The intrinsic square function

Michael Wilson

Abstract

We show that the Lusin area function and essentially all of its real-variable generalizations are pointwise dominated by an "intrinsic" square function, and that this latter function is, for all practical purposes, no larger than a "generic" square function.

0. Introduction

The Lusin area (or "square") function is a familiar object. If $f : \mathbb{R}^d \mapsto \mathbb{R}$ is such that $u(x,y) \equiv P_y * f(x)$, the Poisson integral of f, is defined for all $(x,y) \in \mathbb{R}^{d+1}_+ = \mathbb{R}^d \times (0, \infty)$, then we define the Lusin function $S(f)$ by

$$S(f)(x) \equiv \left(\int_{\Gamma(x)} |\nabla u(t,y)|^2 \frac{dt\,dy}{y^{d-1}} \right)^{1/2},$$

where $\Gamma(x)$ denotes the usual "cone of aperture one",

$$\Gamma(x) = \{(t,y) : \ |x - t| < y\}.$$

We can similarly define a cone of aperture β for any $\beta > 0$:

$$\Gamma_\beta(x) = \{(t,y) : \ |x - t| < \beta y\},$$

and corresponding square function

$$S_\beta(f)(x) \equiv \left(\int_{\Gamma_\beta(x)} |\nabla u(t,y)|^2 \frac{dt\,dy}{y^{d-1}} \right)^{1/2}.$$

There are also "zero-aperture" and (so to speak) "infinite aperture" variations of $S(f)$.

The zero-aperture one is called the Littlewood-Paley g-function and it is defined by

$$g(f)(x) \equiv \left(\int_0^\infty y |\nabla u(x,y)|^2 \, dy \right)^{1/2}.$$

The infinite-aperture object depends on a positive parameter λ and is called the g_λ^*-function. It is given by

$$g_\lambda^*(f)(x) \equiv \left(\int_{\mathbb{R}_+^{d+1}} |\nabla u(t,y)|^2 \left(\frac{y}{y + |x-t|} \right)^\lambda \frac{dt \, dy}{y^{d-1}} \right)^{1/2}.$$

These objects satisfy some well-known pointwise inequalities: If $\beta < \beta'$, $S_\beta(f) \leq S_{\beta'}(f)$; if $\lambda < \lambda'$, $g_\lambda^*(f) \geq g_{\lambda'}^*(f)$; for all $\beta > 0$ and $\lambda > 0$, $S_\beta(f) \leq C(\lambda, \beta) g_\lambda^*(f)$; for all $\beta > 0$, $g(f) \leq C(\beta, d) S_\beta(f)$ (the last one makes use of the sub-mean value property of $|\nabla u|^2$).

They also satisfy some familiar norm inequalities. If $1 < p < \infty$ and $f \in L^p$, then $\|f\|_p \sim \|g(f)\|_p$, with comparability constants that only depend on p and d. Similarly, $\|f\|_p \sim \|S_\beta(f)\|_p$, but now the constants depend on β as well. Finally (and with similar qualifications), if $\lambda > 2d/p$, then $\|f\|_p \sim \|g_\lambda^*(f)\|_p$. These inequalities also hold for many weighted L^p spaces.

What makes these functions useful is that they are "almost invariant" under the action of many Fourier multiplier and singular integral operators. For example, if $d = 1$ and H denotes the Hilbert transform, then $S(Hf) = S(f)$ (because $|\nabla(P_y * Hf)(x)| = |\nabla(P_y * f)(x)|$), immediately implying $\|Hf\|_p \sim \|f\|_p$ for $1 < p < \infty$. Unfortunately, for other operators T, things are not so simple. The best one can usually get is that $S(Tf)$ or $g(Tf)$ is bounded by a constant times the relatively unwieldy $g_\lambda^*(f)$, where λ depends on T. This limits the usefulness of these "classical" Littlewood-Paley functions.

Another serious limitation comes from the very definitions of $S_\beta(f)$, and $g(f)$, and $g_\lambda^*(f)$; namely, that they come from convolutions between f and kernels with unbounded supports. This can make it hard to analyze the local behavior of, say, $S(f)$. Suppose we want to know what $S(f)$ is doing on a cube Q. Here it is natural to consider a "truncated" form of $S(f)$, given by the formula

$$S_{tr}(f)(x) \equiv \left(\int_{(t,y)\in\Gamma(x):\ y\leq\ell(Q)} |\nabla u(t,y)|^2 \frac{dt \, dy}{y^{d-1}} \right)^{1/2}, \qquad (0.1)$$

where $\ell(Q)$ denotes the sidelength of Q. Unfortunately, to understand even this simplified $S(f)$, we must separately consider the contributions of f "near" Q (say, in $\tilde{Q}$, the triple of Q) and "far" from Q (everything else); and the contribution from "everything else" can be substantial.

This problem can be overcome by defining a new square function. Let $\psi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ be real, radial, have support contained in $\{x : |x| \leq 1\}$, and be normalized so that

$$\int_0^\infty |\hat{\psi}(y\xi)|^2 \frac{dy}{y} \equiv 1$$

for all $\xi \neq 0$. The real-variable Lusin function $S_\psi(f)$ is defined by

$$S_\psi(f)(x) \equiv \left(\int_{\Gamma(x)} |f * \psi_y(t)|^2 \frac{dt\, dy}{y^{d+1}} \right)^{1/2},$$

where we are using ψ_y to denote the usual L^1 dilation of ψ:

$$\psi_y(x) \equiv y^{-d}\psi(x/y).$$

Like the classical Lusin function, $S_\psi(f)$ satisfies

$$\|S_\psi(f)\|_p \sim \|f\|_p \qquad \text{for } 1 < p < \infty,$$

but $S_\psi(f)(x)$ has the good property that, for $x \in Q$, its truncated form (analogous to (0.1)) only depends on what f does near Q, which makes it (in many cases) easier to analyze. Also like $S(f)$, it is "almost invariant" under the action of singular integral operators; however, in this case the complications become more acute, because we can no longer use properties of harmonic functions to analyze the resulting square functions.

In this paper we introduce a real-variable version of the Lusin function that, within reason, does everything one would ever want a square function to do: it dominates (modulo constants) $g(f)$ and all the $S_\beta(f)$'s, while being essentially no bigger than a generic $S_\psi(f)$, and its truncated form (see (0.1)) only depends on what f does near Q.

This square function is easy to define. For $0 < \alpha \leq 1$, let $\mathcal{C}_\alpha$ be the family of functions $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ such that ϕ's support is contained in $\{x : |x| \leq 1\}$, $\int \phi\, dx = 0$, and, for all x and x',

$$|\phi(x) - \phi(x')| \leq |x - x'|^\alpha.$$

For $(t, y) \in \mathbb{R}_+^{d+1}$ and $f \in L_{loc}^1(\mathbb{R}^d)$, set

$$A_\alpha(f)(t, y) \equiv \sup_{\phi \in \mathcal{C}_\alpha} |f * \phi_y(t)|.$$

We define the *intrinsic square function of* f (of order α) by the formula

$$G_\alpha(f)(x) \equiv \left(\int_{\Gamma(x)} (A_\alpha(f)(t, y))^2 \frac{dt\, dy}{y^{d+1}} \right)^{1/2}.$$

In section 1 we prove

Theorem 1 *If $0 < \alpha \le 1$ and $1 < p \le 2$, there is a constant $C(p, d, \alpha)$ such that, for all $f \in L^1_{loc}(\mathbb{R}^d)$ and all non-negative $v \in L^1_{loc}(\mathbb{R}^d)$,*

$$\int_{\mathbb{R}^d} (G_\alpha(f))^p \, v \, dx \le C(p, d, \alpha) \int_{\mathbb{R}^d} |f|^p \, M(v) \, dx, \qquad (0.2)$$

where $M(\cdot)$ denotes the Hardy-Littlewood maximal operator.

If $\phi \in \mathcal{C}_\alpha$, and we replace $G_\alpha(f)$ with $S_\phi(f)$ (defined just as we did for $\psi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$), then the inequality,

$$\int_{\mathbb{R}^d} (S_\phi(f))^p \, v \, dx \le C \int_{\mathbb{R}^d} |f|^p \, M(v) \, dx, \qquad (0.3)$$

holds for all $1 < p \le 2$, all f , and all v , with a constant C that only depends on p , d , and α [2]. What is significant is that we can find a $\phi \in \mathcal{C}_\alpha$ such that (0.3) fails for both $p \le 1$ and $p > 2$. (Indeed, if $p \le 1$ and $v \equiv 1$, inequality (0.3) fails if f is the characteristic function of a cube.) It is therefore a little (but only a little) surprising that the inequality remains true for $G_\alpha(f)$.

As we did with $S(f)$, we can define varying-aperture (and even zero-aperture) versions of $G_\alpha(f)$:

$$G_{\alpha,\beta}(f)(x) \equiv \left(\int_{\Gamma_\beta(x)} (A_\alpha(f)(t, y))^2 \, \frac{dt \, dy}{y^{d+1}} \right)^{1/2}$$

$$g_\alpha(f)(x) \equiv \left(\int_0^\infty (A_\alpha(f)(x, y))^2 \, \frac{dy}{y} \right)^{1/2}.$$

We can also define a discretized form:

$$\sigma_\alpha(f)(x) \equiv \left(\sum_{-\infty}^{\infty} (A_\alpha(f)(x, 2^k))^2 \right)^{1/2}.$$

All three of these functions— $G_{\alpha,\beta}(f)$, $g_\alpha(f)$, and $\sigma_\alpha(f)$ —are pointwise comparable to $G_\alpha(f)$, with comparability constants which, for the first function, depend on α , β , and d , and, for other two, only depend on α and d . This result is easy (we prove it in section 1). It is useful because, in estimating $G_\alpha(f)$, it means we can choose a form that requires the least work.

In section 2 we show that $G_\alpha(f)$, which is defined via convolutions with kernels supported in $\{x : |x| \le 1\}$, actually dominates similar-looking square functions that are defined via convolutions with kernels that have unbounded supports. One of these similar-looking functions is bigger than

or equal to a constant times $S(f)$, implying that inequality (0.3) also holds for the classical Lusin function. This yields a positive answer to a question raised by R. Fefferman and E. M. Stein. The result, though expected, is slightly surprising, because the proofs of (0.2) and (0.3) –as we will see– make essential use of the assumption that the ϕ's supports are contained in $\{x : |x| \leq 1\}$.

These similar-looking square functions are also easy to define. For $0 < \alpha \leq 1$ and $\epsilon > 0$, let $\mathcal{C}_{(\alpha,\epsilon)}$ be the family of functions $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ such that, for all x,

$$|\phi(x)| \leq (1 + |x|)^{-d-\epsilon}; \tag{0.4}$$

and, for all x and x',

$$|\phi(x) - \phi(x')| \leq |x - x'|^\alpha \big((1 + |x|)^{-d-\epsilon} + (1 + |x'|)^{-d-\epsilon}\big); \tag{0.5}$$

and also satisfy $\int \phi \, dx = 0$.

It is frequently useful to observe that, if ϕ satisfies (0.4), then inequality (0.5) will hold for all x and x', if and only if it holds for all x and x' such that $|x - x'| \leq 1$.

Let f be such that $|f|(1 + |x|)^{-d-\epsilon} \in L^1$. For every $(t, y) \in \mathbb{R}_+^{d+1}$, set

$$\tilde{A}_{(\alpha,\epsilon)}(f)(t, y) \equiv \sup_{\phi \in \mathcal{C}_{(\alpha,\epsilon)}} |f * \phi_y(t)|.$$

We define

$$\tilde{G}_{(\alpha,\epsilon)}(f)(x) \equiv \left(\int_{\Gamma(x)} (\tilde{A}_{(\alpha,\epsilon)}(f)(t, y))^2 \frac{dt \, dy}{y^{d+1}} \right)^{1/2}.$$

There is a positive constant c, depending only on α, ϵ, and d, such that $c\phi \in \mathcal{C}_{(\alpha,\epsilon)}$ whenever $\phi \in \mathcal{C}_\alpha$. Therefore, $G_\alpha(f) \leq C\tilde{G}_{(\alpha,\epsilon)}(f)$. In section 2 we prove that this inequality has a partial converse; one which, for all practical purposes, might as well be a full converse. We prove:

Theorem 2 *Let $0 < \alpha \leq 1$ and $\epsilon > 0$. For every α' satisfying $0 < \alpha' \leq \alpha$ and $\alpha' < \epsilon$, there is a constant $C(\alpha, \alpha', \epsilon, d)$, so that, for all f such that $|f|(1 + |x|)^{-d-\epsilon} \in L^1$, and all $t \in \mathbb{R}^d$,*

$$\tilde{G}_{(\alpha,\epsilon)}(f)(t) \leq C(\alpha, \alpha', \epsilon, d) G_{\alpha'}(f)(t).$$

The proof of Theorem 2 makes use of a simple observation ("free lunch lemma") which seems to be new.

In section 1 we prove Theorem 1 and the pointwise comparabilities among $G_\alpha(f)$, $G_{\alpha,\beta}(f)$, $g_\alpha(f)$, and $\sigma_\alpha(f)$. In section 2 we prove Theorem 2. In section 3 we discuss some of the implications of these theorems for singular integrals and pointwise convergence of continuous approximations.

This research was carried out while I was spending a delightful sabbatical year at the Universidad de Sevilla. While there my family and I received countless kindnesses and acts of hospitality from friends both inside and outside the Universidad. I am afraid to try to list all of our benefactors, for fear of omitting any one of them. But I must give special thanks to my friend and colleague, Carlos Pérez Moreno, without whose tireless work, both before and after our arrival in Spain, my visit would not have been possible.

1. The basic inequalities

The proof of Theorem 1 has three steps, of which the first is fundamental.
Step One. There is a constant $C(\alpha, d)$ such that, for all $f \in L^1_{loc}(\mathbb{R}^d)$,

$$\|G_\alpha(f)\|_2 \le C(\alpha, d)\|f\|_2.$$

From this the method of [1] will immediately yield:

Step Two. There is a constant $C(\alpha, d)$ such that, for all $f \in L^1_{loc}(\mathbb{R}^d)$ and all non-negative $v \in L^1_{loc}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} (G_\alpha(f))^2 \, v \, dx \le C(\alpha, d) \int_{\mathbb{R}^d} |f|^2 \, M(v) \, dx.$$

Then, mimicking the method from [2], we will get:

Step Three. There is a constant $C(\alpha, d)$ such that, for all $\lambda > 0$, all $f \in L^1_{loc}(\mathbb{R}^d)$ and all non-negative $v \in L^1_{loc}(\mathbb{R}^d)$,

$$v\left(\{x: \ G_\alpha(f)(x) > \lambda\}\right) \le \frac{C(\alpha, d)}{\lambda} \int_{\mathbb{R}^d} |f| \, M(v) \, dx.$$

Theorem 1 will then follow by interpolation.

Step One.   It is an easy exercise to show that $A_\alpha(f)(t, y)$ is a measurable function of (t, y). Therefore the questions we are asking about $A_\alpha(f)(t, y)$ all make sense. An application of Fubini-Tonelli shows that L^2 boundedness of $G_\alpha(f)$ is equivalent to having

$$\int_{\mathbb{R}^{d+1}_+} (A_\alpha(f)(t, y))^2 \frac{dt \, dy}{y} \le C(\alpha, d) \int_{\mathbb{R}^d} |f|^2 \, dx$$

for all $f \in L^2$, which is the same as showing

$$\int_K (A_\alpha(f)(t,y))^2 \frac{dt\,dy}{y} \leq C(\alpha,d) \int_{\mathbb{R}^d} |f|^2 \, dx$$

for all compact $K \subset \mathbb{R}^{d+1}_+$, and this is what we will show.

Let $K \subset \mathbb{R}^{d+1}_+$ be compact and suppose that $\|f\|_2 \leq 1$. If $\phi \in \mathcal{C}_\alpha$ then $f * \phi_y(t)$ is continuous on K. The function $A_\alpha(f)(t,y)$ is also continuous on K. Therefore we can choose, in a measurable (indeed, piecewise constant) fashion, functions $\phi^{(t,y)} \in \mathcal{C}_\alpha$ such that, for every $(t,y) \in K$,

$$|f * \phi_y^{(t,y)}(t)| \geq (1/2) A_\alpha(f)(t,y).$$

Let $g : \mathbb{R}^{d+1}_+ \mapsto \mathbb{R}$ be a measurable function such that

$$\int_{\mathbb{R}^{d+1}_+} |g(t,y)|^2 \frac{dt\,dy}{y} = 1.$$

We will be done if we can show

$$\left| \int_K (f * \phi_y^{(t,y)}(t)) \, g(t,y) \frac{dt\,dy}{y} \right| \leq C(\alpha,d). \tag{1.1}$$

The integral in (1.1) is equal to

$$\int_{\mathbb{R}^d} f(x) \left(\int_K g(t,y) \, \phi_y^{(t,y)}(t-x) \frac{dt\,dy}{y} \right) dx.$$

Set

$$G(x) \equiv \int_K g(t,y) \, \phi_y^{(t,y)}(t-x) \frac{dt\,dy}{y}.$$

Since the integration is over a compact set, this is a bounded, continuous function of x, with bounded support. What we need to show is that

$$\|G\|_2 \leq C(\alpha,d).$$

We will now do this.

For every dyadic cube $Q \subset \mathbb{R}^d$, let $T(Q) = Q \times [\ell(Q)/2, \ell(Q))$, where $\ell(Q)$ is Q's sidelength. (This is the standard "top half" of the Carleson box over Q.) These sets tile $\mathbb{R}^{d+1}_+$, and only finitely many of them intersect K; therefore we can rewrite the integral defining G as

$$\sum_Q \int_{T(Q) \cap K} g(t,y) \, \phi_y^{(t,y)}(t-x) \frac{dt\,dy}{y},$$

and this is in fact a finite sum.

We can rewrite each summand as

$$\int_{T(Q)\cap K} g(t,y)\,\phi_y^{(t,y)}(t-x)\,\frac{dt\,dy}{y} = \lambda_Q a_{(Q)},$$

where

$$\lambda_Q = \left(\int_{T(Q)} |g(t,y)|^2\,\frac{dt\,dy}{y}\right)^{1/2}$$

and $a_{(Q)}$ is a continuous function having certain nice properties. In particular:

1. The support of $a_{(Q)}$ is contained inside $\tilde{Q}$, the triple of Q.

2. $\int a_{(Q)}\,dx = 0$.

3. There is a constant $C(\alpha,d)$ such that, for all x and x',

$$|a_{(Q)}(x) - a_{(Q)}(x')| \le C(\alpha,d)(|x-x'|/\ell(Q))^\alpha |Q|^{-1/2}.$$

Each of these properties follows from a corresponding defining property for $\mathcal{C}_\alpha$. The only moderately tricky one is number 3, for which the crucial inequality to show is:

$$\left(\int_{T(Q)}\left|\phi_y^{(t,y)}(t-x) - \phi_y^{(t,y)}(t-x')\right|^2\,\frac{dt\,dy}{y}\right)^{1/2} \le C(\alpha,d)(|x-x'|/\ell(Q))^\alpha |Q|^{-1/2},$$

which is easy, although the reader should notice how important it is that the smoothness property of functions in $\mathcal{C}_\alpha$ holds pointwise, and not merely in an averaged sense.

It is a theorem (see [5, p. 11]) that a collection of functions $\{a_{(Q)}\}_Q$, indexed over dyadic cubes Q, and satisfying 1, 2, and 3, is *almost orthogonal*: there is a $C(\alpha,d)$ such that, for all finite linear combinations $\sum_Q \gamma_Q a_{(Q)}$,

$$\int_{\mathbb{R}^d}\left|\sum_Q \gamma_Q a_{(Q)}\right|^2\,dx \le C(\alpha,d)\sum_Q |\gamma_Q|^2.$$

But that implies that $\|G\|_2 \le C(\alpha,d)$, and so finishes the proof of Step One.

Step Two. Here and in the sequel, we will be using this definition of the Hardy-Littlewood maximal function:

$$M(h)(x) \equiv \sup_{B:x\in B} \frac{1}{|B|}\int_B |h(t)|\,dt,$$

where the supremum is over all balls that contain x.

Let v be an arbitrary non-negative function in $L^1_{loc}(\mathbb{R}^d)$. When E is a measurable subset of $\mathbb{R}^d$ we will denote $\int_E v\,dx$ by $v(E)$. For $(t,y) \in \mathbb{R}^{d+1}_+$, let $B(t;y)$ be the Euclidean ball in $\mathbb{R}^d$ centered at t with radius y. Fubini-Tonelli implies that

$$\int_{\mathbb{R}^d} (G_\alpha(f))^2\, v\,dx = C(d) \int_{\mathbb{R}^{d+1}_+} (A_\alpha(f)(t,y))^2\, \frac{v(B(t;y))}{|B(t;y)|}\, \frac{dt\,dy}{y}.$$

For each integer k, define

$$F^k = \{(t,y) \in \mathbb{R}^{d+1}_+ :\ 2^k < \frac{v(B(t;y))}{|B(t;y)|} \le 2^{k+1}\}.$$

These sets are disjoint and their union is $\{(t,y) \in \mathbb{R}^{d+1}_+ :\ v(B(t;y)) > 0\}$. Therefore

$$\int_{\mathbb{R}^d} (G_\alpha(f))^2\, v\,dx = C(d) \sum_k \int_{F^k} (A_\alpha(f)(t,y))^2\, \frac{v(B(t;y))}{|B(t;y)|}\, \frac{dt\,dy}{y}$$

$$\le C(d) \sum_k 2^k \int_{F^k} (A_\alpha(f)(t,y))^2\, \frac{dt\,dy}{y}.$$

We now notice two things:

a) if $(t,y) \in F^k$ then $B(t;y) \subset \{x \in \mathbb{R}^d :\ M(v) > 2^k\} \equiv E_k$;

b) if $(t,y) \in F^k$ then $A_\alpha(f)(t,y) = A_\alpha(f\chi_{E_k})(t,y)$.

The first fact is trivial. The second follows because, if $\phi \in \mathcal{C}_\alpha$, the function $\phi_y(t - \cdot)$ has support contained in $B(t;y)$ (this is where compact support plays a vital role). These imply that, for each k,

$$\int_{F^k} (A_\alpha(f)(t,y))^2\, \frac{dt\,dy}{y} = \int_{F^k} (A_\alpha(f\chi_{E_k})(t,y))^2\, \frac{dt\,dy}{y}$$

$$\le C(\alpha,d) \int |f\chi_{E_k}|^2\,dx$$

$$= C(\alpha,d) \int |f|^2\, \chi_{E_k}\,dx,$$

where the second line follows from Step One. Summing on k, we now get

$$\int_{\mathbb{R}^d} (G_\alpha(f))^2\, v\,dx \le C(\alpha,d) \sum_k 2^k \int |f|^2\, \chi_{E_k}\,dx \le C(\alpha,d) \int_{\mathbb{R}^d} |f|^2\, M(v)\,dx,$$

which is Step Two.

Step Three. Following the argument from [2], we will show that there is a constant $C(\alpha, d)$ such that, for all $\lambda > 0$, and all functions f and v,

$$v\left(\{x : \ G_\alpha(f)(x) > \lambda\}\right) \leq \frac{C(\alpha, d)}{\lambda} \int_{\mathbb{R}^d} |f| \, M(v) \, dx,$$

from which Theorem 1 will follow.

Let $\Omega_\lambda = \{x : \ M(f)(x) > \lambda\}$, and let $\{Q_k\}_k$ be a decomposition of Ω_λ into Whitney cubes. These are disjoint dyadic cubes whose union is all of Ω_λ and such that, for all k,

$$\ell(Q_k) \leq d(Q_k, \partial\Omega_\lambda) \leq C(d)\ell(Q_k),$$

where $d(\cdot, \cdot)$ is the Euclidean distance between two sets, and $C(d)$ is a dimensional constant; we refer the reader to [6] for their construction.

A result of Fefferman and Stein [3] says that

$$v(\Omega_\lambda) \leq \frac{C(d)}{\lambda} \int_{\mathbb{R}^d} |f| \, M(v) \, dx.$$

Therefore it will be enough for us to show

$$v\left(\{x \notin \Omega_\lambda : \ G_\alpha(f)(x) > \lambda\}\right) \leq \frac{C(\alpha, d)}{\lambda} \int_{\mathbb{R}^d} |f| \, M(v) \, dx.$$

Following a standard procedure, we split f into two pieces: $f = g + b$, where

$$g(x) = \begin{cases} f(x) & \text{if } x \notin \Omega_\lambda; \\ \frac{1}{|Q_k|} \int_{Q_k} f \, dt & \text{if } x \in Q_k. \end{cases}$$

Because the Whitney cubes are "almost maximal," the function g satisfies

$$\|g\|_\infty \leq C(d)\lambda.$$

Set $b_k \equiv b\chi_{Q_k} = (f - f_{Q_k})\chi_{Q_k}$, where we are using f_{Q_k} to denote f's average over Q_k. Each function b_k is supported in Q_k, has integral equal to 0, and also satisfies $\|b_k\|_1 \leq C(d) \int_{Q_k} |f| \, dx$. It is clear that $b = \sum_k b_k$.

We will separately estimate

$$v\left(\{x \notin \Omega_\lambda : \ G_\alpha(g)(x) > \lambda/2\}\right) \tag{1.2}$$

and

$$v\left(\{x \notin \Omega_\lambda : \ G_\alpha(b)(x) > \lambda/2\}\right). \tag{1.3}$$

Estimation of (1.2).   By Step One,

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} (G_\alpha(g))^2 \, v \, dx = \int_{\mathbb{R}^d} (G_\alpha(g))^2 \, v \chi_{\mathbb{R}^d \setminus \Omega_\lambda} \, dx$$

$$\leq C(\alpha, d) \int_{\mathbb{R}^d} |g|^2 \, M(v \chi_{\mathbb{R}^d \setminus \Omega_\lambda}) \, dx$$

$$\leq C(\alpha, d)\lambda \int_{\mathbb{R}^d \setminus \Omega_\lambda} |f| \, M(v) \, dx + C(\alpha, d)\lambda \sum_k \int_{Q_k} |g| \, M(v \chi_{\mathbb{R}^d \setminus \Omega_\lambda}) \, dx$$

$$\leq C(\alpha, d)\lambda \int_{\mathbb{R}^d} |f| \, M(v) \, dx + C(\alpha, d)\lambda \sum_k |f_{Q_k}| \int_{Q_k} M(v \chi_{\mathbb{R}^d \setminus \Omega_\lambda}) \, dx.$$

However, for each Q_k,

$$\sup_{x \in Q_k} M(v \chi_{\mathbb{R}^d \setminus \Omega_\lambda})(x) \leq C(d) \inf_{x \in Q_k} M(v)(x),$$

and therefore

$$|f_{Q_k}| \int_{Q_k} M(v \chi_{\mathbb{R}^d \setminus \Omega_\lambda}) \, dx \leq C(d) \int_{Q_k} |f| \, M(v) \, dx.$$

Hence,

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} (G_\alpha(g))^2 \, v \, dx \leq C(\alpha, d)\lambda \left(\int_{\mathbb{R}^d} |f| \, M(v) \, dx + \sum_k \int_{Q_k} |f| \, M(v) \, dx \right)$$

$$\leq C(\alpha, d)\lambda \int_{\mathbb{R}^d} |f| \, M(v) \, dx.$$

An application of Chebyshev's Inequality (dividing both sides by $(\lambda/2)^2$) yields the desired estimate.

Estimation of (1.3).   We will show that, for each k,

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} G_\alpha(b_k) \, v \, dx \leq C(\alpha, d) \int_{Q_k} |f| \, M(v) \, dx. \tag{1.4}$$

Summing on k will then yield

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} G_\alpha(b) \, v \, dx \leq C(\alpha, d) \sum_k \int_{Q_k} |f| \, M(v) \, dx,$$

from which the weak $(1, 1)$ estimate will follow at once.

Inequality (1.4) is itself a direct consequence of this simple fact: Suppose that $h \in L^1(Q)$, Q a cube, and $\int_Q h\, dx = 0$. Then, for all x such that $d(x, Q) > \ell(Q)$,

$$G_\alpha(h)(x) \leq C(\alpha, d)\|h\|_1 |Q|^{-1}(1 + |x - x_Q|/\ell(Q))^{-d-\alpha}, \qquad (1.5)$$

where we are using x_Q to denote Q's center. This implies (1.4) because it yields

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} G_\alpha(b_k)\, v\, dx \leq$$
$$C(\alpha, d)\left(\int_{Q_k} |f|\, dx\right)\left(|Q_k|^{-1}\int_{\mathbb{R}^d} v(x)\,(1 + |x - x_Q|/\ell(Q))^{-d-\alpha}\, dx\right).$$

But

$$|Q_k|^{-1}\int_{\mathbb{R}^d} v(x)\,(1 + |x - x_Q|/\ell(Q))^{-d-\alpha}\, dx \leq C(\alpha, d)\inf_{x \in Q_k} M(v)(x),$$

and therefore

$$\int_{\mathbb{R}^d \setminus \Omega_\lambda} G_\alpha(b_k)\, v\, dx \leq C(\alpha, d)\int_{Q_k} |f|\, M(v)\, dx.$$

We finish by proving (1.5).

Suppose $d(x, Q) > \ell(Q)$, $(t, y) \in \Gamma(x)$, and $\phi \in \mathcal{C}_\alpha$. Since ϕ is supported inside $\{x : |x| \leq 1\}$ and $|x - t| < y$, the convolution $h * \phi_y(t)$ will be zero unless $y > c|x - x_Q|$, where c is some positive constant that depends on d. We merely note this fact now; it will become important soon.

We can easily estimate the convolution $h * \phi_y(t)$.

$$|h * \phi_y(t)| = \left|\int_Q \phi_y(t - x)\, h(x)\, dx\right| = \left|\int_Q (\phi_y(t - x) - \phi_y(t - x_Q))\, h(x)\, dx\right|$$

$$\leq C(d)(\ell(Q)/y)^\alpha y^{-d}\|h\|_1 = C(d)\ell(Q)^\alpha\|h\|_1 y^{-d-\alpha},$$

implying that $A_\alpha(h)(t, y) \leq C(\alpha, d)\ell(Q)^\alpha\|h\|_1 y^{-d-\alpha}$ for $(t, y) \in \Gamma(x)$. If we fix a $y > c|x - x_Q|$ then

$$\int_{|t-x|<y} (A_\alpha(f)(t, y))^2\, \frac{dt}{y^{d+1}} \leq C(\alpha, d)\ell(Q)^{2\alpha}\|h\|_1^2 y^{-2d-2\alpha-1}.$$

When we integrate this (in y) from $c|x - x_Q|$ to infinity, and take a square root, we obtain:

$$G_\alpha(h)(x) \leq C(\alpha, d)\ell(Q)^\alpha\|h\|_1 |x - x_Q|^{-d-\alpha}.$$

However, so long as $|x - x_Q| > \ell(Q)$,

$$\frac{\ell(Q)^\alpha}{|x - x_Q|^{d+\alpha}} \leq C\frac{\ell(Q)^\alpha}{(\ell(Q) + |x - x_Q|)^{d+\alpha}}$$
$$= C\ell(Q)^{-d}(1 + |x - x_Q|/\ell(Q))^{-d-\alpha} = C|Q|^{-1}(1 + |x - x_Q|/\ell(Q))^{-d-\alpha},$$

which is exactly what we wanted. Theorem 1 is proved.

To understand the relationships among among $G_\alpha(f)$, $G_{\alpha,\beta}(f)$, $g_\alpha(f)$, and $\sigma_\alpha(f)$, it is convenient to rephrase the definition of $A_\alpha(f)(t, y)$. For $(t, y) \in \mathbb{R}_+^{d+1}$, let $\mathcal{C}_\alpha(t, y)$ be the family of functions $\phi : \mathbb{R}^d \mapsto \mathbb{R}$, supported in $B(t; y)$, such that $\int \phi \, dx = 0$ and, for all x and x' in $\mathbb{R}^d$,

$$|\phi(x) - \phi(x')| \leq y^{-d-\alpha}|x - x'|^\alpha.$$

It is trivial that

$$A_\alpha(f)(t, y) = \sup_{\phi \in \mathcal{C}_\alpha(t,y)} \left| \int f(x)\,\phi(x)\,dx \right|.$$

We make two observations about the families $\mathcal{C}_\alpha(t, y)$.

1. If $0 < y_1 < y_2$ and $\phi \in \mathcal{C}_\alpha(t, y_1)$, then $(y_1/y_2)^{d+\alpha}\phi \in \mathcal{C}_\alpha(t, y_2)$.

2. If $|t-x| < y_1$, $y_2 \geq 2y_1$, and $\phi \in \mathcal{C}_\alpha(t, y_1)$, then $(y_1/y_2)^{d+\alpha}\phi \in \mathcal{C}_\alpha(x, y_2)$. In particular, $2^{d+\alpha}\phi \in \mathcal{C}_\alpha(x, 2y_1)$.

These observations have two easy consequences.

1. If $\beta \geq 1$ then $G_{\alpha,\beta}(f)(x) \leq \beta^{(3/2)d+\alpha}G_\alpha(f)(x)$.

Proof: It is clearly enough to prove this when $x = 0$. We begin with the inequality, valid for all $|t| < \beta y$,

$$A_\alpha(f)(t, y) \leq \beta^{d+\alpha}A_\alpha(f)(t, \beta y),$$

which follows from observation 1. Therefore, for each fixed $y > 0$,

$$\int_{|t|<\beta y} (A_\alpha(f)(t, y))^2\,\frac{dt}{y^{d+1}} \leq \beta^{2d+2\alpha}\int_{|t|<\beta y}(A_\alpha(f)(t, \beta y))^2\,\frac{dt}{y^{d+1}}.$$

Integrating in y from 0 to ∞ yields

$$(G_{\alpha,\beta}(f)(0))^2 \leq \beta^{2d+2\alpha}\int_{|t|<\beta y}(A_\alpha(f)(t, \beta y))^2\,\frac{dt\,dy}{y^{d+1}}.$$

But a change of variable $\theta = \beta y$ shows that the right-hand integral equals $\beta^d(G_\alpha(f)(0))^2$, from which our result follows.

2. The functions $G_\alpha(f)$ and $g_\alpha(f)$ are pointwise comparable, with comparability constants only depending on α and d.

Proof: For all $y > 0$,

$$\int_{|t|<y} (A_\alpha(f)(t,y))^2 \, \frac{dt}{y^d} \le C(\alpha, d)(A_\alpha(f)(0, 2y))^2$$

and

$$(A_\alpha(f)(0,y))^2 \le C(\alpha, d) \int_{|t|<y} (A_\alpha(f)(t, 2y))^2 \, \frac{dt}{y^d}$$
$$\le C(\alpha, d) \int_{|t|<2y} (A_\alpha(f)(t, 2y))^2 \, \frac{dt}{y^d}.$$

Integrating these inequalities in $\frac{dy}{y}$ from 0 to ∞ yields $g_\alpha(f)(0) \sim G_\alpha(f)(0)$, with comparability constants only depending on α and d.

Similar arguments yield the pointwise comparability of $g_\alpha(f)$ and $\sigma_\alpha(f)$; we leave these to the interested reader, as we will also leave the verification of the following observation. Corresponding to $\tilde{G}_{(\alpha,\epsilon)}(f)$, we can define zero-aperture and discretized square functions; respectively,

$$\tilde{g}_{(\alpha,\epsilon)}(f)(x) \equiv \left(\int_0^\infty (\tilde{A}_{(\alpha,\epsilon)}(f)(x,y))^2 \, \frac{dy}{y} \right)^{1/2}$$
$$\tilde{\sigma}_{(\alpha,\epsilon)}(f)(x) \equiv \left(\sum_{-\infty}^\infty (\tilde{A}_{(\alpha,\epsilon)}(f)(x, 2^k))^2 \right)^{1/2}.$$

Having worked through what happens when the kernels have bounded supports, the reader should have no trouble checking that $\tilde{G}_{(\alpha,\epsilon)}(f) \sim \tilde{g}_{(\alpha,\epsilon)}(f) \sim \tilde{\sigma}_{(\alpha,\epsilon)}(f)$, with comparability constants that depend on α, ϵ, and d.

2. Unbounded supports

We need a way to express convolutions with functions in $\mathcal{C}_{(\alpha,\epsilon)}$, in a controlled fashion, in terms of convolutions with (appropriately dilated) functions in some $\mathcal{C}_{\alpha'}$. For this it is convenient to introduce a new class of functions.

Definition. Let $0 < \alpha \le 1$ and $\epsilon > 0$. The Uchiyama class $\mathcal{U}_{(\alpha,\epsilon)}$ is the set of functions $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ such that, for all x,

$$|\phi(x)| \le (1 + |x|)^{-d-\epsilon};$$

and, for all x and x',

$$|\phi(x) - \phi(x')| \le |x - x'|^\alpha \left((1 + |x|)^{-d-\epsilon-\alpha} + (1 + |x'|)^{-d-\epsilon-\alpha} \right); \qquad (2.1)$$

and which also satisfy $\int \phi \, dx = 0$.

The only difference between $\mathcal{C}_{(\alpha,\epsilon)}$ and $\mathcal{U}_{(\alpha,\epsilon)}$ is in (2.1), where we are requiring extra decay in ϕ's smoothness.

This is a real difference. But, as we shall see, it is also an illusion.

The usefulness of $\mathcal{U}_{(\alpha,\epsilon)}$ lies in the following lemma, which is a very slight generalization of a decomposition lemma due to Uchiyama [7].

Lemma 3 *Let $0 < \alpha \leq 1$ and $\epsilon > 0$. There is a constant $C(d,\alpha,\epsilon)$ such that, if $\psi \in \mathcal{U}_{(\alpha,\epsilon)}$, there exists a sequence of functions $\{\phi_k\}_0^\infty$ such that*

$$\psi = C(d,\alpha,\epsilon) \sum_0^\infty 2^{-k\epsilon} \phi_k$$

and every ϕ_k belongs to $\mathcal{C}_\alpha(0, 2^k)$.

Lemma 3 has important consequences. It implies that, if $\psi \in \mathcal{U}_{(\alpha,\epsilon)}$ and $|f|(1+|x|)^{-d-\epsilon} \in L^1$, then

$$|f * \psi(0)| \leq C(d,\alpha,\epsilon) \sum_0^\infty 2^{-k\epsilon} A_\alpha(f)(0, 2^k)$$

$$\leq C(d,\alpha,\epsilon) \left(\sum_0^\infty 2^{-k\epsilon} (A_\alpha(f)(0, 2^k))^2 \right)^{1/2}.$$

It follows from the definition of $\mathcal{C}_\alpha(0, y)$ that, if y and r are positive, then $\phi \in \mathcal{C}_\alpha(0, y)$ if and only if $\phi_r \in \mathcal{C}_\alpha(0, ry)$. Therefore, if $\psi \in \mathcal{U}_{(\alpha,\epsilon)}$ and j is any integer,

$$|f * \psi_{2^j}(0)| \leq C(d,\alpha,\epsilon) \left(\sum_0^\infty 2^{-k\epsilon} (A_\alpha(f)(0, 2^{k+j}))^2 \right)^{1/2}.$$

If $\{\psi^{(j)}\}_{-\infty}^\infty$ is any sequence of functions from $\mathcal{U}_{(\alpha,\epsilon)}$, then, assuming the truth of Lemma 3, we will get

$$\sum_{j=-\infty}^\infty |f * \psi_{2^j}^{(j)}(0)|^2 \leq C(d,\alpha,\epsilon) \sum_{j=-\infty}^\infty \sum_{k=0}^\infty 2^{-k\epsilon} (A_\alpha(f)(0, 2^{k+j}))^2$$

$$= C(d,\alpha,\epsilon) \sum_{l=-\infty}^\infty (A_\alpha(f)(0, 2^l))^2 \sum_{j=-\infty}^l 2^{-(l-j)\epsilon}$$

$$= C(d,\alpha,\epsilon) \sum_{l=-\infty}^\infty (A_\alpha(f)(0, 2^l))^2 \sum_{k=0}^\infty 2^{-k\epsilon} = C(d,\alpha,\epsilon) \sum_{l=-\infty}^\infty (A_\alpha(f)(0, 2^l))^2$$

$$= C(d,\alpha,\epsilon)(\sigma_\alpha(f)(0))^2 \leq C(d,\alpha,\epsilon)(G_\alpha(f)(0))^2,$$

which is almost (but not quite) Theorem 2: all that would be missing is a corresponding decomposition theorem for functions that are only good enough to belong to $\mathcal{C}_{(\alpha,\epsilon)}$.

Proof of Lemma 3. Let $h \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ be real, radial, non-negative, have support contained in $\{x : 1/4 \leq |x| \leq 1\}$, and be normalized so that

$$\sum_{-\infty}^{\infty} h(2^{-k}x) \equiv 1$$

for $x \neq 0$. There exists a $j_0 < 0$ such that

$$\sum_{j_0+1}^{\infty} h(2^{-k}x) \equiv 1$$

when $|x| \geq 1$. Re-index this sum as $\sum_1^\infty h(2^{-(j_0+k)}x)$. Define

$$\rho_0(x) \equiv 1 - \sum_{j_0+1}^{\infty} h(2^{-k}x)$$

and set

$$\rho_k(x) \equiv h(2^{-(j_0+k)}x) \qquad \text{for } k \geq 1.$$

Each ρ_k has support contained in $\{x : |x| \leq 2^k\}$ and satisfies the inequalities:

$$\int \rho_k(x)\,dx \geq C2^{kd}, \qquad |\nabla \rho_k(x)| \leq C2^{-k},$$

where the constants C only depend on d, h, and j_0.

Define, for $k \geq 0$,

$$g_k(x) \equiv \frac{\left(\int (\sum_0^k \rho_j)\,\psi\,dt\right)\rho_k(x)}{\int \rho_k(t)\,dt}.$$

It is easy to see that $g_k \to 0$ uniformly. In fact, this convergence is pretty fast. For any k,

$$\int \left(\sum_0^k \rho_j\right)\psi\,dt = -\int \left(\sum_{j>k} \rho_j\right)\psi\,dt, \qquad (2.2)$$

because $\int \psi\,dt = 0$. But, when $k \geq 1$, the integral on the right-hand side of (2.2) has modulus no bigger than

$$\int_{|x|\geq c2^k} |\psi(t)|\,dt \leq C2^{-k\epsilon}.$$

Therefore, $g_k = C_k \rho_k$, where $|C_k| \leq C2^{-k(d+\epsilon)}$.

Notice that, $\int g_0(t)\, dt = \int \psi\, \rho_0\, dt$; and, when $k \geq 1$,

$$\int \left(g_k(t) - g_{k-1}(t)\right) dt = \int \psi(t)\, \rho_k(t)\, dt.$$

We can now finish the proof of Lemma 3 in a few lines. We write:

$$\psi = \sum_0^\infty \psi\rho_k = \sum_0^\infty \psi\rho_k - \left(g_0 + \sum_1^\infty (g_k - g_{k-1})\right)$$

$$= (\psi\rho_0 - g_0) + \sum_1^\infty \left(\psi\rho_k - (g_k - g_{k-1})\right) \equiv C(d, \alpha, \epsilon) \sum_0^\infty 2^{-k\epsilon} \phi_k,$$

and we claim that, if we take $C(d, \alpha, \epsilon)$ large enough, the ϕ_k's we have just (implicitly) defined will do the trick. This amounts to showing that, for each k, the support of ϕ_k is contained in $\{x : |x| \leq 2^k\}$, $\int \phi_k\, dt = 0$, and, for all x and x',

$$|\phi_k(x) - \phi_k(x')| \leq C|x - x'|^\alpha 2^{-k(d+\alpha)},$$

where C is a constant depending only on h, j_0, α, ϵ, and d. The first two requirements are easy; and the third is also not so bad. Without loss of generality we can assume that $k \geq 2$. Because of the support condition, we can assume that $|x - x'| \leq C2^k$; and, because each ρ_k is supported in an annulus, with inner and outer radii comparable to 2^k, we can assume that $|x|$ and $|x'|$ are both comparable to 2^k. For these x's and x''s we have the following estimates:

$$|\psi(x)| + |\psi(x')| \leq C2^{-k(d+\epsilon)}$$
$$|\psi(x) - \psi(x')| \leq C|x - x'|^\alpha 2^{-k(d+\epsilon+\alpha)}$$
$$|\rho_k(x)| + |\rho_k(x')| \leq C$$
$$|\rho_k(x) - \rho_k(x')| \leq C|x - x'|2^{-k} \leq C|x - x'|^\alpha 2^{-k\alpha},$$

where the last inequality uses the facts that $\alpha \leq 1$ and $|x - x'|/2^k \leq C$. Combining these with our estimates on C_k above, we also get:

$$|g_k(x)| + |g_k(x')| \leq C2^{-k(d+\epsilon)}$$
$$|g_k(x) - g_k(x')| \leq C2^{-k(d+\epsilon+\alpha)}|x - x'|^\alpha,$$

which is exactly the smoothness we require. But now, by mimicking the proof of the product rule, it is also easy to show that

$$|\psi(x)\rho_k(x) - \psi(x')\rho_k(x')| \leq C2^{-k(d+\epsilon+\alpha)}|x - x'|^\alpha,$$

which finishes the proof of Lemma 3.

As we saw, immediately before the proof of Lemma 3, Theorem 2 would follow if we had a "Lemma 3" for functions in $\mathcal{C}_{(\alpha,\epsilon)}$; i.e., for functions that do not have that extra bit of decay in their modulus of Hölder continuity. We do have one.

Free Lunch Lemma. *Let $0 < \alpha \leq 1$ and $\epsilon > 0$, and suppose $0 < \alpha' \leq \alpha$ and $\alpha' < \epsilon$. Define $\epsilon' \equiv \epsilon - \alpha'$. Then:*

$$\mathcal{C}_{(\alpha,\epsilon)} \subset \mathcal{U}_{(\alpha',\epsilon')}.$$

Remark. The moral of the Free Lunch Lemma is that, by sacrificing (actually, by not counting) a little of the decay of a $\phi \in \mathcal{C}_{(\alpha,\epsilon)}$, we get a function (the *same* function!) with "improved" decay in its Hölder modulus, though to a different order.

Proof of the Free Lunch Lemma. Let $\phi \in \mathcal{C}_{(\alpha,\epsilon)}$. It is trivial that

$$|\phi(x)| \leq (1 + |x|)^{-d-\epsilon'},$$

because $\epsilon' < \epsilon$. If $|x - x'| \leq 1$ then

$$|\phi(x) - \phi(x')| \leq |x - x'|^{\alpha} \left((1 + |x|)^{-d-\epsilon} + (1 + |x'|)^{-d-\epsilon} \right)$$
$$\leq |x - x'|^{\alpha'} \left((1 + |x|)^{-d-\epsilon'-\alpha'} + (1 + |x'|)^{-d-\epsilon'-\alpha'} \right),$$

because $|x - x'| \leq 1$, $\alpha' \leq \alpha$, and $\epsilon' + \alpha' = \epsilon$.

On the other hand, if $|x - x'| \geq 1$,

$$|\phi(x) - \phi(x')| \leq (1 + |x|)^{-d-\epsilon} + (1 + |x'|)^{-d-\epsilon}$$
$$\leq |x - x'|^{\alpha'} \left((1 + |x|)^{-d-\epsilon'-\alpha'} + (1 + |x'|)^{-d-\epsilon'-\alpha'} \right).$$

The Free Lunch Lemma is proved, and therefore Theorem 2 is too.

3. Consequences

Continuous approximations. For $0 < \alpha \leq 1$ and $\epsilon > 0$, let $\mathcal{I}_{(\alpha,\epsilon)}$ be the set of functions ϕ such that, for all $x \in \mathbb{R}^d$,

$$|\phi(x)| \leq (1 + |x|)^{-d-\epsilon};$$

and, for all x and x',

$$|\phi(x) - \phi(x')| \leq |x - x'|^{\alpha} \left((1 + |x|)^{-d-\epsilon} + (1 + |x'|)^{-d-\epsilon} \right).$$

The family $\mathcal{I}_{(\alpha,\epsilon)}$ is just $\mathcal{C}_{(\alpha,\epsilon)}$ without the cancellation condition. There is a positive K such that $K\mathcal{I}_{(\alpha,\epsilon)}$ contains infinitely many elements with integral equal to 1. Let $K(\alpha,\epsilon)$ be the infimum of all such K's. (Note that $K(\alpha,\epsilon)$ also depends on d.)

If f is such that $|f|(1+|x|)^{-d-\epsilon} \in L^1$ then, for every integer k and every $x \in \mathbb{R}^d$, the following object makes sense:

$$D_{(\alpha,\epsilon)}(f)(x,2^k) \equiv \sup_{\substack{\phi_1,\phi_2 \in 2K(\alpha,\epsilon)\mathcal{I}_{(\alpha,\epsilon)} \\ \int \phi_1 = \int \phi_2 = 1}} |f * (\phi_1)_{2^k}(x) - f * (\phi_2)_{2^k}(x)|.$$

It measures the maximum possible deviation between two (suitably normalized) continuous "mollifications" of f, at scale 2^k, near the point x. Clearly,

$$D_{(\alpha,\epsilon)}(f)(x,2^k) \le C(\alpha,\epsilon,d)\tilde{A}_{(\alpha,\epsilon)}(f)(x,2^k),$$

and therefore, if $1 < p \le 2$,

$$\int_{\mathbb{R}^d} \left(\sum_k (D_{(\alpha,\epsilon)}(f)(x,2^k))^2 \right)^{p/2} v \, dx \le C(\alpha,\epsilon,p,d) \int_{\mathbb{R}^d} |f|^p \, M(v) \, dx$$

for all f with suitable decay, and all non-negative $v \in L^1_{loc}(\mathbb{R}^d)$. What is interesting about this is that, modulo our normalization, the mollifiers ϕ_1 and ϕ_2 can vary arbitrarily with x and k.

A conjecture of Fefferman and Stein.   R. Fefferman and E. M. Stein [4] conjectured that the following weighted-norm inequality held for all non-negative $v \in L^1_{loc}(\mathbb{R}^d)$ and all f such that the left-hand side makes sense:

$$\int_{\mathbb{R}^d} (S(f))^2 \, v \, dx \le C \int_{\mathbb{R}^d} |f|^2 \, M(v) \, dx. \tag{3.1}$$

The difficulty in proving (3.1) –as opposed to (0.3)– lies in the fact that $S(f)$ is defined via convolutions with non-compactly supported kernels. However, it is trivial that the kernels which "generate" the components of ∇u are all positive multiples of functions in $\mathcal{C}_{(1,1)}$, and therefore

$$y|\nabla u(t,y)| \le C\tilde{A}_{(1,1)}(f)(t,y),$$

implying $S(f) \le CG_\alpha(f)$ pointwise, for all $0 < \alpha < 1$. (In fact, since the generating kernels actually lie in $\mathcal{U}_{(1,1)}$, we can slightly improve the pointwise inequality to $S(f) \le CG_1(f)$.) The conjecture now follows from Theorem 1.

Singular integrals.   Let $\Omega : \mathbb{R}^d \setminus \{0\} \mapsto \mathbb{R}$ be homogeneous of order 0, Hölder continuous of order ϵ on $S^{d-1} \equiv \{x : |x| = 1\}$, and satisfy

$$\int_{S^{d-1}} \Omega(x) \, d\sigma(x) = 0,$$

where $d\sigma$ denotes $(d-1)$-dimensional surface measure. We normalize Ω by requiring that $|\Omega(x) - \Omega(x')| \le |x - x'|^\epsilon$ for all x and x' in S^{d-1}, and of course we assume that $\epsilon \le 1$.

It is well-known that the singular integral operator T_Ω given by

$$T_\Omega(f)(x) \equiv p.v. \int f(y) \frac{\Omega(x-y)}{|x-y|^d} \, dy$$

defines a bounded operator on $L^p(\mathbb{R}^d)$ for all $1 < p < \infty$.

We want to see what the preceding results imply about $S_\psi(T_\Omega f)$. This is an integral involving $(T_\Omega f) * \psi_y(t) = f * (T_\Omega(\psi_y))(t) = f * (T_\Omega \psi)_y(t)$; where the second equation is true because T_Ω commutes with dilations. This can be a very messy object to estimate, because $T_\Omega \psi$ typically does not have compact support. No longer, though. It is a theorem (see [5, p. 84]) that $T_\Omega \psi$ is a positive multiple of a function in $\mathcal{C}_{(\epsilon,\epsilon)}$. Therefore

$$|(T_\Omega f) * \psi_y(t)| \leq C \tilde{A}_{(\epsilon,\epsilon)}(f)(t,y),$$

implying

$$S_\psi(T_\Omega f) \leq C G_\alpha(f)$$

pointwise, for all $0 < \alpha < \epsilon$. As the statement and proof of Theorem 1 show, $G_\alpha(f)$ is essentially no worse than another $S_\phi(f)$.

A final remark.   It is natural to ask how "bad" a function in $\mathcal{C}_{(\alpha,\epsilon)}$ can be. The answer is: very bad. Fix a vector $y \in \mathbb{R}^d$ with unit norm and consider the function defined on $\mathbb{R}^d$,

$$\phi(x) \equiv \frac{\sin(\langle x, y \rangle \cdot |x|^M)}{(1 + |x|)^{d+\epsilon}},$$

where M is a large positive number and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. We claim that if $M\alpha < \epsilon$, then ϕ is a positive multiple of a function in $\mathcal{C}_{(\alpha,\epsilon - M\alpha)}$. Cancellation and the decay condition on $|\phi|$ are trivial. All we need to check is the Hölder modulus, of order α, for large x. But this boils down to showing

$$|\sin(t \cdot |x|^M) - \sin(t' \cdot |x'|^M)| \leq C(1 + |x|)^{M\alpha}|t - t'|^\alpha$$

with $t = \langle x, y \rangle$ and $t' = \langle x', y \rangle$; where we can, without loss of generality, assume that both $|x|$ and $|x'|$ are greater than 1 and that $|x - x'| \leq 1$. Because of the sine function's periodicity, there is t'' such that $|t - t''| \leq C/|x|^M$, $|t - t''| \leq |t - t'|$, where $\sin(t''|x'|^M) = \sin(t' \cdot |x'|^M)$. But now the Mean Value Theorem implies

$$|\sin(t \cdot |x|^M) - \sin(t'' \cdot |x'|^M)| \leq C|x|^M|t - t''| \leq C|x|^M|t - t''|^{1-\alpha}|t - t'|^\alpha$$
$$\leq C|x|^M(|x|^{-M})^{1-\alpha}|t - t'|^\alpha \leq C|x|^{M\alpha}|t - t'|^\alpha,$$

which proves the result.

Note. After this paper had been submitted for publication, the author proved that the intrinsic square function defines a bounded operator on $L^p(\mathbb{R}^d, w)$, for $1 < p < \infty$, whenever w is an A_p weight. The proof of this result (which implies a surprising corollary) will appear elsewhere [8].

References

[1] Chang, S.-Y. A., Wilson, J. M. and Wolff, T. H.: Some weighted norm inequalities concerning the Schrödinger operators. *Comment. Math. Helv.* **60** (1985), 217–246.

[2] Chanillo, S. and Wheeden, R. L.: Some weighted norm inequalities for the area integral. *Indiana Univ. Math. J.* **36** (1987), 277–294.

[3] Fefferman, C. and Stein, E. M.: Some maximal inequalities. *Amer. J. Math.* **93** (1971), 107–115.

[4] Fefferman, R.: Private communication.

[5] Frazier, M., Jawerth, B. and Weiss, G.: *Littlewood-Paley theory and the study of function spaces.* CBMS Regional Conference Series in Mathematics,**79**. American Mathematical Society, Providence, RI, 1991.

[6] Stein, E. M.: *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, **30**. Princeton University Press, Princeton, N.J., 1970.

[7] Uchiyama, A.: A constructive proof of the Fefferman-Stein decomposition of $BMO(\mathbb{R}^n)$. *Acta Math.* **148** (1982), 215–241.

[8] Wilson, M.: *Weighted Littlewood-Paley Theory and Exponential-Square Integrability.* Lecture Notes in Mathematics Series **1924**. Springer-Verlag, New York, 2007.

Michael Wilson
Department of Mathematics
University of Vermont
Burlington, Vermont 05405, USA
wilson@cems.uvm.edu