Non-compact Littlewood–Paley theory for non-doubling measures

by

Michael Wilson (Burlington, VT)

Abstract. We prove weighted Littlewood–Paley inequalities for linear sums of functions satisfying mild decay, smoothness, and cancelation conditions. We prove these for general "regular" measure spaces, in which the underlying measure is not assumed to satisfy any doubling condition. Our result generalizes an earlier result of the author, proved on $\mathbb{R}^d$ with Lebesgue measure. Our proof makes essential use of the technique of random dyadic grids, due to Nazarov, Treil, and Volberg.

0. Introduction. In this note we generalize a slightly non-standard stopping time argument from the usual Euclidean setting on $\mathbb{R}^d$, with Lebesgue measure, to one in which the underlying measure is not assumed to satisfy any doubling condition.

The original motivation for this work came from certain weighted inequalities proved by the author and Richard Wheeden in [WhWi]. They looked for conditions on weights v and non-negative measures μ which ensured that

$$(0.1) \qquad \left(\int_{\mathbb{R}^{d+1}_+} |\nabla u|^q \, d\mu \right)^{1/q} \leq \left(\int_{\mathbb{R}^d} |f|^p \, v \, dx \right)^{1/p}$$

would hold for all f in some reasonable test class. Here we are assuming that p and q lie strictly between 1 and ∞, and that u is the harmonic (Poisson) extension of f into $\mathbb{R}^{d+1}_+ = \mathbb{R}^d \times (0, \infty)$. The approach they used was to consider a dual form of (0.1):

$$(0.2) \qquad \left(\int_{\mathbb{R}^d} |T(g)|^{p'} \sigma \, dx \right)^{1/p'} \leq \left(\int_{\mathbb{R}^{d+1}_+} |g(t,y)|^{q'} \, d\mu \right)^{1/q'}.$$

Here p' and q' are the dual indices to p and q, $\sigma = v^{1-p'}$, and g is an arbitrary bounded, measurable, compactly supported function mapping $\mathbb{R}^{d+1}_+ \to \mathbb{R}$. The operator T is a certain "balayage"-like object, whose precise definition need not concern us here.

The inequality (0.2) turned out, after some juggling, to follow from inequalities like this:

$$\left(\int_{\mathbb{R}^d} \Big| \sum_Q \lambda_Q \phi_{(Q)} \Big|^{p'} \sigma \, dx \right)^{1/p'} \leq C \left(\sum_Q |\lambda_Q|^{q'} w_Q \right)^{1/q'}.$$

The summation is over the dyadic cubes $Q \subset \mathbb{R}^d$; the λ_Q's are arbitrary real numbers, of which all but finitely many are assumed to be 0, and the w_Q's are certain positive numbers whose precise definition need not concern us.

The functions $\phi_{(Q)}$ are, if you will pardon the misnomer, non-compactly supported wavelets. This means: each $\phi_{(Q)}$ is a bump function centered around Q, with size and smoothness decaying at a nice rate as $|x| \to \infty$, and that has some cancelation. To state these conditions precisely, we need to introduce some standard notation.

A *cube* $Q \subset \mathbb{R}^d$ is a Cartesian product of intervals in $\mathbb{R}^1$, all having the same length. We use $\ell(Q)$ to denote this common length (called Q's sidelength) and $|Q| = \ell(Q)^d$ to denote Q's Lebesgue measure; we will also use $|\cdot|$ to denote Lebesgue measure of more general sets. By x_Q we mean the geometric center of Q. The families of functions $\{\phi_{(Q)}\}_Q$ arising in [WhWi] satisfied (after a suitable normalization) the following three conditions:

(1) For all $x \in \mathbb{R}^d$,
$$|\phi_{(Q)}(x)| \leq |Q|^{-1/2}(1 + |x - x_Q|/\ell(Q))^{-M}.$$

Here M is a positive number depending only on the family; in practice, M is always larger than d.

(2) For all x and y in $\mathbb{R}^d$,
$$|\phi_{(Q)}(x) - \phi_{(Q)}(y)| \leq (|x - y|/\ell(Q))|Q|^{-1/2}$$
$$\times((1 + |x - x_Q|/\ell(Q))^{-M} + (1 + |y - x_Q|/\ell(Q))^{-M}).$$

(3) For all finite linear sums $\sum_Q \gamma_Q \phi_{(Q)}$,
$$\int \Big| \sum_Q \gamma_Q \phi_{(Q)} \Big|^2 \, dx \leq \sum_Q |\gamma_Q|^2.$$

This last condition is what we mean by "cancelation."

REMARK. The functions occurring in [WhWi] were not assumed to satisfy (1)–(3), but happened to satisfy conditions that implied them. These stronger conditions were used to prove the results in [WhWi].

In [W2] the author proved that (1)–(3) were enough to imply fairly strong Littlewood–Paley estimates for arbitrary finite linear sums $\sum_Q \lambda_Q \phi_{(Q)}$. These estimates were given in terms of the Muckenhoupt A_∞ weights. Recall that a non-negative $v \in L^1_{\text{loc}}(\mathbb{R}^d)$ is said to belong to A_∞ if, for every $\varepsilon > 0$, there is a $\delta > 0$ so that, for all cubes $Q \subset \mathbb{R}^d$ and measurable subsets $E \subset Q$, $|E|/|Q| < \delta$ implies $v(E) \leq \varepsilon v(Q)$; here, as usual, we are using $v(E)$ to denote $\int_E v \, dx$.

The main result from [W2] is:

THEOREM 0.1. *Let* $\{\phi_{(Q)}\}$ *satisfy* (1)–(3) *for some* $M > d/2$, *and let* $\varrho > d$. *If* $v \in A_\infty$ *and* $0 < p < \infty$ *then there is a constant* $C = C(p, d, \varrho, M, v)$ *such that, for all finite linear sums* $f = \sum_Q \lambda_Q \phi_{(Q)}$,

$$\int_{\mathbb{R}^d} |f|^p \, v \, dx \leq C \int_{\mathbb{R}^d} \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \left(1 + |x - x_Q|/\ell(Q)\right)^{-2M+\varrho} \right)^{p/2} v \, dx.$$

What does this result mean? Let us rewrite the right-hand integral as

$$\int G(f)^p \, v \, dx,$$

where

$$G(f) = \left(\sum_Q \frac{|\lambda_Q|^2}{|Q|} \left(1 + |x - x_Q|/\ell(Q)\right)^{-2M+\varrho} \right)^{1/2}.$$

The function $G(f)$ is simply a discretized version of the familiar g_λ^*-function from classical harmonic analysis. The inequality in Theorem 0.1 controls the size of f (defined by a sum with cancelation in it) with something built from a sum of positive terms. What is interesting is that the long-term decay of $G(f)$ (of order $|x|^{-M+\varrho/2}$) is not much worse than the best possible long-term decay of f (of order $|x|^{-M}$), even against arbitrary A_∞ weights.

The proof of Theorem 0.1 employed a stopping time argument specially adapted to sums of "non-compactly supported wavelets." (Sums of true, compactly supported wavelets can be handled essentially in the same manner as Haar function decompositions; see [W1]). Unfortunately, the argument also relied on the "doubling" property of the underlying Lebesgue measure. We remind the reader that a measure ν is *doubling* if, for every cube Q, the ν-measure of Q's concentric double—called $2Q$—is bounded by a constant times $\nu(Q)$. The doubling property has the following important consequence (and is, in fact, equivalent to this): in any cube Q, the set $\{x \in Q : d(x, Q) < \varepsilon \ell(Q)\}$ has *uniformly* small relative measure in Q as $\varepsilon \to 0$. This was needed in the proof of Theorem 0.1 because the stopping time argument worked by analyzing sums $\sum_Q \lambda_Q \phi_{(Q)}$ on certain dyadic cubes Q', where one loses all control over what happens when x is near $\partial Q'$ ("edge effects"). The doubling property of Lebesgue measure allowed us to discard this set as negligible.

Unfortunately, one cannot do this if the underlying measure does *not* have the doubling property. These include the "regular" measures, which have recently been receiving a great deal of attention ([MMV], [NTV], [T1]–[T3]). A Borel measure μ on $\mathbb{R}^d$ is called *regular* if there exist a positive constant C and a constant $0 < \beta \leq d$ such that, for all cubes Q, $\mu(Q) \leq C\ell(Q)^\beta$; we call such a μ *normalized* if $C = 1$. It is easy to see that regular measures need not have the doubling property. For example, if L is the x-axis in $\mathbb{R}^2$, we could define $\mu(E)$ to be the Lebesgue measure of $E \cap L$ in $\mathbb{R}$. If a square Q is close to the axis but does not touch it, then $\mu(Q) = 0$ but $\mu(2Q) > 0$.

For a long time it was believed that, in order to generalize results in classical harmonic analysis (e.g., those involving singular integrals) from the setting of $(\mathbb{R}^d, dx)$ to that of $(\mathbb{R}^d, \mu)$, it was essential that the underlying measure μ have the doubling property. In many situations, mere regularity turns out to be an adequate substitute for the doubling property.

In this paper we prove a generalization of Theorem 0.1 to the setting of regular measures. How we handle (or avoid) the dreaded edge effects calls for a little bit of explanation.

We use the ingenious trick of "random dyadic grids," created by Nazarov, Treil, and Volberg [NTV]. How these grids work will be described in greater detail below, but we will try to give the basic idea here. We work on $\mathbb{R}^2$. Let us suppose we have tiled the plane with dyadic squares of sidelength 2^{-n_0}, where n_0 is huge. We will call these our base squares. We want to build a grid of squares of sidelength 2^{1-n_0} such that each base square arises from bisecting the sides of a square from this second grid. There are exactly FOUR ways to do this. Pick one of the squares with sidelength 2^{-n_0}. It can be an upper right, upper left, lower right, or lower left quarter of one of the squares from the coarser grid. Obviously, having made such a choice for one base square determines the choices for all the others. Thus, the grid of base squares gives rise to four possible grids that are (so to speak) one degree coarser. We assign a probability of $1/4$ to each choice of grids. We now repeat this procedure on each of the four coarser grids. We obtain sixteen possible grids, each made of squares of sidelength 2^{2-n_0}, and each having a probability of $1/16$. We continue this process forever, making grids of larger and larger squares. In this way we build a probability space (Ω, dP), where Ω is the set of all of the grids built "upward" from the base squares.

Edge effects arise when we have to consider a sum containing many $\phi_{(Q)}$'s on a square R, where the $\ell(Q)$'s are much, much smaller than $\ell(R)$ and the Q's are very close to ∂R. Such Q's are called "bad" relative to R. If we look at a fixed Q arising from the grid-building process described above, it turns out that the (relative) probability of the set of grids for which Q is bad vis-à-vis *any* R is quite small. An averaging technique then allowed [NTV] to, in effect, ignore those grids on which any given square Q was bad, and

thus they could do their analysis on the assumption that *every* square Q was "good." What all of this means is made precise (and, we hope, clear) below. If the reader is already familiar with the random-grid method from [NTV], we should warn him that our situation has required a small (but easy) modification of their approach.

Following [NTV], we shall also work on $\mathbb{R}^2$. We suppose we have a fixed positive measure μ satisfying $\mu(Q) \leq \ell(Q)^\beta$ for all squares Q, where $0 < \beta \leq 2$ is also fixed. All squares are assumed to be Cartesian products of half-open intervals, like $Q = [a, b) \times [c, d)$.

A *dyadic grid* Γ with scale $\aleph$ $(1 \leq \aleph \leq 3/2)$ is a collection of squares Q, with sides parallel to the coordinate axes, satisfying:

1. every Q has sidelength $\ell(Q)$ equal to $2^k \aleph$ for some integer k;
2. for each k, the squares with $\ell(Q) = 2^k \aleph$ tile $\mathbb{R}^2$;
3. the squares Q with $\ell(Q) = 2^k \aleph$ are obtained by partitioning each of the squares Q' with $\ell(Q') = 2^{k+1} \aleph$ into four equal subsquares.

Given a dyadic grid Γ, we say that $Q \in \Gamma$ is *bad* if there is an $R \in \Gamma$ such that $\ell(Q) \leq 2^{-n}\ell(R)$ and $d(Q, \partial R) \leq \ell(Q)^\alpha \ell(R)^{1-\alpha}$, where n is a positive integer and $0 < \alpha < 1$. If Q is not bad, it is called *good*. Obviously, goodness and badness depend on n and α. We will prove some lemmas, valid for certain sums indexed over arbitrary collections of good squares, and at the end we will choose n (large enough) and α (small enough) to serve our purpose.

Let Γ be a dyadic grid and let μ be a normalized regular measure. We say that $\{\phi_{(I)}\}_{I \in \Gamma}$ is a *standard family* of functions (relative to Γ) if, for all $I \in \Gamma$, and all x and y in $\mathbb{R}^2$,

$$|\phi_{(I)}(x)| \leq \frac{\chi_I(x)}{\sqrt{\mu(I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon},$$

$$|\phi_{(I)}(x) - \phi_{(I)}(y)| \leq \left(\frac{|x-y|}{\ell(I)}\right)^\delta \left(\frac{\chi_I(x) + \chi_I(y)}{\sqrt{\mu(I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}\right.$$
$$\left. \times ((1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon} + (1 + |y - x_I|/\ell(I))^{-\beta-\varepsilon}) \right),$$

and, for all finite linear sums $\sum \gamma_I \phi_{(I)}$,

$$\int \left|\sum \gamma_I \phi_{(I)}\right|^2 d\mu \leq \sum |\gamma_I|^2.$$

We should explain where these inequalities come from. The size and smoothness estimates (for x and y *far* from I) are what one gets when one applies a suitable Calderón–Zygmund kernel $K(x_1, x_2)$ to a μ-adapted "Haar function" supported on I (see [NTV]). The estimates for x and y close to I are, we confess, something of a stopgap: they are analogues of

what one would get by applying a classical Calderón–Zygmund kernel to a two-dimensional wavelet supported on I. The third inequality is simply almost-orthogonality.

We say that a standard family $\{\phi_{(I)}\}_{I \in \Gamma}$ is *good* if $\phi_{(I)} \equiv 0$ whenever $I \in \Gamma$ is bad.

We will be considering finite linear sums $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$ where the $\phi_{(I)}$'s belong to some good standard family (relative to Γ). For such a sum, let us define

$$g^*(f)(x) \equiv \left(\sum_{I \in \Gamma} |\lambda_I|^2 \left(\frac{\chi_I}{\mu(I)} + \ell(I)^{-\beta} (1 + |x - x_I|/\ell(I))^{(-\beta - 2\varepsilon + \varrho)(1 - \alpha)} \right) \right)^{1/2},$$

where ϱ and α are assumed to be small and positive. This is a natural translation of $G(f)$ to our new situation. Note that $g^*(f)$ depends on the positive parameters α, β, ε, and ϱ, and that this dependence will normally be suppressed.

Our main theorem is:

THEOREM 1.13. *Let μ be a normalized regular measure on $\mathbb{R}^2$ and let $0 < p < \infty$. There is a constant $C = C(p, \varrho, n, \alpha, \beta, \varepsilon)$ such that, if Γ is any dyadic grid, and $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$, where the $\phi_{(I)}$'s belong to some good standard family (relative to Γ), then*

$$\int |f|^p \, d\mu \leq C \int (g^*(f))^p \, d\mu.$$

Before going further, we should say a few words about the roles played by ϱ, n, α, β, and ε. The relevance of α and n is, we hope, obvious: they are part of the definition of good squares. (Why α shows up in precisely the way it does is another matter, but that will become clear in the proof.) The parameters β and ε come in because of our bounds on the $\phi_{(I)}$'s. But the parameter ϱ is mysterious. It is a "bump" which guarantees that certain integrals and sums have good bounds. In proving Theorem 1.13, we will need to control integrals of the form

$$R^{-\beta} \int_{\mathbb{R}^2} (1 + |x - x_Q|/R)^{-\beta - \varrho} \, d\mu(x),$$

where R is positive but arbitrary. The reader can easily verify that if μ is a normalized regular measure and $\varrho > 0$, the preceding integral has a bound independent of R; while, if $\varrho \leq 0$ and $\mu =$ Lebesgue measure, the integral is infinite. We will repeat this observation below (see Observation 1.5 in the next section; see also the proof of Lemma 1.8). We will also have to control sums of the form

$$\sum_{\substack{I \in \Gamma \\ I \subset Q}} (\ell(I)/\ell(Q))^\varrho \mu(I)$$

for arbitrary squares $Q \in \Gamma$. This sum is bounded by $C(\varrho)\mu(Q)$ if $\varrho > 0$ but can be infinite otherwise.

We say that a Borel measure ν is A_∞ relative to μ if for every positive number τ there is a $\sigma > 0$ such that, for all squares Q and Borel sets $E \subset Q$, if $\mu(E) \leq \sigma\mu(Q)$, then $\nu(E) \leq \tau\nu(Q)$.

We shall see that Theorem 1.13 (with a suitable adjustment in the constant C) also holds with respect to measures ν that are A_∞ relative to μ. Precisely, the definitions of the $\phi_{(I)}$'s and $g^*(f)$ remain unchanged (in particular, they are still expressed in terms of μ), but the integral inequality becomes

$$\int |f|^p \, d\nu \leq C \int (g^*(f))^p \, d\nu.$$

By means of the averaging trick alluded to above, we obtain the promised generalization of Theorem 0.1, one making no assumption of "goodness" on squares. The statement of this corollary requires two new definitions, one of which is standard. The standard one is this: If I is a square and $\lambda > 0$, then λI denotes the square concentric with I and of sidelength equal to $\lambda \ell(I)$.

The next definition is not quite so standard.

Let Γ be a dyadic grid on $\mathbb{R}^2$, and let μ be a normalized regular measure on $\mathbb{R}^2$ (with "parameter" β). Let τ be a small positive number. We say that a family of functions $\{\phi_{(I)}\}_{I \in \Gamma}$, indexed over Γ, is τ-adapted to Γ if, for all $I \in \Gamma$ and x and y in $\mathbb{R}^2$,

$$|\phi_{(I)}(x)| \leq \frac{\chi_I(x)}{\sqrt{\mu((1+\tau)I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon},$$

$$|\phi_{(I)}(x) - \phi_{(I)}(y)| \leq \left(\frac{|x-y|}{\ell(I)}\right)^\delta \left(\frac{\chi_I(x) + \chi_I(y)}{\sqrt{\mu((1+\tau)I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}\right.$$
$$\left. \times \left((1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon} + (1 + |y - x_I|/\ell(I))^{-\beta-\varepsilon}\right)\right),$$

and, for all finite linear sums $\sum \gamma_I \phi_{(I)}$,

$$\int \left|\sum \gamma_I \phi_{(I)}\right|^2 d\mu \leq \sum |\gamma_I|^2.$$

Essentially, a family is τ-adapted if it is a standard family plus a little more. However, we must warn the reader that this "little more" might not be so little. For example, it entails replacing the estimate

$$|\phi_{(I)}(x)| \leq \frac{\chi_I(x)}{\sqrt{\mu(I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon}$$

with

$$|\phi_{(I)}(x)| \leq \frac{\chi_I(x)}{\sqrt{\mu((1+\tau)I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon},$$

and $\mu((1+\tau)I)$ might be infinitely larger than $\mu(I)$. This is unfortunate, but the τ is what (literally) gives us the "wiggle room" to carry out our averaging argument.

We define a corresponding, "τ-adjusted" g^*-function:

$$\tilde{g}^*(f)(x)$$
$$\equiv \left(\sum_I |\lambda_I|^2 \left(\frac{\chi_{(1+\tau)I}}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)} \right) \right)^{1/2},$$

where, as before, ϱ and α are assumed to be small positive numbers. Note that $\tilde{g}^*(f)$ is "a little" larger than $g^*(f)$, but that this "little" might be enormous.

The generalization of Theorem 0.1 is:

THEOREM 2.1. *Let $\{\phi_{(I)}\}_{I \in \Gamma}$ be τ-adapted to a dyadic grid Γ. Let $0 < p < \infty$. There is a constant $C = C(p, \beta, \tau, \varrho, \alpha)$ such that, for all finite linear sums $f = \sum_I \gamma_I \phi_{(I)}$,*

$$\int_{\mathbb{R}^2} |f|^p \, d\mu \leq C \int_{\mathbb{R}^2} (\tilde{g}^*(f))^p \, d\mu.$$

Moreover, if ν is any measure which is A_∞ relative to μ, we also have

$$\int_{\mathbb{R}^2} |f|^p \, d\nu \leq C \int_{\mathbb{R}^2} (\tilde{g}^*(f))^p \, d\nu$$

for all such p and f, where the constant C now also depends on ν.

We prove Theorem 1.13 in Section 1 and we prove Theorem 2.1 in Section 2.

1. When every square is good. The stopping time argument calls for some basic definitions. Until we say otherwise, all squares are assumed

to belong to a fixed dyadic grid Γ; and, until we say otherwise, we shall be working with a fixed finite linear sum $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$, where the $\phi_{(I)}$'s belong to a good standard family.

DEFINITION 1.1. If $Q \in \Gamma$, then $S(Q) = \{I \in \Gamma : I \not\subset Q\}$.

We think of $S(Q)$ as consisting of the squares that "surround" Q.

DEFINITION 1.2. If $Q \in \Gamma$, then $N(Q) = \{I \in \Gamma : I \subset Q, \ell(I) = (1/2)\ell(Q)\}$.

We think of $N(Q)$ as the "next generation" begotten by Q.

DEFINITION 1.3. For $Q \in \Gamma$, define

$$F(Q) = \frac{1}{\mu(Q)} \int_Q \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x) \Big) \, d\mu(x),$$

and set $F(Q)$ to be zero if $\mu(Q) = 0$. We define a corresponding maximal operator

$$F^*(x) = \sup_{Q : x \in Q} |F(Q)|.$$

We also define, for $x \in Q$,

$$F(Q, x) = \sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x),$$

and note that

$$F(Q) = \frac{1}{\mu(Q)} \int_Q F(Q, x) \, d\mu(x).$$

DEFINITION 1.4. For $Q \in \Gamma$, we define a *partial Littlewood–Paley operator*:

$$G(Q) =$$
$$\left(\frac{1}{\mu(Q)} \int_Q \Big(\sum_{I \in S(Q)} |\lambda_I|^2 \Big(\frac{\chi_I}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho} \Big) \Big) \, d\mu(x) \right)^{1/2}$$

if $\mu(Q) > 0$ and zero otherwise. We define a corresponding *maximal Littlewood–Paley operator*:

$$G^*(x) = \sup_{Q : x \in Q} G(Q).$$

As the reader can probably guess, we will prove Theorem 1.13 via a good-λ inequality. If he has read [W2], it is only natural for him to think that this inequality will have the form

$$\mu(\{x : F^*(x) > 2\lambda, \ G^*(x) \leq \gamma\lambda\}) \leq \tilde{\gamma}\mu(\{x : F^*(x) > \lambda\}).$$

But it will not. The actual inequality we will prove has the form

$$\mu(\{x : F^*(x) > 2\lambda, \ g^*(f)(x) \leq \gamma\lambda\}) \leq \tilde{\gamma}\mu(\{x : F^*(x) > \lambda\}).$$

However, we will *use* $G^*(x)$ in an essential way in our stopping time arguments.

As in [W2], the stopping time argument will follow from a series of lemmas, which we have laid out below, so as to correspond *fairly* closely to the order in [W2]. However, before beginning these, we wish to repeat an observation we made in the introduction, and which we invited the reader to prove for himself.

OBSERVATION 1.5. *Let μ be a normalized regular measure on $\mathbb{R}^2$, with parameter $\beta > 0$, and let $\gamma > \beta$. There is an absolute constant $C = C(\gamma, \beta)$ such that, for all $R > 0$ and all $x_0 \in \mathbb{R}^2$,*

$$R^{-\beta} \int (1 + |x - x_0|/R)^{-\gamma} \, d\mu(x) \leq C.$$

Now we prove the lemmas.

LEMMA 1.6. *For μ-a.e. x, $|f(x)| \leq F^*(x)$.*

Proof. Trivial.

LEMMA 1.7. *For every x, $G^*(x) \leq c g^*(f)(x)$, where c depends only on the "natural" parameters.*

Proof. We take x and y to be arbitrary points in Q, and consider $I \in S(Q)$, with I a good square. We wish to compare

(1.1) $$\frac{\chi_I}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho}$$

and

(1.2) $$\frac{\chi_I}{\mu(I)} + \ell(I)^{-\beta}(1 + |y - x_I|/\ell(I))^{(-\beta - 2\varepsilon + \varrho)(1-\alpha)},$$

and show that $(1.1) \leq c(1.2)$, with a constant independent of x and y.

If $Q \subset I$, there is nothing to prove. The inequality is also easy to prove if $\ell(Q) \leq \ell(I)$ or if $d(Q, I) > \ell(Q)$. Therefore we only need to consider the case where $\ell(I) < \ell(Q)$, I is close to Q, but I is disjoint from Q.

Since I is good, we know that either $\ell(I) > 2^{-n}\ell(Q)$ or $d(I, \partial Q) > \ell(I)^\alpha \ell(Q)^{1-\alpha}$. In the former case our inequality is easy to prove, with a constant that depends on n. The reason is that then $|x - x_I|$ and $|y - x_I|$ are comparable, with comparability constants depending on n. By itself this fact would yield

$$(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho} \leq C(1 + |y - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho},$$

and the extra $1 - \alpha$ in the exponent just makes the right-hand side bigger.

So now we suppose that $I \not\subset Q$, $\ell(I)$ is small, and $d(I, \partial Q) > \ell(I)^\alpha \ell(Q)^{1-\alpha}$.

On the one hand, we have

$$|x - x_I| > \ell(I)^\alpha \ell(Q)^{1-\alpha},$$

implying

$$\frac{|x - x_I|}{\ell(I)} > \left(\frac{\ell(Q)}{\ell(I)}\right)^{1-\alpha},$$

and therefore

$$1 + \frac{|x - x_I|}{\ell(I)} > \left(\frac{\ell(Q)}{\ell(I)}\right)^{1-\alpha}.$$

On the other hand,

$$1 + \frac{|y - x_I|}{\ell(I)} \le C\,\frac{\ell(Q)}{\ell(I)}.$$

Combining the inequalities, and inverting, we get

$$\left(1 + \frac{|x - x_I|}{\ell(I)}\right)^{-1} \le \left(\frac{\ell(I)}{\ell(Q)}\right)^{1-\alpha} \le C\left(1 + \frac{|y - x_I|}{\ell(I)}\right)^{-(1-\alpha)},$$

from which the inequality $(1.1) \le c(1.2)$ follows, by raising both sides to the positive power $\beta + 2\varepsilon - \varrho$.

The inequality we have just proved shows that, for all y in Q,

$$\frac{1}{\mu(Q)} \int_Q \left(\sum_{I \in S(Q)} |\lambda_I|^2 \left(\frac{\chi_I}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho}\right)\right) d\mu(x)$$
$$\le C g^*(f)(y),$$

which gives the lemma. ∎

We will encounter the *style* of the preceding proof repeatedly. (As John Garnett likes to say, "Five hundred theorems—one proof.")

LEMMA 1.8. *Let $Q \in \Gamma$, and let x, y, and z be points in Q. Then*

$$\left|\left(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x)\right) - \left(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(y)\right)\right| \le c g^*(f)(z).$$

Proof. Define $h = \sum_{I \in S(Q)} \lambda_I \phi_{(I)}$, and set

$$h_1 = \sum_{\substack{I \in S(Q) \\ \ell(I) \ge \ell(Q)}} \lambda_I \phi_{(I)}, \qquad h_2 = \sum_{\substack{I \in S(Q) \\ \ell(I) < \ell(Q)}} \lambda_I \phi_{(I)}.$$

We will show that (i) $|h_1(x) - h_1(y)| \le c g^*(f)(z)$ and (ii) $|h_2(x)| \le c g^*(f)(z)$.

(i) By the smoothness bound on the $\phi_{(I)}$'s,

$$|h_1(x) - h_1(y)| \le \sum_{\substack{I \in S(Q) \\ \ell(I) \ge \ell(Q)}} |\lambda_I| \left(\frac{|x - y|}{\ell(I)}\right)^\delta \left(\frac{\chi_I(x) + \chi_I(y)}{\sqrt{\mu(I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}\right.$$
$$\times \left.\left((1 + |x - x_I|/\ell(I))^{-\beta - \varepsilon} + (1 + |y - x_I|/\ell(I))^{-\beta - \varepsilon}\right)\right),$$

which we split into two pieces:

$$(1.3) \qquad \sum_{I\,:\,Q\subset I,\,Q\neq I} |\lambda_I|(\frac{|x-y|}{\ell(I)})^\delta \frac{1}{\sqrt{\mu(I)}}$$

and

$$(1.4) \qquad \sum_{\substack{I\in S(Q)\,:\,Q\cap I=\emptyset \\ \ell(I)\geq\ell(Q)}} |\lambda_I|\left(\frac{|x-y|}{\ell(I)}\right)^\delta$$

$$\times\, (\ell(I)^{-\beta}\sqrt{\mu(I)}((1+|x-x_I|/\ell(I))^{-\beta-\varepsilon}+(1+|y-x_I|/\ell(I))^{-\beta-\varepsilon})).$$

Now, $|x-y| \leq c\ell(Q)$ and, for each positive k, there is a unique I containing Q such that $\ell(I) = 2^k\ell(Q)$. Therefore, by the Cauchy–Schwarz inequality, the first piece (1.3) is less than or equal to

$$c\left(\sum_{I\,:\,Q\subset I} \frac{|\lambda_I|^2}{\mu(I)}\right)^{1/2} \leq cg^*(z)$$

for any $z \in Q$.

The second piece (1.4) is less than or equal to a constant times

$$\sum_{\substack{I\in S(Q)\,:\,Q\cap I=\emptyset \\ \ell(I)\geq\ell(Q)}} |\lambda_I|\left(\frac{\ell(Q)}{\ell(I)}\right)^\delta \ell(I)^{-\beta}\sqrt{\mu(I)}(1+|z-x_I|/\ell(I))^{-\beta-\varepsilon},$$

because $1+|x-x_I|/\ell(I)$ and $1+|y-x_I|\ell(I)$ are both comparable to $1+|z-x_I|/\ell(I)$. Using Cauchy–Schwarz again, the second piece is less than or equal to a constant times

$$\left(\sum_{\substack{I\in S(Q)\,:\,Q\cap I=\emptyset \\ \ell(I)\geq\ell(Q)}} |\lambda_I|^2\ell(I)^{-\beta}(1+|z-x_I|/\ell(I))^{-\beta-2\varepsilon+\varrho}\right)^{1/2}$$

$$\times \left(\sum_{\substack{I\in S(Q)\,:\,Q\cap I=\emptyset \\ \ell(I)\geq\ell(Q)}} \left(\frac{\ell(Q)}{\ell(I)}\right)^{2\delta} \mu(I)\ell(I)^{-\beta}(1+|z-x_I|/\ell(I))^{-\beta-\varrho}\right)^{1/2}.$$

The first factor is less than or equal to $g^*(f)(z)$ (again, the $1-\alpha$ in the definition of $g^*(f)$ only helps things).

We claim that the second factor is less than or equal to a constant. Our argument will be in the spirit of that on page 36 of [W2].

We split the sum up according to the sizes of the $\ell(I)$'s:

$$\sum_{k=0}^{\infty} 2^{-2k\delta} \sum_{I\in S(Q)\,:\,\ell(I)=2^k\ell(Q)} \mu(I)\ell(I)^{-\beta}(1+|z-x_I|/\ell(I))^{-\beta-\varrho};$$

and it is clearly enough to show that

$$\sum_{I \in S(Q)\,:\,\ell(I)=2^k\ell(Q)} \mu(I)\ell(I)^{-\beta}(1+|z-x_I|/\ell(I))^{-\beta-\varrho}$$

is less than or equal to a constant independent of k. However, a moment's thought shows that this last sum is bounded by a constant times

$$R^{-\beta} \int_{\mathbb{R}^2} (1+|x-x_Q|/R)^{-\beta-\varrho}\,d\mu(x),$$

where $R = 2^k\ell(Q)$, which, by our observation, is $\leq C(\beta, \varrho)$. That takes care of (i).

(ii) This argument will closely follow that on pages 36–38 of [W2], with a couple of twists.

Write $\mathbb{R}^2 = Q \cup \bigcup_j Q_j$, where the Q_j's are congruent, disjoint copies of Q. Every I occurring in our sum lies inside a unique Q_j. Thus,

$$
\begin{aligned}
|h_2(x)| &\leq \sum_j \sum_{I\,:\,I \subset Q_j} |\lambda_I \phi_{(I)}(x)| \\
&= \sum_{j\,:\,d(Q,Q_j)=0} \sum_{I\,:\,I \subset Q_j} |\lambda_I \phi_{(I)}(x)| + \sum_{j\,:\,d(Q,Q_j)>0} \sum_{I\,:\,I \subset Q_j} |\lambda_I \phi_{(I)}(x)| \\
&\equiv \text{(I)} + \text{(II)}.
\end{aligned}
$$

We will handle (II) first. The Q_j's in this sum satisfy $d(Q,Q_j) \geq \ell(Q)$, making the terms especially easy to handle. By the size bound on the $\phi_{(I)}$'s and the Cauchy–Schwarz inequality,

$$
\begin{aligned}
\text{(II)} &\leq \Bigg(\sum_{j\,:\,d(Q,Q_j)\geq\ell(Q)} \sum_{I\,:\,I \subset Q_j} |\lambda_I|^2 \ell(I)^{-\beta}(1+|x-x_I|/\ell(I))^{-\beta-2\varepsilon+\varrho} \Bigg)^{1/2} \\
&\quad \times \Bigg(\sum_{j\,:\,d(Q,Q_j)\geq\ell(Q)} \sum_{I\,:\,I \subset Q_j} \ell(I)^{-\beta}\mu(I)(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \Bigg)^{1/2}.
\end{aligned}
$$

The first factor is easily to seen to be $\leq cg^*(f)(z)$, for any $z \in Q$, with c independent of z. We claim that the second factor is less than or equal to a constant.

Consider, for fixed j,

$$\sum_{I\,:\,I \subset Q_j} \ell(I)^{-\beta}\mu(I)(1+|x-x_I|/\ell(I))^{-\beta-\varrho}.$$

Take $I \subset Q_j$. For some positive c, an absolute constant, we have

$$|x-x_I| > c|x_Q - x_{Q_j}|,$$

implying (using the fact that $|x_Q - x_{Q_j}|/\ell(Q) \geq 1$)

$$1 + |x - x_I|/\ell(I)) > c'(1 + |x_Q - x_{Q_j}|/\ell(Q))(\ell(Q)/\ell(I));$$

and thus

$$\ell(I)^{-\beta}\mu(I)(1 + |x - x_I|/\ell(I))^{-\beta-\varrho}$$
$$\leq \tilde{c}\ell(I)^{-\beta}\mu(I)(1 + |x_Q - x_{Q_j}|/\ell(Q))^{-\beta-\varrho}(\ell(Q)/\ell(I))^{-\beta-\varrho}$$
$$\leq \tilde{c}\ell(Q)^{-\beta}(1 + |x_Q - x_{Q_j}|/\ell(Q))^{-\beta-\varrho}\sum_{I \subset Q_j}(\ell(I)/\ell(Q))^{\varrho}\mu(I).$$

We estimate the last expression by writing the sum as

$$\sum_{k=0}^{\infty}\sum_{\substack{I\,:\,I \subset Q_j \\ \ell(I)=2^{-k}\ell(Q_j)}} 2^{-k\varrho}\mu(I) = C_{\varrho}\mu(Q_j),$$

where we recall that $\ell(Q_j) = \ell(Q)$.

Thus, the second factor bounding (II) is dominated by a constant times the square root of

$$\ell(Q)^{-\beta}\sum_j \mu(Q_j)(1 + |x_Q - x_{Q_j}|/\ell(Q))^{-\beta-\varrho},$$

which, following the reasoning above, is bounded by a constant times

$$R^{-\beta}\int_{\mathbb{R}^2}(1 + |x - x_Q|/R)^{-\beta-\varrho}\,d\mu(x),$$

for $R = \ell(Q)$—and is therefore less than a constant.

Now we deal with (I). The cubes I in this sum all satisfy $d(I,Q) \leq c\ell(Q)$ and (from our "good square" hypothesis) either $\ell(I) > 2^{-n}\ell(Q)$ or $d(I,Q) > \ell(I)^{\alpha}\ell(Q)^{1-\alpha}$. The upshot is that, for any $x \in Q$,

$$c_1\ell(I)^{\alpha}\ell(Q)^{1-\alpha} < |x - x_I| \leq c_2\ell(Q).$$

Let x and z be arbitrary points in Q. The preceding inequalities imply

$$1 + |x - x_I|/\ell(I) > c(\ell(Q)/\ell(I))^{1-\alpha}$$

and

$$1 + |z - x_I|/\ell(I) < c\ell(Q)/\ell(I),$$

where the second inequality uses the fact that $1 \leq \ell(Q)/\ell(I)$. Combining these, we get

$$(1 + |x - x_I|/\ell(I))^{-\beta-2\varepsilon+\varrho} \leq c(1 + |z - x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)}$$

with c an absolute constant.

By the Cauchy–Schwarz inequality and the preceding inequality,

$$(\mathrm{I}) \leq \Big(\sum_{j \,:\, d(Q,Q_j)=0} \sum_{I \,:\, I \subset Q_j} |\lambda_I|^2 \ell(I)^{-\beta}(1+|x-x_I|/\ell(I))^{-\beta-2\varepsilon+\varrho} \Big)^{1/2}$$

$$\times \Big(\sum_{j \,:\, d(Q,Q_j)=0} \sum_{I \,:\, I \subset Q_j} \ell(I)^{-\beta}\mu(I)(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \Big)^{1/2}$$

$$\leq c\Big(\sum_{j \,:\, d(Q,Q_j)=0} \sum_{I \,:\, I \subset Q_j} |\lambda_I|^2 \ell(I)^{-\beta}(1+|z-x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)} \Big)^{1/2}$$

$$\times \Big(\sum_{j \,:\, d(Q,Q_j)=0} \sum_{I \,:\, I \subset Q_j} \ell(I)^{-\beta}\mu(I)(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \Big)^{1/2}$$

$$\leq cg^*(f)(z)\Big(\sum_{j \,:\, d(Q,Q_j)=0} \sum_{I \,:\, I \subset Q_j} \ell(I)^{-\beta}\mu(I)(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \Big)^{1/2}.$$

We will finish the proof by showing that the second factor in the last expression is bounded by a constant.

This is where it becomes essential to make α very small.

By the inequality proved in the last paragraph but one,

$$(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \leq c(\ell(I)/\ell(Q))^{(\beta+\varrho)(1-\alpha)},$$

which implies

$$\ell(I)^{-\beta}(1+|x-x_I|/\ell(I))^{-\beta-\varrho} \leq \ell(Q)^{-\beta}(\ell(I)/\ell(Q))^{(\beta+\varrho)(1-\alpha)-\beta}.$$

We now choose α so small that $\eta \equiv (\beta+\varrho)(1-\alpha)-\beta > 0$. Then the second factor in the last expression is bounded by a constant times the square root of

$$\sum_{j \,:\, d(Q,Q_j)=0} \ell(Q_j)^{-\beta} \sum_{I \,:\, I \subset Q_j} \mu(I)(\ell(I)/\ell(Q_j))^{\eta},$$

which is easily seen (consult the earlier arguments) to be less than a constant. Lemma 1.8 is proved. ∎

REMARK. The proof of our bound on h_2 has the following consequence, which will be useful later. Suppose that our finite linear sum $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$ is such that $\lambda_I = 0$ unless I is a subset of some fixed $Q_0 \in \Gamma$. Then, for all $z \notin Q_0$, $F^*(z) \leq cg^*(f)(z)$, where c is an absolute constant. To see this, let $z \notin Q_0$ and let $Q \in \Gamma$ be such that $z \in Q$. Either $Q_0 \subset Q$ or $Q \cap Q_0 = \emptyset$. In the former case, $F(Q) = 0$. In the latter case, $|F(Q)|$ is less than or equal to the supremum over Q of $\sum_{I \in \Gamma} |\lambda_I| |\phi_{(I)}|$, which is less than or equal to $cg^*(f)(z)$.

LEMMA 1.9. *If $Q' \in N(Q)$, $\mu(Q') > 0$, and $z \in Q'$ then $|F(Q)-F(Q')| \leq cg^*(f)(z)$.*

Proof.

$$|F(Q) - F(Q')|$$

$$= \left| \frac{1}{\mu(Q)} \int_Q \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x) \Big) \, d\mu(x) - \frac{1}{\mu(Q')} \int_{Q'} \Big(\sum_{I \in S(Q')} \lambda_I \phi_{(I)}(x') \Big) \, d\mu(x') \right|$$

$$\leq \left| \frac{1}{\mu(Q)} \int_Q \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x) \Big) \, d\mu(x) - \frac{1}{\mu(Q')} \int_{Q'} \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x') \Big) \, d\mu(x') \right|$$

$$+ \left| \frac{1}{\mu(Q')} \int_{Q'} \Big(\sum_{I \in S(Q') \setminus S(Q)} \lambda_I \phi_{(I)}(x') \Big) \, d\mu(x') \right|$$

$$= (\mathrm{I}) + (\mathrm{II}).$$

Because the $\phi_{(I)}$'s are continuous, term (I) is no bigger than

$$\left| \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(x) \Big) - \Big(\sum_{I \in S(Q)} \lambda_I \phi_{(I)}(y) \Big) \right|$$

for some points x and y in Q. By Lemma 1.8, this is

$$\leq c \inf_{z \in Q} g^*(f)(z) \leq c \inf_{z \in Q'} g^*(f)(z),$$

so (I) is all right.

We may bound the integrand in (II) the same way we bounded h_2 in the proof of Lemma 1.8 (strictly speaking, the way we bounded (I) in that proof); details are left to the reader. Lemma 1.9 is proved. ∎

LEMMA 1.10. *Let $Q' \in N(Q)$. There is a $c > 0$ such that, for all $z \in Q$, $G(Q') \leq c g^*(f)(z)$.*

REMARK. The purpose of this lemma is to give a *lower* bound on $g^*(f)$.

Proof. We may assume that $\mu(Q') > 0$ and take $z \in Q \setminus Q'$. Take $I \in S(Q')$ and $x \in Q'$. Then it is enough to show

$$\frac{\chi_I(x)}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho}$$

$$\leq c \left[\frac{\chi_I(z)}{\mu(I)} + \ell(I)^{-\beta}(1 + |z - x_I|/\ell(I))^{(-\beta - 2\varepsilon + \varrho)(1 - \alpha)} \right].$$

If $Q' \subset I$ then (since $I \neq Q'$) $\chi_I \equiv 1$ across Q and the inequality is trivial. If $z \in I$ and $\ell(I) \leq \ell(Q')$, then $|z - x_I| \leq c|x - x_I|$, and the inequality is once again trivial. If $\ell(I) \geq \ell(Q')$ and $I \cap Q' = I \cap Q = \emptyset$ then $|z - x_I|$ and $|x - x_I|$ are comparable, and the inequality is trivial. If $\ell(I) \leq \ell(Q')$ and $d(I, Q') > \ell(Q')$ then $|z - x_I|$ and $|x - x_I|$ are still comparable, and the inequality is still trivial.

The nub of the matter is: $\ell(I) \le \ell(Q')$, $z \notin I$, $I \cap Q' = \emptyset$, $d(I, Q') \le \ell(Q')$. Now we use the fact that I is good, which implies that either $\ell(I) > 2^{-n}\ell(Q')$ or $d(I, Q') > \ell(I)^\alpha \ell(Q')^{1-\alpha}$. These imply that, for all $x \in Q'$,

$$|x - x_I| > c\ell(I)^\alpha \ell(Q')^{1-\alpha},$$

where c depends on n. Thus

$$1 + |x - x_I|/\ell(I) > c(\ell(Q')/\ell(I))^{1-\alpha}.$$

On the other hand,

$$|z - x_I| \le c\ell(Q'),$$

implying

$$1 + |z - x_I|/\ell(I) \le c\ell(Q')/\ell(I),$$

where we use the fact that $\ell(Q')/\ell(I) \ge 1$. Raising the last two inequalities to the appropriate negative powers, we get

$$(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho} \le c(1 + |z - x_I|/\ell(I))^{(-\beta - 2\varepsilon + \varrho)(1-\alpha)},$$

which proves our result. ∎

LEMMA 1.11. *Let $Q_0 \in \Gamma$ and $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$ where the $\phi_{(I)}$'s belong to some good standard family relative to Γ. Assume that $\lambda_I = 0$ for $I \in S(Q_0)$ (i.e., the only I's that count are properly contained in Q_0). For every $\tilde{\gamma} > 0$ there is a $\gamma > 0$ such that*

$$\mu(\{x \in Q_0 : F^*(x) > 1,\, g^*(f)(x) \le \gamma\}) \le \tilde{\gamma}\mu(Q_0).$$

Proof. Let $\mathcal{J}$ be the family of maximal $Q \subset Q_0$ with the property that there exists a $Q' \in N(Q)$ such that $G(Q') > A\gamma$, where A is a large positive constant to be chosen presently—in fact, right now. By Lemma 1.10, we have $g^*(f) > cA\gamma$ on all of any $Q \in \mathcal{J}$. Pick $A = c^{-1}$.

Before proceeding, let us observe that, for every $Q \in \mathcal{J}$, $G(Q) \le A\gamma$, and also that, if $x \notin \bigcup_{\mathcal{J}} Q$, then $G^*(x) \le A\gamma$.

Let $\mathcal{K}$ be the family of maximal $I \subset Q_0$ that are not properly contained in any $Q \in \mathcal{J}$ and which satisfy $|F(I)| > 1$.

Take the union $\mathcal{J} \cup \mathcal{K}$ and consider its family of maximal cubes $\mathcal{L}$. We claim

(1.5) $\quad \{x \in Q_0 : F^*(x) > 1,\, g^*(f)(x) \le \gamma\}$

$$\subset \bigcup_{\substack{Q\,:\,Q \in \mathcal{L} \\ |F(Q)| > 1}} \{x \in Q : g^*(f)(x) \le \gamma\}.$$

The proof is easy. Suppose $F^*(x) > 1$. Then $|F(I)| > 1$ for some maximal I containing x. But if I is properly contained in any $Q \in \mathcal{J}$, then $g^*(f)(x) > \gamma$. So x must belong to the right-hand union of (1.5).

Because the $Q \in \mathcal{L}$ are disjoint, our problem reduces to bounding

$$\sum_{\substack{Q \in \mathcal{L} \\ |F(Q)| > 1}} \mu(\{x \in Q : g^*(f)(x) \leq \gamma\}).$$

Define $\mathcal{F}_1 \equiv \{I \subset Q_0 : \forall J \in \mathcal{L} \ (I \not\subset J)\}$ and $\mathcal{F}_2 \equiv \{I \subset Q_0 : I \subset J$ for some $J \in \mathcal{L}\}$, and set

$$f_l = \sum_{I \in \mathcal{F}_l} \lambda_I \phi_{(I)}$$

for $l = 1, 2$. Similarly, define analogous partial sum and maximal operators $F_l(Q)$ and $F_l^*(x)$; and Littlewood–Paley operators $G_l(Q)$, $G_l^*(x)$, and $g^*(f_l)$. Let us note in passing that both of the $G_l(Q)$'s are $\leq A\gamma$ for all $Q \in \mathcal{L}$ and that $G_1^*(x) \leq A\gamma$ if $x \notin \bigcup_{\mathcal{L}} Q$.

Then our problem has reduced (again!) to bounding

$$\sum_{\substack{Q \in \mathcal{L} \\ |F_l(Q)| > 1/2}} \mu(\{x \in Q : g^*(f)(x) \leq \gamma\}).$$

for $l = 1, 2$.

We take $l = 1$ first. We observe that, for $x \in Q \in \mathcal{L}$,

$$f_1(x) = \sum_{I \in \mathcal{F}_1} \lambda_I \phi_{(I)}(x) = \sum_{\substack{I \in \mathcal{F}_1 \\ I \not\subset Q}} \lambda_I \phi_{(I)}(x) r = F_1(Q, x).$$

(Think about it.)

Because of the $\phi_{(I)}$s' almost-orthogonality,

$$\int |f_1(x)|^2 \, d\mu(x) \leq \sum_{I \in \mathcal{F}_1} |\lambda_I|^2 = \int \left(\sum_{I \in \mathcal{F}_1} \frac{|\lambda_I|^2}{\mu(I)} \chi_I \right) d\mu.$$

If $x \in Q \in \mathcal{L}$, then the last integrand is $\leq (G_1(Q))^2 \leq (A\gamma)^2$. If $x \notin \bigcup_{\mathcal{L}} Q$, then it is $\leq (G_1^*(x))^2 \leq (A\gamma)^2$. Therefore,

$$\int |f_1(x)|^2 \, d\mu(x) \leq (A\gamma)^2 \mu(Q_0).$$

Now the weak-type $(2, 2)$ bound for the μ-weighted *dyadic* maximal function implies

$$\sum_{\substack{Q \in \mathcal{L} \\ |F_1(Q)| > 1/2}} \mu(Q) \leq 4(A\gamma)^2 \mu(Q_0),$$

yielding our estimate for F_1.

Now take $l = 2$. Enumerate the members of $\mathcal{L}$ as $\{Q_k\}$, and temporarily fix $x \in Q_j \in \mathcal{L}$. Without loss of generality, we can assume that

$\inf_{Q_j} g^*(f)(z) \leq \gamma$. Our first task is to estimate $F_2(Q_j, x)$, which is

$$\sum_{k\,:\,k\neq j} \sum_{I\subset Q_k} \lambda_I \phi_{(I)}(x)$$

where $x \in Q_j$ and—important fact!—the I's are all good.

By the Cauchy–Schwarz inequality,

$$(1.6) \quad |F_2(Q_j, x)|^2 \leq \Big(\sum_{k\,:\,k\neq j} \sum_{I\subset Q_k} |\lambda_I|^2 \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{-\beta - 2\varepsilon + \varrho} \Big)$$

$$\times \Big(\sum_{k\,:\,k\neq j} \sum_{I\subset Q_k} \ell(I)^{-\beta}\mu(I)(1 + |x - x_I|/\ell(I))^{-\beta - \varrho} \Big).$$

By an argument which is by now familiar, the first factor on the right-hand side of (1.6) is bounded by a constant times

$$\inf_{Q_j} g^*(f)(z) \leq \gamma^2.$$

We claim the second is bounded by a constant times

$$\sum_{k\,:\,k\neq j} \ell(Q_k)^{-\beta} \sum_{I\subset Q_k} (\ell(I)/\ell(Q_k))^\eta \mu(I)(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho},$$

where $\eta = (\beta + \varrho)(1 - \alpha) - \beta > 0$.

To see this, fix k and consider $x \notin Q_k$ and $I \subset Q_k$, with I good.
If $d(x, Q_k) \leq \ell(Q_k)$, then

$$|x - x_I| \geq c\ell(I)^\alpha \ell(Q_k)^{1-\alpha},$$

implying

$$1 + |x - x_I|/\ell(I) > c(\ell(Q_k)/\ell(I))^{1-\alpha}(1 + |x - x_{Q_k}|/\ell(Q_k)),$$

and thus

$$\ell(I)^{-\beta}\mu(I)(1 + |x - x_I|/\ell(I))^{-\beta - \varrho}$$
$$\leq c\ell(Q_k)^{-\beta}(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho}(\ell(I)/\ell(Q_k))^\eta \mu(I).$$

On the other hand, if $d(x, Q_k) > \ell(Q_k)$, then

$$|x - x_I|/\ell(I) > c(|x - x_{Q_k}|/\ell(Q_k))(\ell(Q_k)/\ell(I)),$$

which implies

$$1 + |x - x_I|/\ell(I) > c(1 + |x - x_{Q_k}|/\ell(Q_k))(\ell(Q_k)/\ell(I)),$$

yielding

$$\ell(I)^{-\beta}\mu(I)(1 + |x - x_I|/\ell(I))^{-\beta - \varrho}$$
$$\leq c\ell(Q_k)^{-\beta}(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho}(\ell(I)/\ell(Q_k))^\varrho \mu(I)$$
$$\leq c\ell(Q_k)^{-\beta}(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho}(\ell(I)/\ell(Q_k))^\eta \mu(I),$$

since $\eta < \varrho$. This proves the claim.

Now, for each fixed $k \neq j$,

$$\sum_{I \subset Q_k} (\ell(I)/\ell(Q_k))^{\eta} \mu(I) \leq C\mu(Q_k),$$

and so the second factor is dominated by

$$\sum_{k \,:\, k \neq j} \ell(Q_k)^{-\beta} \mu(Q_k)(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho}.$$

Therefore,

$$|F_2(Q_j, x)|^2 \leq C\gamma^2 \Big(\sum_{k \,:\, k \neq j} \ell(Q_k)^{-\beta} \mu(Q_k)(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho} \Big)$$

$$\leq C\gamma^2 \Big(\sum_k \ell(Q_k)^{-\beta} \mu(Q_k)(1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho} \Big).$$

Therefore

$$\sum_j \int_{Q_j} |F_2(Q_j, x)|^2 \, d\mu(x)$$

$$\leq C\gamma^2 \sum_k (\ell(Q_k)^{-\beta} \mu(Q_k) \int (1 + |x - x_{Q_k}|/\ell(Q_k))^{-\beta - \varrho} \, d\mu(x))$$

$$\leq C\gamma^2 \sum_k \mu(Q_k) \leq C\gamma^2 \mu(Q_0),$$

and now our bound follows from a weak $(2, 2)$ estimate as before. This proves Lemma 1.11. ∎

THEOREM 1.12. *Let* $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$ *where the* $\phi_{(I)}$*'s belong to some good standard family relative to* Γ*. Furthermore, suppose that all of the* I*'s occurring in* f*'s sum are subsets of some* $Q' \in \Gamma$*. For every* $\tilde{\gamma} > 0$ *there is a* $\gamma > 0$ *such that, for all* $\lambda > 0$*,*

$$\mu(\{x : F^*(x) > 2\lambda, \, g^*(f)(x) \leq \gamma\lambda\}) \leq \tilde{\gamma}\mu(\{x : F^*(x) > \lambda\}).$$

Proof. Because of the remark following the proof of Lemma 1.8, we have $F^* \leq cg^*(f)$ outside Q'. Therefore, for sufficiently small γ,

$$\mu(\{x \notin Q' : F^*(x) > 2\lambda, \, g^*(f)(x) \leq \gamma\lambda\}) = 0.$$

The remark also points out that $F(Q)$ will be 0 for all Q's that contain Q'. Therefore we only need to consider x's inside Q' and $F(Q)$'s for $Q \subset Q'$. For $\lambda > 0$, let $\{Q_j\}$ be the maximal $Q_j \subset Q'$ such that $|F(Q_j)| > \lambda$. Such maximal Q_j's exist because $F(Q') = 0$. Note also that each Q_j is *strictly* contained in Q'. Our problem now reduces to showing, for each Q_j,

(1.7) $$\mu(\{x \in Q_j : F^*(x) > 2\lambda, \, g^*(f)(x) \leq \gamma\lambda\}) \leq \tilde{\gamma}\mu(Q_j).$$

It will turn out that $\tilde{\gamma}$ can be taken to be $c\gamma^2$ where c is an absolute constant.

Let $Q_j \in N(Q^*)$. Then $|F(Q^*)| \leq \lambda$. By Lemma 1.9, $|F(Q_j) - F(Q^*)| \leq c \inf_{Q_j} g^*(f)$. If $|F(Q_j)| > 1.1\lambda$, then, by taking γ small enough, the left-hand side of (1.7) is 0. We do take γ to be "small enough," which allows us to assume $|F(Q_j)| \leq 1.1\lambda$. Now Lemma 1.8 implies that, for all $x \in Q_j$, $|F(Q_j, x)| \leq 1.1\lambda + c\gamma\lambda$, which will be $\leq 1.2\lambda$ if we take γ small enough (and we do). Therefore we can *ignore* the $F(Q_j, x)$ portion of f and concentrate on

$$\tilde{f} \equiv \sum_{I \subset Q_j} \lambda_I \phi_{(I)}.$$

Define an analogous partial sum and Littlewood–Paley objects $\tilde{F}(Q)$, $\tilde{G}(Q)$, $\tilde{g}^*(\tilde{f})$, etc., for this "truncated" sum. Then our problem has (at last!) reduced to showing

$$\mu(\{x \in Q_j : \tilde{F}^*(x) > .8\lambda, \, \tilde{g}^*(\tilde{f})(x) \leq \gamma\lambda\}) \leq \tilde{\gamma}\mu(Q_j).$$

But, after dividing $\tilde{f}$ by $.8\lambda$, this becomes the statement of Lemma 1.11, with $\tilde{\gamma} \approx \gamma^2$. Theorem 1.12 is proved. ∎

Proof of Theorem 1.13. There exist congruent squares Q_1', Q_2', Q_3', and Q_4', all in Γ, such that all of the I's in f's sum are contained in $\bigcup Q_i'$. Define

$$f_i = \sum_{\substack{I \in \Gamma \\ I \subset Q_i'}} \lambda_I \phi_{(I)},$$

with corresponding partial sum and Littlewood–Paley objects F_i, F_i^*, G_i, $g^*(f_i)$, etc. Notice that, since we are dealing with finite linear sums, all of these objects are bounded. It is now important to recall a basic fact about the method of good-λ inequalities. Suppose that h_1 and h_2 are two non-negative functions defined on some measure space (X, m), and that for every $\varepsilon > 0$ there is a $\delta > 0$ such that, for all $\lambda > 0$,

$$m(\{x \in X : h_1(x) > 2\lambda, \, h_2(x) \leq \delta\lambda\}) \leq \varepsilon m(\{x \in X : h_1(x) > \lambda\}).$$

Then, given $0 < p < \infty$, there is a constant C such that

$$\int_X (h_1(x))^p \, dm \leq C \int_X (h_2(x))^p \, dm$$

provided that the left-hand integral is finite. By Theorem 1.12, a good-λ inequality holds between $h_1 = F_i^*$ and $h_2 = g^*(f_i)$. Therefore our desired inequality,

$$\int |f|^p \, d\mu \leq C \int (g^*(f))^p \, d\mu,$$

will follow if we can show that

$$\int (F_i^*)^p \, d\mu < \infty$$

for each i, under the assumption that

$$\int (g^*(f_i))^p \, d\mu < \infty.$$

Now, we have just observed that F_i^* is bounded. Therefore the integral

$$\int_{Q_i'} (F_i^*)^p \, d\mu$$

presents no problem. On the other hand, as we noted in the proof of Theorem 1.12, we have $F_i^* \leq cg^*(f_i)$ outside Q_i'. That was all we needed to finish the proof of Theorem 1.13. If instead we consider integrals with respect to a measure ν which is A_∞ relative to μ, then the μ-based good-λ inequality immediately implies the corresponding ν-based good-λ inequality. At the risk of being redundant (but to avoid a possible confusion), recall that, in this case, our maximal functions F_i^*, etc., and Littlewood–Paley objects G_i^*, g^*, etc., are still the ones based on μ. The boundedness of F_i^* and the pointwise inequality $F_i^* \leq cg^*(f_i)$ outside Q_i' now let us conclude that

$$\int |f|^p \, d\nu \leq C \int (g^*(f))^p \, d\nu$$

holds as well, the only difference being that now the constant C carries an additional dependence on ν's A_∞ specifications.

2. When some squares are bad: the averaging trick. Now we come to an averaging trick that will allow us to extend the Main Theorem, with a little modification, to arbitrary (not necessarily good) finite sums $f = \sum_{I \,:\, I \in \Gamma} \lambda_I \phi_{(I)}$. Here we will depend heavily on the method of random dyadic grids due to Nazarov, Treil, and Volberg.

Let us assume that, for every $I \in \Gamma$, there is an integer k such that $\ell(I) = 2^k$. Fix a positive number $0 < \tau < 1/2$. We will now construct a family of random dyadic grids $\Gamma'(\omega)$, where ω belongs to a probability space Ω with probability measure P (to be specified shortly). The squares in these grids will have sidelengths $(1 + \tau/2)2^k$; i.e., for each $Q \in \Gamma'(\omega)$, there will exist an integer k such that $\ell(Q) = (1 + \tau/2)2^k$.

We construct our random dyadic grids $\Gamma'(\omega)$ this way. Let n_0 be so large that 2^{-n_0} is much, much smaller than $\tau\ell(I)$ for any I occurring in the sum that defines f, and fix n_0. Build a "base grid" of squares of the form $[j(1+\tau/2)2^{-n_0}, (j+1)(1+\tau/2)2^{-n_0}) \times [j'(1+\tau/2)2^{-n_0}, (j'+1)(1+\tau/2)2^{-n_0})$, where j and j' are arbitrary integers. Following [NTV], we subdivide these squares into fourths, sixteenths, etc., in the usual fashion, to obtain the smaller squares of the grid $\Gamma'(\omega)$. We follow [NTV] also in working "upward", letting the square $[0, (1+\tau/2)2^{-n_0}) \times [0, (1+\tau/2)2^{-n_0})$ have a probability $1/4$ of being the northeast, northwest, southwest, or southeast corner of a square of sidelength $(1+\tau/2)2^{-n_0+1}$, drawing the other squares of

sidelength $(1 + \tau/2)2^{-n_0+1}$ consistent with this, and repeating this random choice forever on successively larger squares.

Let us say that $I \in \Gamma$ is *enveloped* (relative to $\Gamma'(\omega)$) if there exists a $Q \in \Gamma'(\omega)$ such that $I \subset Q \subset (1 + \tau)I$. By drawing pictures the reader can quickly convince himself that the probability that any I in f's sum is enveloped is bounded below by $c\tau^2$, where c is an absolute constant. Note that any I is enveloped by at most one $Q \in \Gamma'(\omega)$.

Now, given that I is enveloped, the conditional probability that the enveloping Q is *bad* relative to $\Gamma'(\omega)$ is less than or equal to a constant times

$$\frac{2^{-n\alpha}}{1 - 2^{-\alpha}},$$

which we can make $< 1/2$ by taking n large (see [NTV]). Let us assume that n is so fixed.

Define, for I in f's sum,

$$\chi(I, \omega) = \begin{cases} 1 & \text{if } I \text{ is enveloped by a good } Q \in \Gamma'(\omega), \\ 0 & \text{otherwise.} \end{cases}$$

We have observed that, for such I,

$$c_I \equiv \int_\Omega \chi(I, \omega)\, dP(\omega) > c\tau^2,$$

for an absolute constant c.

We write

$$f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)} = \int_\Omega \Big(\sum_{I \in \Gamma} c_I^{-1} \lambda_I \phi_{(I)} \chi(I, \omega) \Big)\, dP(\omega)$$

$$\equiv \int_\Omega \Big(\sum_{I \in \Gamma} \gamma(I, \omega) \phi_{(I)} \Big)\, dP(\omega).$$

We have written f as an average of sums $\sum_{I \in \Gamma} \gamma(I, \omega) \phi_{(I)}$, where each $|\gamma(I, \omega)| \leq C\tau^{-2}|\lambda_I|$ and is zero if I is not enveloped by a good $Q \in \Gamma'(\omega)$. For such Q and I, the following inequalities are trivial:

$$\frac{\chi_I(x)}{\sqrt{\mu((1 + \tau)I)}} \leq \frac{\chi_Q(x)}{\mu(Q)} \leq \frac{\chi_{(1+\tau)I}(x)}{\sqrt{\mu(I)}};$$

also

$$\ell(I)^{-\beta} \sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-r}$$

$$\leq C_1(\beta, r)\ell(Q)^{-\beta} \sqrt{\mu(Q)}(1 + |x - x_Q|/\ell(Q))^{-r}$$

$$\leq C_2(\beta, r)\ell(I)^{-\beta} \sqrt{\mu((1 + \tau)I)}(1 + |x - x_I|/\ell(I))^{-r},$$

independent of x, where the constants only depend on β and the positive number r.

Whenever $\gamma(I,\omega) \neq 0$, let us denote I's enveloping, good $Q \in \Gamma(\omega)$ by $Q(I,\omega)$.

Let us recall that a family $\{\phi_{(I)}\}_I$, indexed over $I \in \Gamma$, is τ-*adapted* to Γ if, for all $I \in \Gamma$ and all x and y in $\mathbb{R}^2$,

$$|\phi_{(I)}(x)| \leq \frac{\chi_I(x)}{\sqrt{\mu((1+\tau)I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}(1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon},$$

$$|\phi_{(I)}(x) - \phi_{(I)}(y)| \leq \left(\frac{|x-y|}{\ell(I)}\right)^{\delta}\left(\frac{\chi_I(x) + \chi_I(y)}{\sqrt{\mu((1+\tau)I)}} + \ell(I)^{-\beta}\sqrt{\mu(I)}\right.$$
$$\left. \times ((1 + |x - x_I|/\ell(I))^{-\beta-\varepsilon} + (1 + |y - x_I|/\ell(I))^{-\beta-\varepsilon})\right),$$

and, for all finite linear sums $\sum \gamma_I \phi_{(I)}$,

$$\int \left|\sum \gamma_I \phi_{(I)}\right|^2 d\mu \leq \sum |\gamma_I|^2.$$

If $\{\phi_{(I)}\}_I$ is τ-adapted to Γ, then, for every $\omega \in \Omega$, $\{\phi_{(Q(I,\omega))}\}_I$ is, modulo an absolute positive constant, a good standard family relative to $\Gamma(\omega)$, where we are setting $\tilde{\phi}_{(Q(I,\omega))} \equiv \phi_{(I)}$ when $\gamma(I,\omega) \neq 0$, and setting it equal to 0 otherwise. By Theorem 1.13, for every $\omega \in \Omega$,

$$\int_{\mathbb{R}^2} \left|\sum_I \gamma(I,\omega)\tilde{\phi}_{(Q(I,\omega))}\right|^p d\mu$$

$$\leq C \int_{\mathbb{R}^2} \left(\sum_I |\gamma(I,\omega)|^2\right.$$
$$\left. \times \left(\frac{\chi_{Q(I,\omega)}}{\mu(Q(I,\omega))} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)}\right)\right)^{p/2} d\mu,$$

which is less than or equal to

$$C \int_{\mathbb{R}^2} \left(\sum_I |\lambda_I|^2 \left(\frac{\chi_{(1+\tau)I}}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)}\right)\right)^{p/2} d\mu$$

for *every* $\omega \in \Omega$. Define

$$\tilde{g}^*(f)(x)$$
$$\equiv \left(\sum_I |\lambda_I|^2 \left(\frac{\chi_{(1+\tau)I}}{\mu(I)} + \ell(I)^{-\beta}(1 + |x - x_I|/\ell(I))^{(-\beta-2\varepsilon+\varrho)(1-\alpha)}\right)\right)^{1/2}.$$

By taking averages in $\omega \in \Omega$, we obtain, for $1 \leq p < \infty$,

$$(2.1) \qquad \int_{\mathbb{R}^2} |f|^p d\mu \leq C \int_{\mathbb{R}^2} (\tilde{g}^*(f))^p d\mu,$$

for all finite linear sums from a τ-adapted family. We also immediately obtain the inequality

$$(2.2) \qquad \int_{\mathbb{R}^2} |f|^p \, d\sigma \le C \int_{\mathbb{R}^2} (\tilde{g}^*(f))^p \, d\sigma,$$

valid for measures σ that are A_∞ relative to μ. This gives Theorem 2.1 for $1 \le p < \infty$.

An ingenious trick, which the author learned from F. Nazarov, allows us to extend (2.1) and (2.2) to p's below 1. Let us suppose we have a fixed grid Γ and a finite, τ-adapted family $\{\phi_{(I)}\}_I$. We consider the family of linear combinations $f = \sum_{I \in \Gamma} \lambda_I \phi_{(I)}$ that satisfy

$$\int (g^*(f))^p \, d\mu \le 1,$$

and we let R be the maximum value of

$$\int |f|^p \, d\mu$$

for all such f. It is obvious that R is finite and that the maximum is attained by some f. What we need to show is that R is bounded by some absolute constant, independent of the family $\{\phi_{(I)}\}_I$.

For every $\omega \in \Omega$, we write

$$f = f_g(\omega) + f_b(\omega) = \sum_I \lambda_I \chi(I, \omega) \phi_{(I)} + \sum_I \lambda_I (1 - \chi(I, \omega)) \phi_{(I)}.$$

So, for each $\omega \in \Omega$, f_g is the "good" part of f (it is indexed only over squares I that are enveloped by squares that are good relative to $\Gamma'(\omega)$) and f_b is the "bad" part. It is trivial that

$$(g^*(f))^2 = (g^*(f_g(\omega)))^2 + (g^*(f_b(\omega)))^2$$

for each $\omega \in \Omega$. Also, there is a $\delta > 0$ such that, for every I in f's sum,

$$\int_\Omega \chi(I, \omega) \, dP(\omega) \ge \delta,$$

and therefore

$$\int_\Omega (g^*(f_b)(\omega))^2 \, dP(\omega) \le (1 - \delta)(g^*(f))^2.$$

Note, by the way, that the preceding is a *pointwise* inequality (we have suppressed its dependence on x).

Let us assume that $\int (g^*(f))^p \, d\mu(x) \le 1$, and let us choose an optimal f, i.e., one for which $\int |f|^p \, d\mu(x) = R$; such an f exists by compactness. Since $0 < p \le 1$, we have, for each $\omega \in \Omega$,

$$R = \int |f|^p \, d\mu(x) \le \int |f_g(\omega)|^p \, d\mu(x) + \int |f_b(\omega)|^p \, d\mu(x)$$

$$\le C + R \int (g^*(f_b(\omega)))^p \, d\mu(x).$$

The "C" (an absolute constant) in the second inequality follows from the fact that f_g's decomposition only involves "good" squares. We claim that, for some $\omega_0 \in \Omega$, $\int (g^*(f_b(\omega_0)))^p \, d\mu$ is less than or equal to $1 - \delta'$, for some δ' that is independent of f. That will prove the result, because it will yield

$$R \leq C + (1 - \delta')R,$$

implying $R \leq C/\delta'$.

We write

$$\iint_\Omega (g^*(f_b(\omega)))^p \, d\mu(x) \, dP(\omega) = \iint_\Omega (g^*(f_b(\omega)))^2)^{p/2} \, dP(\omega) \, d\mu(x)$$

$$\leq \int \left(\int_\Omega (g^*(f_b(\omega)))^2 \, dP(\omega) \right)^{p/2} d\mu(x)$$

$$\leq (1 - \delta)^{p/2} \int (g^*(f))^p \, d\mu(x)$$

$$\leq (1 - \delta)^{p/2} \equiv 1 - \delta'.$$

This implies that $\int (g^*(f_b(\omega)))^p \, d\mu(x)$ must be less than or equal to $1 - \delta'$ for at least one $\omega_0 \in \Omega$, and that is what we wanted. Theorem 2.1 is proved.

A final remark. These arguments work equally well, with obvious modifications, in $\mathbb{R}^d$ when $d > 2$. A normalized regular measure μ is now defined to be one such that $\mu(Q) \leq \ell(Q)^\beta$ for all cubes Q and for some fixed $0 < \beta \leq d$. The definitions of good and bad squares, good standard family, and the maximal functions and Littlewood–Paley objects are unchanged, except that now β can be as big as d. The proofs of Lemmas 1.1–1.11 and Theorems 1.12 and 1.13 are unchanged, except that now some of the multiplicative constants depend on d. The definitions of "enveloped" and τ-adapted do not change, but the lower bound on c_I, which was $c\tau^2$, becomes $c\tau^d$. This introduces an additional dependence in the constant C in the $\mathbb{R}^d$ generalization of Theorem 2.1. And, of course, if ν is a d-dimensional measure that is A_∞ relative to μ, the corresponding extension of Theorem 2.1 holds for it as well.

References

[MMV]   P. Mattila, M. Melnikov and J. Verdera, *The Cauchy integral, analytic capacity, and uniform rectifiability*, Ann. of Math. 144 (1996), 127–137.

[NTV]   F. Nazarov, S. Treil and A. Volberg, *Cauchy integral and Calderón–Zygmund operators on nonhomogeneous spaces*, Int. Math. Res. Not. 1997, no. 15, 703–726.

[T1]   X. Tolsa, L^2 *boundedness of the Cauchy integral operator for continuous measures*, Duke Math. J. 98 (1999), 269–304.

[T2]   —, *Cotlar's inequality and existence of principal values for the Cauchy integral without the doubling condition*, J. Reine Angew. Math. 502 (1998), 199–235.

[T3]   —, *Weighted norm inequalities for Calderón–Zygmund operators without doubling conditions*, preprint, 2001.

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

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

[W2] —, *Global orthogonality implies local almost-orthogonality*, Rev. Mat. Iberoamericana 16 (2000), 29–48.

Department of Mathematics
University of Vermont
Burlington, VT 05405, U.S.A.
E-mail: wilson@cem.uvm.edu