Appro imation and Spanning in the ardy Space, by Affine Systems

H.-Q. Bui [1] and R.S. Laugesen [2]

[1] *Department of Mathematics & Statistics,*
University of Canterbury,
Private Bag 4800, Christchurch, New Zealand.

[2] *Department of Mathematics,*
University of Illinois,
Urbana, IL 61801, U.S.A.

Report Number: UCDMS2005/6 July 2005

Keywords: Completeness, quasi-interpolation, approximate identity, atomic, scale averaging.

APPROXIMATION AND SPANNING IN THE HARDY SPACE, BY AFFINE SYSTEMS

H.-Q. BUI AND R. S. LAUGESEN

ABSTRACT. We find weak conditions on $\psi \in L^1(\mathbb{R}^d)$ with $\widehat{\psi}(0) = 1$ such that every function in the Hardy space is a linear combination of translates and dilates of ψ. More precisely, we prove for each $f \in H^1(\mathbb{R}^d)$ the scale averaged approximation formula

$$f(x) = \lim_{J\to\infty} \frac{1}{J} \sum_{j=1}^{J} \sum_{k\in\mathbb{Z}^d} c_{j,k}\psi(a_j x - k) \qquad \text{in } H^1(\mathbb{R}^d),$$

where $\{a_j\}$ is an arbitrary lacunary sequence (such as $a_j = 2^j$) and the coefficients $c_{j,k}$ are local averages of f. This holds in particular if ψ is Schwartz class, or if $\psi \in L^p$ (for some $1 < p < \infty$) has compact support. A corollary is a new affine decomposition of H^1 in terms of differences of ψ.

1. Introduction

We recently proved [2] a scale averaged, discretized approximation to the identity formula for $L^p = L^p(\mathbb{R}^d)$. Precisely, if $\psi \in L^p \cap L^1$, $1 \le p < \infty$, $\int_{\mathbb{R}^d} \psi\, dx = 1$, $\sum_{k\in\mathbb{Z}^d} |\psi(x-k)| \in L^p_{loc}$ and $\{a_j\}_{j=1}^{\infty}$ is a sequence of real numbers that grows exponentially (*i.e.*, is lacunary), then

$$f(x) = \lim_{J\to\infty} \frac{1}{J} \sum_{j=1}^{J} \sum_{k\in\mathbb{Z}^d} c_{j,k}\psi(a_j x - k) \qquad \text{in } L^p \tag{1}$$

for all $f \in L^p$. The coefficients $c_{j,k} = \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-k)\, dy$ are sampled average values of f; here the analyzer ϕ has integral 1 and satisfies some other conditions. The scale averaging over $j = 1, \ldots, J$, in formula (1), cannot generally be omitted [2, §1].

Theorem 1 in this paper extends (1) to the Hardy space $H^1 = H^1(\mathbb{R}^d)$. Special cases of Theorem 1 say that (1) holds in H^1 if $\int_{\mathbb{R}^d} \psi\, dx = 1$ and either $\psi \in L^p$ ($1 < p < \infty$) has compact support or else $\psi \in L^1$ has gradient $D\psi \in H^1$ (which holds for example for all Schwartz functions ψ).

Our conditions on ψ and ϕ, as well as our proof, must be substantially modified from the L^p case to deal with the Riesz transform. To hint at the difficulties, observe in formula (1) that $\psi(a_j x - k) \notin H^1$ because $\int_{\mathbb{R}^d} \psi(a_j x - k)\, dx \neq 0$, but that the infinite sum $\sum_{k\in\mathbb{Z}^d} c_{j,k}\psi(a_j x - k)$ can still belong to H^1 provided $\sum_{k\in\mathbb{Z}^d} c_{j,k} = 0$. We further discuss the modifications needed for H^1 in Section 8, after the proof of Theorem 1.

Corollary 2 shows the Hardy space is spanned by an affine system of *differences* of ψ, somewhat in the spirit of atomic and molecular decompositions of function spaces (for which see [5, 6, 7, 14]). It particularly recalls the work of J. E. Gilbert *et al.* [7], who obtained frame decompositions for Triebel–Lizorkin spaces using affine systems generated by "$\mathcal{M}_\delta$-molecules". Their theorem immediately implies a spanning result for H^1. In Section 3.6 we construct an example to show that their spanning result and our Corollary 2 are independent.

Date: July 7, 2005.

1991 *Mathematics Subject Classification.* Primary 42A50,42B30,42C30.

Key words and phrases. Completeness, quasi-interpolation, approximate identity, atomic, scale averaging.

Laugesen was partially supported by N.S.F. Award DMS–0140481, a Maclaurin Fellowship from the New Zealand Institute of Mathematics and its Applications, and a Visiting Erskine Fellowship from the University of Canterbury.

2. **Definitions and Notation**

1. Fix the *dimension* $d \in \mathbb{N}$ and write $\mathcal{C} = [0,1)^d$ for the unit cube in $\mathbb{R}^d$. Write $L^p = L^p(\mathbb{R}^d)$ for the class of complex valued functions with finite L^p-norm $\|f\|_p = (\int_{\mathbb{R}^d} |f(x)|^p \, dx)^{1/p}$. Occasionally we consider $\mathbb{C}^d$-valued functions, especially the gradient Df and the Riesz transform Rf, defined below. If a function F is $\mathbb{C}^d$-valued then we interpret its L^p norm in the obvious way, with $|F(x)|$ denoting the euclidean length of the vector $F(x)$.

2. Define the Fourier transform with 2π in the exponent: $\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i \xi x} \, dx$, for row vectors $\xi \in \mathbb{R}^d$.

3. Write $Rf = (R_1 f, \ldots, R_d f)$ for the Riesz transform of $f \in L^1$, where

$$R_s f(x) = c_d \ \text{p.v.} \int_{\mathbb{R}^d} f(x-y) \frac{y_s}{|y|^{d+1}} \, dy \qquad \text{for } s = 1, \ldots, d,$$

with normalizing constant $c_d = \Gamma((d+1)/2)\pi^{-(d+1)/2}$. Then $R_s f$ is finite a.e., and is a measurable function of $x \in \mathbb{R}^d$. Recall $R_s : L^p \to L^p$ for $1 < p < \infty$.

The *Hardy space* is

$$H^1 = H^1(\mathbb{R}^d) = \{f \in L^1 : Rf \in L^1\}, \quad \text{with the norm } \|f\|_{H^1} = \|f\|_1 + \|Rf\|_1.$$

A vector valued function is said to belong to H^1 if each of its components does. In particular, $Df \in H^1$ means $D_t f \in H^1$ for each $t = 1, \ldots, d$, where D is the gradient operator and $D_t = \partial/\partial x_t$.

Notice

$$\widehat{Rf}(\xi) = -i \frac{\xi}{|\xi|} \hat{f}(\xi).$$

If $f \in H^1$ then $Rf \in L^1$ and so $\widehat{Rf}$ is continuous, which implies

$$\hat{f}(0) = \int_{\mathbb{R}^d} f(x) \, dx = 0 \quad \text{and} \quad \widehat{Rf}(0) = \int_{\mathbb{R}^d} Rf(x) \, dx = 0. \tag{2}$$

Recall too that the Riesz transform commutes with dilations and translations: $R(f(\alpha x - \beta)) = \text{sign}(\alpha)(Rf)(\alpha x - \beta)$ when $\alpha \in \mathbb{R} \setminus \{0\}, \beta \in \mathbb{R}^d$. Dilation invariance fails when α is a matrix, which is why we restrict to isotropic dilations $a_j \in \mathbb{R}$ throughout this paper. (Our L^p results do hold for general dilation matrices [2].) If $f \in H^1$ and $g \in L^1$, then $f * g \in H^1$ and $R_s(f * g) = R_s f * g$.

See [10, 11] for all these facts about Riesz transforms and H^1.

4. Let the *dilations* a_j for $j > 0$ be nonzero real numbers with $|a_j| \to \infty$ as $j \to \infty$.

In some results we will further assume the a_j *grow exponentially*, meaning $|a_{j+1}| \geq \gamma |a_j|$ for all $j > 0$, for some growth factor $\gamma > 1$ (so that the dilation sequence is lacunary).

5. Fix a *translation matrix* b, assumed to be an invertible $d \times d$ real matrix.

6. For $\theta \in L^1$, define

$$\theta_{j,k}(x) = |a_j|^d \theta(a_j x - bk), \qquad j > 0, \quad k \in \mathbb{Z}^d, \quad x \in \mathbb{R}^d.$$

Notice we have put an L^1 normalization on $\theta_{j,k}$ (namely $\|\theta_{j,k}\|_1 = \|\theta\|_1$) instead of the L^2 normalization that is customary in wavelet theory.

7. We will use the *periodization* operator

$$Pf(x) = |\det b| \sum_{k \in \mathbb{Z}^d} f(x - bk) \qquad \text{for } x \in \mathbb{R}^d.$$

If $f \in L^1$, then the series for Pf converges absolutely for almost every x, and Pf is locally integrable.

8. The *first difference* operator $\Delta_z f(x) = f(x) - f(x-z)$ commutes with the Riesz transform.

3. Statements of the results

3.1. The results. We define an *approximation to f at scale j* by

$$
\begin{aligned}
f_j(x) &= |\det b||a_j|^{-d} \sum_{k\in\mathbb{Z}^d} \langle f, \overline{\phi}_{j,k}\rangle \psi_{j,k} \\
&= |\det b| \sum_{k\in\mathbb{Z}^d} \left(\int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)\,dy \right) \psi(a_j x - bk), \qquad j>0, \quad x\in\mathbb{R}^d, \qquad (3)
\end{aligned}
$$

where f is the *signal*, ϕ is the *analyzer* and ψ is the *synthesizer.* To understand f_j, consider ϕ to be a delta function (although admittedly this extreme case is not covered by our theorem); then with $b = I$ we get the quasi-interpolant $f_j(x) = \sum_{k\in\mathbb{Z}^d} f(a_j^{-1}k)\psi(a_j x - k)$.

Our theorem finds conditions under which the f_j provide a good approximation to f.

Theorem 1 (Approximation)**.** *Assume $\psi \in L^1$ with $\int_{\mathbb{R}^d} \psi\,dx = 1$. Let $\phi \in L^\infty$ be compactly supported with $P\phi = 1$ a.e. (so that $\int_{\mathbb{R}^d} \phi\,dx = 1$). Consider $f \in H^1$. Then* (a), (b) *and* (c) *hold.*

(a) [Stability] *The series* (3) *defining f_j converges unconditionally in L^1, so that $f_j \in L^1$. If*

$$\sup_{|z|\le 1} \|\Delta_z\psi\|_{H^1} < \infty \qquad (4)$$

then $f_j \in H^1$ with

$$\|f_j\|_{H^1} \le C(\phi,\psi)\|f\|_{H^1}. \qquad (5)$$

(b) [Constant periodization] *If*

$$\|\Delta_z\psi\|_{H^1} \to 0 \qquad \text{as } z \to 0 \qquad (6)$$

and ψ has constant periodization $P\psi = 1$ a.e., then the stability bound (5) *holds and*

$$f = \lim_{j\to\infty} f_j \qquad \text{in } H^1. \qquad (7)$$

(c) [Scale averaging] *If* (6) *holds and the dilations a_j grow exponentially, then the stability bound* (5) *holds and*

$$f = \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} f_j \qquad \text{in } H^1. \qquad (8)$$

If the dilations do not grow exponentially, then one can always pass to a subsequence that does, before applying part (c).

A spanning corollary follows from Theorem 1. Write e_t for the unit vector in the t^{th} coordinate direction and let $\theta^{(t)} = \psi - \psi(\cdot - be_t)$, for $t = 1,\dots,d$.

Corollary 2 (Spanning)**.** *Assume $\psi \in L^1$ with $\int_{\mathbb{R}^d} \psi\,dx \neq 0$ and $\|\Delta_z\psi\|_{H^1} \to 0$ as $z \to 0$.*

Then $\theta^{(t)} \in H^1$ for each t, and the system $\{\theta^{(t)}_{j,k} : j > 0, k \in \mathbb{Z}^d, t = 1,\dots,d\}$ spans H^1.

Spanning means the finite linear combinations of the functions

$$\theta^{(t)}_{j,k} = \psi_{j,k} - \psi_{j,k+e_t} \qquad (9)$$

are dense in H^1.

We discuss these results before proving them.

3.2. **Properties of f_j.**

- Theorem 1(a) shows $f_j \in H^1$. This is plausible because

$$\begin{aligned}\int_{\mathbb{R}^d} f_j(x)\,dx &= |\det b||a_j|^{-d} \sum_{k\in\mathbb{Z}^d} \langle f, \overline{\phi}_{j,k}\rangle && \text{using } \int_{\mathbb{R}^d} \psi\,dx = 1 \\ &= \int_{\mathbb{R}^d} f(y) P\phi(a_j y)\,dy \\ &= \int_{\mathbb{R}^d} f(y)\,dy = 0 && \text{since } P\phi = 1 \text{ a.e. and } f \in H^1.\end{aligned}$$

This calculation demonstrates that our assumption $P\phi \equiv 1$ is natural, in Theorem 1.

- f_j is related to a classical approximation to the identity:

$$\begin{aligned}f(x) &= \lim_{j\to\infty} (f * \psi_{a_j^{-1}})(x) \qquad \text{in } H^1 \\ &= \lim_{j\to\infty} \int_{\mathbb{R}^d} f(z)|a_j|^d \psi(a_j(x-z))\,dz \\ &= \lim_{j\to\infty} \int_{\mathbb{R}^d} f(a_j^{-1}y)\psi(a_j x - y)\,dy \qquad \text{by } z = a_j^{-1}y \\ &\approx \lim_{j\to\infty} \sum_{k\in\mathbb{Z}^d} \left(\int_{k+\mathcal{C}} f(a_j^{-1}y)\,dy \right) \psi(a_j x - k)\end{aligned} \tag{10}$$

by a Riemann sum approximation. This last line (10) is exactly $\lim_{j\to\infty} f_j$, with $\phi = \mathbb{1}_{\mathcal{C}}$ and $b = I$. Caution is required in the Riemann sum approximation step, because we discretize with fixed step size 1. Theorem 1 nonetheless shows the approximation (10) is exact in the H^1-norm as $j \to \infty$ provided either ψ has constant periodization or else we average over all dilation scales.

- In terms of integral kernels, $f_j(x) = \int_{\mathbb{R}^d} K_j(x,y) f(y)\,dy$ where

$$K_j(x,y) = |a_j|^d K(a_j x, a_j y) \quad \text{and} \quad K(x,y) = |\det b| \sum_{k\in\mathbb{Z}^d} \psi(x - bk)\phi(y - bk).$$

Thus Theorem 1(a) says $K_j : H^1 \to H^1$ with a norm estimate that is independent of j.

- The coefficients in f_j are controlled by the L^1-norm of f:

$$|\det b||a_j|^{-d} \sum_{k\in\mathbb{Z}^d} |\langle f, \overline{\phi}_{j,k}\rangle| \le \|f\|_1 \|P|\phi|\|_\infty.$$

3.3. **Examples for ψ.** Many functions satisfy the hypotheses in Theorem 1 and Corollary 2:

- Lemmas 3 and 4 show that

$$\psi \in L^1,\ D\psi \in H^1 \quad \Longrightarrow \quad \|\Delta_z \psi\|_{H^1} \to 0 \text{ as } z \to 0. \tag{11}$$

In the notation for Triebel–Lizorkin spaces the subclass $\{\psi \in L^1 : D\psi \in H^1\}$ equals $L^1 \cap \dot{F}^1_{1,2} = F^1_{1,2}$, by [6, (5.30)] or [15]. This subclass certainly contains all Schwartz functions (see [15, Theorem 2.2.3]).

- Lemma 5 shows that

$$\psi \in L^p \text{ compactly supported for some } p > 1 \quad \Longrightarrow \quad \|\Delta_z \psi\|_{H^1} \to 0 \text{ as } z \to 0.$$

In particular, in dimension 1 with $\psi = \mathbb{1}_{[0,1)}, a_j = 2^j$ and $b = 1$, Corollary 2 says the collection $\{\theta_{j,k} : j > 0, k \in \mathbb{Z}\}$ spans the Hardy space $H^1(\mathbb{R})$, where $\theta = \mathbb{1}_{[0,1)} - \mathbb{1}_{[1,2)}$ is a Haar-like function. Note this collection is oversampled by a factor of 2 compared with the usual Haar system, since $\theta_{j,k}$ overlaps $\theta_{j,k+1}$ and so on.

3.4. **Hypotheses on ψ: some finer points.**

• By Lemma 6,

$$\|\Delta_z\psi\|_{H^1} \to 0 \text{ as } z \to 0 \quad \Longrightarrow \quad \sup_{|z|\leq 1} \|\Delta_z\psi\|_{H^1} < \infty. \tag{12}$$

That is, (6) implies (4).

• The *constant periodization* condition $P\psi = 1$ a.e. in Theorem 1(b) says that the integer translates of ψ form a partition of unity. The condition is equivalent to

$$\hat{\psi}(0) = 1, \quad \hat{\psi}(nb^{-1}) = 0 \quad \text{for } n \in \mathbb{Z}^d \setminus \{0\},$$

which is the first Strang–Fix condition in approximation theory [1, 12]. All B-splines satisfy it. The simplest examples with $P\psi \equiv 1$ in one dimension (for $b = 1$) are the characteristic function $\psi = \mathbb{1}_{[0,1)}$ and the tent function $\psi(x) = 1 - |x|$ for $|x| < 1$.

• Corollary 2 fails for some $\psi \in L^1$, because $\theta = \Delta_{be_t}\psi$ need not belong to H^1 even though it integrates to zero; see the example in the remarks after Lemma 5.

3.5. **Connection to MRA scaling functions.** Suppose $\phi \in L^\infty$ has compact support and constant periodization $P\phi = 1$ a.e. Then for all $f \in H^1$, Theorem 1(b) with $\psi = \phi$ gives

$$f_j = |\det b||a_j|^{-d} \sum_{k\in\mathbb{Z}^d} \langle f, \overline{\phi}_{j,k}\rangle \phi_{j,k} \to f \qquad \text{in } H^1 \text{ as } j \to \infty. \tag{13}$$

(The hypothesis $\lim_{z\to 0} \|\Delta_z\psi\|_{H^1} \to 0$ in Theorem 1(b) is ensured by Lemma 5.)

Approximations like (13) in L^p arise in wavelet theory, when ϕ is a scaling (or refinable) function for a multiresolution analysis (MRA) in one dimension. There the $\phi_{j,k}$ are assumed orthogonal for $k \in \mathbb{Z}$, for each j, and so f_j in (13) represents the L^2-projection of f onto the span of the $\phi_{j,k}$. We are not aware of (13) having been proved previously for H^1.

Incidentally, if a scaling function ϕ is integrable then it automatically satisfies the constant periodization condition $P\phi \equiv 1$ by [8, Proposition 5.3.14].

3.6. **Comparison with molecular affine systems.**

• Let $0 < \delta < 1$. A continuous function $\theta : \mathbb{R} \to \mathbb{C}$ is called an $\mathcal{M}_\delta$-*molecule* if $\int_{\mathbb{R}} \theta(x)\,dx = 0$ and it satisfies the two conditions

$$|\theta(x)| \leq C(1 + |x|)^{-1-\delta},$$
$$|\theta(x + y) - \theta(x)| \leq C|y|^\delta(1 + |x|)^{-1-2\delta},$$

for all $x, y \in \mathbb{R}$ with $|y| \leq (1 + |x|)/2$, for some $C > 0$. The "frame decomposition" theorem of Gilbert *et al.* [7, Theorem 1.5] immediately implies the following spanning result for $H^1(\mathbb{R})$: if θ is an $\mathcal{M}_\delta$-molecule and there are $0 < A \leq B < \infty$ such that

$$A \leq \int_0^\infty |\hat{\theta}(t\xi)|^2 \,\frac{dt}{t} \leq B$$

for all $\xi \in \mathbb{R} \setminus \{0\}$, then there exist numbers $a > 1, b > 0$ such that the full affine system

$$\{\theta(a^j x - bk) : j \in \mathbb{Z}, k \in \mathbb{Z}\}$$

spans $H^1(\mathbb{R})$.

The allowable values of a and b are unknown, unlike in our Corollary 2 where the dilations a_j and translation step b can be arbitrary (subject only to $|a_j| \to \infty$). Also, Corollary 2 uses only the small scales $j > 0$, rather than all scales $j \in \mathbb{Z}$. Moreover the example below gives a function ψ satisfying the assumptions of Corollary 2 for which $\theta(x) = \psi(x) - \psi(x - 1)$ is not an $\mathcal{M}_\delta$-molecule.

On the other hand, $\theta(x) = \frac{d}{dx}[e^{-x^2}]$ gives an example of an $\mathcal{M}_\delta$-molecule that cannot be expressed as a difference like $\psi(x) - \psi(x - b)$ in Corollary 2, because $\hat{\theta}$ vanishes only at the origin. Such examples are plentiful. And the work in [7] gives more than just spanning for $H^1(\mathbb{R})$: it provides norm convergent expansions of the form $f = \sum_{j,k}\langle f, \rho_{(j,k)}\rangle\theta_{j,k}$ for a whole scale of homogeneous Triebel–Lizorkin spaces including H^1, and it does so in $\mathbb{R}^d$ for all $d \geq 1$.

In any event, the spanning results deduced from our Corollary 2 and the work of Gilbert *et al.* in [7] are independent of each other.

• *Example.* Let $I = [-1/2, 1/2]$ and let g be the function supported in I with $g(\pm 1/2) = 0$ and $g(\pm 1/4) = \mp 1$ and with g being linear between those points. Let $g_n(x) = n^4 g(n^4(x - 3n)), n \in \mathbb{N}$, so that g_n is supported in $I_n = [3n - 1/(2n^4), 3n + 1/(2n^4)]$ with $g_n(3n \pm 1/(2n^4)) = 0$ and $g_n(3n \pm 1/(4n^4)) = \mp n^4$. Since $\|g_n\|_\infty = n^4 \leq |I_n|^{-1}$ and $\int_{\mathbb{R}} g_n\, dx = 0$, each g_n is an H^1-atom (see [4],[11, pp. 91–92]). Let $c_n = 1/(3n(\log 3n)^2)$. Then $\|\sum_n c_n g_n\|_{H^1} \leq \sum_n c_n \cdot \|g\|_{H^1} < \infty$ because $\|g_n\|_{H^1} = \|g\|_{H^1}$ by translation and dilation invariance of the Riesz transform (or Hilbert transform, since we are in one dimension). Define $\psi(x) = \int_{-\infty}^{x}(\sum_n c_n g_n(y))\, dy$. Then $\psi' = \sum_n c_n g_n \in H^1$. Observe the graph of ψ consists of infinitely many disjoint nonnegative bumps supported in $\cup I_n$; the bump supported in I_n peaks at $x = 3n$, at which $\psi(3n) = c_n/4$. We deduce $0 < \int_{\mathbb{R}} \psi\, dx < \sum_n (c_n/4)/n^4 < \infty$. It follows now from (11) that ψ satisfies all the assumptions in Corollary 2.

Writing $\theta(x) = \psi(x) - \psi(x-1)$, notice the graph of θ consists of infinitely many disjoint bumps, with positive bumps around $x = 3n$ and negative ones around $x = 3n + 1$. Moreover, $|\theta(x)| \sim 1/(x(\log x)^2)$ for x near $3n$ (or near $3n+1$), and so θ does not decay at infinity like $|x|^{-1-\delta}$ for any $\delta > 0$. Thus θ fails the first condition for an $\mathcal{M}_\delta$-molecule. Clearly θ fails the other condition too, since $|\theta(x+1) - \theta(x)| \sim 1/(x(\log x)^2)$ for x near $3n - 1$, which does not decay at infinity like $|x|^{-1-2\delta}$.

3.7. **Unconditional bases for H^1.** Corollary 2 generates (small-scale) affine spanning sets for H^1. A basis would be stronger than a spanning set, since bases require unique representations. Unconditional bases for H^1 certainly do exist with affine structure: Strömberg [13] showed this using the Franklin wavelet system, and by now it is known that every wavelet basis is unconditional for H^1 provided the wavelet possesses sufficient smoothness and decay [8, Theorem 5.6.19]. Of course, requiring that the generating function ψ be a wavelet is a very strong assumption. In this paper we try instead to assume as little as possible about ψ, when obtaining spanning sets.

3.8. **Open problems.** The idea underlying Theorem 1 is to discretize the translation step in an approximate identity formula. We have succeeded in doing this for H^1, and also for L^p in the earlier paper [2]. We will treat Sobolev spaces in a forthcoming paper. But a number of interesting spaces remain, such as H^p for $0 < p < 1$.

Regarding Corollary 2 and spanning questions, even $H^1(\mathbb{R})$ presents simple questions we cannot yet answer. Consider for example the Mexican hat function $\theta(x) = (1 - x^2)e^{-x^2/2}$. Does its dyadic system $\{\theta(2^j x - k) : j \in \mathbb{Z}, k \in \mathbb{Z}\}$ span $H^1(\mathbb{R})$? And if so, then does it span also using only the small scales $j > 0$?

The Mexican hat has integral zero, $\int_{\mathbb{R}} \theta\, dx = 0$, because it is the second derivative of $-e^{-x^2/2}$. But the Mexican hat cannot be written as a difference of some ψ like in Corollary 2, since the Fourier transform of the Mexican hat vanishes only at the origin.

It is an open problem raised by Y. Meyer [9, p. 137] to determine whether the Mexican hat system spans L^p for $1 < p < \infty$. This is known to be true for $p = 2$, but the question remains open for all other p-values. See [2, §4] for more discussion of such spanning problems.

4. **Preparatory lemmas**

The next four sections develop tools for understanding and proving Theorem 1.

We start with a simple result on Riesz transforms and derivatives, already used in Section 3.3.

Lemma 3. *If $\psi \in W^{1,1}$ and $D\psi \in H^1$, then $R\psi \in L^p$ for all $1 < p < d/(d-1)$, and $R\psi \in W^{1,1}_{loc}$ with weak derivatives $D_t(R_s\psi) = R_s(D_t\psi) \in L^1$ for each $s, t = 1, \dots, d$.*

Proof. The Sobolev imbedding [10, p. 124] gives $\psi \in W^{1,1} \subset L^p$ for $1 < p < d/(d-1)$, and so $R\psi \in L^p$. Now we show $R_s\psi$ is weakly differentiable, with $D_t(R_s\psi) = R_s(D_t\psi)$ (which is integrable by the hypothesis $D\psi \in H^1$). Indeed for all test functions $v \in C_c^\infty$,

$$\begin{aligned}\int_{\mathbb{R}^d} (R_s\psi)\overline{D_t v}\, dx &= -\int_{\mathbb{R}^d} i\frac{\xi_s}{|\xi|}\widehat{\psi}(\xi)\,\overline{2\pi i\xi_t\hat{v}(\xi)}\, d\xi \\ &= \int_{\mathbb{R}^d} i\frac{\xi_s}{|\xi|}\widehat{D_t\psi}(\xi)\overline{\hat{v}(\xi)}\, d\xi \\ &= -\int_{\mathbb{R}^d} (R_s D_t\psi)\overline{v}\, dx.\end{aligned}$$

□

Next is a lemma used in proving (11).

Lemma 4. *If $\psi \in W^{1,1}_{loc}$ with $D\psi \in L^1$, then $\|\Delta_z\psi\|_1 \le |z| \cdot \|D\psi\|_1$ for all $z \in \mathbb{R}^d$.*

Proof. Fix $z \in \mathbb{R}^d$. Then for almost every $x \in \mathbb{R}^d$,

$$\Delta_z\psi(x) = \psi(x) - \psi(x-z) = \int_{-1}^{0} D\psi(x+uz)\cdot z\, du.$$

Now integrate with respect to x. □

The third lemma shows a way to satisfy hypothesis (6).

Lemma 5. *If $\psi \in L^p$ for some $p > 1$ and ψ has compact support, then $\|\Delta_z\psi\|_{H^1} \to 0$ as $z \to 0$.*

Proof. First, $\psi \in L^1$ and so $\|\Delta_z\psi\|_1 \to 0$ as $z \to 0$, by continuity of translation in L^1.

To handle the Riesz transform of $\Delta_z\psi$, we introduce a cut-off function $\chi \in C_c^\infty$ such that $\chi \equiv 1$ on a neighborhood of the support of ψ, and decompose $R\psi = h_1 + h_2$ where

$$h_1 = \chi \cdot R\psi, \qquad h_2 = (1-\chi)\cdot R\psi.$$

Because $\psi \in L^p$ we get $R\psi \in L^p$. Hence $h_1 \in L^1$, so that $\|\Delta_z h_1\|_1 \to 0$ as $z \to 0$. Further, h_2 is smooth because the Riesz transform

$$R\psi(x) = c_d \int_{\mathrm{spt}(\psi)} \psi(y)\frac{x-y}{|x-y|^{d+1}}\, dy$$

is smooth off the support of ψ. And $Dh_2 \in L^1$, because near infinity one has $h_2 \equiv R\psi$ and

$$|D(R\psi)(x)| \le C\int_{\mathrm{spt}(\psi)} \frac{|\psi(y)|}{|x-y|^{d+1}}\, dy \le C\|\psi\|_1 |x|^{-d-1} \qquad \text{as } |x| \to \infty.$$

Now Lemma 4 says $\|\Delta_z h_2\|_1 \to 0$ as $z \to 0$. Therefore $\|\Delta_z R\psi\|_1 \to 0$ as $z \to 0$, as desired. □

Side remarks (not needed later).

1. Lemma 5 fails for $p = 1$, because $\Delta_z\psi$ need not even belong to H^1 when $\psi \in L^1$ has compact support, as the following one dimensional example shows. Let

$$\psi(x) = \begin{cases} \frac{1}{x(\log x)^2} & \text{if } 0 < x < 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

Put $f = \Delta_z\psi \in L^1$, where $z > 0$ is arbitrary. Then it is easy to see $f \ge 0$ on the interval $(-\infty, z)$, but that $f\log(1+f)$ is not locally integrable around $x = 0$. Hence $f \notin H^1$ by [11, §III.5.3].

2. Lemma 5 and its proof do hold for $p = 1$ under the additional assumption that $R\psi \in L^1_{loc}$.

3. The decomposition in the proof can be used to show that $\psi \in H^1$ if $\psi \in L^p$ for some $1 < p < \infty$ and ψ has compact support and vanishing moment $\int_{\mathbb{R}^d} \psi \, dx = 0$. That result is known, of course [11, §III.1.2.4].

The fourth lemma shows that the supremum taken over $|z| \le 1$ in assumption (4) of Theorem 1 could just as well be taken over any other ball of z-values.

Lemma 6. *Let $\alpha, \beta > 0$ and $x, y \in \mathbb{R}^d$. Then*

$$\begin{aligned} \psi \text{ measurable and finite a.e.} \quad &\Longrightarrow \quad \sup_{z \in B(x,\alpha)} \|\Delta_z \psi\|_1 \le C(\alpha,\beta,x,y) \sup_{z \in B(y,\beta)} \|\Delta_z \psi\|_1, \\ \psi \in L^1 \quad &\Longrightarrow \quad \sup_{z \in B(x,\alpha)} \|\Delta_z \psi\|_{H^1} \le C(\alpha,\beta,x,y) \sup_{z \in B(y,\beta)} \|\Delta_z \psi\|_{H^1}. \end{aligned}$$

Proof. Let τ denote the translation operator: $\tau_z \psi(x) = \psi(x - z)$. Then for all $z, w \in \mathbb{R}^d$ one has

$$\Delta_{z-w}\psi = \tau_{-w}(\Delta_z \psi - \Delta_w \psi), \qquad \Delta_{z+w}\psi = \tau_w(\Delta_z \psi) + \Delta_w \psi.$$

The lemma now follows, using translation invariance of the L^1- and H^1-norms. □

The next lemma develops a simple H^1-density result, to be used in the proof of Theorem 1. Define the *local supremum* of f by

$$Qf(x) = \text{ess. sup}_{|z| \le \sqrt{d}} |f(x+z)| = \|f\|_{L^\infty(B(x,\sqrt{d}))},$$

which is a lower semicontinuous function of x.

Lemma 7. *The collection $\{f \in H^1 : Qf \in L^1\}$ is dense in H^1.*

Proof. Finite linear combinations of H^1 atoms are dense by the atomic decomposition of H^1 [11, §III.2.2]. Each such finite linear combination f is a bounded function with compact support, and hence Qf is also bounded with compact support. This more than proves the lemma.

We prefer to avoid calling on heavy machinery, though, and so we now present a direct, elementary proof of the lemma. Let η be a smooth, compactly supported mollifier. Then $Q\eta$ is bounded with compact support, and so $Q\eta \in L^1$. Let $f \in H^1$, so that $f * \eta \in H^1$, and

$$\begin{aligned} Q(f*\eta)(x) &\le \text{ess. sup}_{|z| \le \sqrt{d}} \int_{\mathbb{R}^d} |f(y)||\eta(x+z-y)| \, dy \\ &\le \int_{\mathbb{R}^d} |f(y)| Q\eta(x-y) \, dy \\ &= (|f| * Q\eta)(x) \in L^1. \end{aligned}$$

Thus $Q(f*\eta) \in L^1$. Obviously $Q(f*\eta_\varepsilon) \in L^1$ by the same reasoning, for each $\varepsilon > 0$. Since $f * \eta_\varepsilon \to f$ in H^1 as $\varepsilon \to 0$ (noting $R(f*\eta_\varepsilon) = (Rf)*\eta_\varepsilon \to Rf$ in L^1), we conclude the collection of $f \in H^1$ with $Qf \in L^1$ is dense in H^1. □

Now we show that convolution interacts well with the Riesz transform and differences. Fix a compact set $E \subset \mathbb{R}^d$ and define a local modulus of continuity operator by

$$Sf(x) = \text{ess. sup}_{z \in E} |f(x) - f(x+z)| = \text{ess. sup}_{z \in E} |\Delta_{-z} f(x)|, \qquad x \in \mathbb{R}^d.$$

Sf is measurable whenever f is measurable and finite a.e.

Lemma 8. *If $\psi \in L^1$ and $\eta \in C_c^\infty$, then $S(\psi * \eta) \in L^1$ and $SR(\psi*\eta) \in L^1$.*

Proof. $S\eta \in L^1$ because η is bounded with compact support. Thus $S(\psi * \eta) \le |\psi| * S\eta \in L^1$.

To treat $SR(\psi * \eta)$, we first observe that

$$\begin{aligned}\Delta_{-z}R(\psi * \eta) &= R\Delta_{-z}(\psi * \eta) \\ &= R(\psi * \Delta_{-z}\eta) \\ &= \psi * (R\Delta_{-z}\eta),\end{aligned}$$

where it is valid to move the Riesz transform inside the convolution because $\Delta_{-z}\eta \in H^1$ by [11, §III.1.2.4] (noting $\Delta_{-z}\eta$ is bounded with compact support and has integral zero). Next,

$$\Delta_{-z}\eta(x) = \eta(x) - \eta(x+z) = -\sum_{t=1}^{d} z_t \int_0^1 D_t\eta(x+uz)\,du$$

by the chain rule, so that

$$R\Delta_{-z}\eta(x) = -\sum_{t=1}^{d} z_t \int_0^1 RD_t\eta(x+uz)\,du.$$

Hence for $x \in \mathbb{R}^d$, $z \in E$,

$$|R\Delta_{-z}\eta(x)| \le \rho \sum_{t=1}^{d} \text{ess. sup}_{|y|\le\rho} |RD_t\eta(x+y)| =: \sigma(x),$$

where $\rho = \max_{z\in E}|z|$.

This new function σ is integrable, because $RD_t\eta$ is locally bounded and it decays at infinity like $|x|^{-d-1}$ (by integrating by parts in the formula for the Riesz transform $RD_t\eta$, using that η is smooth with compact support).

Combining the above estimates gives

$$\begin{aligned}SR(\psi * \eta)(x) &\le \text{ess. sup}_{z\in E}\,(|\psi| * |R\Delta_{-z}\eta|)(x) \\ &\le |\psi| * \sigma(x) \\ &\in L^1.\end{aligned}$$

□

The final lemma of the section controls H^1-convergence for an approximate identity.

Lemma 9. *If $f \in H^1$, and $\eta \in L^1$ is supported in the unit ball, then*

$$\|f - \eta_\varepsilon * f\|_{H^1} \le \sup_{|y|\le\varepsilon} \|\Delta_y f\|_{H^1} \cdot \|\eta\|_1, \qquad \varepsilon > 0.$$

Proof.

$$\begin{aligned}\|f - \eta_\varepsilon * f\|_1 &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \eta_\varepsilon(y)\Delta_y f(x)\,dy \right| dx \\ &\le \int_{\mathbb{R}^d} |\eta_\varepsilon(y)| \|\Delta_y f\|_1\,dy.\end{aligned}$$

A similar estimate applies to $\|R(f - \eta_\varepsilon * f)\|_1 = \|Rf - \eta_\varepsilon * Rf\|_1$, and so $\|f - \eta_\varepsilon * f\|_{H^1} \le \int_{\mathbb{R}^d} |\eta_\varepsilon(y)| \|\Delta_y f\|_{H^1}\,dy$, which implies the lemma. □

5. **Riesz transforms through the integral**

Given a function $F(x, y)$, write RF for the Riesz transform of F with respect to x. The Riesz transform can be taken through an integral with respect to y, as the next lemma shows.

Lemma 10. *Let $E \subset \mathbb{R}^d$ be compact. Suppose $F \in L^1(\mathbb{R}^d \times E)$ and $RF \in L^1(\mathbb{R}^d \times E)$.*
Then the function $f(x) = \int_E F(x, y)\, dy$ belongs to H^1, with Riesz transform

$$Rf(x) = \int_E RF(x, y)\, dy \tag{14}$$

and norm

$$\|f\|_{H^1} \leq \int_E \|F(\cdot, y)\|_{H^1}\, dy.$$

Proof. By Fubini's theorem, $f \in L^1$ with

$$\|f\|_1 \leq \int_E \|F(\cdot, y)\|_1\, dy = \|F\|_1 \quad \text{and} \quad \hat{f}(\xi) = \int_E \hat{F}(\xi, y)\, dy, \quad \xi \in \mathbb{R}^d.$$

Similarly the function $G(x) = \int_E RF(x, y)\, dy$ belongs to L^1 with $\|G\|_1 \leq \int_E \|RF(\cdot, y)\|_1\, dy = \|RF\|_1$, and G has Fourier transform

$$\begin{aligned}\hat{G}(\xi) &= \int_E \widehat{RF}(\xi, y)\, dy \\ &= -\int_E i\frac{\xi}{|\xi|}\hat{F}(\xi, y)\, dy \\ &= -i\frac{\xi}{|\xi|}\hat{f}(\xi).\end{aligned}$$

Thus $Rf = G$, which is (14). Moreover,

$$\begin{aligned}\|f\|_{H^1} &= \|f\|_1 + \|Rf\|_1 \\ &\leq \|F\|_1 + \|RF\|_1 \\ &= \int_E \|F(\cdot, y)\|_{H^1}\, dy.\end{aligned}$$

□

6. **Averages of rapidly oscillating functions**

This section investigates norm convergence of scale averages of rapidly oscillating functions, in preparation for the proof of Lemma 13.

Lemma 11. *Let $g \in L^1_{loc}$ be $b\mathbb{Z}^d$-periodic with mean value zero, and assume the dilations a_j grow exponentially.*

(i) *Then*

$$\lim_{J \to \infty} \frac{1}{J} \sum_{j=1}^{J} g(a_j x) = 0 \qquad \text{in } L^1_{loc}.$$

(ii) *If $f \in L^1$ with $Qf \in L^1$, then*

$$\lim_{J \to \infty} f(x) \frac{1}{J} \sum_{j=1}^{J} g(a_j x) = 0 \qquad \text{in } L^1.$$

Proof. Part (i) is proved in [2, Lemma 6].

To prove part (ii), first notice $f \in L^\infty$ because for each $y \in \mathbb{R}^d$,

$$\begin{aligned}
\|f\|_\infty &\le \sup_{k\in\mathbb{Z}^d} \|f\|_{L^\infty(B(y+k,\sqrt{d}))} \\
&\le \sum_{k\in\mathbb{Z}^d} \|f\|_{L^\infty(B(y+k,\sqrt{d}))} \\
&= \sum_{k\in\mathbb{Z}^d} (Qf)(y+k),
\end{aligned}$$

and integrating over $y \in \mathcal{C} = [0,1)^d$ gives $\|f\|_\infty \le \|Qf\|_1 < \infty$.

Let $K > 0$ be arbitrary. Then

$$\begin{aligned}
\int_{B(0,K)} \left| f(x)\frac{1}{J}\sum_{j=1}^{J} g(a_j x)\right| dx &\le \|f\|_\infty \int_{B(0,K)} \left| \frac{1}{J}\sum_{j=1}^{J} g(a_j x)\right| dx \\
&\to 0 \qquad \text{as } J \to \infty,
\end{aligned}$$

by the L^1_{loc} convergence in part (i). Furthermore, $|f(x)| \le Qf(k)$ for almost every $x \in k + \mathcal{C} \subset B(k,\sqrt{d})$ by definition of Q. Thus for each J,

$$\begin{aligned}
\int_{\mathbb{R}^d\setminus B(0,K)} \left| f(x)\frac{1}{J}\sum_{j=1}^{J} g(a_j x)\right| dx &\le \sum_{|k|>K-\sqrt{d}} \int_{k+\mathcal{C}} Qf(k) \left| \frac{1}{J}\sum_{j=1}^{J} g(a_j x)\right| dx \\
&\le \sum_{|k|>K-\sqrt{d}} Qf(k) \frac{1}{J}\sum_{j=1}^{J} |a_j|^{-d} \int_{a_j(k+\mathcal{C})} |g(x)|\, dx \\
&\le \sum_{|k|>K-\sqrt{d}} Qf(k)\cdot C\|g\|_{L^1(b\mathcal{C})} \qquad (15)
\end{aligned}$$

since the integral of the $b\mathbb{Z}^d$-periodic function $|g|$ over the set $a_j(k+\mathcal{C})$ is bounded by $C|a_j|^d$ times the integral of $|g|$ over $b\mathcal{C}$ (see for example [3, Lemma 25]; the constant C here depends on b and on the dilation sequence $\{a_j\}$). The expression (15) can be made as small as we like by choosing K sufficiently large, because $B(x,\sqrt{d}) \subset \cup_{|\ell|<3\sqrt{d}} B(\ell,\sqrt{d})$ whenever $|x| < \sqrt{d}$ and translating these balls by $k - x$ gives $Qf(k) \le \sum_{|\ell|<3\sqrt{d}} Qf(\ell + k - x)$, which implies by integrating over $x \in \mathcal{C}$ that

$$\begin{aligned}
\sum_{k\in\mathbb{Z}^d} Qf(k) &\le \sum_{k\in\mathbb{Z}^d} \int_{\mathcal{C}} \sum_{|\ell|<3\sqrt{d}} Qf(\ell+k-x)\, dx \\
&= \sum_{|\ell|<3\sqrt{d}} \|Qf\|_1 < \infty.
\end{aligned}$$

This proves part (ii). $\square$

7. The T-operator

Recall the periodization operator P defined in Section 2. Let $\phi \in L^\infty$ have compact support, and define a new operator T by

$$Th(x) = T_\phi h(x) = \int_{\mathbb{R}^d} (P\Delta_z h)(x)\phi(z)\, dz.$$

We show this operator is well defined. Later it plays a role in proving Theorem 1: see formula (26).

Lemma 12. *Assume h is measurable and finite a.e. and $\sup_{|z|\leq 1}\|\Delta_z h\|_1 < \infty$. Let $\phi \in L^\infty$ have compact support. Then Th is $b\mathbb{Z}^d$-periodic and locally integrable.*

Proof. We have

$$\begin{aligned}
\int_{b\mathcal{C}}\int_{\mathbb{R}^d} |(P\Delta_z h)(x)||\phi(z)|\,dzdx &\leq |\det b|\int_{\operatorname{spt}\phi}\int_{b\mathcal{C}}\sum_{k\in\mathbb{Z}^d}|\Delta_z h(x-bk)|\,dxdz\,\|\phi\|_\infty \\
&= |\det b|\int_{\operatorname{spt}\phi}\|\Delta_z h\|_1\,dz\,\|\phi\|_\infty \\
&\leq C(\phi,b)\sup_{|z|\leq 1}\|\Delta_z h\|_1\|\phi\|_\infty < \infty,
\end{aligned}$$

where the last step uses Lemma 6 to replace a supremum over $z \in \operatorname{spt}\phi$ with one over $|z| \leq 1$.

Hence the series defining $(P\Delta_z h)(x)\phi(z)$ converges absolutely for almost every $(x,z) \in b\mathcal{C}\times\mathbb{R}^d$, and defines an integrable function there. Hence the a.e. defined function $Th(x) = \int_{\mathbb{R}^d}(P\Delta_z h)(x)\phi(z)\,dz$ is measurable, with $\|Th\|_{L^1(b\mathcal{C})} < \infty$. Clearly Th is $b\mathbb{Z}^d$-periodic. □

The scale averages of T converge to zero, by the next lemma.

Lemma 13. *Assume $\psi \in W^{1,1}$ with $D\psi \in H^1$. Let $\phi \in L^\infty$ have compact support. Suppose $f \in L^1$ with $Qf \in L^1$.*

If the dilations a_j grow exponentially then

$$\lim_{J\to\infty} f(x)\frac{1}{J}\sum_{j=1}^{J}(T\psi)(a_j x) = 0 \qquad \text{in } L^1, \quad \text{and}$$

$$\lim_{J\to\infty} f(x)\frac{1}{J}\sum_{j=1}^{J}(TR_s\psi)(a_j x) = 0 \qquad \text{in } L^1, \ \text{for each } s = 1,\dots,d.$$

Proof. $T\psi$ belongs to L^1_{loc} by Lemma 12, and so does $TR_s\psi$ by (11) and (12) and Lemma 12. We need only show these functions have mean value zero, because then Lemma 11(ii) can be applied.

The mean value of $T\psi$ is

$$\begin{aligned}
|\det b|^{-1}\int_{b\mathcal{C}} T\psi(x)\,dx &= \int_{\mathbb{R}^d}|\det b|^{-1}\int_{b\mathcal{C}}(P\Delta_z\psi)(x)\,dx\,\phi(z)\,dz \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\Delta_z\psi)(x)\,dx\,\phi(z)\,dz \\
&= 0
\end{aligned}$$

since $\Delta_z\psi(x)$ integrates to zero over $x \in \mathbb{R}^d$, using $\psi \in L^1$. The mean value of $TR_s\psi$ is

$$\begin{aligned}
|\det b|^{-1}\int_{b\mathcal{C}} TR_s\psi(x)\,dx &= \int_{\mathbb{R}^d}|\det b|^{-1}\int_{b\mathcal{C}}(P\Delta_z R_s\psi)(x)\,dx\,\phi(z)\,dz \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(R_s\Delta_z\psi)(x)\,dx\,\phi(z)\,dz \\
&= 0,
\end{aligned}$$

since $\Delta_z\psi \in H^1$ for each z by Lemmas 3 and 4 so that the integral of $R_s(\Delta_z\psi)$ equals zero by property (2).

Note. The use of Fubini's theorem above is justified by Lemma 12 and its proof. □

The T-operator vanishes when ψ has constant periodization, as the next lemma shows.

Lemma 14. *Assume $\psi \in W^{1,1}$ with $D\psi \in H^1$. Let $\phi \in L^\infty$ have compact support.*

If $P\psi = 1$ a.e. then $T\psi = 0$ and $TR_s\psi = 0$ a.e., for each $s = 1, \ldots, d$.

Proof. $T\psi$ and $TR_s\psi$ are well defined and locally integrable, as in Lemma 13. The integrability of ψ implies that $P\psi$ is well defined pointwise a.e., and so $P\Delta_z\psi = \Delta_z P\psi$. Hence if $P\psi \equiv 1$ then $\Delta_z P\psi \equiv 0$, leading to $T\psi \equiv 0$.

To get $TR_s\psi \equiv 0$ we must argue more carefully, since $R_s\psi \notin L^1$. We do have by Lemma 3 that $R_s\psi \in L^p$ for $1 < p < d/(d-1)$ and $R_s\psi \in W^{1,1}_{loc}$ with gradient $D(R_s\psi) \in L^1$. This implies $\Delta_z R_s\psi(x) = \int_{-1}^0 D(R_s\psi)(x+uz)\cdot z\,du$ for almost every (x,z), so that

$$(P\Delta_z R_s\psi)(x) = \int_{-1}^0 PD(R_s\psi)(x+uz)\cdot z\,du.$$

We will show $PD(R_s\psi) \equiv 0$, so that $P\Delta_z R_s\psi \equiv 0$ and hence $TR_s\psi \equiv 0$, by definition of the T operator.

We need the simple fact that if $h \in L^1$ then the $b\mathbb{Z}^d$-periodic function $Ph(x)$ has Fourier series $\sum_{n\in\mathbb{Z}^d} \hat{h}(nb^{-1})e^{2\pi i nb^{-1}x}$. Applying this observation to $h = D(R_s\psi)$ yields that the Fourier coefficients of $PD(R_s\psi)$ are

$$\begin{aligned}
[D(R_s\psi)]\widehat{\ }(nb^{-1}) &= 2\pi i nb^{-1}(R_s\psi)\widehat{\ }(nb^{-1}) \\
&= 2\pi nb^{-1}\frac{(nb^{-1})_s}{|nb^{-1}|}\widehat{\psi}(nb^{-1}) \\
&= 2\pi nb^{-1}\frac{(nb^{-1})_s}{|nb^{-1}|}[n\text{th Fourier coefficient of } P\psi] \\
&= 0 \qquad \text{for all } n \neq 0,
\end{aligned}$$

because the constant function $P\psi \equiv 1$ must have all its Fourier coefficients equalling zero for $n \neq 0$. Further, the zero-th Fourier coefficient of $PD(R_s\psi)$ is $[D(R_s\psi)]\widehat{\ }(0) = [R_s(D\psi)]\widehat{\ }(0) = 0$ by (2).

We have shown all the Fourier coefficients of $PD(R_s\psi)$ are zero, and so $PD(R_s\psi) \equiv 0$. □

8. **Proof of Theorem 1**

Suppose $\psi \in L^1$. Take a compactly supported L^∞-function ϕ, so that $P|\phi| \in L^\infty$. Let $f \in H^1$.

8.1. **Proof of Part (a).**

We first prove $f_j \in L^1$, where we recall the definition (3):

$$f_j(x) = |\det b| \sum_{k\in\mathbb{Z}^d} \left(\int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)\,dy \right) \psi(a_j x - bk). \tag{16}$$

The sum defining f_j converges absolutely a.e. to a function in L^1 because

$$\begin{aligned}
\|f_j\|_1 &\leq \int_{\mathbb{R}^d} |\det b| \sum_{k\in\mathbb{Z}^d} \left| \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)\,dy \right| |\psi(a_j x - bk)|\,dx \\
&\leq |\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |f(a_j^{-1}y)||\phi(y-bk)|\,dy \cdot |a_j|^{-d}\|\psi\|_1 \\
&= \int_{\mathbb{R}^d} |f(a_j^{-1}y)|(P|\phi|)(y)\,dy \cdot |a_j|^{-d}\|\psi\|_1 \\
&\leq \|f\|_1 \|P|\phi|\|_\infty \|\psi\|_1 < \infty.
\end{aligned}$$

It follows that the series defining f_j converges unconditionally in L^1.

Next assume $P\phi = 1$ a.e. (so that $\int_{\mathbb{R}^d} \phi\, dx = 1$, by integrating $P\phi$ over the period cell $b\mathcal{C}$) and suppose ψ satisfies hypothesis (4). We will show $f_j \in H^1$. We have

$$\begin{aligned} f_j(x) &= |\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)|a_j|^{-d}\psi_{j,k}(x)\, dy \\ &= |a_j|^{-d}|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)\psi_{j,b^{-1}y}(x)\, dy + E_j\psi(x) \qquad \text{with } E_j \text{ defined below,} \\ &= |a_j|^{-d} \int_{\mathbb{R}^d} f(a_j^{-1}y)\psi_{j,b^{-1}y}(x)\, dy + E_j\psi(x) \end{aligned} \tag{17}$$

by interchanging sum and integral and using that $P\phi = 1$ a.e.; here $\psi_{j,b^{-1}y}(x) = |a_j|^d\psi(a_jx-y)$ and the error (or remainder) term is

$$E_j\psi(x) = |a_j|^{-d}|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)[\psi_{j,k}(x) - \psi_{j,b^{-1}y}(x)]\, dy. \tag{18}$$

Hence

$$f_j(x) = (f * \psi_{a_j^{-1}})(x) + E_j\psi(x), \tag{19}$$

by putting $y \mapsto a_jy$ in (17) to get the convolution $f * \psi_{a_j^{-1}}$ in (19).

Since $f * \psi_{a_j^{-1}} \in H^1$, to verify that $f_j \in H^1$ we need only show $E_j\psi \in H^1$. So temporarily fix j and write $E_j\psi = \sum_{k\in\mathbb{Z}^d} F_k$ where

$$\begin{aligned} F_k(x) &= |a_j|^{-d}|\det b| \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)[\psi_{j,k}(x) - \psi_{j,b^{-1}y}(x)]\, dy \\ &= |a_j|^{-d}|\det b| \int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)(\Delta_{y-bk}\psi)_{a_j^{-1}}(x - a_j^{-1}bk)\, dy. \end{aligned}$$

It follows from Lemma 10 that $F_k \in H^1$ with

$$\begin{aligned} \|F_k\|_{H^1} &\le |a_j|^{-d}|\det b| \int_{\mathbb{R}^d} |f(a_j^{-1}y)||\phi(y-bk)|\|(\Delta_{y-bk}\psi)_{a_j^{-1}}(\cdot - a_j^{-1}bk)\|_{H^1}\, dy \\ &\le |a_j|^{-d}|\det b| \left(\int_{\mathbb{R}^d} |f(a_j^{-1}y)||\phi(y-bk)|\, dy\right)\left(\sup_{z\in \operatorname{spt}\phi} \|\Delta_z\psi\|_{H^1}\right), \end{aligned}$$

where we used the translation and dilation invariance of the H^1-norm. (Note $\sup_{z\in\operatorname{spt}\phi} \|\Delta_z\psi\|_{H^1} < \infty$ by assumption (4) and Lemma 6.) Summing over $k \in \mathbb{Z}^d$ gives that

$$E_j\psi = \sum_{k\in\mathbb{Z}^d} F_k \in H^1$$

with norm estimate

$$\|E_j\psi\|_{H^1} \le \sum_{k\in\mathbb{Z}^d} \|F_k\|_{H^1} \le \|P|\phi|\|_\infty \|f\|_1 \left(\sup_{z\in\operatorname{spt}\phi} \|\Delta_z\psi\|_{H^1}\right) < \infty.$$

Moreover, since

$$\|f * \psi_{a_j^{-1}}\|_{H^1} \le \|f\|_{H^1}\|\psi\|_1$$

we deduce from (19) that $f_j \in H^1$ with the stability estimate

$$\|f_j\|_{H^1} \le \left(\|\psi\|_1 + \|P|\phi|\|_\infty \sup_{z\in\operatorname{spt}\phi} \|\Delta_z\psi\|_{H^1}\right)\|f\|_{H^1}. \tag{20}$$

Hence part (a) of Theorem 1 is proved. (*Aside.* The proof gives unconditional convergence in H^1 for the series defining $E_j\psi$. But we cannot claim the series defining f_j converges in H^1, because $\psi_{j,k}$ need not belong to H^1.)

For use later in the proof, we pause here to define $I_j[\psi,\phi]f := f_j$, emphasizing by this notation the fact that f_j in (16) arises from applying to f a linear operator $I_j[\psi,\phi]$ depending on ψ and ϕ. In this new terminology, the stability bound (20) says

$$\|I_j[\psi,\phi]f\|_{H^1} \leq \left(\|\psi\|_1 + \|P|\phi|\|_\infty \sup_{z\in \operatorname{spt}\phi} \|\Delta_z\psi\|_{H^1} \right) \|f\|_{H^1}. \tag{21}$$

8.2. **Proof of Parts (b) and (c).** Assume $\psi \in L^1$ with $\int_{\mathbb{R}^d} \psi\,dx = 1$. Let $\phi \in L^\infty$ be compactly supported with $P\phi = 1$ a.e. Let $f \in H^1$. Assume hypothesis (6) holds, that is $\|\Delta_z\psi\|_{H^1} \to 0$ as $z \to 0$. In particular $\|\Delta_z\psi\|_{H^1}$ is bounded for all small z and so is bounded for all $|z| \leq 1$ by Lemma 6. Hence part (a) of the theorem holds, including the stability bound (20).

We can suppose $Qf \in L^1$ when proving parts (b) and (c), because of the density of such functions in H^1 (by Lemma 7) and the stability $\|f_j\|_{H^1} \leq C(\phi,\psi)\|f\|_{H^1}$ in (20).

Next we reduce to proving the theorem for a dense class of synthesizers ψ. Specifically, we will show we can reduce to $\psi \in W^{1,1} \cap C^\infty$ with $D\psi \in H^1$ and $S\psi \in L^1, SR\psi \in L^1$, where the S-operator was defined before Lemma 8 (take $E = \operatorname{spt}\phi$ in that definition).

Choose a smooth, nonnegative mollifier $\eta(x)$ supported in the unit ball and define $\psi^{(\varepsilon)} = \eta_\varepsilon * \psi$. Then $\psi^{(\varepsilon)} \in W^{1,1} \cap C^\infty$ satisfies $\int_{\mathbb{R}^d} \psi^{(\varepsilon)}\,dx = 1$ and $\|\psi - \psi^{(\varepsilon)}\|_1 \to 0$ as $\varepsilon \to 0$. By Lemma 8, $S\psi^{(\varepsilon)} \in L^1$ and $SR\psi^{(\varepsilon)} \in L^1$.

Also $D\psi^{(\varepsilon)} = (D\eta_\varepsilon) * \psi \in H^1$, noting $D\eta_\varepsilon \in H^1$ because $D\eta_\varepsilon$ is bounded, compactly supported and has integral zero [11, §III.5.5]. Hence $\psi^{(\varepsilon)}$ satisfies hypothesis (6) by implication (11). Furthermore

$$\begin{aligned}
\sup_{z\in\mathbb{R}^d} \|\Delta_z(\psi - \psi^{(\varepsilon)})\|_{H^1} &= \sup_{z\in\mathbb{R}^d} \|\Delta_z\psi - \eta_\varepsilon * \Delta_z\psi\|_{H^1} \\
&\leq \sup_{z\in\mathbb{R}^d} \sup_{|y|\leq\varepsilon} \|\Delta_y\Delta_z\psi\|_{H^1} && \text{by Lemma 9 applied to } f = \Delta_z\psi \in H^1 \\
&\leq 2 \sup_{|y|\leq\varepsilon} \|\Delta_y\psi\|_{H^1} && \text{after commuting } \Delta_y\Delta_z = \Delta_z\Delta_y \\
&\to 0 && \text{as } \varepsilon \to 0, \text{ by hypothesis (6).}
\end{aligned}$$

Hence

$$\begin{aligned}
&\|I_j[\psi,\phi]f - I_j[\psi^{(\varepsilon)},\phi]f\|_{H^1} \\
&= \|I_j[\psi - \psi^{(\varepsilon)},\phi]f\|_{H^1} \\
&\leq \left(\|\psi - \psi^{(\varepsilon)}\|_1 + \|P|\phi|\|_\infty \sup_{z\in\operatorname{spt}\phi} \|\Delta_z(\psi - \psi^{(\varepsilon)})\|_{H^1} \right) \|f\|_{H^1} \qquad \text{by the stability estimate (21)} \\
&\to 0 \qquad \text{as } \varepsilon \to 0,
\end{aligned}$$

and this estimate is uniform with respect to j. It follows that Theorem 1(b)(c) need only be proved for $\psi^{(\varepsilon)}$, for each fixed $\varepsilon > 0$; regarding part (b), note also that if ψ has constant periodization $P\psi = 1$ a.e. then so does $\psi^{(\varepsilon)}$, in fact $P(\psi^{(\varepsilon)}) = \eta_\varepsilon * P\psi \equiv 1$.

This completes the reduction step on ψ, and so from now on we may assume $\psi \in W^{1,1} \cap C^\infty$ with $D\psi \in H^1$ and $S\psi \in L^1, SR\psi \in L^1$, and $\int_{\mathbb{R}^d} \psi\,dx = 1$.

Notice

$$f * \psi_{a_j^{-1}} \to f \qquad \text{in } H^1 \text{ as } j \to \infty. \tag{22}$$

In view of decomposition (19), then, our task is to understand the error term $E_j\psi$ as $j \to \infty$. We decompose it as

$$E_j\psi = E_j^{(1)}\psi + E_j^{(2)}\psi$$

where

$$E_j^{(1)}\psi(x) = |a_j|^{-d}|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} [f(a_j^{-1}y) - f(x)]\phi(y-bk)\,[\psi_{j,k}(x) - \psi_{j,b^{-1}y}(x)]\,dy,$$

$$E_j^{(2)}\psi(x) = f(x)\,|a_j|^{-d}|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \phi(y-bk)[\psi_{j,k}(x) - \psi_{j,b^{-1}y}(x)]\,dy.$$

(Convergence of these two series is justified by our work below.) These formulas can be expressed more usefully as

$$E_j^{(1)}\psi(x) = |\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} [f(x) - f(a_j^{-1}y)]\phi(y-bk)\,(\Delta_{bk-y}\psi)(a_jx-y)\,dy, \tag{23}$$

$$E_j^{(2)}\psi(x) = f(x)\,|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \phi(y-bk)(\Delta_{y-bk}\psi)(a_jx-bk)\,dy. \tag{24}$$

We first show $E_j^{(1)}$ converges to zero in L^1. Note by definition of the S operator that

$$|(\Delta_{bk-y}\psi)(a_jx-y)| \le S\psi(a_jx-y)$$

when $y - bk \in \operatorname{spt}\phi$. Therefore

$$\begin{aligned}
|E_j^{(1)}\psi(x)| &\le |\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |f(x) - f(a_j^{-1}y)||\phi(y-bk)|S\psi(a_jx-y)\,dy \\
&\le \|P|\phi|\|_\infty \int_{\mathbb{R}^d} |f(x) - f(a_j^{-1}y)|S\psi(a_jx-y)\,dy \\
&= \|P|\phi|\|_\infty \int_{\mathbb{R}^d} |f(x) - f(x - a_j^{-1}y)|S\psi(y)\,dy \qquad \text{by } y \mapsto a_jx - y.
\end{aligned}$$

Thus

$$\begin{aligned}
\|E_j^{(1)}\psi\|_1 &\le \|P|\phi|\|_\infty \int_{\mathbb{R}^d} \|f - f(\cdot - a_j^{-1}y)\|_1\,S\psi(y)\,dy \\
&\to 0 \qquad \text{as } j \to \infty,
\end{aligned} \tag{25}$$

by dominated convergence with respect to the y-integral. (Here we use $S\psi \in L^1$.)

Next we evaluate

$$\begin{aligned}
E_j^{(2)}\psi(x) &= f(x)\,|\det b| \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \Delta_z\psi(a_jx-bk)\phi(z)\,dz \qquad \text{by letting } z = y - bk \text{ in (24)} \\
&= f(x) \int_{\mathbb{R}^d} (P\Delta_z\psi)(a_jx)\phi(z)\,dz \qquad \text{by interchanging sum and integral,} \\
&\qquad \text{which is valid for almost every } x \text{ by the proof of Lemma 12,} \\
&= f(x)T\psi(a_jx).
\end{aligned} \tag{26}$$

In part (b) of the theorem, if $P\psi = 1$ a.e. then $T\psi = 0$ a.e. by Lemma 14 (using that $\psi \in W^{1,1}, D\psi \in H^1$), so that

$$E_j^{(2)}\psi = 0 \quad \text{a.e.} \quad \text{for each } j > 0. \tag{27}$$

In part (c) of the theorem, if the dilations a_j grow exponentially then

$$\begin{aligned}\frac{1}{J}\sum_{j=1}^{J} E_j^{(2)}\psi(x) &= f(x)\frac{1}{J}\sum_{j=1}^{J}(T\psi)(a_j x) \\ &\to 0 \qquad \text{in } L^1 \text{ as } J\to\infty \end{aligned} \tag{28}$$

by Lemma 13 (using that $\psi \in W^{1,1}, D\psi \in H^1$ and $Qf \in L^1$). From (25), (27) and (28) we deduce "L^1 error estimates", namely for part (b) that $E_j\psi \to 0$ in L^1 as $j\to\infty$, and for part (c) that $\frac{1}{J}\sum_{j=1}^{J} E_j\psi \to 0$ in L^1 as $J\to\infty$.

We still need to prove analogous "H^1 error estimates", namely for part (b) that $R_sE_j\psi \to 0$ in L^1 as $j\to\infty$, and for part (c) that $\frac{1}{J}\sum_{j=1}^{J} R_sE_j\psi \to 0$ in L^1 as $J\to\infty$; then the L^1 and H^1 error estimates together with (19) and (22) will prove parts (b) and (c) of the theorem.

We have $R_sE_j\psi = R_s\sum_{k\in\mathbb{Z}^d} F_k = \sum_{k\in\mathbb{Z}^d} R_sF_k$ in L^1, and so

$$(R_sE_j\psi)(x) = |a_j|^{-d}|\det b|\sum_{k\in\mathbb{Z}^d}\int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)R_s[\psi_{j,k}-\psi_{j,b^{-1}y}](x)\,dy,$$

where the validity of taking R_s through the integral defining F_k is justified by Lemma 10 (cf. the proof of part (a) above). Now the translation and dilation invariance of the Riesz transform gives

$$\begin{aligned}(R_sE_j\psi)(x) &= |a_j|^{-d}|\det b|\sum_{k\in\mathbb{Z}^d}\int_{\mathbb{R}^d} f(a_j^{-1}y)\phi(y-bk)\,\mathrm{sign}(a_j)[(R_s\psi)_{j,k}(x)-(R_s\psi)_{j,b^{-1}y}(x)]\,dy \\ &= \mathrm{sign}(a_j)(E_jR_s\psi)(x) \\ &= \mathrm{sign}(a_j)[(E_j^{(1)}R_s\psi)(x)+(E_j^{(2)}R_s\psi)(x)].\end{aligned}$$

We find that $\mathrm{sign}(a_j)E_j^{(1)}R_s\psi \to 0$ in L^1 as $j\to\infty$, by modifying the earlier proof for $E_j^{(1)}\psi$ (and remembering that $SR_s\psi \in L^1$ by construction).

Next we consider $\mathrm{sign}(a_j)E_j^{(2)}R_s\psi$. For part (b), if $P\psi = 1$ a.e. then $TR_s\psi = 0$ a.e. by Lemma 14, so that $E_j^{(2)}R_s\psi = 0$ for all j by replacing ψ with $R_s\psi$ in (26). For part (c), if the dilations a_j grow exponentially then $\frac{1}{J}\sum_{j=1}^{J}\mathrm{sign}(a_j)E_j^{(2)}R_s\psi \to 0$ in L^1 as $J\to\infty$, simply by modifying the proof above (and splitting the sum over j into two parts, treating the terms with $a_j>0$ and those with $a_j<0$ separately). This completes the proof of the H^1 error estimates and hence of Theorem 1.

Remark. Our corresponding L^1-approximation result in [2, Theorem 1] holds under less restrictive assumptions on ψ and ϕ, namely $\psi\in L^1$ and $P|\phi|\in L^\infty$. The proof uses a different decomposition of f_j, in [2, (5.6)]. The cancellation property $\hat{f}(0)=0$ of functions in H^1 requires us, in this paper, to use the more elaborate decompositions (18), (19), (23) and (24), which then necessitate stronger assumptions on ψ and ϕ in our H^1 results.

9. **Proof of Corollary 2**

First, $\theta^{(t)}(x) = \psi(x)-\psi(x-be_t) = \Delta_{be_t}\psi(x)$ is integrable and belongs to H^1 by Lemma 6, since $\|\Delta_z\psi\|_{H^1}$ is bounded for all small z by the hypothesis that $\|\Delta_z\psi\|_{H^1}\to 0$ as $z\to 0$.

We can assume the dilations a_j grow exponentially, after passing to a subsequence if necessary. And we may normalize ψ by $\int_{\mathbb{R}^d}\psi\,dx = 1$, since multiplying ψ by a nonzero constant does not affect the span of the $\theta_{j,k}^{(t)}$.

We first prove the corollary in one dimension, and then sketch the extension to higher dimensions.

Suppose $b > 0$ (the case $b < 0$ being similar). Take $\phi = b^{-1}\mathbb{1}_{(-b,0]}$, so that ϕ is bounded with compact support and $P\phi \equiv 1$. Consider $f \in H^1$ with $(1 + |\cdot|)f \in L^1$. Such functions are dense in H^1; in fact the Schwartz functions in H^1 are already dense by [10, p. 231].

The series (3) defining f_j converges in L^1 by Theorem 1(a), with

$$f_j(x) = \sum_{k\in\mathbb{Z}} \left(\int_{a_j^{-1}b(k-1,k]} f(y)\,dy \right) \psi_{j,k}(x) \qquad \text{by putting } y \mapsto a_j y \text{ in (3)} \tag{29}$$

$$= \sum_{k\in\mathbb{Z}} \left(\int_{(-\infty,a_j^{-1}bk]} f(y)\,dy \right) \theta_{j,k}(x), \tag{30}$$

where one recovers (29) by substituting $\theta(x) = \psi(x) - \psi(x-b)$ into (30), noting that the coefficient sequence belongs to ℓ^1:

$$\sum_{k\in\mathbb{Z}} \left| \int_{(-\infty,a_j^{-1}bk]} f(y)\,dy \right| \le C\|(1 + |a_j x|)f\|_1 < \infty \qquad \text{for each } j > 0, \tag{31}$$

as we prove below.

The new series (30) for f_j converges not only in L^1 but also in H^1, by the coefficient bound (31) and because $\|\theta_{j,k}\|_{H^1} = \|\theta\|_{H^1}$ is independent of j and k. Hence f_j lies in the H^1-span of $\{\theta_{j,k} : k \in \mathbb{Z}\}$. Theorem 1(c) shows how to approximate f in H^1 using linear combinations of the f_j with $j > 0$, and so f too lies in the H^1-span of the $\theta_{j,k}$. Thus the $\theta_{j,k}$ span H^1.

It remains to prove the coefficient bound (31), which we shall do for all $f \in L^1$ satisfying $(1 + |x|)f \in L^1$ and $\int_{\mathbb{R}} f\,dx = 0$. Suppose $a_j > 0$; the case $a_j < 0$ is similar. We have

$$\begin{aligned}
\sum_{k\le 0} \left| \int_{-\infty}^{a_j^{-1}bk} f(y)\,dy \right| &\le \sum_{k\le 0} (1 + |k|) \int_{a_j^{-1}b(k-1)}^{a_j^{-1}bk} |f(y)|\,dy \\
&\le \sum_{k\le 0} \int_{a_j^{-1}b(k-1)}^{a_j^{-1}bk} (1 + |b^{-1}a_j y|)|f(y)|\,dy \\
&= \int_{-\infty}^{0} (1 + |b^{-1}a_j y|)|f(y)|\,dy.
\end{aligned}$$

A similar estimate holds for the sum over $k > 0$, after substituting $\int_{-\infty}^{a_j^{-1}bk} f(y)\,dy = -\int_{a_j^{-1}bk}^{\infty} f(y)\,dy$ (recalling $\int_{\mathbb{R}} f\,dy = 0$). These two estimates imply (31), completing the proof in one dimension.

In higher dimensions the analogue of (30) fails because its coefficient sequence is generally not in ℓ^1: essentially, we cannot expect f to integrate to zero on every line parallel to a coordinate axis.

A less elegant argument still gives spanning in higher dimensions, as we now show. Take $\phi = |\det b|^{-1}\mathbb{1}_{-b\mathcal{C}}$, so that ϕ is bounded with compact support and $P\phi \equiv 1$. Again consider $f \in H^1$ with $(1+|\cdot|)f \in L^1$. The analogue of series (29) for f_j is $f_j = \sum_{k\in\mathbb{Z}^d} c_{j,k}\psi_{j,k}$ where the coefficients are $c_{j,k} = \int_{a_j^{-1}b(k-\mathcal{C})} f(y)\,dy$. They satisfy $\sum_{k\in\mathbb{Z}^d} |k||c_{j,k}| < \infty$ since $|x|f \in L^1$. Also $\sum_{k\in\mathbb{Z}^d} c_{j,k} = \int_{\mathbb{R}^d} f\,dy = 0$. Hence we can subtract $\psi_{j,0}$ to obtain

$$f_j = \sum_{k\in\mathbb{Z}^d} c_{j,k}[\psi_{j,k} - \psi_{j,0}].$$

By reverse telescoping $\psi_{j,k} - \psi_{j,0}$, it can be expressed as a sum of at most $|k_1| + \cdots + |k_d| = O(|k|)$ functions $\theta_{j,\ell}^{(t)}$, with coefficients ± 1 that depend on the signs of the entries in $k = (k_1, \ldots, k_d)$. For example, when $k = -e_1 + e_2$ we can express

$$\psi_{j,k} - \psi_{j,0} = \psi_{j,-e_1+e_2} - \psi_{j,e_2} + \psi_{j,e_2} - \psi_{j,0} = \theta^{(1)}_{j,-e_1+e_2} - \theta^{(2)}_{j,0}$$

by formula (9) for the $\theta^{(t)}_{j,k}$. Such reverse telescoping yields a formula for f_j in terms of the $\theta^{(t)}_{j,k}$. The coefficient sequence of this formula belongs to ℓ^1 since $\sum_{k\in\mathbb{Z}^d} |k||c_{j,k}| < \infty$. Hence the formula converges to f_j in H^1, leading to the desired spanning result by Theorem 1(c).

Remarks on the one dimensional proof.

1. The factor of $|a_j| \to \infty$ on the righthand side of the coefficient bound (31) suggests "instability" of the representation (30) for f_j. Such instabilities when the generating function θ has integral zero have been known in the L^p setting since at least Strang and Fix [12, p. 827].

2. Formula (30) was derived for $f \in H^1$ with $(1 + |\cdot|)f \in L^1$. One might nonetheless hope it would hold for all $f \in H^1$, so that every H^1 function could be explicitly approximated by our linear combinations of the $\theta_{j,k}$. This seems unlikely with our approach, because the coefficient series on the lefthand side of (31) can diverge, as follows. Choose $F \in C^\infty$ with $F = 0$ on $(-\infty, 1)$ and $F(x) = (x \log x)^{-1}$ for $x \in [2, \infty)$. Then $f(x) = F'(x) = O(1/x^2)$ at infinity and $\int_{\mathbb{R}} f\, dx = 0$, so that $f \in H^1$ (see [11, §III.5.7], [5] or [14]). However for dilations $a_j > 0$ we see

$$\sum_{k\in\mathbb{Z}} \left| \int_{-\infty}^{a_j^{-1}bk} f(y)\, dy \right| \geq \sum_{k\geq 2b^{-1}a_j} (a_j^{-1}bk \log a_j^{-1}bk)^{-1} = \infty.$$

3. One would like an "atomic" representation of the form $f = \sum_{j>0} \sum_{k\in\mathbb{Z}} c_{j,k}\theta_{j,k}$ for each $f \in H^1$, with the coefficients $c_{j,k}$ given explicitly and having ℓ^1-norm comparable to the H^1-norm of f. We do not see how to achieve this with our quasi-interpolants f_j.

Acknowledgments

We thank the Institute for Mathematical Sciences, National University of Singapore, for supporting our participation in the program on "Mathematics and Computation in Imaging Science and Information Processing". We also appreciate the hospitality of the Erwin Schrödinger Institut, Vienna, during the program on "Modern Methods of Time–Frequency Analysis". This paper was improved by suggestions of participants at both programs.

References

[1] C. de Boor and R.-Q. Jia. *Controlled approximation and a characterization of the local approximation order.* Proc. AMS 95:547–553, 1985.

[2] H.-Q. Bui and R. S. Laugesen. *Affine systems that span Lebesgue spaces.* J. Fourier Anal. Appl., to appear.

[3] H.-Q. Bui and R. S. Laugesen. *Spanning and sampling in Lebesgue and Sobolev spaces.* University of Canterbury Research Report UCDMS2004/8, 2004.

[4] R. R. Coifman. *A real-variable characterization of H^p.* Studia Math. 51:269–274, 1974.

[5] R. R. Coifman and G. Weiss. *Extensions of Hardy spaces and their use in analysis.* Bull. AMS. 83:569–645, 1977.

[6] M. Frazier, B. Jawerth and G. Weiss. Littlewood-Paley theory and the study of function spaces. CBMS Reg. Conf. Ser. in Math., No. 79, Amer. Math. Soc., Providence, RI, 1991.

[7] J. E. Gilbert, Y. S. Han, J. A. Hogan, J. D. Lakey, D. Weiland and G. Weiss. *Smooth molecular decompositions of functions and singular integral operators.* Memoirs AMS 156, no. 742, 74 pp., 2002.

[8] E. Hernández and G. Weiss. A first course on wavelets. CRC Press, Boca Raton, Florida, 1996.

[9] Y. Meyer. Wavelets and operators. Cambridge University Press, Cambridge, 1992.

[10] E. M. Stein. Singular integrals and differentiability properties of functions. Princeton University Press, Princeton, New Jersey, 1970.

[11] E. M. Stein. Harmonic analysis, real-variable methods, orthogonality, and oscillatory integrals. Princeton University Press, Princeton, New Jersey, 1993.

[12] G. Strang and G. Fix. A Fourier analysis of the finite element variational method. In: Constructive aspects of functional analysis (ed. G. Greymonat), pp. 793–840. C.I.M.E., 1973.

[13] J.-O. Strömberg. *A modified Franklin system and higher-order spline systems on R^n as unconditional bases for Hardy spaces.* Conference on harmonic analysis in honor of Antoni Zygmund, Vol. II (Chicago, Ill., 1981), 475–494, Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983.

[14] M. H. Taibleson and G. Weiss. *The molecular characterization of certain Hardy spaces.* Representation theorems for Hardy spaces, pp. 67–149, Astérisque, 77, Soc. Math. France, Paris, 1980.

[15] H. Triebel. Theory of functions spaces, Birkhäuser, Basel, 1983.

Department of Mathematics, University of Canterbury, Christchurch 8020, New Zealand,
E-mail address: `Q.Bui@math.canterbury.ac.nz`

Department of Mathematics, University of Illinois, Urbana, IL 61801, U.S.A. ,
E-mail address: `Laugesen@math.uiuc.edu`