Bounded Variation, Convexity, and Almost-Orthogonality

Michael Wilson

Abstract Let $N \geq 2$ be fixed. Suppose that, for every dyadic cube Q in $\mathbf{R}^d$, we have: N convex regions $\{R_i(Q)\}_1^N$, subsets of Q; and N complex numbers $\{c_i(Q)\}_1^N$ such that $|c_i(Q)| \leq 1$ and $\sum_1^N c_i(Q)|R_i(Q)| = 0$. Define $\tilde{h}_{(Q)}(x) \equiv |Q|^{-1/2}(\sum_1^N c_i(Q)\chi_{R_i(Q)}(x))$. We prove a technical theorem which implies that, for all such collections $\{\tilde{h}_{(Q)}\}_{Q \in \mathcal{D}}$ and all finite linear combinations $\sum \lambda_Q \tilde{h}_{(Q)}(x)$,

$$\left\| \sum \lambda_Q \tilde{h}_{(Q)} \right\|_2 \leq (2 + \sqrt{2}) Nd \left(\sum |\lambda_Q|^2 \right)^{1/2}.$$

We show that, if $\{\tilde{h}1_{(Q)}\}_{Q \in \mathcal{D}}$ and $\{\tilde{h}2_{(Q)}\}_{Q \in \mathcal{D}}$ are two such families, the L^2 bounded linear operator T defined by

$$T(f) \equiv \sum_Q \langle f, \tilde{h}1_{(Q)} \rangle h2_{(Q)}$$

is, in a natural sense, stable with respect to small dilation and translation errors in the kernel functions $\{\tilde{h}1_{(Q)}\}_{Q \in \mathcal{D}}$ and $\{\tilde{h}2_{(Q)}\}_{Q \in \mathcal{D}}$.

AMS Subject Classification (2000): 42B25 (primary); 42C15, 42C40 (secondary)

1 Introduction

A family $\{\psi_\gamma\}_{\gamma \in \Gamma} \subset L^2(\mathbf{R}^d)$ is called *almost-orthogonal* if there is finite R so that, for all finite subsets $\mathcal{F} \subset \Gamma$ and all linear sums $\sum_{\gamma \in \mathcal{F}} \lambda_\gamma \psi_\gamma$,

$$\left\| \sum_{\gamma \in \mathcal{F}} \lambda_\gamma \psi_\gamma \right\|_2 \leq R \left(\sum_{\mathcal{F}} |\lambda_\gamma|^2 \right)^{1/2}. \tag{1}$$

M. Wilson (✉)
Department of Mathematics, University of Vermont, Burlington, VT 05405, USA
e-mail: jmwilson@uvm.edu

If R is the least such constant for which (1) holds, we say the family is almost-orthogonal with constant R.

"Almost-orthogonal" is a mild misnomer: "almost-orthonormal" may be more accurate. We recall that a family $\{\psi_\gamma\}_{\gamma \in \Gamma}$ is orthonormal if, for all γ and γ' in Γ,

$$\langle \psi_\gamma, \psi_{\gamma'} \rangle \equiv \int_{\mathbf{R}^d} \psi_\gamma(x)\, \overline{\psi_{\gamma'}(x)}\, dx = \begin{cases} 1 & \text{if } \gamma = \gamma'; \\ 0 & \text{otherwise.} \end{cases}$$

The family $\{\psi_\gamma\}_{\gamma \in \Gamma}$ is orthonormal if and only if, for all finite sums as in (1), we have equality, with $R = 1$.

A duality argument shows that $\{\psi_\gamma\}_{\gamma \in \Gamma} \subset L^2(\mathbf{R}^d)$ satisfies (1) if and only if, for all $f \in L^2(\mathbf{R}^d)$,

$$\left(\sum_\Gamma |\langle f, \psi_\gamma \rangle|^2 \right)^{1/2} \leq R \|f\|_{L^2}. \tag{2}$$

Combining (1) and (2), we see that, if $\{\psi_\gamma^{(1)}\}_{\gamma \in \Gamma}$ and $\{\psi_\gamma^{(2)}\}_{\gamma \in \Gamma}$ are two almost-orthogonal families in $L^2(\mathbf{R}^d)$, with respective constants R_1 and R_2, then, for all $f \in L^2(\mathbf{R}^d)$,

$$\sum_\Gamma \langle f, \psi_\gamma^{(1)} \rangle \psi_\gamma^{(2)}$$

converges unconditionally[1] to define a linear operator $T : L^2 \to L^2$ with bound $\leq R_1 R_2$. The canonical example of such an operator is the identity, where $\{\psi_\gamma^{(1)}\}_{\gamma \in \Gamma}$ and $\{\psi_\gamma^{(2)}\}_{\gamma \in \Gamma}$ are both the same complete orthonormal family, such as the classical Haar functions [3]. Recall that an interval I is dyadic if $I = [j2^k, (j+1)2^k)$ for some integers j and k. For each such I we set

$$h^{(I)}(x) \equiv \chi_{I_l}(x) - \chi_{I_r}(x),$$

where I_l is I's left half and I_r is I's right half. (We also use this notation for non-dyadic intervals.) The *Haar function* associated to I is $h^{(I)}(x)/|I|^{1/2}$, where, here and henceforth, $|E|$ is a set E's Lebesgue measure (of varying dimension!).

One can define "Haar functions" adapted to dyadic cubes in $\mathbf{R}^d$. A *cube* $Q \subset \mathbf{R}^d$ is a cartesian product of d intervals $I_i(Q)$ of equal length: $Q = \prod_1^d I_i(Q)$. We call their common length Q's *sidelength*, denoted $\ell(Q)$. The cube is *dyadic* if each $I_i(Q)$ is a dyadic interval. The set of all dyadic cubes in $\mathbf{R}^d$ is $\mathcal{D}$. The dimension d will vary but be clear from the context. We get d-dimensional Haar functions for the Qs

---

[1] We state our precise meaning of "unconditional convergence" in Definition 1.

in $\mathcal{D}$ by taking products

$$\mu_1(x_1) \times \mu_2(x_2) \times \cdots \times \mu_d(x_d)$$

where $x = (x_1, x_2, \ldots, x_d) \in \mathbf{R}^d$ and each μ_j is $h^{(I_j(Q))}$ or $\chi_{I_j(Q)}$. We run over all such products *except* the one for which every μ_j equals $\chi_{I_j(Q)}$. This yields, for each $Q \in \mathcal{D}$, an orthogonal set of $2^d - 1$ functions, $\{h_{(i)}^{(Q)}\}_1^{2^d-1}$. Each $h_{(i)}^{(Q)}$ is supported on Q (where it only takes on the values ± 1), has integral equal to 0, and is constant on Q's immediate dyadic subcubes. We normalize the set by dividing each $h_{(i)}^{(Q)}$ by $|Q|^{1/2}$. The resulting "Haar functions",

$$\left\{ \frac{h_{(i)}^{(Q)}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{D}, \, 1 \le i < 2^d} \tag{3}$$

make up a complete orthonormal family for $L^2(\mathbf{R}^d)$, letting us write

$$f = \sum_{Q,i} \frac{\langle f, h_{(i)}^{(Q)} \rangle}{|Q|} h_{(i)}^{(Q)}, \tag{4}$$

for any $f \in L^2(\mathbf{R}^d)$.

Formula (4) is true, but is it *stable*? If we want to use (4) to investigate f, we have to estimate integrals

$$\langle f, h_{(i)}^{(Q)} \rangle = \int_{\mathbf{R}^d} f(x) \, h_{(i)}^{(Q)}(x) \, dx,$$

which are likely to have small errors. We might make translation errors: instead of $f(x)$ we have $f(x + \vec{\tau}_1^i(Q))$, where (we hope) $|\vec{\tau}_1^i(Q)| < \ell(Q)$, the computed inner product is

$$\int_{\mathbf{R}^d} f(x) \, h_{(i)}^{(Q)}(x - \vec{\tau}_1^i(Q)) \, dx \equiv \langle f, h1_{(i)}^{(Q)} \rangle.$$

We can expect similar translation errors—call them $\vec{\tau}_2^i(Q)$—in the other $h_{(i)}^{(Q)}$s occurring in (4), resulting in "perturbed" Haar functions $h2_{(i)}^{(Q)}$. If we try to add up part of (4), we face

$$\sum_{Q,i} \frac{\langle f, h1_{(i)}^{(Q)} \rangle}{|Q|} h2_{(i)}^{(Q)}. \tag{5}$$

If the $\vec{\tau}_k^i(Q)$s have norms $\leq \eta\ell(Q)$, where η is small, then we hope that

$$\left\| f - \sum_{Q,i} \frac{\langle f, h1_{(i)}^{(Q)} \rangle}{|Q|} h2_{(i)}^{(Q)} \right\|_2 \leq C(\eta)\|f\|_2 \tag{6}$$

for some function $C(\eta)$ going to 0 as $\eta \to 0$.

But it is not clear that the families $\{hk_{(i)}^{(Q)}/|Q|^{1/2}\}_{Q,i}$ ($k = 1, 2$) are even almost-orthogonal. The problem comes from the Haar functions' jumps. We can fix this by working with a smoother family. Let $0 < \alpha \leq 1$. Suppose that, for each $Q \in \mathcal{D}$, we have a function $\phi^{(Q)} : \mathbf{R}^d \to \mathbf{C}$ such that:

(a) supp $\phi^{(Q)} \subset Q$;
(b) $|\phi^{(Q)}(x) - \phi_{(Q)}(x')| \leq (|x - x'|/\ell(Q))^\alpha$ for all x and x';
(c) $\int \phi^{(Q)}\, dx = 0$.

It is well known that $\{\phi^{(Q)}/|Q|^{1/2}\}_{Q\in\mathcal{D}}$ is almost-orthogonal in $L^2(\mathbf{R}^d)$ [3, 4]. If $\{\phi_{(1)}^{(Q)}/|Q|^{1/2}\}_{Q\in\mathcal{D}}$ and $\{\phi_{(2)}^{(Q)}/|Q|^{1/2}\}_{Q\in\mathcal{D}}$ are two such families then the unconditionally convergent sum

$$\sum_{Q\in\mathcal{D}} \frac{\langle f, \phi_{(1)}^{(Q)} \rangle}{|Q|} \phi_{(2)}^{(Q)}(x), \tag{7}$$

defines bounded linear operator $T : L^2 \to L^2$. This sum is also *stable*. Let $0 < \eta < 1/2$ and let $\{\vec{\tau}_i(Q)\}_{Q\in\mathcal{D}}$ ($i = 1, 2$) be two families of vectors in $\mathbf{R}^d$ such that $|\vec{\tau}_i(Q)| \leq \eta\ell(Q)$. Define $\widetilde{\phi_{(i)}^{(Q)}}(x) = \phi_{(i)}^{(Q)}(x - \vec{\tau}_1(Q))$ ($i = 1, 2$). The families $\{\widetilde{\phi_{(i)}^{(Q)}}/|Q|^{1/2}\}_{Q\in\mathcal{D}}$ are almost-orthogonal, with constants $\leq C(\alpha, d)$ [3, 4], implying that

$$\widetilde{T}(f) \equiv \sum_{Q\in\mathcal{D}} \frac{\langle f, \widetilde{\phi_{(1)}^{(Q)}} \rangle}{|Q|} \widetilde{\phi_{(2)}^{(Q)}}$$

defines a bounded linear operator on L^2. Moreover, for every $0 < r < \alpha$, there is a constant $C = C(\alpha, r, d)$ so that, for all $f \in L^2(\mathbf{R}^d)$ [4],

$$\left\| T(f) - \widetilde{T}(f) \right\|_2 \leq C\eta^r \|f\|_2; \tag{8}$$

and analogous results hold in $L^p(\mathbf{R}^d)$ if $1 < p < \infty$ [4]. The $\phi_{(i)}^{(Q)}$s' smoothness is crucial here. But with the $hk_{(i)}^{(Q)}$s, "α is 0", and the Hölder smooth $\phi^{(Q)}$s seem better for working with wavelet representations of operators. This superiority is somewhat specious. In the real world, (7) is discretized: the $\phi^{(Q)}$s are replaced by discontinuous, piecewise constant functions. Sums like (4) provide a model to understand their sensitivity to errors.

It turns out that the perturbed Haar systems *are* almost-orthogonal in $L^2(\mathbf{R}^d)$ (Theorems 1 and 2) and series like (4) are stable: they satisfy (6) with $C(\eta)$ equal to a dimensional constant times $\eta^{1/2}$ (Theorem 3). The almost-orthogonality and stability results hold for much more general systems, perturbations, and operators than those discussed above, and the exponent on η is sharp.

Our proofs of these facts start from a familiar concept. A function $f : [a, b] \to \mathbf{C}$ is said to be *of bounded variation on* $[a, b]$ (written $f \in BV[a, b]$) [1] if there is a finite M so that, for all partitions $P = \{a = x_0 < \cdots < x_n = b\}$ of $[a, b]$,

$$\sum_1^n |f(x_k) - f(x_{k-1})| \le M.$$

The supremum over all such sums is called f's total variation over $[a, b]$ and is denoted $V_f[a, b]$. (When we write $V_f(I)$ and $I = [a, b]$, we mean $V_f[a, b]$.) If $f \in BV[a, b]$ then $f \in BV[c, d]$ for every $[c, d] \subset [a, b]$, and, for every partition P as above,

$$\sum_1^n V_f[x_{k-1}, x_k] = V_f(\cup_1^n [x_{k-1}, x_k]) = V_f[a, b].$$

We say that a function is of bounded variation on $\mathbf{R}$ if the supremum of the preceding expression, over all closed bounded intervals, is finite; and we call that supremum the function's total variation on $\mathbf{R}$.

For every cube $Q \subset \mathbf{R}^d$, let $NBV(Q)$ be the set of $f : \mathbf{R}^d \to \mathbf{C}$ such that: (a) f is measurable; (b) f's support is a subset of $\overline{Q}$ (the closure of Q); (c) for each $1 \le i \le d$, f is of bounded variation with respect to x_i on $\mathbf{R}$, with total variation on $\mathbf{R}$ being ≤ 1; (d) $\int f \, dx = 0$.

Condition (c) means: If we fix the $x_2, x_3, \ldots, x_d$ components of $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$, then the function $f(\cdot, x_2, x_3, \ldots, x_d)$ has total variation ≤ 1 on $\mathbf{R}$, with the analogous statements for x_2, x_3, etc.

Our fundamental result is:

Theorem 1 *If $f^{(Q)} \in NBV(Q)$ for every $Q \in \mathcal{D}$ then*

$$\left\{ \frac{f^{(Q)}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{D}}$$

is almost-orthogonal in $L^2(\mathbf{R}^d)$, with constant $\le \left(1 + \frac{1}{\sqrt{2}}\right) d.$

Theorem 1 immediately implies the fact stated in the abstract:

Corollary 1 *Let $N \ge 2$. Suppose that, for every dyadic cube $Q \subset \mathbf{R}^d$, we have N convex regions $\{R_i(Q)\}_1^N$, subsets of Q, and N complex numbers $\{c_i(Q)\}_1^N$ such that*

$|c_i(Q)| \le 1$ and $\sum_1^N c_i(Q)|R_i(Q)| = 0$. Define, for every $Q \in \mathcal{D}$,

$$\tilde{h}_{(Q)}(x) \equiv |Q|^{-1/2}\left(\sum_1^N c_i(Q)\chi_{R_i(Q)}(x)\right).$$

Then, for every finite linear combination $\sum_{Q \in \mathcal{D}} \lambda_Q \tilde{h}_{(Q)}$,

$$\left\|\sum_{Q \in \mathcal{D}} \lambda_Q \tilde{h}_{(Q)}\right\|_2 \le \left(2 + \sqrt{2}\right) Nd \left(\sum_{Q \in \mathcal{D}} |\lambda_Q|^2\right)^{1/2}.$$

Proof Each function $\sum_1^N c_i(Q)\chi_{R_i(Q)}$ equals $2N$ times some $f^{(Q)} \in NBV(Q)$.

Corollary 1 holds no matter what the convex bodies are (cones, spheres, parallelepipeds, cylinders, etc.) or how they are placed (overlapping, disjoint, etc.). Careful placement gives a better constant.

Corollary 2 *Suppose that, for every dyadic cube $Q \subset \mathbf{R}^d$, we have 2^d convex regions $\{R_i(Q)\}_1^{2^d}$, where each $R_i(Q)$ is a subset of a **unique** immediate dyadic subcube of Q, and that we have complex numbers $\{c_i(Q)\}_1^{2^d}$ such that $|c_i(Q)| \le 1$ and $\sum_1^{2^d} c_i(Q)|R_i(Q)| = 0$. Define, for every $Q \in \mathcal{D}$,*

$$\tilde{h}_{(Q)}(x) \equiv |Q|^{-1/2}\left(\sum_1^{2^d} c_i(Q)\chi_{R_i(Q)}(x)\right).$$

For every finite linear combination $\sum_{Q \in \mathcal{D}} \lambda_Q \tilde{h}_{(Q)}$,

$$\left\|\sum_{Q \in \mathcal{D}} \lambda_Q \tilde{h}_{(Q)}\right\|_2 \le \left(4 + 2\sqrt{2}\right) d \left(\sum_{Q \in \mathcal{D}} |\lambda_Q|^2\right)^{1/2}.$$

Again, it's simple: because of how we placed the $R_i(Q)$s, each function $\sum_1^{2^d} c_i(Q)\chi_{R_i(Q)}$ equals 4 times some $f^{(Q)} \in NBV(Q)$.

After proving Theorem 1 we look at the stability of almost-orthogonal expansions of the form

$$T(g) \equiv \sum_{Q \in \mathcal{D}} \frac{\langle g, f_1^{(Q)}\rangle}{|Q|} f_2^{(Q)}, \tag{9}$$

where each $f_i^{(Q)} \in NBV(Q)$. Corollary 3 shows that, for any $g \in L^2$, the series in (9) converges unconditionally to define T as a bounded linear operator on L^2. In Theorem 3 we show that the operator defined by (9) is L^2-stable with respect to

small dilation and translation errors in the functions $f_i^{(Q)}$. We now say precisely what those small errors are.

Given a family of functions $\{f^{(Q)}\}_{Q \in \mathcal{D}}$, where each $f^{(Q)} \in NBV(Q)$, we suppose we have two sequences of vectors $\{\vec{\delta}(Q)\}_{Q \in \mathcal{D}}$ and $\{\vec{\tau}(Q)\}_{Q \in \mathcal{D}}$ in $\mathbf{R}^d$. The vectors $\vec{\tau}(Q)$ are assumed to be small and the vectors $\vec{\delta}(Q)$ are assumed to be close to $\vec{1} \equiv (1, 1, 1, \ldots, 1)$. Precisely, for some $0 < \eta < 1/2$, $|\vec{1} - \vec{\delta}(Q)| + |\vec{\tau}(Q)| < \eta$ for all $Q \in \mathcal{D}$. If $\vec{\delta}(Q) = (\delta_1, \delta_2, \ldots, \delta_d)$ and $x = (x_1, x_2, \ldots, x_d) \in \mathbf{R}^d$ we shall set $\vec{\delta}(Q)x \equiv (\delta_1 x_1, \delta_2 x_2, \delta_3 x_3, \ldots, \delta_d x_d)$.

We define the *perturbed* form of $f^{(Q)}$ by

$$\widetilde{f^{(Q)}}(x) \equiv f^{(Q)}\left(\vec{\delta}(Q)(x - x_Q + \ell(Q)\vec{\tau}(Q)) + x_Q\right). \tag{10}$$

The effect of replacing x with $\vec{\delta}(Q)(x - x_Q + \ell(Q)\vec{\tau}(Q)) + x_Q$ is to shift $f^{(Q)}$'s "center" a bit and dilate it slightly "relative to x_Q". For example, if

$$g(x) = \chi_{B(x_Q; \ell(Q))}(x),$$

the characteristic function of a ball roughly comparable to Q, and $\vec{\delta}(Q) = (\delta, \delta, \delta, \ldots, \delta)$, then

$$g(\vec{\delta}(Q)(x - x_Q + \ell(Q)\vec{\tau}(Q)) + x_Q) = \chi_{B(x_Q - \ell(Q)\vec{\tau}(Q); \ell(Q)/\delta)}(x):$$

the center shifts by a small multiple of $\ell(Q)$ and the radius gets multiplied by δ^{-1}. (For a general $\delta(Q)$, the ball becomes an ellipsoid.) For an operator T like (9) built from two families $\{f_j^{(Q)}\}_{Q \in \mathcal{D}}$ ($j = 1, 2$), we assume we have sequences of vectors $\{\vec{\delta}_j(Q)\}_{Q \in \mathcal{D}}$ and vectors $\{\vec{\tau}_j(Q)\}_{Q \in \mathcal{D}}$ such that $|\vec{1} - \vec{\delta}_j(Q)| + |\vec{\tau}_j(Q)| < \eta$, from which we define the analogous $\widetilde{f_j^{(Q)}}$s as given by formula (10). We define a *perturbation* of T in the obvious way:

$$\widetilde{T}(g) \equiv \sum_{Q \in \mathcal{D}} \frac{\langle g, \widetilde{f_1^{(Q)}} \rangle}{|Q|} \widetilde{f_2^{(Q)}}. \tag{11}$$

In Sect. 3 we prove:

Theorem 2 *The operator defined by (11) is L^2 bounded, with norm $\leq C(d)$.*

Theorem 3 *There is a constant $C = C(d)$, independent of η, so that, for all operators T and $\widetilde{T}$ (as defined by (9) and (11), respectively), and all $g \in L^2(\mathbf{R}^d)$,*

$$\|T(g) - \widetilde{T}(g)\|_2 \leq C(d)\eta^{1/2}\|g\|_2.$$

The exponent $1/2$ is the best possible. Let $\{f_j^{(Q)}/|Q|^{1/2}\}_{Q \in \mathcal{D}}$ ($j = 1, 2$) be the Haar functions on $\mathbf{R}$ and let $g = h_{[0,1)}$. Leave the $f_1^{(Q)}$s alone but shift $h_{[0,1)}$ in the

$f_2^{(Q)}$ system to the right by $0 < \eta < 1/10$. Then $T(g)(x) = h_{[0,1)}(x)$, $\widetilde{T}(g)(x) = h_{[0,1)}(x - \eta)$, and $\|T(g) - \widetilde{T}(g)\|_2 \sim \eta^{1/2}$.

At two places the reader may wonder why we are doing things certain ways when others seem simpler. Remarks there (labeled "Point 1" and "Point 2") direct the reader to an appendix (Sect. 3) for motivations. Originally we tried to put these in the introduction, but attempts to motivate the motivations (before stating the proofs) made the paper too long and confusing. We removed them, thinking nobody would care about them anyway, but the referee asked about precisely those issues. We then had the idea of addressing them in an appendix. We are grateful to the referee for getting us to explain ourselves, and helping to make the paper not too long and just confusing enough. The "points" remarks occur, respectively, after the proofs of Lemma 1 and Theorem 1.

We write $A \sim B$—where A and B are positive quantities depending on some parameters—to mean that there are positive numbers c_1 and c_2 ("comparability constants") so that

$$c_1 A \le B \le c_2 A; \tag{12}$$

and, if c_1 and c_2 do happen to depend on parameters, (12) does not become trivial. We will often use 'C' to denote a constant which might change to occurrence. We will not always state the parameters C depends on. If E and F are sets, we write $E \subset F$ to express $E \subseteq F$.

We indicate the end of the proof with the symbol ♠.

2   The Proof of Theorem 1

We begin with two lemmas.

Lemma 1 *Let I be a closed, bounded interval. Suppose that $f : I \to \mathbf{C}$ is of bounded variation, with $V_f(I) \le 1$, $b : I \to \mathbf{R}$ is integrable, and $\int b\, dx = 0$. Then:*

$$\left| \int_I f\, b\, dx \right| \le (1/2)\|b\|_1. \tag{13}$$

Proof of Lemma 1 Take $\|b\|_1 = 1$. Assume first that f is real. If b^+ and b^- are b's positive and negative parts then $\int b^+ dx = \int b^- dx = 1/2$, implying

$$\int f(x)\, b^+(x)\, dx = (1/2)s_1$$

and

$$\int f(x)\, b^-(x)\, dx = (1/2)s_2,$$

where s_1 and s_2 are two numbers lying in $[\inf_I f, \sup_I f]$. Therefore

$$\left| \int f(x)\,(b^+(x) - b^-(x))\,dx \right| = (1/2)|s_1 - s_2| \leq (1/2) \sup_{x,y \in I} |f(x) - f(y)| \leq 1/2.$$

If f is not real, let α be a complex number with modulus equal to 1 such that

$$\left| \int_I f\,b\,dx \right| = \int (\alpha f(x))\,b(x)\,dx = \int (\Re(\alpha f(x)))\,b(x)\,dx,$$

and apply the same argument to $\Re(\alpha f)$. ♠

Point 1. Using bounded variation seems like overkill. For f defined on I we can set

$$\Omega_f(I) \equiv \sup\{|f(x) - f(y)| : x, y \in I\}.$$

If $\Omega_f(I) \leq 1$ we'll get

$$\left| \int_I f\,b\,dx \right| \leq (1/2)\|b\|_1.$$

Why use $V_f(I)$? See the appendix.

The second lemma lets us prove Theorem 1 by induction on d.

Lemma 2 *Suppose that $d \geq 2$, $Q \subset \mathbf{R}^d$ is a cube, and $f : \mathbf{R}^d \to \mathbf{C}$ lies in NBV(Q). Write $Q \equiv I_1(Q) \times K(Q)$, where $K(Q) = \prod_2^d I_i(Q)$. For $y \in \mathbf{R}^{d-1}$ define*

$$\phi(y) \equiv \ell(Q)^{-1} \int_{I_1(Q)} f(t, y)\,dt.$$

Then $\phi \in NBV(K(Q))$.

Proof of Lemma 2 It is trivial that supp $\phi \subset \overline{K(Q)}$ and $\int \phi\,dy = 0$. For $2 \leq j \leq d$, let $\{y_k\}_0^n$ be points in $\mathbf{R}^{d-1}$ differing only in their x_j coordinates, where these increase with k. Then:

$$\sum_1^n |\phi(y_k) - \phi(y_{k-1})| \leq \ell(Q)^{-1} \int_{I_1(Q)} \left(\sum_1^n |f(t, y_k) - f(t, y_{k-1})| \right) dt \leq 1,$$

because $f \in NBV(Q)$. ♠

We now prove Theorem 1.

Let $d = 1$. Take Q and I, dyadic intervals. Consider the inner product

$$\left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle, \tag{14}$$

where $f^{(Q)} \in NBV(Q)$ and $h^{(I)}/|I|^{1/2}$ is the classical Haar function associated to I. If $Q \cap I = \emptyset$ or Q is properly contained in I then (14) is 0. If $I \subset Q$ then, by Lemma 1,

$$\left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right| \le (1/2) V_{f^{(Q)}}(\bar{I}) \left(\frac{|I|}{|Q|} \right)^{1/2}.$$

Therefore, for each $j \ge 0$,

$$\sum_{\substack{I \subset Q \\ \ell(I) = 2^{-j}\ell(Q)}} \left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right| \le (1/2) 2^{-j/2} \sum_{\substack{I \subset Q \\ \ell(I) = 2^{-j}\ell(Q)}} V_{f^{(Q)}}(\bar{I})$$

$$= (1/2) 2^{-j/2} V_{f^{(Q)}}(\bar{Q}) \tag{15}$$

$$\le (1/2) 2^{-j/2}.$$

For each $Q \in \mathcal{D}$,

$$\sum_{I \in \mathcal{D}} \left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right| = \sum_{\substack{I \in \mathcal{D} \\ I \subset Q}} \left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right|$$

$$\le (1/2) \sum_{0}^{\infty} 2^{-j/2}$$

$$= 1 + \frac{1}{\sqrt{2}}.$$

For every $I \in \mathcal{D}$,

$$\sum_{Q \in \mathcal{D}} \left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right| = \sum_{\substack{Q \in \mathcal{D} \\ I \subset Q}} \left| \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle \right|$$

$$\le (1/2) \sum_{0}^{\infty} 2^{-j/2}$$

$$= 1 + \frac{1}{\sqrt{2}}.$$

By the Schur Test, the linear mapping $L : \ell^2(\mathcal{D}) \to \ell^2(\mathcal{D})$ defined by

$$L\left(\{\lambda_Q\}_{Q\in\mathcal{D}}\right) \equiv \left\{\sum_{Q\in\mathcal{D}} \lambda_Q \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle\right\}_{I\in\mathcal{D}}$$

has a bound less than or equal to $1 + \frac{1}{\sqrt{2}}$. Let $g = \sum_{Q\in\mathcal{D}} \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}}$ be a finite linear sum. The classical Haar functions form a complete orthonormal set in $L^2(\mathbf{R})$. Therefore,

$$\int |g|^2 \, dx = \sum_I \left|\left\langle g, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle\right|^2$$

$$= \sum_I \left|\sum_{Q\in\mathcal{D}} \lambda_Q \left\langle \frac{f^{(Q)}}{|Q|^{1/2}}, \frac{h^{(I)}}{|I|^{1/2}} \right\rangle\right|^2$$

$$\leq \left(1 + \frac{1}{\sqrt{2}}\right)^2 \sum_{Q\in\mathcal{D}} |\lambda_Q|^2,$$

proving the Theorem 1 when $d = 1$.

Assume the result for $d - 1 \geq 1$, with constant $C(d-1)$; i.e., assume that if $f^{(Q)} \in NBV(Q)$ for every $(d-1)$-dimensional dyadic cube Q, and $\sum_{Q\in\mathcal{D}} \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}}$ is any finite linear combination, then

$$\left(\int_{\mathbf{R}^{d-1}} \left|\sum_{Q\in\mathcal{D}} \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}}\right|^2 dx\right)^{1/2} \leq C(d-1) \left(\sum_{Q\in\mathcal{D}} |\lambda_Q|^2\right)^{1/2}.$$

Consider the family

$$\left\{\frac{f^{(Q)}}{|Q|^{1/2}}\right\}_{Q\in\mathcal{D}},$$

where every Q is a d-dimensional dyadic cube and each $f^{(Q)} \in NBV(Q)$. Put $f^{(Q)}(x) = f^{(Q)}(x', y)$, where $x' \in \mathbf{R}$ and $y \in \mathbf{R}^{d-1}$. Write

$$f^{(Q)}(x', y) = f_1^{(Q)}(x', y) + f_2^{(Q)}(x', y),$$

where

$$f_1^{(Q)}(x', y) = \left(f^{(Q)}(x', y) - \ell(Q)^{-1} \int_{I_1(Q)} f^{(Q)}(t, y) \, dt \right) \chi_{\overline{I_1(Q)}}(x') \chi_{\overline{K(Q)}}(y)$$

$$f_2^{(Q)}(x', y) = \left(\ell(Q)^{-1} \int_{I_1(Q)} f^{(Q)}(t, y) \, dt \right) \chi_{\overline{I_1(Q)}}(x') \chi_{\overline{K(Q)}}(y),$$

and $I_1(Q)$ and $K(Q)$ are as in the statement of Lemma 2, so that $Q = I_1(Q) \times K(Q)$.

By our $d = 1$ result, for each fixed $y \in \mathbf{R}^{d-1}$, the family $\{\ell(Q)^{-1/2} f_1^{(Q)}(x', y)\}_{Q \in \mathcal{D}}$ is almost-orthogonal in $L^2(\mathbf{R})$, with constant $\leq C(1)$. This is because, for each fixed y, the function $\ell(Q)^{-1/2} f_1^{(Q)}(x', y)$ is either identically 0 (with respect to x') or it's a suitably scaled, uniformly bounded-variation function, with integral 0, adapted to a unique dyadic interval $I_1(Q)$. Note that subtracting a term of the form $c \chi_{\overline{I_1(Q)}}(x')$ does not change $f^{(Q)}$'s total variation in x' on $\overline{I_1(Q)}$, and so does not affect the relevant Schur Test estimates. (See the proof of Lemma 1.)

For each fixed $y \in \mathbf{R}^{d-1}$,

$$\int_{\mathbf{R}} \left| \sum \lambda_Q |Q|^{-1/2} f_1^{(Q)}(x', y) \right|^2 \, dx' \leq C(1)^2 \sum |\lambda_Q|^2 \ell(Q)^{-(d-1)} \chi_{\overline{K(Q)}}(y).$$

Since $|\overline{K(Q)}| = \ell(Q)^{d-1}$, integrating in y yields

$$\int_{\mathbf{R}^d} \left| \sum \lambda_Q |Q|^{-1/2} f_1^{(Q)}(x', y) \right|^2 \, dx' \, dy \leq C(1)^2 \sum |\lambda_Q|^2. \qquad (16)$$

By induction (and because of Lemma 2), for each fixed $x' \in \mathbf{R}$, the family $\{\ell(Q)^{-(d-1)/2} f_2^{(Q)}(x', y)\}_{Q \in \mathcal{D}}$ is almost-orthogonal in $L^2(\mathbf{R}^{d-1})$, with constant $\leq C(d-1)$. (As with the $f_1^{(Q)}$s, for some x', $f_2^{(Q)}(x', y)$ is identically 0 in y—which is fine.) Hence, for each fixed $x' \in \mathbf{R}$,

$$\int_{\mathbf{R}^{d-1}} \left| \sum \lambda_Q |Q|^{-1/2} f_2^{(Q)}(x', y) \right|^2 \, dy \leq C(d-1)^2 \sum |\lambda_Q|^2 \ell(Q)^{-1} \chi_{\overline{I_1(Q)}}(x').$$

Now integrating in x' yields:

$$\int_{\mathbf{R}^d} \left| \sum \lambda_Q |Q|^{-1/2} f_2^{(Q)}(x', y) \right|^2 \, dx' \, dy \leq C(d-1)^2 \sum |\lambda_Q|^2. \qquad (17)$$

Combining (16) and (17) yields

$$\left\| \sum \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}} \right\|_2 \leq (C(1) + C(d-1)) \left(\sum |\lambda_Q|^2 \right)^{1/2},$$

which implies

$$C(d) \le d\, C(1).$$

The Schur Test gives $C(1) \le 1 + \frac{1}{\sqrt{2}}$. We have Theorem 1. ♠

Point 2. Why do induction? We have d-dimensional Haar functions (3). Why not get Schur test estimates directly from inner products between them and the functions in

$$\left\{ \frac{f^{(Q)}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{D}} ?$$

See the appendix.

Theorem 1 implies the L^2 boundedness of a certain "rough" operator (see the introduction), defined as a limit of finite sums. We need to specify in what way this limit is taken.

Definition 1 We say that a sequence $\{\mathcal{E}_k\}_1^\infty$ of *finite* subsets of $\mathcal{D}$ *fills up* $\mathcal{D}$ if every $Q \in \mathcal{D}$ is in all but finitely many $\mathcal{E}_k$s. (This holds if the $\mathcal{E}_k$s are increasing and $\cup_k \mathcal{E}_k = \mathcal{D}$.) Let $\{\lambda_Q\}_{Q \in \mathcal{D}}$ be a sequence of complex numbers, and $\{g_{(Q)}\}_{Q \in \mathcal{D}}$ a sequence of functions in $L^2(\mathbf{R}^d)$, each indexed over the family of dyadic cubes $\mathcal{D}$. We say that

$$\sum_{Q \in \mathcal{D}} \lambda_Q g_{(Q)}$$

converges unconditionally to $h \in L^2(\mathbf{R}^d)$ if, for every sequence of finite subsets $\{\mathcal{E}_k\}_1^\infty$ that fills up $\mathcal{D}$,

$$\lim_{k \to \infty} \left\| h - \sum_{Q \in \mathcal{E}_k} \lambda_Q g_{(Q)} \right\|_2 = 0.$$

Corollary 3 *Let $\{f_1^{(Q)}\}_{Q \in \mathcal{D}}$ and $\{f_2^{(Q)}\}_{Q \in \mathcal{D}}$ be two families such that $f_i^{(Q)} \in NBV(Q)$ for all $Q \in \mathcal{D}$ and $i = 1, 2$. If $g \in L^2(\mathbf{R}^d)$ then the series*

$$\sum_{Q \in \mathcal{D}} \frac{\langle g, f_1^{(Q)} \rangle}{|Q|} f_2^{(Q)} \tag{18}$$

converges unconditionally to some h in L^2. Moreover,

$$\|h\|_2 \le \left(\left(1 + \frac{1}{\sqrt{2}} \right) d \right)^2 \|g\|_2.$$

In other words, (18) defines a linear operator $T : L^2 \rightarrow L^2$ with norm $\leq ((1 + \frac{1}{\sqrt{2}})d)^2$.

Proof of Corollary 3 Let $g \in L^2(\mathbf{R}^d)$, and suppose that $\mathcal{E} \subset \mathcal{D}$ is a finite subset. Define

$$T_{\mathcal{E}}(g) \equiv \sum_{\mathcal{E}} \frac{\langle g, f_1^{(Q)} \rangle}{|Q|} f_2^{(Q)}.$$

By Theorem 1,

$$\|T_{\mathcal{E}}(g)\|_2 \leq \left(1 + \frac{1}{\sqrt{2}}\right) d \left(\sum_{\mathcal{E}} \frac{|\langle g, f_1^{(Q)} \rangle|^2}{|Q|}\right)^{1/2} \leq \left(\left(1 + \frac{1}{\sqrt{2}}\right) d\right)^2 \|g\|_2 < \infty. \tag{19}$$

If $\{\mathcal{E}_k\}_1^{\infty}$ is a sequence of finite subsets that fills up $\mathcal{D}$ then, for any m and n,

$$\|T_{\mathcal{E}_m}(g) - T_{\mathcal{E}_n}(g)\|_2 \leq \left(1 + \frac{1}{\sqrt{2}}\right) d \left(\sum_{\mathcal{E}_m \Delta \mathcal{E}_n} \frac{|\langle g, f_1^{(Q)} \rangle|^2}{|Q|}\right)^{1/2};$$

which, because of (19), goes to 0 as m and n go to infinity. (Apply dominated convergence to the sums over the symmetric differences $\mathcal{E}_m \Delta \mathcal{E}_n$.) Therefore $\{T_{\mathcal{E}_k}(g)\}_k$ is Cauchy in $L^2(\mathbf{R}^d)$ and converges to an h with norm $\leq ((1 + \frac{1}{\sqrt{2}})d)^2 \|g\|_2$. The function h is unique because, if $\{\mathcal{E}_k\}_1^{\infty}$ and $\{\mathcal{E}_k'\}_1^{\infty}$ fill up $\mathcal{D}$, so does $\{\mathcal{E}_1, \mathcal{E}_1', \mathcal{E}_2, \mathcal{E}_2', \ldots\}$. ♠

3 The Proofs of Theorems 2 and 3

As with Theorem 1, we will first work in one dimension, where we will sometimes call dyadic intervals I or J, and sometimes Q.

Both proofs make use of a simple fact whose proof can be found in [2] and [3].

Lemma 3 *If $\tilde{\mathcal{D}}$ denotes the family of concentric triples of dyadic intervals in $\mathbf{R}$ then $\tilde{\mathcal{D}}$ can be decomposed into 3 disjoint families,*

$$\tilde{\mathcal{D}} = \cup_1^3 \mathcal{G}_i,$$

such that, for each $1 \leq i \leq 3$: a) $\forall I, J \in \mathcal{G}_i$, either $I \cap J = \emptyset$ or one is a subset of the other; b) every $I \in \mathcal{G}_i$ is the right or left half of a $J \in \mathcal{G}_i$; c) $\forall I \in \mathcal{G}_i$, I's right and left halves belong to $\mathcal{G}_i$; d) $\mathbf{R}$ is covered by the set of $I \in \mathcal{G}_i$ of length 3; and therefore, for any k, $\mathbf{R}$ is covered by the set of $I \in \mathcal{G}_i$ of length $3 \cdot 2^k$.

As an immediate corollary of Lemma 3, the set of concentric triples of dyadic *cubes* in $\mathbf{R}^d$ (also denoted $\tilde{\mathcal{D}}$) can be split into 3^d disjoint families, each one having the analogous inclusion/exclusion and relative size properties as the set of dyadic cubes. The proof is trivial: for every $\vec{a} = (a_1, \ldots, a_d) \in \{1, 2, 3\}^d$, let $\mathcal{G}_{\vec{a}}$ be the set of *cubes* $Q = \prod_1^d I_i(Q)$ such that each $I_i(Q) \in \mathcal{G}_{a_i}$.

The Proof of Theorem 2 If I is a dyadic interval, we use $\tilde{I}$ to denote I's concentric triple, and we define $h^{(\tilde{I})}$ by

$$h^{(\tilde{I})}(x) = \chi_{\tilde{I}_l}(x) - \chi_{\tilde{I}_r}(x).$$

Then $h^{(\tilde{I})}/|\tilde{I}|^{1/2}$ is the "Haar function" associated to $\tilde{I}$. Because of Lemma 3, for each $1 \leq i \leq 3$, $\{h^{(\tilde{I})}/|\tilde{I}|^{1/2} : \tilde{I} \in \mathcal{G}_i\}$ forms a complete orthonormal basis for $L^2(\mathbf{R})$.

For each $1 \leq i \leq 3$ we let $\mathcal{F}_i$ be the set of dyadic intervals Q such that $\tilde{Q} \in \mathcal{G}_i$. We note that if $Q \in \mathcal{F}_i$ and $f^{(Q)} \in NBV(Q)$ then $\widetilde{f^{(Q)}} \in NBV(\tilde{Q})$, where $\tilde{Q} \in \mathcal{G}_i$. We claim that if $1 \leq i \leq 3$ and $\{f^{(Q)}\}_{Q \in \mathcal{F}_i}$ is any family such that each $f^{(Q)} \in NBV(Q)$ then

$$\left\{ \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{F}_i}$$

is almost-orthogonal in $L^2(\mathbf{R})$, with a constant less than or equal to an absolute C. The proof is easy. We only need to bound

$$\left| \left\langle \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}}, \frac{h^{(\tilde{I})}}{|\tilde{I}|^{1/2}} \right\rangle \right| \tag{20}$$

for Q and I both lying in $\mathcal{F}_i$. But we have already seen this sort of thing. If $\tilde{Q} \cap \tilde{I} = \emptyset$ or $\tilde{Q}$ is strictly contained in $\tilde{I}$ then the inner product is 0. Otherwise $\tilde{I} \subset \tilde{Q}$, with $|\tilde{I}| = 2^{-j}|\tilde{Q}|$ for some $j \geq 0$, and (20) is less than or equal to

$$\left(\frac{|\tilde{I}|}{|Q|} \right)^{1/2} V_{\widetilde{f^{(Q)}}}(\tilde{I}) = 3^{1/2} 2^{-j/2} V_{\widetilde{f^{(Q)}}}(\tilde{I}).$$

For every $Q \in \mathcal{F}_i$ and $j \geq 0$,

$$\sum_{\substack{I \in \mathcal{F}_i: \ \tilde{I} \subset \tilde{Q} \\ |\tilde{I}| = 2^{-j}|\tilde{Q}|}} \left| \left\langle \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}}, \frac{h^{(\tilde{I})}}{|\tilde{I}|^{1/2}} \right\rangle \right|$$

is less than or equal to a constant times

$$2^{-j/2} V_{\widetilde{f^{(Q)}}}(\tilde{Q}) \le 2^{-j/2},$$

implying that, for every $Q \in \mathcal{F}_i$,

$$\sum_{I \in \mathcal{F}_i} \left| \left\langle \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}}, \frac{h^{(\tilde{I})}}{|\tilde{I}|^{1/2}} \right\rangle \right| \le C(1 + \frac{1}{\sqrt{2}}) \le C.$$

Similarly, for every $I \in \mathcal{F}_i$,

$$\sum_{Q \in \mathcal{F}_i} \left| \left\langle \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}}, \frac{h^{(\tilde{I})}}{|\tilde{I}|^{1/2}} \right\rangle \right| \le C(1 + \frac{1}{\sqrt{2}}) \le C.$$

Combining the two inequalities proves our claim.

For every $\vec{a} \in \{1, 2, 3\}^d$, let $\mathcal{F}_{\vec{a}}$ be the family of dyadic cubes Q such that $\tilde{Q} \in \mathcal{G}_{\vec{a}}$. Fix an $\vec{a} \in \{1, 2, 3\}^d$. If $Q \in \mathcal{F}_{\vec{a}}$ then $\widetilde{f^{(Q)}} \in NBV(\tilde{Q})$. We can now repeat the inductive argument from the proof of Theorem 1 to get that

$$\left\{ \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{F}_{\vec{a}}}$$

is almost-orthogonal in $L^2(\mathbf{R}^d)$, with constant $\le Cd$, where C is the constant we get for $d = 1$. We get the same estimate for every $\vec{a} \in \{1, 2, 3\}^d$, implying that

$$\left\{ \frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{D}}$$

is almost-orthogonal in $L^2(\mathbf{R}^d)$, with constant $\le C3^d d \equiv C(d)$. A repetition of the argument in the proof of Corollary 3 shows that, for any $g \in L^2(\mathbf{R}^d)$,

$$\sum_{Q \in \mathcal{D}} \frac{\langle g, \widetilde{f_1^{(Q)}} \rangle}{|Q|} \widetilde{f_2^{(Q)}}$$

converges unconditionally in L^2 to define a bounded linear operator $\widetilde{T} : L^2 \to L^2$ with norm $\le C(d)^2$. ♠

The Proof of Theorem 3 Write $T(g) - \widetilde{T}(g)$ as $S_1(g) + S_2(g)$, where

$$S_1(g) = \sum_{Q \in \mathcal{D}} \frac{\langle g, f_1^{(Q)} - \widetilde{f_1^{(Q)}} \rangle}{|Q|} f_2^{(Q)}$$

$$S_2(g) = \sum_{Q \in \mathcal{D}} \frac{\langle g, \widetilde{f_1^{(Q)}} \rangle}{|Q|} \left(f_2^{(Q)} - \widetilde{f_2^{(Q)}} \right).$$

Because of (2) and Theorem 2, Theorem 3 will follow once we show that, for all finite linear sums

$$\sum_{Q \in \mathcal{D}} \lambda_Q \left(\frac{f_i^{(Q)} - \widetilde{f_i^{(Q)}}}{|Q|^{1/2}} \right)$$

($i = 1, 2$), we have

$$\left\| \sum_{Q \in \mathcal{D}} \lambda_Q \left(\frac{f_i^{(Q)} - \widetilde{f_i^{(Q)}}}{|Q|^{1/2}} \right) \right\|_2 \leq C \eta^{1/2} \left(\sum_{Q \in \mathcal{D}} |\lambda_Q|^2 \right)^{1/2} \tag{21}$$

for a constant C only depending on d. Inequality (21) will follow from Theorem 2 and a technical, one-dimensional lemma (Lemma 4). We prove Lemma 4 first. We warn the reader that its proof requires an additional (fortunately very easy) lemma (Lemma 5).

Since the $f_i^{(Q)}$s' subscripts are now irrelevant, we no longer write them.

Until otherwise stated, $\mathcal{D}$, $\tilde{\mathcal{D}}$, $\mathcal{F}_i$, and $\mathcal{G}_i$ refer to families of *intervals*.

Lemma 4 *For each $Q \in \mathcal{F}_i$, let $g^{(Q)} : \mathbf{R} \to \mathbf{R}$ have support contained in $\overline{Q}$ and be of bounded variation, with total variation ≤ 1. (Note: we do* not *require that $\int g^{(Q)} \, dx = 0$.) Let $\{\delta(Q)\}_{Q \in \mathcal{D}}$ and $\{\tau(Q)\}_{Q \in \mathcal{D}}$ be two sequences of real numbers indexed over $\mathcal{D}$, such that $|1 - \delta(Q)| + |\tau(Q)| < \eta < 1/2$ for all $Q \in \mathcal{D}$. Define $\widetilde{g^{(Q)}}(x) \equiv g^{(Q)}(\delta(Q)(x - x_Q + \ell(Q)\tau(Q)) + x_Q)$ and, for each $\tilde{I} \in \mathcal{G}_i$, set*

$$a(Q, \tilde{I}) \equiv \left\langle \frac{g^{(Q)} - \delta(Q)\widetilde{g^{(Q)}}}{|Q|^{1/2}}, \frac{h^{(\tilde{I})}}{|\tilde{I}|^{1/2}} \right\rangle.$$

There is an absolute C such that, for all $Q \in \mathcal{F}_i$ and $\tilde{I} \in \mathcal{G}_i$,

$$\sum_{\tilde{I} \in \mathcal{G}_i} |a(Q, \tilde{I})| \tag{22}$$

and

$$\sum_{Q \in \mathcal{F}_i} |a(Q, \tilde{I})| \tag{23}$$

are both bounded by $C\eta^{1/2}$.

Proof of Lemma 4 If $Q \in \mathcal{F}_i$ and $\tilde{I} \in \mathcal{G}_i$ then

$$\delta(Q) \int \widetilde{g^{(Q)}}(x) \, h^{(\tilde{I})}(x) \, dx$$

$$= \delta(Q) \int g^{(Q)} \left(\delta(Q)(x - x_Q + \ell(Q)\tau(Q)) + x_Q \right) h^{(\tilde{I})}(x) \, dx$$

$$= \int g^{(Q)}(u) \, h^{(\tilde{I})} \left(\delta(Q)^{-1}(u - x_Q - \delta(Q)\ell(Q)\tau(Q)) + x_Q \right) \, du,$$

after substituting $u = \delta(Q)(x - x_Q + \ell(Q)\tau(Q)) + x_Q$. Therefore,

$$\int \left(g^{(Q)}(x) - \delta(Q)\widetilde{g^{(Q)}}(x) \right) h^{(\tilde{I})}(x) \, dx = \int g^{(Q)}(x) \, \gamma^{(\tilde{I})}(x) \, dx, \tag{24}$$

where $\gamma^{(\tilde{I})}(x)$ equals

$$h^{(\tilde{I})}(x) - h^{(\tilde{I})} \left(\delta(Q)^{-1}(x - x_Q - \delta(Q)\ell(Q)\tau(Q)) + x_Q \right). \tag{25}$$

We note a fact which will be important soon. Although we do not assume that $\int g^{(Q)} \, dx = 0$, we do have $\int \left(g^{(Q)}(x) - \delta(Q)\widetilde{g^{(Q)}}(x) \right) dx = 0$, ensuring that (24) equals 0 if $\tilde{I} \not\subset \tilde{Q}$: if $\tilde{I} \not\subset \tilde{Q}$ and $\tilde{I} \cap \tilde{Q} \neq \emptyset$, the support of $g^{(Q)}(x) - \delta(Q)\widetilde{g^{(Q)}}$ is entirely contained in either the right or the left half of $\tilde{I}$, across which $h^{(\tilde{I})}$ is constant.

The key to the proof of Lemma 4 is a good estimate for the right-hand side of (24), which follows from Lemma 1 and a bound on $\|\gamma^{(\tilde{I})}\|_1$. For the latter we need the simple lemma mentioned above.

Lemma 5 *If I is a bounded interval, with endpoints* $a < b$, *and I′ is another bounded interval, with endpoints* $a' < b'$, *then*

$$\int |\chi_I(x) - \chi_{I'}(x)| \, dx \leq |a - a'| + |b - b'|. \tag{26}$$

Proof of Lemma 5 Assume that $b - a \leq b' - a'$. If $b \leq a'$ the left-hand side of (26) is $b - a + b' - a'$, while $|a - a'| \geq b - a$ and $|b - b'| \geq b' - a'$. If $a \leq a' < b$ the left-hand side of (26) is exactly $a' - a + b' - b$ (because $b \leq b'$), and if $a' < a < b \leq b'$ it is $a - a' + b' - b$. The other cases follow from symmetry. ♠

We continue the proof of Lemma 4. Recall that $h^{(\tilde{I})}$ has the form $\chi_{[a,b)} - \chi_{[a',b')}$, where $[a, b) = \tilde{I}_l$ and $[a', b') = \tilde{I}_r$. If $\phi : \mathbf{R} \to \mathbf{R}$ is a strictly increasing bijection

then

$$\chi_{[\alpha,\beta)}(\phi(x)) = \chi_{[\phi^{-1}(\alpha),\phi^{-1}(\beta))}(x).$$

Set $\phi(x) = \delta(Q)^{-1}(x - x_Q - \delta(Q)\ell(Q)\tau(Q)) + x_Q$. Then $\phi^{-1}(x) = \delta(Q)(x - x_Q + \ell(Q)\tau(Q)) + x_Q$. For ease of reading we will refer to ϕ^{-1} as ψ. We can write

$$h^{(\tilde{I})}\left(\delta(Q)^{-1}(x - x_Q - \delta(Q)\ell(Q)\tau(Q)) + x_Q\right)$$
$$= \chi_{[\psi(a),\psi(b))}(x) - \chi_{[\psi(a'),\psi(b'))}(x),$$

and therefore the L^1 norm of

$$\gamma^{(\tilde{I})}(x) \equiv h^{(\tilde{I})}(x) - h^{(\tilde{I})}\left(\delta(Q)^{-1}(x - x_Q - \delta(Q)\ell(Q)\tau(Q)) + x_Q\right)$$

is less than or equal to

$$|a - \psi(a)| + |b - \psi(b)| + |a' - \psi(a')| + |b' - \psi(b')|. \tag{27}$$

A quick calculation yields

$$a - \psi(a) = (a - x_Q)(1 - \delta(Q)) - \delta(Q)\ell(Q)\tau(Q), \tag{28}$$

with similar expressions for the other terms.

We recall that $Q \in \mathcal{F}_i$, $\tilde{I} \in \mathcal{G}_i$, and that the inner product (24) is zero unless $\tilde{I} \subset \tilde{Q}$; thus, for the only cases of interest, $\ell(\tilde{I}) = 2^{-j}\ell(\tilde{Q})$ for some $j \geq 0$. Given $0 < \eta < 1/2$, let N be the unique natural number such that $\eta \in [2^{-N-1}, 2^{-N})$. For such $\tilde{I}$, the absolute value of (28)—and thus $\|\gamma^{(\tilde{I})}\|_1$—is less than or equal to a constant times $2^{-N}\ell(Q)$.

We will give two bounds on the absolute value of (24), depending on whether $j \leq N$ or $j > N$. We only use (28) for the $j \leq N$ estimate.

If $j \leq N$ (so that $\tilde{I}$ is not too small compared to $\tilde{Q}$), then the absolute value of (24) is less than or equal to a constant times $2^{-N}\ell(\tilde{Q})V_{g^{(Q)}}(\bar{\tilde{I}})$.

If $j > N$ (meaning that $\tilde{I}$ is very small compared to $\tilde{Q}$) then the absolute value of (24) is less than or equal to

$$\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right)\|h^{(\tilde{I})}\|_1,$$

which is the same as

$$2^{-j}\ell(\tilde{Q})\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right). \tag{29}$$

Of course, what we need to bound is not the absolute value of (24), but the same divided by $|Q|^{1/2}|\tilde{I}|^{1/2} \sim 2^{-j/2}\ell(Q)$. (Recall that we are still working in $d = 1$.) If

$j \leq N$, the quotient is less than or equal to a constant times $2^{j/2}2^{-N}V_{g^{(Q)}}(\bar{\tilde{I}})$. If $j > N$ the corresponding estimate is $2^{-j/2}\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\delta(Q)\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right)$. Therefore, if $Q \in \mathcal{F}_i$, $\tilde{I} \in \mathcal{G}_i, \tilde{I} \subset \tilde{Q}$, and $\ell(\tilde{I}) = 2^{-j}\ell(\tilde{Q})$, then

$$|a(Q,\tilde{I})| \leq \begin{cases} C2^{j/2}2^{-N}V_{g^{(Q)}}(\bar{\tilde{I}}) & \text{if } j \leq N; \\ C2^{-j/2}\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\delta(Q)\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right) & \text{if } j > N; \end{cases}$$

while $a(Q,\tilde{I}) = 0$ if $\tilde{I} \not\subset \tilde{Q}$.

We now estimate (22)

$$\sum_{\tilde{I}}|a(Q,\tilde{I})| = \sum_{\tilde{I}:\,\tilde{I}\subset\tilde{Q}}|a(Q,\tilde{I})|$$

and (23)

$$\sum_{Q}|a(Q,\tilde{I})| = \sum_{Q:\,\tilde{I}\subset\tilde{Q}}|a(Q,\tilde{I})|.$$

Estimate of (22):

$$\sum_{\tilde{I}:\,\tilde{I}\subset\tilde{Q}}|a(Q,\tilde{I})| \leq C\sum_{j\geq 0}\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}|a(Q,\tilde{I})|$$

$$= C2^{-N}\sum_{0\leq j\leq N}2^{j/2}\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}V_{g^{(Q)}}(\bar{\tilde{I}})$$

$$+ C\sum_{j>N}2^{-j/2}\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\delta(Q)\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right)$$

$$= (I) + (II),$$

where

$$(I) = C2^{-N}\sum_{0\leq j\leq N}2^{j/2}\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}V_{g^{(Q)}}(\bar{\tilde{I}})$$

$$(II) = C\sum_{j>N}2^{-j/2}\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}\left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\delta(Q)\widetilde{g^{(Q)}}}(\bar{\tilde{I}})\right).$$

For each Q and $j \geq 0$,

$$\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})}V_{g^{(Q)}}(\bar{\tilde{I}}) \leq V_{g^{(Q)}}(\overline{\tilde{Q}}) \leq 1$$

and

$$\sum_{\tilde{I}:\ell(\tilde{I})=2^{-j}\ell(\tilde{Q})} \left(V_{g^{(Q)}}(\bar{\tilde{I}}) + V_{\delta(Q)g^{\widetilde{(Q)}}}(\bar{\tilde{I}}) \right) \leq V_{g^{(Q)}}(\bar{\tilde{Q}}) + V_{\delta(Q)g^{\widetilde{(Q)}}}(\bar{\tilde{Q}}) \leq 5/2,$$

because the change of variable does not affect the total variation and $|\delta(Q)| \leq 3/2$. Therefore

$$(I) \leq C2^{-N} \sum_{0 \leq j \leq N} 2^{j/2} \leq C2^{-N/2}$$

$$(II) \leq C \sum_{j>N} 2^{-j/2} \leq C2^{-N/2},$$

implying that $\sum_{\tilde{I}:\,\tilde{I}\subset\tilde{Q}} |a(Q,\tilde{I})| \leq C2^{-N/2}$.

Estimate of (23): This is like case a), but simpler, because, for each $j \geq 0$ and $\tilde{I}$, there is only one $\tilde{Q}$ such that $\tilde{I} \subset \tilde{Q}$ and $\ell(\tilde{I}) = 2^{-j}\ell(\tilde{Q})$. We get the same estimate: $\leq C2^{-N/2}$.

That proves Lemma 4, since $\eta \sim 2^{-N}$. ♠

The Schur Test and Lemma 4 imply that if $\{g^{(Q)}\}_{Q\in\mathcal{D}}$ and $\{\widetilde{g^{(Q)}}\}_{Q\in\mathcal{D}}$ are two families as given in Lemma 4's hypotheses, then, for any finite linear sum

$$\sum_{Q\in\mathcal{D}} \lambda_Q \left(\frac{g^{(Q)} - \delta(Q)\widetilde{g^{(Q)}}}{|Q|^{1/2}} \right),$$

we have

$$\left\| \sum_{Q\in\mathcal{D}} \lambda_Q \left(\frac{g^{(Q)} - \delta(Q)\widetilde{g^{(Q)}}}{|Q|^{1/2}} \right) \right\|_2 \leq C\eta^{1/2} \left(\sum_{Q\in\mathcal{D}} |\lambda_Q|^2 \right)^{1/2}, \tag{30}$$

with C an absolute constant. This implies Theorem 3 when $d = 1$. We can write

$$f^{(Q)} - \widetilde{f^{(Q)}} = \left(f^{(Q)} - \delta(Q)\widetilde{f^{(Q)}} \right) + (\delta(Q) - 1)\widetilde{f^{(Q)}}.$$

Lemma 4 implies that

$$\left\| \sum_{Q\in\mathcal{D}} \lambda_Q \left(\frac{f^{(Q)} - \delta(Q)\widetilde{f^{(Q)}}}{|Q|^{1/2}} \right) \right\|_2 \leq C\eta^{1/2} \left(\sum_{Q\in\mathcal{D}} |\lambda_Q|^2 \right)^{1/2}; \tag{31}$$

while, by Theorem 2,

$$\left\| \sum_{Q \in \mathcal{D}} \lambda_Q \left(\frac{\widetilde{f^{(Q)}}}{|Q|^{1/2}} \right) \right\|_2 \leq C \left(\sum_{Q \in \mathcal{D}} |\lambda_Q|^2 \right)^{1/2}. \tag{32}$$

Since $|1 - \delta(Q)| \leq \eta \leq \eta^{1/2}$ for all Q, we get (21) and thus Theorem 3 *in one dimension.*

We now prove (21) for general d. From here on we work in $\mathbf{R}^d$: $\mathcal{F}_{\vec{a}}$, $\mathcal{G}_{\vec{a}}$, $\mathcal{D}$, and $\tilde{\mathcal{D}}$ are families of *cubes*; $\vec{\delta}(Q)$ and $\vec{\tau}(Q)$ are *vectors*.

Fix $\vec{a} \in \{1, 2, 3\}^d$. For $Q \in \mathcal{F}_{\vec{a}}$ we write $\vec{\delta}(Q)$ as $(\delta_1(Q), \delta_2(Q), \ldots, \delta_d(Q))$ and $\vec{\tau}(Q)$ as $(\tau_1(Q), \tau_2(Q), \ldots, \tau_d(Q))$. Associated to each $\vec{\delta}(Q)$ and $\vec{\tau}(Q)$ will be two finite sequences of vectors $\{\tilde{\delta}_j(Q)\}_0^d$ and $\{\tilde{\tau}_j(Q)\}_0^d$, defined by

$$\begin{aligned} \tilde{\delta}_0(Q) &\equiv \vec{1} \\ \tilde{\delta}_1(Q) &\equiv (\delta_1(Q), 1, 1, \ldots, 1) \\ \tilde{\delta}_2(Q) &\equiv (\delta_1(Q), \delta_2(Q), 1, 1, \ldots, 1) \\ \tilde{\delta}_3(Q) &\equiv (\delta_1(Q), \delta_2(Q), \delta_3(Q), 1, 1, \ldots, 1) \\ &\cdots \\ \tilde{\delta}_d(Q) &= \delta(Q) \end{aligned}$$

and

$$\begin{aligned} \tilde{\tau}_0(Q) &= 0 \\ \tilde{\tau}_1(Q) &= (\tau_1(Q), 0, 0, \ldots, 0) \\ \tilde{\tau}_2(Q) &= (\tau_1(Q), \tau_2(Q), 0, 0, \ldots, 0) \\ \tilde{\tau}_3(Q) &= (\tau_1(Q), \tau_2(Q), \tau_3(Q), 0, 0, \ldots, 0) \\ &\cdots \\ \tilde{\tau}_d(Q) &= \tau(Q). \end{aligned}$$

In other words, considered as a dilation operator, $\tilde{\delta}_0(Q)$ starts as the identity, and then, as j advances, morphs—one variable at a time—into $\vec{\delta}(Q)$; while $\tilde{\tau}_j(Q)$ similarly morphs from the identity into $\vec{\tau}(Q)$, but now considered as a sequence of *translation* operators. *Keep in mind that $\delta_j(Q)$ and $\tau_j(Q)$ are numbers (components of the vectors $\vec{\delta}(Q)$ and $\vec{\tau}(Q)$) while $\tilde{\delta}_j(Q)$ and $\tilde{\tau}_j(Q)$ are vectors.*

Define $\zeta_0^{(Q)}(x) \equiv f^{(Q)}(x)$ and, for $1 \le k \le d$,

$$\zeta_k^{(Q)}(x) = \left(\prod_1^k \delta_j(Q) \right) f^{(Q)}(\tilde{\delta}_k(Q)(x - x_Q + \ell(Q)\tilde{\tau}_k(Q)) + x_Q).$$

After noticing that $\zeta_d^{(Q)}(x) = \left(\prod_1^d \delta_k(Q) \right) \widetilde{f^{(Q)}}(x)$, we write

$$f^{(Q)}(x) - \widetilde{f^{(Q)}}(x) = f^{(Q)}(x) - \left(\prod_1^d \delta_k(Q) \right) \widetilde{f^{(Q)}}(x) + \left(\left(\prod_1^d \delta_k(Q) \right) - 1 \right) \widetilde{f^{(Q)}}(x)$$

$$= \zeta_0^{(Q)}(x) - \zeta_d^{(Q)}(x) + \left(\left(\prod_1^d \delta_k(Q) \right) - 1 \right) \widetilde{f^{(Q)}}(x)$$

$$= \left[\sum_{k=1}^d \left(\zeta_{k-1}^{(Q)}(x) - \zeta_k^{(Q)}(x) \right) \right] + \left[\left(\left(\prod_1^d \delta_k(Q) \right) - 1 \right) \widetilde{f^{(Q)}}(x) \right]$$

$$\equiv [I] + [II].$$

The term $[II]$ is no problem, because

$$\left| \left(\left(\prod_1^d \delta_k(Q) \right) - 1 \right) \right| \le C(d)\eta$$

and Theorem 2 controls the almost-orthogonal "norm" of $\{ \widetilde{f^{(Q)}} / |Q|^{1/2} \}_{\mathcal{F}_{\tilde{a}}}$.

To see what is going on with $[I]$, we look at the first term in the sum,

$$\zeta_0^{(Q)}(x) - \zeta_1^{(Q)}(x) = f^{(Q)}(x) - \delta_1(Q)f^{(Q)}(\tilde{\delta}_1(Q)(x - x_Q + \ell(Q)\tilde{\tau}_1(Q)) + x_Q). \quad (33)$$

Write $x = (x_1, x_2, \ldots, x_d)$ as (x_1, x^*), where $x_1 \in \mathbf{R}$ and $x^* \in \mathbf{R}^{d-1}$. For fixed x^*, (33) is

$$f^{(Q)}(x_1, x^*) - \delta_1(Q)f^{(Q)}(\delta_1(Q)(x_1 - (x_Q)_1 + \ell(Q)\tau_1(Q)) + (x_Q)_1, x^*) \quad (34)$$

(note the absence of tildes), because the (respective) dilation and translation operators $\tilde{\delta}_1(Q)$ and $\tilde{\tau}_1(Q)$ do not affect the x^* components at all.

To ease reading we refer to (34) as $\omega^{(Q)}(x)$.

For $\tilde{Q} \in \mathcal{G}_{\tilde{a}}$, write $\tilde{Q} = I_1(\tilde{Q}) \times K(\tilde{Q})$, as in the statement of Lemma 2. Then

$$\omega^{(Q)}(x_1, x^*) = \omega^{(Q)}(x_1, x^*)\chi_{I_1(\tilde{Q})}(x_1)\chi_{K(\tilde{Q})}(x^*)$$

and, for every fixed $x^* \in \mathbf{R}^{d-1}$ and every finite linear sum

$$\sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}}{|Q|^{1/2}} \right) = \sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}(x_1, x^*) \chi_{I_1(\tilde{Q})}(x_1) \chi_{K(\tilde{Q})}(x^*)}{|Q|^{1/2}} \right),$$

we have, by the one-dimensional version of Theorem 3,

$$\int_{\mathbf{R}} \left| \sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}(x_1, x^*)}{|Q|^{1/2}} \right) \right|^2 dx_1 \leq C\eta \sum_{Q \in \mathcal{F}_{\vec{a}}} |\lambda_Q|^2 |K(\tilde{Q})|^{-1} \chi_{K(\tilde{Q})}(x^*). \quad (35)$$

Here we are arguing just as we did in estimating (16), but incorporating the '$\leq C\eta$' bound we have from the one-dimensional Theorem 3 (see (30)–(32)). We get η this time, and not $\eta^{1/2}$, because we are not taking the square root of the integral. When we integrate (35) in x^* we get, for every $\vec{a} \in \{1, 2, 3\}^d$,

$$\int_{\mathbf{R}^d} \left| \sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}(x)}{|Q|^{1/2}} \right) \right|^2 dx = \int_{\mathbf{R} \times \mathbf{R}^{d-1}} \left| \sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}(x_1, x^*)}{|Q|^{1/2}} \right) \right|^2 dx_1 \, dx^*$$

$$\leq C\eta \sum_{Q \in \mathcal{F}_{\vec{a}}} |\lambda_Q|^2.$$

The other summands in $[I]$ are handled in a similar fashion, successively treating the variables $x_2, \dots, x_d$ as we did x_1. For example, $\zeta_1^{(Q)}(x) - \zeta_2^{(Q)}(x)$ equals $\delta_1(Q)$ times

$$f^{(Q)}(\tilde{\delta}_1(Q)(x - x_Q + \ell(Q)\tilde{\tau}_1(Q)) + x_Q) - \delta_2(Q)f^{(Q)}(\tilde{\delta}_2(Q)(x - x_Q + \ell(Q)\tilde{\tau}_2(Q)) + x_Q),$$

where the functions' two arguments, respectively

$$(\tilde{\delta}_1(Q)(x - x_Q + \ell(Q)\tilde{\tau}_1(Q)) + x_Q \quad (36)$$

and

$$\tilde{\delta}_2(Q)(x - x_Q + \ell(Q)\tilde{\tau}_2(Q)) + x_Q, \quad (37)$$

differ only in their second components. The second component of (36) is x_2, and that of (37) is

$$\delta_2(Q)(x_2 - (x_Q)_2 + \ell(Q)\tau_2(Q)) + (x_Q)_2.$$

But their first components both equal

$$\delta_1(Q)(x_1 - (x_Q)_1 + \ell(Q)\tau_1(Q)) + (x_Q)_1;$$

and, for $3 \leq k \leq d$, each kth component for both functions equals x_k.

If we now define, more or less as before,

$$\omega^{(Q)}(x) \equiv \zeta_1^{(Q)}(x) - \zeta_2^{(Q)}(x),$$

then the preceding argument applies virtually verbatim to yield

$$\int_{\mathbf{R}^d} \left| \sum_{Q \in \mathcal{F}_{\vec{a}}} \lambda_Q \left(\frac{\omega^{(Q)}(x)}{|Q|^{1/2}} \right) \right|^2 dx \leq C\eta \sum_{Q \in \mathcal{F}_{\vec{a}}} |\lambda_Q|^2$$

for every $\vec{a} \in \{1, 2, 3\}^d$. (Recall that $\delta_1(Q)$ is essentially 1.) The same argument applies to the other summands $\zeta_{k-1}^{(Q)} - \zeta_k^{(Q)}$ for $3 \leq k \leq d$ to yield the same estimates. When we add up over all k and all $\vec{a} \in \{1, 2, 3\}^d$, and include the term [II], we get

$$\left\| \sum_{Q \in \mathcal{D}} \lambda_Q \left(\frac{f^{(Q)} - \widetilde{f^{(Q)}}}{|Q|^{1/2}} \right) \right\|_2 \leq C\eta^{1/2} \left(\sum_{Q \in \mathcal{D}} |\lambda_Q|^2 \right)^{1/2}$$

for all finite linear sums,

$$\sum_{Q \in \mathcal{D}} \lambda_Q \left(\frac{f^{(Q)} - \widetilde{f^{(Q)}}}{|Q|^{1/2}} \right),$$

where C depends on d. That's (21). Theorem 3 is proved. ♠

Appendix

Point 1. The total variation $V_f(I)$ adds like a measure; $\Omega_f(I)$ doesn't. In particular, if $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ then

$$\sum_1^n V_f[x_{k-1}, x_k] \le V_f[a, b],$$

which we need (see (15)); but

$$\sum_1^n \Omega_f[x_{k-1}, x_k] \le \Omega_f[a, b]$$

fails. (As with V_f, we use $\Omega_f[x, y]$ to mean $\Omega_f([x, y])$.)

Point 2. We give two answers; the second makes the first redundant.

a) Theorem 1 yields

$$\left\| \sum \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}} \right\|_2 \le C(d) \left(\sum |\lambda_Q|^2 \right)^{1/2},$$

with $C(d)$ growing at worst linearly in d. If we apply Schur's test to

$$\left\{ \frac{f^{(Q)}}{|Q|^{1/2}} \right\}_{Q \in \mathcal{D}}$$

and the orthonormal family (3) then, for each J and Q in $\mathcal{D}$ such that $J \subset Q$, we have to consider the inner products

$$|Q|^{-1/2} |J|^{-1/2} \langle f^{(Q)}, h_i^{(J)} \rangle$$

for **every** $1 \le i < 2^d$, and the Schur bound grows exponentially in d.

But it's worse than that.

b) For fixed $d > 1$ set $Q_0 = [0, 1)^d$ and let B be the ball (open or closed) of radius $1/3$ centered at $(1/2, 1/2, 1/2, \ldots, 1/2)$ (the center of Q_0). Define $f(x) = \chi_B(x) - |B| \chi_{Q_0}(x)$. Then f is a bounded multiple of a function in $NBV(Q_0)$. We look at the terms appearing in the Schur test,

$$|J|^{-1/2} |Q_0|^{-1/2} \langle f, h_i^{(J)} \rangle = |J|^{-1/2} \langle f, h_i^{(J)} \rangle, \tag{38}$$

for $J \subset Q_0$ and a fixed i (the value of i doesn't matter: say it's 1). Let 2^{-k} be J's sidelength, where $k > 0$. The inner product (38) equals 0 if $J \subset B$ or $J \cap B = \emptyset$. It's possibly non-zero if J straddles B's boundary, and when that happens our best estimate for the absolute value of (38) is roughly

$$|J|^{-1/2} |J| = |J|^{1/2} = 2^{-kd/2}.$$

The $(d-1)$-dimensional measure of B's boundary is ~ 1 (it depends on d but not on k). The cubes J have diameters $\sim 2^{-k}$. For large k, $\sim 2^{k(d-1)}$ of them can straddle B's boundary. (The reader might want to sketch this.) Therefore, when we add up the estimated absolute values of (38) for these J's, we get $\sim 2^{-kd/2}2^{k(d-1)} = 2^{k(d/2-1)}$, which sums (over k) to infinity.

Philosophical Remark We think Theorem 1 holds because of subtle cancelation in the sums

$$\sum \lambda_Q \frac{f^{(Q)}}{|Q|^{1/2}},$$

which the bounded-variation-plus-induction argument lets us exploit without really understanding.

References

1. T. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, 1974)
2. J.B. Garnett, *Bounded Analytic Functions* (Academic, New York, 1981)
3. M. Wilson, *Weighted Littlewood-Paley Theory and Exponential-Square Integrability*. Springer Lecture Notes in Mathematics, vol. 1924 (Springer, New York, 2007)
4. M. Wilson, Invariance and stability of almost-orthogonal systems. Trans. Am. Math. Soc. **368**, 2515–2546 (2016)