_1)^{d_1 + 1 - (d_1 + 2 + m_1 - \epsilon)/q'} \ell(Q_2)^{d_2 + 1 - (d_2 + 2 + m_2 - \epsilon)/q'}$$

for all rectangles $R = Q_1 \times Q_2$, then (1.4) holds for all $f \in \cup_{1 \leq r < \infty} L^r(\mathbf{R}^d, dx)$.

Remark. Note the absence of log's in the numerator and the extra ϵ's in the denominator of the two-parameter condition (1.10).

The rest of the paper is laid out as follows. In Section 2 we state and prove certain results from weighted Littlewood-Paley theory which we will need in the proof of Theorem 1.3. In Section 3 we state a technical result from [**WhWi**], concerning convolutions of smooth functions with specified amounts of decay and cancellation, and we apply this result to prove a lemma (Lemma 3.2). Lemma 3.2 is a pointwise substitute for a series of integral inequalities used in [**WhWi**] to prove Theorem 1.1. This pointwise result is part of what lets us prove our two-parameter theorem without having a full-blooded, two-parameter weighted-norm theory of the Littlewood-Paley square function; it is also where the extra ϵ's in (1.10) will come from. In Section 4 we prove Theorem 1.3. In Section 5 we state and prove a sufficiency result for the biharmonic Poisson kernel.

2. Littlewood-Paley theory

The basis of all of our arguments is the Calderón-Torchinsky decomposition lemma. Let ψ_i ($i = 1, 2$) be real, radial, $\mathcal{C}^\infty$ functions defined on $\mathbf{R}^{d_i}$, that satisfy:

1) $\int \psi_i = 0$;
2) $\operatorname{supp} \psi_i \subset \{x_i : |x_i| \leq 1\} \subset \mathbf{R}^{d_i}$;
3) for any $\xi \in \mathbf{R}^{d_i} \setminus \{0\}$,

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

For $y_i > 0$, we let $(\psi_i)_{y_i}(x_i) = y_i^{-d_i} \psi_i(x_i/y_i)$. If y_1 and y_2 are positive numbers and $x = (x_1, x_2) \in \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, we define $y = (y_1, y_2)$ and set $\Psi_y(x) = (\psi_1)_{y_1}(x_1) \cdot (\psi_2)_{y_2}(x_2)$. The Calderón-Torchinsky lemma consists in the following observation: if $f \in L^2(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$, then, by Fourier inversion,

$$(2.1) \qquad f(x) = \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} (f * \Psi_y(t)) \cdot \Psi_y(x - t) \frac{dt_1\, dt_2\, dy_1\, dy_2}{y_1 y_2}$$

as a distribution [**CF**].

It is easy to show that, for $f \in L^2$, the (vector-valued) integral (2.1) actually converges to f in the L^2 norm. If f is smooth and decays rapidly at infinity, then the integral (2.1) converges uniformly and pointwise, and can be cut up and rearranged at will. We will use this freedom in the following way. Let $R = Q_1 \times Q_2 \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ be a double-dyadic rectangle, that is, a Cartesian product of dyadic cubes $Q_i \subset \mathbf{R}^{d_i}$, and let $T(R) = T(Q_1) \times T(Q_2)$ be the corresponding "top half" of its two-parameter Carleson box, as defined above. (It is important to note that the family $\{T(R)\}_R$ tiles $\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$.) With suitable (and quite weak) hypotheses on f, we may re-write the integral formula (2.1) as a sum:

$$f = \sum_R \int_{T(R)} (f * \Psi_y(t)) \cdot \Psi_y(x - t) \frac{dt_1\, dt_2\, dy_1\, dy_2}{y_1 y_2} = \sum_R b_R(x).$$

Each of these functions b_R has support contained in $\tilde{R}$ (the concentric triple of R), is smooth (it inherits this from the ψ_i's), and has cancellation in the x_1 and x_2 directions; that is to say, for each fixed $x_1^* \in \mathbf{R}^{d_1}$,

$$\int_{\mathbf{R}^{d_2}} b(x_1^*, t)\, dt = 0,$$

and analogously for each fixed $x_2^* \in \mathbf{R}^{d_2}$; this cancellation property is also inherited from the ψ_i's.

The meaning of the Calderón-Torchinsky lemma is that any (essentially arbitrary) function can be written as a sum of smooth, compactly supported functions that have cancellation. We can go further. Let us say that a function $a_R(x)$ is *adapted* to a rectangle $R = Q_1 \times Q_2 \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ if:

a) $\operatorname{supp} a_R \subset R$;

b) a_R is infinitely differentiable;

c) for each $x_1^* \in \mathbf{R}^{d_1}$, $\|\nabla_{x_2} a_R(x_1^*, \cdot)\|_\infty \leq \ell(Q_2)^{-1}|R|^{-1/2}$;

d) for each $x_2^* \in \mathbf{R}^{d_2}$, $\|\nabla_{x_1} a_R(\cdot, x_2^*)\|_\infty \leq \ell(Q_1)^{-1}|R|^{-1/2}$;

e) $\|\nabla_{x_1}\nabla_{x_2} a_R\|_\infty \leq \ell(Q_1)^{-1}\ell(Q_2)^{-1}|R|^{-1/2}$;

f) for each $x_1^* \in \mathbf{R}^{d_1}$, $\int_{\mathbf{R}^{d_2}} a_R(x_1^*, t)\, dt = 0$;

g) for each $x_2^* \in \mathbf{R}^{d_2}$, $\int_{\mathbf{R}^{d_1}} a_R(t, x_2^*)\, dt = 0$.

Each of the functions b_R obtained above can be expressed as $\lambda_{\tilde{R}} a_{\tilde{R}}$, where each $a_{\tilde{R}}$ is adapted to $\tilde{R}$, and the $\lambda_{\tilde{R}}$'s are complex numbers satisfying:

$$|\lambda_{\tilde{R}}| \leq C \left(\int_{T(R)} |f * \Psi_y(t)|^2 \frac{dt_1\, dt_2\, dy_1\, dy_2}{y_1 y_2} \right)^{1/2},$$

for some constant C that depends on Ψ (which, recall, depends on d_1 and d_2) but not on f.

Let us say that a function f is in *standard form* if there is a finite family, $\mathcal{G}$, of *triples of double-dyadic rectangles*, such that

$$f(x) = \sum_{R \in \mathcal{G}} \lambda_R a_R(x),$$

where the λ_R's are real numbers and each a_R is adapted to R. (Notice that the 'tildes' have been "absorbed" into the R's.)

We will use Littlewood-Paley theory to bound certain functions in standard form on weighted spaces. We will measure the "badness" of our weights via the Orlicz-type norm $\sigma(R, \eta)$ given in Definition 1.2. When $\eta > 0$, the ratio $\sigma(R, \eta) / \int_R \sigma$ measures the extent to which σ's mass gets concentrated on a small part of R (that is, a subset with small Lebesgue measure compared to $|R|$). In particular, the ratio is uniformly bounded (for any $\eta > 0$) if and only if σ is a two-parameter A_∞ weight.

The proof of Theorem 1.3 depends on this result from [**W2**]:

Theorem 2.1. *Let $\eta > 2$. There is a constant $C = C(\eta, d_1, d_2)$ so that the following holds: If $\sigma \in L^1_{\mathrm{loc}}(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$ is any non-negative weight and $f = \sum_{R \in \mathcal{G}} \lambda_R a_R$ is any function in standard form, then*

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^2 \sigma \, dx \leq C \sum_{R \in \mathcal{G}} \frac{|\lambda_R|^2}{|R|} \sigma(R, \eta).$$

Theorem 2.1 has an immediate consequence. For $f = \sum_{R \in \mathcal{G}} \lambda_R a_R$ in standard form, set

$$\tilde{S}(f)(x) = \left(\sum_{R \in \mathcal{G}} \frac{|\lambda_R|^2}{|R|} \chi_R(x) \right)^{1/2}.$$

(This is one of many variants of the Lusin square function.) The next corollary follows by rearranging sums.

Corollary 2.2. *Let σ and w be weights such that, for some $\eta > 2$, $\sigma(R, \eta) \leq \int_R w$ for all rectangles $R \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$. For any f in standard form,*

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^2 \sigma \, dx \leq C(\eta, d_1, d_2) \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \tilde{S}^2(f) w \, dx.$$

In the one-parameter setting, both Theorem 2.1 and Corollary 2.2 have L^p analogues for $p \neq 2$ [**W1**]. Precisely, by applying the one-parameter version of the Calderón-Torchinsky lemma, we can write an essentially arbitrary f as a sum $f = \sum_Q \lambda_Q a_{(Q)}$, indexed over the dyadic cubes $Q \subset \mathbf{R}^d$, where the λ_Q's are numbers and the $a_{(Q)}$'s are smooth functions satisfying:

   a) $\mathrm{supp} \, a_{(Q)} \subset \tilde{Q}$, the concentric triple of Q;

   b) $\|\nabla a_{(Q)}\|_\infty \leq \ell(Q)^{-1} |Q|^{-1/2}$;

   c) $\int a_{(Q)} = 0$.

We define an analogous one-parameter square function:

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

Now let $\eta > p/2$ $(0 < p < \infty)$, and suppose that σ and w are two weights in $L^1_{\mathrm{loc}}(\mathbf{R}^d)$ satisfying

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

for all cubes $Q \subset \mathbf{R}^d$. Then, for all reasonable f (say, $f \in \cup_{1<r<\infty} L^r$),

$$\int |f|^p \sigma \, dx \leq C(p, \eta, d) \int (S(f))^p w \, dx.$$

In the two parameter setting, the appropriate theorem would be that if $\eta > p$, and σ and w are two weights such that $\sigma(R, \eta) \leq w(R)$ for all rectangles R, then

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p \sigma \, dx \leq C(\eta, p, d_1, d_2) \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} (\tilde{S}(f))^p w \, dx$$

for all f in standard form.

Unfortunately, this is not known to be the case (yet) in the two-parameter context. Now, we obviously need *some* L^p estimates to prove our main result. However, the L^p estimates we need do not *have* to be the precise analogues of the one given in Corollary 2.2. This saving fact lets us get around the hole in our theory by means of a **trick**. The author introduced this device in the context of two-parameter martingales in [**W3**], and we apply it with essentially no change here. The only difference is that, in [**W3**], we used it to estimate linear sums of two-parameter Haar functions, whereas here we are applying it to linear sums of two-parameter (i.e., rectangle) adapted functions, as defined above.

The **trick** yields:

Theorem 2.3. *Let $r \geq 2$ and let $\eta > r$. There is a constant $C = C(\eta, r, d_1, d_2)$ such that the following holds: If $\sigma \in L^1_{\mathrm{loc}}(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$ is a non-negative weight and $f = \sum_{R \in \mathcal{G}} \lambda_R a_R$ is a function in standard form, then*

$$\int |f|^r \sigma \, dx \leq C \left(\sum_{R \in \mathcal{G}} \frac{|\lambda_R|^2}{|R|} \sigma(R, \eta)^{2/r} \right)^{r/2}.$$

Proof of Theorem 2.3: Let $\tilde{s} = (r/2)'$, the dual exponent to $r/2$ (which, recall, is ≥ 1). Let h be a non-negative, measurable function defined on $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, such that $\|h\|_{L^{\tilde{s}}(\sigma)} = 1$ and

$$(2.2) \qquad \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^r \sigma \, dx = \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^2 h \, \sigma \, dx \right)^{r/2}.$$

Let $\tilde{\alpha} = 2\eta/r > 2$. Define $w \equiv h\sigma$. According to Theorem 2.1, the right-hand side of (2.2) is less than or equal to

$$(2.3) \qquad C \left(\sum_{R \in \mathcal{G}} \frac{|\lambda_R|^2}{|R|} w(R, \tilde{\alpha}) \right)^{r/2}.$$

Let us now consider one of the terms $w(R, \tilde{\alpha})$. By following the argument from [**W3**] (which is essentially Young's Inequality[1]) we see that $w(R, \tilde{\alpha})$ is bounded above by a positive constant times

$$(2.4) \qquad \int_R w(x)\phi^{(R)}(x)\,dx,$$

where $\phi^{(R)}$ is a positive function satisfying

$$(2.5) \qquad \frac{1}{|R|} \int_R \exp([\phi^{(R)}(x)]^{1/\tilde{\alpha}}) \le 6.$$

Now let's apply Hölder's Inequality to (2.4). We get:

$$\int_R w(x)\phi^{(R)}(x)\,dx = \int_R h(x)\sigma(x)\phi^{(R)}(x)\,dx$$

$$\le \|h\|_{L^{\tilde{s}}(\sigma)} \cdot \left(\int_{\mathbf{R}} [\phi^{(R)}(x)]^{r/2}\sigma\,dx \right)^{2/r}$$

$$\le \left(\int_{\mathbf{R}} [\phi^{(R)}(x)]^{r/2}\sigma\,dx \right)^{2/r}.$$

(What we just did is the beginning of the **trick** we mentioned before the statement of the theorem.) Define $\psi(x) = [\phi^{(R)}(x)]^{r/2}$. Because of (2.5), the function ψ satisfies

$$\frac{1}{|R|} \int_R \exp(\psi^{2/(r\tilde{\alpha})})\,dx = \frac{1}{|R|} \int_R \exp(\psi^{(1/\eta)})\,dx$$

$$\le 6.$$

---

[1] Here we are applying the Young's Inequality that deals with pairs of so-called "complementary" functions, not the more familiar theorem on L^p estimates for convolutions. We refer the reader to [**St**, p. 358] for a fuller discussion of this topic.

But now, a second application of our Young's Inequality argument implies that

$$w(R, \tilde{\alpha}) \leq \left(\int_R \psi(x)\sigma \, dx \right)^{2/r}$$

(2.6)

$$\leq C\sigma(R, \eta)^{2/r}.$$

This finishes the **trick**.

Plugging (2.6) into (2.3) yields the result.   □

It might be helpful here if we explain how we will use Theorem 2.3. The reader of [**WhWi**] will recall how, in that paper, inequality (1.3) was treated by writing the kernel ϕ as a sum of a ψ_1 and a ψ_2, where ψ_1 had compact support and integral equal to 0; and ψ_2, while not compactly supported, had many moments of cancellation. The analogous inequalities (1.3) for ψ_1 and ψ_2 were treated by different arguments. In the two-parameter setting, we get *four* inequalities like (1.4). One of these —that in which both kernels are compactly supported— will be handled as a direct consequence of Theorem 2.3. The other three will require more subtlety, but their treatment will follow the basic idea of Theorem 2.3.

3. Two technical estimates

The proof of Theorem 1.3 depends on certain precise estimates on the convolutions of smooth kernels that have cancellation. These estimates are stated in the following (highly technical) lemma, whose proof can be found in [**WhWi**, pp. 939–941].

Lemma 3.1. *Let ψ_i and ϕ_i belong to $\mathcal{C}^\infty(\mathbf{R}^{d_i})$ $(i = 1, 2)$. Assume that each ψ_i has support contained in $\{|x| \leq 1\}$ and satisfies $\int \psi_i = 0$. Furthermore, suppose that, for some non-negative integer m_i, and for all $x_i \in \mathbf{R}^{d_i}$,*

$$|\phi_i(x_i)| \leq (1 + |x_i|)^{-d_i - m_i - 2}$$

$$|\nabla \phi_i(x_i)| \leq (1 + |x_i|)^{-d_i - m_i - 3},$$

and that $\int_{\mathbf{R}^{d_i}} \phi_i(x_i)P(x_i) \, dx_i = 0$ for all polynomials of degree $\leq m_i + 1$. Then the following estimates hold for the convolutions $(\psi_i)_y * (\phi_i)_\eta(x_i)$, for all $x_i \in \mathbf{R}^{d_i}$ and positive numbers y and η:

a) *if* $\eta \geq y$,

$$|(\psi_i)_y * (\phi_i)_\eta(x_i)| \leq \frac{Cy\eta^{m_i+2}}{(\eta + |x_i|)^{d_i+m_i+3}};$$

b) *if* $\eta \leq y$ *and* $|x_i| \geq 5y$,

$$|(\psi_i)_y * (\phi_i)_\eta(x_i)| \leq \frac{Cy\eta^{m_i+2}}{(\eta + |x_i|)^{d_i+m_i+3}};$$

c) *if* $\eta \leq y$ *and* $|x_i| \leq 5y$,

$$|(\psi_i)_y * (\phi_i)_\eta(x_i)| \leq \frac{C\eta^{m_i+2}\log(e + y/\eta)}{y^{d_i+m_i+2}};$$

for constants $C = C_i$ *only depending on the* ϕ_i's*,* ψ_i's*,* m_i's*, and* d_i's*.*

We will be applying Lemma 3.1 in the following, very specific way. Let Q_i and Q_i' be dyadic cubes in $\mathbf{R}^{d_i}$. Call the center of Q_i (resp., Q_i'), x_{Q_i} (resp., $x_{Q_i'}$). Fix $(x_i, y_i) \in T(Q_i')$ and $(t_i, \eta_i) \in T(Q_i)$ (these assumptions force $y_i \sim \ell(Q_i')$ and $\eta_i \sim \ell(Q_i)$). If ψ_i and ϕ_i satisfy the hypotheses of Lemma 3.1 for some m_i, then

$$(3.1) \quad \eta_i^{-1}|(\psi_i)_{y_i} * (\phi_i)_{\eta_i}(t_i - x_i)|$$

$$\leq C\left(\frac{\ell(Q_i)^{m_i+1}}{\ell(Q_i')^{d_i+m_i+2}}\right) \cdot \log(e + \ell(Q_i')/\ell(Q_i)),$$

if $Q_i \subset \tilde{Q}_i'$; and

$$(3.2) \quad \eta_i^{-1}|(\psi_i)_{y_i} * (\phi_i)_{\eta_i}(t_i - x_i)| \leq C\frac{\ell(Q_i')\ell(Q_i)^{m_i+1}}{(\ell(Q_i) + |x_{Q_i} - x_{Q_i'}|)^{d_i+m_i+3}},$$

if $Q_i \not\subset \tilde{Q}_i'$. Inequality (3.1) follows from statement c) in Lemma 3.1 and inequality (3.2) follows from a) and b). We will be seeing a lot of (3.1) and (3.2). Therefore, let us define, for dyadic cubes Q_i' and Q_i in $\mathbf{R}^{d_i}$, and non-negative integers m_i:

$$a_i(Q_i', Q_i) = \begin{cases} \left(\dfrac{\ell(Q_i)^{m_i+1}}{\ell(Q_i')^{d_i+m_i+2}}\right) \cdot \log(e + \ell(Q_i')/\ell(Q_i)) & \text{if } Q_i \subset \tilde{Q}_i'; \\[2ex] \dfrac{\ell(Q_i')\ell(Q_i)^{m_i+1}}{(\ell(Q_i) + |x_{Q_i} - x_{Q_i'}|)^{d_i+m_i+3}} & \text{if } Q_i \not\subset \tilde{Q}_i'. \end{cases}$$

The next lemma is important:

Lemma 3.2. *With $a_i(Q'_i, Q_i)$ $(i = 1, 2)$ as defined above, let*

$$A_i(Q'_i, Q_i) = a_i(Q'_i, Q_i) \cdot |\hat{Q}_i|.$$

Let $\gamma > 0$, $0 < \epsilon < 1$, and let k be an integer. There is a constant $C = C(\gamma, \epsilon, d_i, m_i)$ such that, for all $x_i \in \mathbf{R}^{d_i}$, all cubes $Q_i \subset \mathbf{R}^{d_i}$, and all k,

$$(3.3) \qquad \sum_{Q'_i : \ell(Q'_i) = 2^k \ell(Q_i)} [A_i(Q'_i, Q_i)]^\gamma \, \chi_{\tilde{Q}'_i}(x_i)$$

$$\leq C(1 + |k|)^\gamma 2^{-|k|\epsilon\gamma} \ell(Q_i)^{(d_i + m_i + 2 - \epsilon)\gamma} \left(\frac{1}{(\ell(Q_i) + |x_i - x_{Q_i}|)^{d_i + m_i + 2 - \epsilon}} \right)^\gamma.$$

Proof of Lemma 3.2: For each fixed k, no point is in more than $C(d_i)$ cubes $\tilde{Q}'_i$. Thus, it is enough to show that, if $\ell(Q'_i) = 2^k \ell(Q_i)$, then

$$(3.4) \quad A_i(Q'_i, Q_i) \chi_{\tilde{Q}'_i}(x_i)$$

$$\leq C(1 + |k|) 2^{-|k|\epsilon} \ell(Q_i)^{d_i + m_i + 2 - \epsilon} \left(\frac{1}{(\ell(Q_i) + |x_i - x_{Q_i}|)^{d_i + m_i + 2 - \epsilon}} \right).$$

We consider two cases: $k \leq 0$ and $k > 0$.

$k \leq 0$. In this case, $\ell(Q_i) + |x_{Q'_i} - x_{Q_i}|$ is comparable to $\ell(Q_i) + |x_i - x_{Q_i}|$, and we have that

$$A(Q'_i, Q_i) = \frac{\ell(Q'_i)\ell(Q_i)^{d_i + m_i + 2}}{(\ell(Q_i) + |x_{Q_i} - x_{Q'_i}|)^{d_i + m_i + 3}}$$

$$= 2^{-|k|} \frac{\ell(Q_i)^{d_i + m_i + 3}}{(\ell(Q_i) + |x_{Q_i} - x_{Q'_i}|)^{d_i + m_i + 3}},$$

so the inequality is trivial.

$k > 0$. If $Q_i \subset \tilde{Q}'_i$ then

$$A(Q'_i, Q_i) = \left(\frac{\ell(Q_i)}{\ell(Q'_i)} \right)^{d_i + m_i + 2} \log(e + \ell(Q'_i)/\ell(Q_i))$$

$$\leq C 2^{-k\epsilon} \left(\frac{\ell(Q_i)}{\ell(Q'_i)} \right)^{d_i + m_i + 2 - \epsilon} \log(e + \ell(Q'_i)/\ell(Q_i)).$$

It is obvious that

$$e + \ell(Q_i')/\ell(Q_i) \leq C2^k;$$

therefore

$$\log(e + \ell(Q_i')/\ell(Q_i)) \leq C(1 + k).$$

On the other hand, if $x_i \in \tilde{Q}_i'$, then $\ell(Q_i)$ and $|x_{Q_i} - x_i|$ are both $\leq C\ell(Q_i')$. Therefore

$$A(Q_i', Q_i) \leq C(1 + k)2^{-k\epsilon} \left(\frac{\ell(Q_i)}{(\ell(Q_i) + |x_{Q_i} - x_i|)} \right)^{d_i + m_i + 2 - \epsilon},$$

which is what we wanted.

If $Q_i \not\subset \tilde{Q}_i'$ and $x_i \in \tilde{Q}_i'$, then $\ell(Q_i) + |x_i - x_{Q_i}| \leq C(\ell(Q_i) + |x_{Q_i'} - x_{Q_i}|)$, and the latter quantity is $\geq C2^k \ell(Q_i) = C\ell(Q_i')$. Thus:

$$A(Q_i', Q_i) = \frac{\ell(Q_i')\ell(Q_i)^{d_i + m_i + 2}}{(\ell(Q_i) + |x_{Q_i} - x_{Q_i'}|)^{d_i + m_i + 3}}$$

$$\leq C \frac{\ell(Q_i)^{d_i + m_i + 2}}{(\ell(Q_i) + |x_{Q_i} - x_{Q_i'}|)^{d_i + m_i + 2}}$$

$$\leq C2^{-k\epsilon} \frac{\ell(Q_i)^{d_i + m_i + 2 - \epsilon}}{(\ell(Q_i) + |x_{Q_i} - x_{Q_i'}|)^{d_i + m_i + 2 - \epsilon}}$$

$$\leq C2^{-k\epsilon} \frac{\ell(Q_i)^{d_i + m_i + 2 - \epsilon}}{(\ell(Q_i) + |x_{Q_i} - x_i|)^{d_i + m_i + 2 - \epsilon}}. \qquad \square$$

Definition 3.3. For $R = Q_1 \times Q_2$ and $R' = Q_1' \times Q_2'$, rectangles in $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, set

$$\beta(R', R) = a_1(Q_1', Q_1) \cdot a_2(Q_2', Q_2)$$

$$B(R', R) = \beta(R', R) \cdot |\hat{Q}_1| \cdot |\hat{Q}_2|$$

$$= \beta(R', R) \cdot |\hat{R}|$$

$$= A_1(Q_1', Q_1) \cdot A_2(Q_2', Q_2).$$

The next corollary follows by iterating Lemma 3.2:

Corollary 3.4. *Let $\gamma > 0$ and $0 < \epsilon < 1$. There is a constant $C = C(\gamma, \epsilon, d_1, d_2, m_1, m_2)$ so that, for all rectangles $R = Q_1 \times Q_2 \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ and all integers k_1 and k_2,*

$$\sum_{R' = Q_1' \times Q_2' \times \ell(Q_i') = 2^{k_i}\ell(Q_i)} B(R', R)^\gamma \chi_{\tilde{R}'}(x_1, x_2)$$

$$\leq C \times (1 + |k_1|)^\gamma 2^{-|k_1|\epsilon\gamma} \times (1 + |k_2|)^\gamma 2^{-|k_2|\epsilon\gamma}$$

$$\times \ell(Q_1)^{(d_1 + m_1 + 2 - \epsilon)\gamma} \times \ell(Q_2)^{(d_2 + m_2 + 2 - \epsilon)\gamma}$$

$$\times \left(\frac{1}{(\ell(Q_1) + |x_1 - x_{Q_1}|)^{d_1 + m_1 + 2 - \epsilon}(\ell(Q_2) + |x_2 - x_{Q_2}|)^{d_2 + m_2 + 2 - \epsilon}} \right)^\gamma.$$

4. Proof of Theorem 1.3

We rephrase our weighted norm inequality in a dual form. Set $\sigma = v^{1-p'}$. Let ϕ_1 and ϕ_2 satisfy the respective hypotheses of Theorem 1.3. If $g \colon \mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1} \mapsto \mathbf{C}$ is bounded, Borel measurable, and compactly supported, we define:

$$\tilde{T}g(x_1, x_2) = \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} g(t_1, t_2, \eta_1, \eta_2)$$

$$\times [\eta_1^{-1}(\phi_1)_{\eta_1}(t_1 - x_1)\eta_2^{-1}(\phi_2)_{\eta_2}(t_2 - x_2)] \, d\mu(t_1, t_2, \eta_1, \eta_2).$$

This integral converges absolutely for all $x \in \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ because of our special assumptions on g (note that the support of g stays away from $\partial(\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1})$). The operator $\tilde{T}$ is the adjoint of the operator that takes f into

$$f * \left[(y_1^{-1}(\phi_1)_{y_1}) \cdot (y_2^{-1}(\phi_2)_{y_2}) \right] (x_1, x_2).$$

Inequality (1.4) will hold for all $f \in \cup_{1 \leq r < \infty} L^r(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}, dx)$ if

$$\left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}g(x)|^{p'} \sigma \, dx \right)^{1/p'} \leq \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |g(t, y)|^{q'} \, d\mu(t, y) \right)^{1/q'}$$

for all these g. We will prove Theorem 1.3 by showing that that is what happens (given hypotheses (1.9) and (1.10)).

For $i = 1, 2$, we can write

(4.1) $$\phi_i = \rho_i^{(1)} + \rho_i^{(2)},$$

where $\operatorname{supp} \rho_i^{(1)} \subset \{x_i : |x_i| \le 1\} \subset \mathbf{R}^{d_i}$, each $\int_{\mathbf{R}^{d_i}} \rho_i^{(j)} = 0$, and $\int_{R^{d_i}} \rho_i^{(2)} P_i(x) \, dx = 0$ for all polynomials P_i (in the x_i variables) of degree $\le m_i + 1$; we do this by, essentially, throwing $m_i + 1$ of ϕ_i's moments "onto" $\rho_i^{(1)}$. When we do this, the functions $\rho_i^{(1)}$ get one good property (compact support), while the non-compactly supported $\rho_i^{(2)}$'s get lots of cancellation. Using our decompositon (4.1), we may write $\tilde{T}g$ as a sum of four terms:

$$\tilde{T}g = \sum_{k,j=1}^{2} \tilde{T}^{(k,j)}g,$$

where

$$\tilde{T}^{(k,j)}g(x_1, x_2) = \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} g(t_1, t_2, \theta_1, \theta_2)$$

$$\times [\theta_1^{-1}(\rho_1^{(k)})_{\theta_1}(t_1 - x_1)\theta_2^{-1}(\rho_2^{(j)})_{\theta_2}(t_2 - x_2)] \, d\mu(t_1, t_2, \theta_1, \theta_2).$$

Now, the piece $\tilde{T}^{(1,1)}g$, from its very formulation, is equal to a function in standard form. We can dispose of it quickly. We write:

(4.2)
$$\tilde{T}^{(1,1)}g(x_1, x_2) = \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} g(t_1, t_2, \theta_1, \theta_2)$$

$$\times [\theta_1^{-1}(\rho_1^{(1)})_{\theta_1}(t_1 - x_1)\theta_2^{-1}(\rho_2^{(1)})_{\theta_2}(t_2 - x_2)] \, d\mu(t_1, t_2, \theta_1, \theta_2)$$

$$= \sum_R \int_{T(R)} g(t_1, t_2, \theta_1, \theta_2)$$

$$\times [\theta_1^{-1}(\rho_1^{(1)})_{\theta_1}(t_1 - x_1)\theta_2^{-1}(\rho_2^{(1)})_{\theta_2}(t_2 - x_2)] \, d\mu(t_1, t_2, \theta_1, \theta_2).$$

The sum is over all double-dyadic rectangles R, but only finitely many terms are not identically zero, because g and the $\rho_i^{(1)}$'s have compact supports. It is clear that each summand in (4.2), as a function of x, has support contained in its respective $\tilde{R}$. These functions also inherit smoothness and cancellation from the $\rho_i^{(1)}$'s. Thus we may write the

sum as

$$\sum_R \lambda_{\tilde{R}} b_{\tilde{R}}(x_1, x_2),$$

where each $b_{\tilde{R}}$ is adapted to $\tilde{R}$ and the $\lambda_{\tilde{R}}$'s satisfy

$$|\lambda_{\tilde{R}}| \le C \left(\int_{T(R)} |g| \, d\mu(t, y) \right) \ell(Q_1)^{-1} \ell(Q_2)^{-1} |R|^{-1/2}$$

$$\le C \left(\int_{T(R)} |g|^{q'} \, d\mu(t, y) \right)^{1/q'} \mu(T(R))^{1/q} \ell(Q_1)^{-1} \ell(Q_2)^{-1} |R|^{-1/2}$$

with a constant C that depends on the d_i's and the $\rho_i^{(1)}$'s.

Take $\eta > p'$, as in the hypotheses of Theorem 1.3, and suppose that w is a weight satisfying (1.9) for all rectangles R. By Theorem 2.3,

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(1,1)} g|^{p'} \sigma \, dx \le C \left(\sum_R \frac{|\lambda_{\tilde{R}}|^2}{|R|} \sigma(\tilde{R}, \eta)^{2/p'} \right)^{p'/2}$$

$$\le C \left(\sum_R \frac{|\lambda_{\tilde{R}}|^2}{|R|} w(\tilde{R})^{2/p'} \right)^{p'/2}.$$

Since $q' \le 2$, the last quantity is less than or equal to

$$(4.3) \qquad C \left(\sum_R \frac{|\lambda_{\tilde{R}}|^{q'}}{|R|^{q'/2}} w(\tilde{R})^{q'/p'} \right)^{p'/q'}.$$

The hypothesis (1.10) on w implies (after an elementary estimate)

$$\mu(T(R))^{q'/q} w(\tilde{R})^{q'/p'} \ell(Q_1)^{-q'} \ell(Q_2)^{-q'} |\tilde{R}|^{-q'} \le C.$$

Therefore, our bound on $\lambda_{\tilde{R}}$ implies that (4.3) is less than or equal to

$$C \left(\sum_R \left[\int_{T(R)} |g|^{q'} \, d\mu(t, y) \right] \right)^{p'/q'} = C \left(\int_{\mathbf{R}^{d_1+1}_+ \times \mathbf{R}^{d_2+1}_+} |g|^{q'} \, d\mu(t, y) \right)^{p'/q'},$$

which is exactly what we want. Thus, the $\tilde{T}^{(1,1)} g$ term is okay.

The terms $\tilde{T}^{(1,2)} g$, $\tilde{T}^{(2,1)} g$, and $\tilde{T}^{(2,2)} g$ involve non-compactly-supported kernels, and require different arguments. This is where we will use Lemma 3.1. It is obvious that $\tilde{T}^{(1,2)} g$ and $\tilde{T}^{(2,1)} g$ are the same kind of animal, and so we need only treat one of them. It will turn out

that the argument that handles $\tilde{T}^{(2,2)}g$ can also be used, with minor modifications, on $\tilde{T}^{(1,2)}g$. Therefore we shall deal with $\tilde{T}^{(2,2)}g$ first.

Our argument is modeled closely on that of [**WhWi**]. Let κ be the dual exponent to $p'/2$ (which, recall, is ≥ 1), and let $h \in L^\kappa(\sigma)$ be non-negative, satisfy $\|h\|_{L^\kappa(\sigma)} = 1$, and be chosen so that

$$(4.4) \qquad \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(2,2)}g|^{p'}\sigma \, dx = \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(2,2)}g|^2 h\,\sigma \, dx \right)^{p'/2}.$$

We seek a good *a priori* bound, independent of h, for the right-hand side of (4.4).

The function $\tilde{T}^{(2,2)}g$ is bounded, smooth, and has good decay at infinity. If we let Ψ_y be as defined at the beginning of Section 2, then by a standard approximation argument (essentially Fatou's Lemma), combined with Theorem 2.1, we may write:

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(2,2)}g|^2 h\,\sigma \, dx \leq C \sum_R \frac{|\Lambda_R|^2}{|R|}(h\sigma)(\tilde{R},\eta),$$

where η is any number larger than 2, and

$$\Lambda_R = \left(\int_{T(R)} |\tilde{T}^{(2,2)}g * \Psi_y(t)|^2 \frac{dt_1\, dt_2\, dy_1\, dy_2}{y_1 y_2} \right)^{1/2}.$$

As in the proof of Theorem 2.3, we can dominate $(h\sigma)(\tilde{R},\eta)$ by a constant times

$$\int_{\tilde{R}} h\,\sigma\,\phi^{(\tilde{R})}\, dx,$$

where $\phi^{(\tilde{R})}$ is positive and satisfies

$$\frac{1}{|\tilde{R}|} \int_{\tilde{R}} \exp([\phi^{(\tilde{R})}]^{1/\eta})\, dx \leq 6.$$

With the $\phi^{(\tilde{R})}$'s now fixed, let us define

$$\nu(R) = \int_{\tilde{R}} h\,\sigma\,\phi^{(\tilde{R})}\, dx.$$

Then:

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(2,2)}g|^2 h\,\sigma \, dx \leq C \sum_R \frac{|\Lambda_R|^2}{|R|}\nu(R),$$

and it is this last object which we must bound.

We need to know how big $\Lambda(R)$ can get (or doesn't get). Let us make the convention that "$(x,y) \in \mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$" means "$x = (x_1, x_2)$; $x_i \in \mathbf{R}^{d_i}$; $y = (y_1, y_2)$; $y_i > 0$"; and analogously, when we write "$(x,y) \in$

$T(R)$", with $R = Q_1 \times Q_2$, we mean that $(x_i, y_i) \in T(Q_i)$. For $(t, \theta) \in \mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$, set

$$\Pi_\theta(t) = (\rho_1^{(2)})_{\theta_1}(t_1) \cdot (\rho_2^{(2)})_{\theta_2}(t_2).$$

The "$\rho^{(2)}$" functions satisfy the cancellation and decay hypotheses required of the ϕ_i's in the statement of Lemma 3.1. The discussion following the lemma shows that if $(t, \theta) \in T(R) = T(Q_1) \times T(Q_2)$ and $(x, y) \in T(R') = T(Q_1') \times T(Q_2')$, then

$$\theta_1^{-1}\theta_2^{-1}|\Psi_y * \Pi_\theta(t - x)| \leq C a_1(Q_1', Q_1) \cdot a_2(Q_2', Q_2)$$

$$= C\beta(R', R).$$

Since

$$|\tilde{T}^{(2,2)}g * \Psi_y(x)| \leq \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |g(t, \theta)|\theta_1^{-1}\theta_2^{-1}|\Psi_y * \Pi_\theta(t - x)| \, d\mu(t, \theta)$$

$$= \sum_R \int_{(T(R))} |g(t, \theta)|\theta_1^{-1}\theta_2^{-1}|\Psi_y * \Pi_\theta(t - x)| \, d\mu(t, \theta),$$

we at once get that

$$\Lambda(R') = \left(\int_{T(R')} |\Psi_y * T^{(2,2)}g(x)|^2 \frac{dx_1 \, dx_2 \, dy_1 \, dy_2}{y_1 y_2} \right)^{1/2}$$

$$\leq C|R'|^{1/2} \left[\sum_R \beta(R', R)G(R) \right],$$

where we have set

$$G(R) = \int_{T(R)} |g(t, \theta)| \, d\mu(t, \theta).$$

(We refer the reader to [**WhWi**, pp. 942–943] for a detailed discussion of this argument in the one-parameter setting.)

If we now define

$$\Gamma(R) = \mu(T(R))^{1-q'},$$

then Hölder's inequality implies

$$\left(\sum_R G(R)^{q'}\Gamma(R) \right)^{1/q'} \leq \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |g|^{q'} \, d\mu(t, \theta) \right)^{1/q'}.$$

On the other hand, the preceding discussion implies that

$$
\left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |\tilde{T}^{(2,2)} g|^{p'} \sigma \, dx \right)^{1/p'} \leq C \left(\sum_{R'} \frac{|\Lambda(R')|^2}{|R'|} \nu(R') \right)^{1/2}
$$

(4.5)

$$
\leq C \left(\sum_{R'} \left[\sum_{R} \beta(R', R) G(R) \right]^2 \nu(R') \right)^{1/2}.
$$

Our goal now is to show that, under the hypotheses of Theorem 1.3, the inequality

$$
\left(\sum_{R'} \left[\sum_{R} \beta(R', R) G(R) \right]^2 \nu(R') \right)^{1/2} \leq C \left(\sum_{R} G(R)^{q'} \Gamma(R) \right)^{1/q'}
$$

obtains for all *non-negative, finitely-supported* sequences $\{G(R)\}_R$.

In other words, we have reduced our problem to showing that the "kernel" $\beta(R', R)$ maps boundedly from the sequence space $\ell^{q'}(\Gamma(R))$ into the sequence space $\ell^2(\nu(R))$. We shall prove this boundedness in the same way as in [**WhWi**], i.e., by means of the Riesz-Thorin Interpolation Theorem.

We shall need two endpoint estimates, $\ell^\infty \mapsto \ell^\infty$ and $\ell^1 \mapsto \ell^{2/q'}$ (recall that $2/q' \geq 1$). In order to make these estimates (particularly the first) go through smoothly, let us redefine our problem, by setting $G(R) = Y(R)|\hat{R}|$, and having $\{Y(R)\}$ be the sequence that is acted on. This change-of-variable requires that we replace the kernel $\beta(R', R)$ with $B(R', R)$. In addition, we must replace the "weight" $\Gamma(R)$ by $|\hat{R}|^{q'} \Gamma(R)$. This done, we now need to show that the kernel $B(R', R)$ maps boundedly $\ell^\infty \mapsto \ell^\infty$ and $\ell^1(|\hat{R}|^{q'} \Gamma(R)) \mapsto \ell^{2/q'}(\nu(R))$.

$\ell^\infty \mapsto \ell^\infty$. This is equivalent to having $\sum_R B(R', R) \leq C$ for all R', and this inequality will follow if we have, for $i = 1, 2$, and all dyadic cubes $Q'_i \subset \mathbf{R}^{d_i}$,

$$
\sum_{Q_i} A_i(Q'_i, Q_i) \leq C.
$$

This is proved in [**WhWi**], though with slightly different notation from what we have here. For the sake of completeness (and ease of reading), we shall give a proof that uses our present notation.

Let us write the sum as $(I)_i + (II)_i + (III)_i$, where

$$(I)_i = \sum_{\substack{Q_i \\ Q_i \subset \tilde{Q}'_i}} A_i(Q'_i, Q_i)$$

$$(II)_i = \sum_{\substack{Q_i : Q_i \not\subset \tilde{Q}'_i \\ \ell(Q_i) \leq \ell(Q'_i)}} A_i(Q'_i, Q_i)$$

$$(III)_i = \sum_{\substack{Q_i : Q_i \not\subset \tilde{Q}'_i \\ \ell(Q_i) > \ell(Q'_i)}} A_i(Q'_i, Q_i).$$

Now:

$$(I)_i \leq C \sum_{\substack{Q_i \\ Q_i \subset \tilde{Q}'_i}} \left(\frac{\ell(Q_i)}{\ell(Q'_i)} \right)^{d_i + m_i + 2} \log(e + \ell(Q'_i)/\ell(Q_i))$$

$$\leq C_\delta \sum_{\substack{Q_i \\ Q_i \subset \tilde{Q}'_i}} \left[\frac{|Q_i|}{|Q'_i|} \right]^{1+\delta},$$

for some $\delta > 0$, since $d_i + m_i + 2 > d_i$. But it is easy to see [**WhWi**] that this last sum is $\leq C_{\delta, d_i}$. So much for $(I)_i$.

$(II)_i$:

$$(II)_i = \sum_{k=0}^{\infty} \sum_{\substack{Q_i : Q_i \not\subset \tilde{Q}'_i \\ \ell(Q_i) = 2^{-k}\ell(Q'_i)}} A_i(Q'_i, Q_i)$$

$$\leq C \sum_{k=0}^{\infty} \int_{x_i \notin \tilde{Q}'_i} \frac{\ell(Q'_i)(2^{-k}\ell(Q'_i))^{m_i+2}}{(2^{-k}\ell(Q'_i) + |x - x_{Q'_i}|)^{d_i+m_i+3}} \, dx_i$$

$$\leq C \sum_{k=0}^{\infty} \ell(Q'_i)(2^{-k}\ell(Q'_i))^{m_i+2}\ell(Q'_i)^{-m_i-3}$$

$$\leq C \sum_{k=0}^{\infty} 2^{-k(m_i+2)}$$

$$\leq C.$$

$(III)_i$:

$$(III)_i = \sum_{k=1}^{\infty} \sum_{\substack{Q_i : Q_i \not\subset \tilde{Q}'_i \\ \ell(Q_i) = 2^k \ell(Q'_i)}} A_i(Q'_i, Q_i)$$

$$\leq C \sum_{k=1}^{\infty} \int_{\mathbf{R}^{d_i}} \frac{\ell(Q'_i)(2^k \ell(Q'_i))^{m_i+2}}{(2^k \ell(Q'_i) + |x - x_{Q'_i}|)^{d_i+m_i+3}} \, dx_i$$

$$\leq C \sum_{k=1}^{\infty} 2^{-k}$$

$$\leq C.$$

The $\ell^\infty \mapsto \ell^\infty$ bound has been proved.

Now for the $\ell^1 \mapsto \ell^{2/q'}$ bound. By Minkowki's inequality for double integrals,

$$\left(\sum_{R'} \left[\sum_R B(R', R) Y(R) \right]^{2/q'} \nu(R') \right)^{q'/2}$$

$$\leq \sum_R \left[\sum_{R'} B(R', R)^{2/q'} \nu(R') \right]^{q'/2} Y(R).$$

Therefore, the $\ell^1 \mapsto \ell^{2/q'}$ bound will follow if

(4.6) $$\left(\sum_{R'} B(R', R)^{2/q'} \nu(R') \right)^{q'/2} \leq C |\hat{R}|^{q'} \Gamma(R),$$

holds for all R, for some constant C.

Inequality (4.6) will turn out to be an easy consequence of Lemma 3.2 and the hypotheses of Theorem 1.3.

Proof of Inequality (4.6):

(4.7)
$$\sum_{R'} B(R',R)^{2/q'} \nu(R')$$

$$= \sum_{k_1,k_2} \sum_{\substack{R'=Q'_1 \times Q'_2 \\ \ell(Q'_j)=2^{k_i}\ell(Q_j)}} B(R',R)^{2/q'} \int_{\tilde{R}'} h\phi^{(\tilde{R}')} \sigma\, dx$$

$$\leq C \sum_{k_1,k_2} \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\sum_{\substack{R'=Q'_1 \times Q'_2 \\ \ell(Q'_j)=2^{k_j}\ell(Q_j)}} B(R',R)^{2/q'} \chi_{\tilde{R}'}(x) \right) h\phi^{(\tilde{R}')} \sigma\, dx$$

$$\leq C \sum_{k_1,k_2} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\sum_{\substack{R'=Q'_1 \times Q'_2 \\ \ell(Q'_j)=2^{k_j}\ell(Q_j)}} B(R',R)^{p'/q'} (\phi^{(\tilde{R}')})^{p'/2} \chi_{\tilde{R}'}(x) \right) \sigma\, dx \right)^{2/p'}.$$

Inequality (4.7) is true because of Hölder's inequality (recall the normalization on h) and the fact that, for each fixed pair (k_1,k_2), no point of $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ lies in more than $C(d_1,d_2)$ rectangles $\tilde{R}'$ with the specified dimensions. The reasoning from Theorem 2.3 tells us that, for each $\tilde{R}'$,

$$\int (\phi^{(\tilde{R}')})^{p'/2} \chi_{\tilde{R}'}(x)\sigma\, dx \leq C\sigma(\tilde{R}', \eta p'/2),$$

which, by taking η sufficiently close to 2, we may assume is $\leq Cw(\tilde{R}')$. Thus, we may dominate (4.7) by

$$C \sum_{k_1,k_2} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\sum_{\substack{R'=Q'_1 \times Q'_2 \\ \ell(Q'_j)=2^{k_j}\ell(Q_j)}} B(R',R)^{p'/q'} \chi_{\tilde{R}'}(x) \right) w\, dx \right)^{2/p'}.$$

Because of Corollary 3.4, this is less than or equal to:

$$C \sum_{k_1,k_2} \left[(1+|k_1|)(1+|k_2|)2^{-(|k_1|+|k_2|)\epsilon}\right]^{2/q'}$$

$$\times \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\frac{1}{(\ell(Q_1)+|x_1-x_{Q_1}|)^{d_1+m_1+2-\epsilon}}\right.\right.$$

$$\times \frac{1}{(\ell(Q_2)+|x_2-x_{Q_2}|)^{d_2+m_2+2-\epsilon}}$$

$$\left.\left.\times \ell(Q_1)^{(d_1+m_1+2-\epsilon)}\ell(Q_2)^{(d_2+m_2+2-\epsilon)}\right)^{p'/q'} w\,dx\right)^{2/p'}$$

$$\leq C \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\frac{1}{(\ell(Q_1)+|x_1-x_{Q_1}|)^{d_1+m_1+2-\epsilon}}\right.\right.$$

$$\times \frac{1}{(\ell(Q_2)+|x_2-x_{Q_2}|)^{d_2+m_2+2-\epsilon}}$$

$$\left.\left.\times \ell(Q_1)^{(d_1+m_1+2-\epsilon)}\ell(Q_2)^{(d_2+m_2+2-\epsilon)}\right)^{p'/q'} w\,dx\right)^{2/p'},$$

which (see again the hypotheses of Theorem 1.3) is assumed to be less than or equal to

$$C\mu(T(R))^{-2/q}\ell(Q_1)^{2(d_1+1)}\ell(Q_2)^{2(d_2+1)}.$$

When we raise this to the power $q'/2$, the result is less than or equal to

$$C\mu(T(R))^{-q'/q}|\hat{R}|^{q'} = C\Gamma(R)|\hat{R}|^{q'},$$

which is what we wanted. Therefore, the $T^{(2,2)}$ term is okay.

We can handle the term $\tilde{T}^{(1,2)}g$ by modifying the preceding argument just a little. First, observe that, if $f \in L^2(\mathbf{R}^{d_1} \times \mathbf{R}^{d_2})$, then

$$(4.8) \quad f(x_1,x_2) = \int_{\mathbf{R}_+^{d_2+1}} [f(x_1,\cdot) * (\psi_2)_{y_2}(t_2) \cdot (\psi_2)_{y_2}(x_2-t_2)]\frac{dt_2\,dy_2}{y_2}$$

in L^2. The meaning of (4.8) is that we take the convolution of f with $(\psi_2)_{y_2}$ "in the x_2 variable" (leaving x_1 fixed), and then convolve that with $(\psi_2)_{y_2}$ again, much as we do in the original Calderón-Torchinsky formula (2.1). The proof of (4.8) comes by Fourier inversion, where we

take the Fourier transform only with respect to the x_2 variable. If we let $f = \tilde{T}^{(1,2)}g$ in (4.8), we get:

$$(4.9) \quad \tilde{T}^{(1,2)}g(x_1, x_2) = \int_{\mathbf{R}_+^{d_2+1}} H(x_1, t_2, y_2)(\psi_2)_{y_2}(x_2 - t_2)\frac{dt_2\, dy_2}{y_2},$$

where

$$(4.10) \quad H(x_1, t_2, y_2) = \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} g(s, \eta)$$

$$\times \left[\eta_1^{-1}(\rho_1^{(1)})_{\eta_1}(s_1 - x_1)\left[\eta_2^{-1}(\psi_2)_{y_2} * (\rho_2^{(2)})_{\eta_2}(t_2 - s_2)\right] \right]\, d\mu(s, \eta).$$

Let us define

$$(4.11) \quad F(x_1, s, \eta, t_2, y_2) \equiv \eta_1^{-1}(\rho_1^{(1)})_{\eta_1}(s_1 - x_1)$$

$$\times \left[\eta_2^{-1}(\psi_2)_{y_2} * (\rho_2^{(2)})_{\eta_2}(t_2 - s_2)\right].$$

If we plug (4.11) into (4.10), and then substitute that into (4.9), we get

$$\tilde{T}^{(1,2)}g(x_1, x_2) = \int_{\mathbf{R}_+^{d_2+1}} \left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} g(s, \eta)F(x_1, s, \eta, t_2, y_2)\, d\mu(s, \eta) \right)$$

$$\times (\psi_2)_{y_2}(x_2 - t_2)\frac{dt_2\, dy_2}{y_2}$$

$$= \sum_{Q_2 : Q_2 \subset \mathbf{R}^{d_2}} \int_{T(Q_2)} \sum_{Q_1 : Q_1 \subset \mathbf{R}^{d_1}} \left(\int_{T(Q_1) \times \mathbf{R}_+^{d_2+1}} g(s, \eta)F(x_1, s, \eta, t_2, y_2)\, d\mu(s, \eta) \right)$$

$$\times (\psi_2)_{y_2}(x_2 - t_2)\frac{dt_2\, dy_2}{y_2},$$

where the Q_i are, as usual, dyadic cubes. For $R' = Q_1' \times Q_2' \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, define

$$b_{R'}(x_1, x_2) \equiv \int_{T(Q_2')} \left(\int_{T(Q_1') \times \mathbf{R}_+^{d_2+1}} g(s, \eta)F(x_1, s, \eta, t_2, y_2)\, d\mu(s, \eta) \right)$$

$$\times (\psi_2)_{y_2}(x_2 - t_2)\frac{dt_2\, dy_2}{y_2}.$$

It is important to note that the integration over $T(Q_2')$ is done in the (t_2, y_2) variables and that the integration over $T(Q_1') \times \mathbf{R}_+^{d_2+1}$ is done in the (s, η) variables: failure to observe this hung the author up for some time.

It is easy to see that, if $(s_1, \eta_1) \in T(Q_1')$, then $F(x_1, s, \eta, t_2, y_2)$, considered as a function of x_1, is supported in $\tilde{Q}_1'$. Similarly, if $(t_2, y_2) \in T(Q_2')$, then $(\psi_2)_{y_2}(x_2 - t_2)$, as a function of x_2, is supported in $\tilde{Q}_2'$. Therefore, $b_{R'}(x_1, x_2)$ is supported in $\tilde{R}'$. The function F inherits smoothness and cancellation (in x_1) from $\rho_1^{(1)}$, and therefore so does $b_{R'}$. In the same fashion, $b_{R'}$ inherits smoothness and cancellation (in x_2) from $(\psi_2)_{y_2}(x_2 - t_2)$. Thus, we may write $b_{R'}(x_1, x_2) = \lambda_{R'} a_{\tilde{R}'}(x_1, x_2)$, where $\lambda_{R'}$ is a number and $a_{\tilde{R}'}(x_1, x_2)$ is *adapted* to $\tilde{R}'$.

We need a good bound on $|\lambda_{R'}|$, which we get, as usual, by controlling $\|b_{R'}\|_\infty$. Let $x = (x_1, x_2) \in \tilde{R}' \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$. Note that, for any $x_2 \in \mathbf{R}^{d_2}$,

$$\int_{T(Q_2')} |(\psi_2)_{y_2}(x_2 - t_2)| \frac{dt_2 \, dy_2}{y_2} \leq C$$

for some constant $C = C(\psi_2, d_2)$. Therefore,

$$(4.12) \quad |b_{R'}(x)| \leq C \sup_{(t_2, y_2) \in T(Q_2')} \int_{T(Q_1') \times \mathbf{R}_+^{d_2+1}} |g(s, \eta)|$$

$$\times [|F(x_1, s, \eta, t_2, y_2)|] \, d\mu(s, \eta),$$

$$= C \sup_{(t_2, y_2) \in T(Q_2')} \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |g(s, \eta)|$$

$$\times \left[|F(x_1, s, \eta, t_2, y_2)| \chi_{T(Q_1')}(s_1, \eta_1) \right] \, d\mu(s, \eta),$$

$$\equiv C \sup_{(t_2, y_2) \in T(Q_2')} \int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} |g(s, \eta)|$$

$$\times [P(x_1, s, \eta, t_2, y_2)] \, d\mu(s, \eta),$$

$$(4.13) \quad = C \sup_{(t_2, y_2) \in T(Q_2')} \sum_{R = Q_1 \times Q_2} \int_{T(R)} |g(s, \eta)|$$

$$\times [P(x_1, s, \eta, t_2, y_2)] \, d\mu(s, \eta).$$

The key to estimating (4.12) is to bound P. Now, P has two factors, one depending on (s_1, η_1) and one depending on (s_2, η_2) and (t_2, y_2). The absolute value of the first factor, $\eta_1^{-1}(\rho_1^{(1)})_{\eta_1}(x_1 - s_1) \cdot \chi_{T(Q_1')}(s_1, \eta_1)$, is less than or equal to $C\ell(Q_1')^{-d_1-1}$ and is zero if $(s_1, \eta_1) \notin T(Q_1')$. We note that, if $(s_1, \eta_1) \in T(Q_1)$, then this factor is less than or equal to a constant times what we have called $a(Q_1', Q_1)$, *for any value of m_1.* The meaning of that last sentence is so simple-minded that it might appear to result from a typographical error. What we mean is this: Take $(s_1, \eta_1) \in T(Q_1)$. Either Q_1 equals Q_1' or it doesn't. In the first case, our factor is less than or equal to $C\ell(Q_1')^{-d_1-1} \le Ca(Q_1', Q_1') = Ca(Q_1', Q_1)$; and in the second case, it's zero.

The function P's second factor, $|\eta_2^{-1}(\psi_2)_{y_2} * (\rho_2^{(2)})_{\eta_2}(t_2 - s_2)|$ is treated similarly. If $(t_2, y_2) \in T(Q_2')$ and $(s_2, \eta_2) \in T(Q_2)$ (which is the situation we have), then, by Lemma 3.1, this second factor is less than or equal to a constant times $a(Q_2', Q_2)$. Thus, in each of the separate integrals in (4.13), the function P is dominated by a constant times what we have called $\beta(R', R)$. Recalling our earlier convention,

$$G(R) = \int_{T(R)} |g(s, \eta)| \, d\mu(s, \eta),$$

we see that

$$\|b_{R'}\|_\infty \le C \sum_R \beta(R', R) G(R);$$

implying

$$|\lambda_{R'}| \le C|R'|^{1/2} \left[\sum_R \beta(R', R) G(R) \right].$$

But this (see (4.5) and preceding) is *precisely* the bound we got earlier! Our result now follows from the arguments that took care of the $\tilde{T}^{(2,2)}$ term. Theorem 1.3 is proved. $\square$

5. The Poisson kernel

The Poisson kernel presents us with some new difficulties. One of them concerns notation; something which, in this context, is non-trivial. For $i = 1, 2$, we let $P^{(i)}$ denote the Poisson kernel for $\mathbf{R}^{d_i}$. If $f \colon \mathbf{R}^{d_1} \times \mathbf{R}^{d_2} \mapsto \mathbf{R}$ is measurable and satisfies

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \frac{|f(x)|}{(1 + |x_1|)^{d_1+1}(1 + |x_2|)^{d_2+1}} \, dx < \infty,$$

then

$$u(x, y) = [(P^{(1)})_{y_1} \cdot (P^{(2)})_{y_2}] * f(x_1, x_2)$$

$$\equiv \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} (P^{(1)})_{y_1}(x_1 - t_1)(P^{(2)})_{y_2}(x_2 - t_2) f(t_1, t_2) \, dt_1 \, dt_2$$

is defined, and is called f's bi-harmonic extension into $\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}$. Let us (very temporarily) set $x_i = (x_i^{(1)}, \ldots, x_i^{(d_i)})$ and $y_i = x_i^{(0)}$ ($i = 1, 2$). As stated in the introduction, the full gradient $\nabla_1 \nabla_2 u$ is a $(d_1 + 1) \times (d_2 + 1)$ matrix whose entries are the second mixed partials

$$\frac{\partial^2 u}{\partial x_1^{(j_1)} \partial x_2^{(j_2)}},$$

where the j_i's run respectively from 0 to d_i. These entries are given by double convolutions:

$$\frac{\partial^2 u}{\partial x_1^{(j_1)} \partial x_2^{(j_2)}} = \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} y_1^{-1}(\phi_1^{(j_1)})_{y_1}(x_1 - t_1) y_2^{-1}(\phi_2^{(j_2)})_{y_2}(x_2 - t_2)$$

$$\times f(t_1, t_2) \, dt_1 \, dt_2,$$

where

$$\phi_i^{(j_i)}(x_i) = \begin{cases} \dfrac{\partial P^{(i)}}{\partial x_i^{(j_i)}}(x_i) & \text{if } j_i \neq 0; \\[2ex] -d_i P^{(i)}(x_i) - \sum_{k=1}^{d_i} x_i^{(k)} \dfrac{\partial P^{(i)}}{\partial x_i^{(k)}}(x_i) & \text{if } j_i = 0. \end{cases}$$

It is easy to see that all of these kernels satisfy $\int_{\mathbf{R}^{d_i}} \phi_i^{(j_i)}(x_i) \, dx_i = 0$. If we also have $j_i \neq 0$, then $\phi_i^{(j_i)}$ satisfies

$$|\phi_i^{(j_i)}(x_i)| \leq C_i (1 + |x_i|)^{-d_i - 2}$$

$$|\nabla_i \phi_i^{(j_i)}(x_i)| \leq C_i (1 + |x_i|)^{-d_i - 3}.$$

These are the hypotheses of Theorem 1.3 (with $m_i = 0$). Therefore, the following theorem is immediate:

Theorem 5.1. *Let* $1 < p \le 2 \le q < \infty$ *and let* v, μ, *and* σ *be as in the hypotheses of Theorem 1.3. Let* $\eta > p'$ *and* $\epsilon > 0$. *There is a positive constant*

$$C = C(\eta, \epsilon, p, q, d_1, d_2)$$

such that the following is true: If there exists a weight w *satisfying*

$$\sigma(R, \eta) \le \int_R w$$

and

$$\mu(T(R))^{1/q} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left[\frac{1}{(\ell(Q_1) + |x_1 - x_{Q_1}|)^{(d_1 + 2 - \epsilon)p'/q'}} \right. \right.$$

$$\times \frac{1}{(\ell(Q_2) + |x_2 - x_{Q_2}|)^{(d_2 + 2 - \epsilon)p'/q'}}$$

(5.1)

$$\left. \left. \times w(x) \right] dx \right)^{1/p'}$$

$$\le C\ell(Q_1)^{d_1 + 1 - (d_1 + 2 - \epsilon)/q'} \ell(Q_2)^{d_2 + 1 - (d_2 + 2 - \epsilon)/q'}$$

for all rectangles $R = Q_1 \times Q_2$, *then*

$$\left(\int_{\mathbf{R}^{d_1 + 1}_+ \times \mathbf{R}^{d_2 + 1}_+} |\frac{\partial^2 u}{\partial x_1^{(j_1)} \partial x_2^{(j_2)}}|^q \, d\mu(x, y) \right)^{1/q} \le \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p v \, dx \right)^{1/p}$$

holds for all mixed partials such that neither $j_i = 0$, *and for all* $f \in \cup_{1 \le r < \infty} L^r(\mathbf{R}^d, dx)$.

Unfortunately, the kernels $\phi_i^{(0)}$ only decay like $(1 + |x_i|)^{-d_i - 1}$, which is not quite good enough for Theorem 1.3. In [**WhWi**], the authors circumvented this by a trick that exploited harmonicity and the Poisson kernel's semigroup property. We refer the reader to [**WhWi**, pp. 957–959] for the details of this argument. Its upshot is that, in obtaining our sequence space estimates when $j_i = 0$, it is sufficient to replace $a_i(Q'_i, Q_i)$ by (see the top of p. 959):

$$a_i^*(Q'_i, Q_i) = \begin{cases} \dfrac{\ell(Q'_i)^2}{(\ell(Q'_i) + |x_{Q_i} - x_{Q'_i}|)^{d_i + 3}} & \text{if } \ell(Q_i) \le \ell(Q'_i); \\ 0 & \text{otherwise.} \end{cases}$$

We analogously define $A_i^*(Q'_i, Q_i) = a_i^*(Q'_i, Q_i) \cdot |\hat{Q}_i|$. The proof of the following lemma is like that of Lemma 3.2, and we omit it.

Lemma 5.2. *Let $\gamma > 0$, $0 < \epsilon < 1$, and let k be a positive integer. There is a constant $C = C(\gamma, \epsilon, d_i)$ such that, for all all $x_i \in \mathbf{R}^{d_i}$, all cubes $Q_i \subset \mathbf{R}^{d_i}$, and all k,*

$$(5.2) \qquad \sum_{Q_i': \ell(Q_i') = 2^k \ell(Q_i)} [A_i^*(Q_i', Q_i)]^\gamma \chi_{\tilde{Q}_i'}(x_i)$$

$$\leq C 2^{-k\epsilon\gamma} \ell(Q_i)^{(d_i + 1 - \epsilon)\gamma} \left(\frac{1}{(\ell(Q_i) + |x - x_{Q_i}|)^{d_i + 1 - \epsilon}} \right)^\gamma.$$

The proof of the following theorem is essentially identical to that of Theorem 5.1, and we omit it.

Theorem 5.3. *Let $1 < p \leq 2 \leq q < \infty$ and let v, μ, and σ be as in the hypotheses of Theorem 1.3. Let $\eta > p'$ and $\epsilon > 0$. There is a positive constant*

$$C = C(\eta, \epsilon, p, q, d_1, d_2)$$

such that the following is true: If there exists a weight w satisfying

$$\sigma(R, \eta) \leq \int_R w$$

and

$$\mu(T(R))^{1/q} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\frac{1}{(\ell(Q_1) + |x_1 - x_{Q_1}|)^{(d_1 + 1 - \epsilon)p'/q'}} \right. \right.$$

$$\times \frac{1}{(\ell(Q_2) + |x_2 - x_{Q_2}|)^{(d_2 + 2 - \epsilon)p'/q'}}$$

$$(5.3) \qquad \qquad \left. \left. \times w(x) \right] dx \right)^{1/p'}$$

$$\leq C \ell(Q_1)^{d_1 + 1 - (d_1 + 1 - \epsilon)/q'} \ell(Q_2)^{d_2 + 1 - (d_2 + 2 - \epsilon)/q'}$$

for all rectangles $R = Q_1 \times Q_2$, then

$$\left(\int_{\mathbf{R}_+^{d_1+1} \times \mathbf{R}_+^{d_2+1}} \left| \frac{\partial^2 u}{\partial x_1^{(0)} \partial x_2^{(j_2)}} \right|^q d\mu(x, y) \right)^{1/q} \leq \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^p v \, dx \right)^{1/p}$$

holds for all mixed partials such that $j_2 \neq 0$, and for all
$f \in \cup_{1 \leq r < \infty} L^r(\mathbf{R}^d, dx)$. *The symmetric result holds for $j_2 = 0$ and*
$j_1 \neq 0$. *When $j_1 = j_2 = 0$, the inequality analogous to (5.3) is:*

$$
\mu(T(R))^{1/q} \left(\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left[\frac{1}{(\ell(Q_1) + |x_1 - x_{Q_1}|)^{(d_1+1-\epsilon)p'/q'}} \right. \right.
$$

$$
\times \frac{1}{(\ell(Q_2) + |x_2 - x_{Q_2}|)^{(d_2+1-\epsilon)p'/q'}}
$$

$$
\left. \left. \times\, w(x) \right] dx \right)^{1/p'}
$$

$$
\leq C\ell(Q_1)^{d_1+1-(d_1+1-\epsilon)/q'} \ell(Q_2)^{d_2+1-(d_2+1-\epsilon)/q'}
$$

$$
= C|\hat{R}|^{1/q} \ell(Q_1)^{-\epsilon/q'} \ell(Q_2)^{-\epsilon/q'}.
$$

References

[CF]   S.-Y. A. CHANG AND R. A. FEFFERMAN, A continuous version of duality of H^1 with BMO on the bidisc, *Ann. of Math. (2)* **112(1)** (1980), 179–201.

[St]   E. M. STEIN, *"Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals"*, Princeton Mathematical Series **43**, Monographs in Harmonic Analysis III, Princeton University Press, Princeton, NJ, 1993.

[WhWi]   R. L. WHEEDEN AND J. M. WILSON, Weighted norm estimates for gradients of half-space extensions, *Indiana Univ. Math. J.* **44(3)** (1995), 917–969.

[W1]   J. M. WILSON, Weighted norm inequalities for the continuous square function, *Trans. Amer. Math. Soc.* **314(2)** (1989), 661–692.

[W2]   J. M. WILSON, Some two-parameter square function inequalities, *Indiana Univ. Math. J.* **40(2)** (1991), 419–442.

[W3]   J. M. WILSON, A two-parameter "Bergman space" inequality, *Proc. Amer. Math. Soc.* **125(3)** (1997), 755–762.

[W4]   J. M. WILSON, A semi-discrete Littlewood-Paley inequality, *Studia Math.* (to appear).

Department of Mathematics and Statistics
University of Vermont
Burlington, VT 05405
U.S.A.
E-mail address: `wilson@emba.uvm.edu`