Multiresolution Analysis, Haar Bases, and Self-Similar Tilings of R^n

K. Gröchenig and W. R. Madych

Abstract—**Orthonormal bases for $L^2(R^n)$ are constructed that have properties that are similar to those enjoyed by the classical Haar basis for $L^2(R)$. For example, each basis consists of appropriate dilates and translates of a finite collection of "piecewise constant" functions. The construction is based on the notion of multiresolution analysis and reveals an interesting connection between the theory of compactly supported wavelet bases and the theory of self-similar tilings.**

Index Terms—**Multiresolution analysis, multivariate Haar basis, self-similar tilings, wavelets, fractals.**

I. Introduction

Recall that the Haar system on $L^2(R)$ is the collection of functions

$$2^{k/2}\psi(2^k x - j), \qquad j, k \in Z, \tag{1}$$

where

$$\psi(x) = \begin{cases} 1, & \text{if } 0 \le x < 1/2, \\ -1, & \text{if } 1/2 \le x < 1, \\ 0, & \text{otherwise,} \end{cases}$$

where Z denotes the set of integers. Note the role played by the dilation $x \to 2x$ and the translations $x \to x - j$. It is well known that this collection is a complete orthonormal system for $L^2(R)$.

The point of this paper is to construct analogous systems for $L^2(R^n)$, $n \ge 2$, where the dilation noted above is replaced by appropriate linear transformations of R^n and the integers Z are replaced by an appropriate lattice in R^n. The motivation and framework for our construction is outlined in Section II.

We remind the reader that there are obvious generalizations of the Haar basis to higher dimensions. However, the general case is by no means obvious and offers some interesting surprises.

The plan of this paper is as follows. In Section II, we briefly review the concepts of multiresolution analysis and wavelet basis, introduced by Mallat [7] and Meyer [9], for $L^2(R^n)$ and explain how the classical Haar system and our construction fit into this scheme. In short, these bases are simply wavelets whose corresponding scaling functions are characteristic functions of appropriate sets. In Section III, we show how such scaling functions are related to certain self-similar tilings of R^n and indicate how to construct such tilings. Essentially the celebrated two scale functional equations reduce to simple iterated function systems in this case. Thurston [10, Sections 8–10] considers self-similar tilings generated by similarities, that is matrices which are constant multiples of rotations, which do not necessarily preserve some lattice. The construction presented here is different because it requires matrices which leave a lattice invariant but includes many cases that are not similarities. In Section IV, we construct the promised bases from appropriate scaling functions or, equivalently, certain self-similar tilings of R^n. Representative examples in R and R^2 together with several general observations are presented in Section V. We conclude the paper with miscellaneous remarks and citations to the literature.

Manuscript received February 1, 1991; revised July 10, 1991. This work was supported in part by Grant AFOSR-90-0311.

The authors are with the Department of Mathematics U-9, University of Connecticut, Storrs, CT 06269-3009.

IEEE Log Number 9105033.

II. Multiresolution Analysis and Wavelet Bases

In what follows Γ is a lattice in R^n, that is, Γ is the image of the integer lattice Z^n under some nonsingular linear transformation. We say that a linear transformation A on R^n is an *acceptable dilation for* Γ if it satisfies the following properties:

- A leaves Γ invariant. In other words, $A\Gamma \subset \Gamma$. Here
$$A\Gamma = \{y : y = Ax \quad \text{and} \quad x \in \Gamma\};$$
- all the eigenvalues, λ_i, of A satisfy $|\lambda_i| > 1$.

These properties imply that $|\det A|$ is an integer q which is ≥ 2.

Such an A induces a unitary dilation operator $U_A : f \to U_A f$ on $L^2(R^n)$, defined by

$$U_A f(x) = |\det A|^{-1/2} f(A^{-1}x). \tag{2}$$

If V is a subspace of $L^2(R^n)$ we use the customary notation $U_A V$ to denote the image of V under U_A, that is, $U_A V = \{f : f = U_A g,\ g \in V\}$. The translation operator τ_y is defined by $\tau_y f(x) = f(x - y)$.

A *wavelet basis* associated to (Γ, A) is a complete orthonormal basis of $L^2(R^n)$ whose members are A dilates of Γ translates of a finite collection $\psi_1, \cdots, \psi_m$ of orthonormal functions. More specifically, the members of the basis are the functions

$$U_A^{-j}\tau_\gamma\psi_i(x) = |\det A|^{j/2}\psi_i(A^j x - \gamma), \qquad j \in Z,\ \gamma \in \Gamma, \qquad i = 1, \cdots, m. \tag{3}$$

The ψ_i's are called the basic wavelets.

The classical Haar system defined by (1) is the simplest example of such a basis for $L^2(R)$. In this case $m = 1$, $\Gamma = Z$, and A is the dyadic dilation $Ax = 2x$.

There is a generic recipe, due to Y. Meyer, for the construction of wavelet bases. The main ingredient is the notion of multiresolution analysis.

A *multiresolution analysis* $\mathscr{V}$ associated with (Γ, A) is an increasing family $\cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \cdots$, $j \in Z$, of closed subspaces of $L^2(R^n)$ with the following properties:

1) $\cup_{j \in Z} V_j$ is dense in $L^2(R^n)$, and $\cap_j V_j = \{0\}$,
2) $f(x) \in V_j$, if and only if $f(Ax) \in V_{j+1}$. In other words
$$V_j = U_A^{-j} V_0, \qquad j \in Z. \tag{4}$$
3) V_0 is invariant under τ_γ. More specifically, if $f(x)$ is in V_0 then so is $f(x - \gamma)$ for all γ in Γ.
4) There is a function $\phi \in V_0$, called the scaling function, such that $\{\tau_\gamma \phi, \gamma \in \Gamma\}$ is a complete orthonormal basis for V_0.

From the definition it should be clear that a multiresolution analysis is determined by the scaling function ϕ. Since $V_0 \subset V_1$ there is a sequence $\{a_\gamma\}$ in $l^2(\Gamma)$ such that
$$\begin{aligned}\phi(x) &= \sum_{\gamma \in \Gamma} a_\gamma U_A^{-1} \tau_\gamma \phi(x) \\ &= \sum_{\gamma \in \Gamma} a_\gamma |\det A|^{1/2} \phi(Ax - \gamma).\end{aligned} \tag{5}$$
It is known that, under certain conditions, these coefficients determine the scaling function ϕ uniquely. On the other hand, in spite of the fact that the orthogonality relations
$$\langle \tau_\gamma \phi, \phi \rangle = \delta_{\gamma 0} \tag{6}$$
impose certain restrictions on the sequence $\{a_\gamma\}$, these conditions are not sufficient to guarantee that the a_γ's are acceptable scaling coefficients; in short, the nature of the scaling coefficients is not completely understood in the general case.

The simplest example of a multiresolution analysis in one dimension with $(\Gamma, A) = (Z, 2I)$, where I is the identity, is given by the scaling function $\phi(x) = \chi_{[0,1)}(x)$. Then V_0 is the closed subspace of $L^2(R)$ consisting of all functions that are constant on the intervals $[j, j+1)$, $j \in Z$, and the subspaces V_k consist of those functions that are constant on subintervals $[j2^{-k}, (j+1)2^{-k})$, $j \in Z$. The scaling relation is given by
$$\phi(x) = \phi(2x) + \phi(2x - 1). \tag{7}$$
Given a multiresolution analysis $\mathscr{V}$, define W_j as the orthogonal complement of V_j in V_{j+1}, thus $V_j \oplus W_j = V_{j+1}$ for all j. It follows that
$$W_j = U_A^{-j} W_0 \quad \text{and} \quad L^2(R^n) = \bigoplus_{j \in Z} W_j. \tag{8}$$
If $q = |\det A|$ then a result which can be found in Meyer's paper [8] says the following: There exist $q - 1$ functions $\psi_1, \cdots, \psi_{q-1}$ such that $\{\tau_\gamma \psi_i; \gamma \in \Gamma, i = 1, \cdots, q-1\}$ *is a complete orthonormal basis of* W_0.

In view of (8) this theorem implies that the collection $\{U_A^j \tau_\gamma \psi_i : j \in Z, \gamma \in \Gamma, i = 1, \cdots, q-1\}$ is a wavelet basis of $L^2(R^n)$. Furthermore, since $W_0 \subset V_1$ there are sequences $\{a_{i\gamma}\}$ in $l^2(\Gamma)$ such that
$$\psi_i(x) = \sum_{\gamma \in \Gamma} a_{i\gamma} |\det A|^{1/2} \phi(Ax - \gamma), \qquad i = 1, \cdots, q-1. \tag{9}$$
For example, in the specific case of multiresolution analysis mentioned above where the scaling relation is given by (7), $q = 2$ and the corresponding basic wavelet is given by
$$\begin{aligned}\psi(x) &= \phi(2x) - \phi(2x - 1) \\ &= \chi_{[0,1)}(2x) - \chi_{[0,1)}(2x - 1).\end{aligned}$$
This is the basic Haar wavelet.

Thus, the generic recipe to construct a wavelet basis can be briefly summarized as follows: Start with a multiresolution analysis with scaling relation (5) and look for basic wavelets which are of form (9).

Due to the work of Cohen [2], Daubechies [3], Mallat [7], Meyer [9] and others, the algorithm outlined above is well understood in the case when $n = 1$ and $Ax = 2x$. In this case, multiresolution analyses can be constructed that have desired continuity and support properties. The coefficients for the basic wavelet can always be expressed in terms of the original scaling coefficients via a simple formula.

The construction of the basic wavelets in the general case is not so clear. For example, the structure of scaling sequences which will produce multiresolution analyses with desired properties is not well understood. Also, although it is clear that the coefficients of (9) should have some relationship to the coefficients in the basic scaling relation (5), except for certain examples, there are no known formulas for the coefficients in (9) in the general case.

We are now ready to state the questions addressed in this paper precisely: Given (Γ, A), what are the multiresolution analyses whose scaling functions are characteristic functions of measurable sets Q? What are the corresponding basic wavelets whose support is in Q and how can they be constructed explicitly? In what follows we will refer to such multiresolution analyses as *simple* and to such wavelets as *elementary*.

Before we start, let us make the following simplification: Since every lattice $\Gamma \subset R^n$ is of the form $\Gamma = EZ^n$ for some invertible real-valued $n \times n$ matrix E, without loss of generality, we may and do restrict our attention to the case $\Gamma = Z^n$. In this case the matrices A must have integer entries.

III. Self-Similar Tilings and Scaling Functions

In this section, we establish a connection between self-similar tilings and multiresolution analyses that have a characteristic function for a scaling function.

Given a measurable set S, χ_S denotes its characteristic or indicator function and $|S|$ denotes its Lebesgue measure. The notation $S \simeq T$ means that the sets S and T are equal up to a set of measure zero, in other words, $|S \setminus T| = |T$

$\setminus S| = 0$. If $S \cap T \simeq \emptyset$ we say that S and T are essentially disjoint. Also recall that $q = |\det A|$.

We begin with two technical lemmas which are elementary and are probably folklore.

Lemma 1: Suppose Q is a measurable subset of R^n such that

$$\bigcup_{k \in Z^n} (Q + k) \simeq R^n.$$

Then the following are equivalent:

1) $Q \cap (Q + k) \simeq \emptyset$ whenever k is a nonzero element in Z^n,
2) $|Q| = 1$.

Proof: Let $f(x) = \Sigma_{j \in Z^n} \chi_Q(x - j)$. Then if Q satisfies property 1) it follows that $f \equiv 1$ and we may write

$$|Q| = \int_{R^n} \chi_Q(x)\,dx = \int_{Q_0} f(x)\,dx = |Q_0| = 1,$$

where $Q_0 = [0, 1]^n$.

To see the converse let f and Q_0 be as previously defined. Note that assumption implies $f(x) \geq 1$. Also observe that

$$\int_{Q_0} f(x)\,dx = \int_{R^n} \chi_Q(x)\,dx = |Q| = 1$$

implies that $f \equiv 1$, which in turn implies the desired result. □

Lemma 2: The number of disjoint cosets in Z^n/AZ^n is q.

Proof: Let $Q_0 = [0, 1]^n$ and let $k_1 + AZ^n, \cdots, k_m + AZ_n$ be an enumeration of the cosets in Z^n/AZ^n. Express R^n as a union of essentially disjoint subsets as follows

$$\bigcup_{k \in Z^n} \left\{ Ak + \bigcup_{i=1}^{m} (k_i + Q_0) \right\}$$

$$= \bigcup_{k \in Z^n} \left\{ \bigcup_{i=1}^{m} (k_i + Ak + Q_0) \right\} = \bigcup_{k \in Z^n} (k + Q_0) = R^n.$$

Since $A^{-1}R^n = R^n$, applying A^{-1} to both sides of the last equality results in

$$\bigcup_{k \in Z^n} \left\{ k + \bigcup_{i=1}^{m} A^{-1}(k_i + Q_0) \right\} = R^n.$$

It should now be clear that

$$Q = \bigcup_{i=1}^{m} A^{-1}(k_i + Q_0)$$

satisfies the hypothesis and 1) in Lemma 1. Hence, $|Q| = 1$, and, since Q is the union of disjoint subsets $A^{-1}(k_1 + Q_0), \cdots, A^{-1}(k_m + Q_0)$ each of which has measure $1/q$, it follows that $m = q$. □

Theorem 1: Suppose $\phi = c\chi_Q$ is a scaling function for a multiresolution analysis associated with (Z^n, A). Here χ_Q is the characteristic function of a measurable set Q and $c = |Q|^{-1/2}$. Then Q satisfies the following properties.

1)
$$Q \cup (Q + k) \simeq \emptyset, \qquad \text{for } k \neq 0,\ k \in Z^n. \tag{10}$$

2) There is a collection of q lattice points $k_1, \cdots, k_q$ that are representatives of distinct cosets in Z^n/AZ^n such that
$$AQ \simeq \bigcup_{i=1}^{q} (k_i + Q). \tag{11}$$

3)
$$\bigcup_{k \in Z^n} (Q + k) \simeq R^n. \tag{12}$$

4) There is a compact set K such that $Q \simeq K$.

Conversely, the characteristic function of a bounded measurable set Q that satisfies properties 1), 2), and 3) is the scaling function of a multiresolution analysis associated with (Z^n, A).

Remarks:

- Properties (10) and (12) mean that translates of Q by the integer lattice form a tiling of R^n. Sets Q that enjoy property (11) are sometimes said to be self-similar in the affine sense.
- In view of Lemma 1 properties (10) and (12) imply that $|Q| = 1$. Thus $c = 1$, $\phi = \chi_Q$ and satisfies the functional equation
$$\phi(x) = \sum_{i=1}^{q} \phi(Ax - k_i), \tag{13}$$
where the k_i's are those lattice points whose existence is implied by property 2).
- In view of Lemma 2 the k_i's in (11) are a full set of coset representatives, namely,
$$\bigcup_{i=1}^{q} (k_i + AZ^n) = Z^n. \tag{14}$$

Proof: Suppose $\phi = c\chi_Q$ is a scaling function for a multiresolution analysis associated with (Z^n, A).

The disjointness of the translates of Q (10) follows from (6): $\langle L_k\phi, \phi \rangle = c^2 |Q \cap (Q + k)| = \delta_{k0}$.

The second property follows from the scaling relation for χ_Q,

$$\chi_Q(x) = \sum_k a_k q^{1/2} \chi_Q(Ax - k)$$
$$= \sum_k a_k q^{1/2} \chi_{A^{-1}(Q+k)}(x),$$

where $a_{k_i} q^{1/2} = 1$ for exactly q lattice points k_i and $a_k = 0$ for the remaining coefficients. The fact concerning the coefficients a_k, which of course immediately implies (11), follows from (10) and the formula $|AQ| = q|Q|$.

That the k_i's are representatives of distinct cosets of Z^n/AZ^n follows from the orthogonality of $\phi(A^{-1}x - k)$, $k \in Z^n$. To wit, suppose k_1 and k_2 are not in distinct cosets. Then there is a lattice point k so that $k_1 = k_2 + Ak$ and this in turn implies that

$$AQ = \bigcup_{i=1}^{q} (k_i + Q) \quad \text{and}$$

$$A(Q + k) = \bigcup_{i=1}^{q} (k_i + Ak + Q)$$

are not disjoint which contradicts the orthogonality of $\phi(A^{-1}x - k)$ and $\phi(A^{-1}x)$.

The covering property (12) is a consequence of the density of $\bigcup_j V_j$ in $L^2(R^n)$ and 2). To see this, let $P_j f$ be the projection of $f \in L^2(R^n)$ onto V_j, in other words, $P_j f = \sum_k |Q|^{-j} \langle U_A^{-j} \tau_k \chi_Q, f \rangle U_A^{-j} \tau_k \chi_Q$. Then $P_j f \to f$ as $j \to \infty$ and

$$\| P_j f \|_2^2 = \sum_k |Q|^{-j^0} |\langle U_A^{-j} \tau_k \chi_Q, f \rangle|^2 \to \| f \|_2^2 \quad (15)$$

as $j \to \infty$. Hence, if B is any measurable set of finite measure, if follows from (15) that

$$\sum_k |Q|^{-1} q^j | A^{-j}(k + Q) \cap B|^2$$
$$= \sum_k |Q|^{-1} q^{-j} |(k + Q) \cap A^j B|^2 \to |B| \quad (16)$$

as $j \to \infty$. Now, in view of 2),

$$\cup_k (Q + k) = \cup_k (A^{-j}(Q + k)),$$

for all j in Z^n. Hence, we may write

$$|B| \geq |\cup_k (Q + k) \cap B|$$
$$= |\cup_k (A^{-j}(Q + k)) \cap B|$$
$$= \sum_k q^{-j} |(k + Q) \cap A^j B|$$
$$\geq \sum_k q^{-j} |Q| \left(\frac{|(k + Q) \cap A^j B|}{|Q|} \right)^2$$
$$= \sum_k |Q|^{-1} q^{-j} |(k + Q) \cap A^j B|^2.$$

By virtue of (16) the last expression converges to $|B|$ as $j \to \infty$ and, as a consequence, we may conclude that $|B| = |\cup_k (Q + k) \cap B|$ and, since B was arbitrary, the desired result follows.

Property 4) is a consequence of Lemmas 3 and 5.

To see the converse let

$$V_0 = \left\{ f \in L^2(R^n) : f(x) = \sum_{k \in Z^n} c_k \chi_Q(x - k) \right\}$$

and let $V_j = U_A^{-j} V_0$ for each integer j. Then $\mathscr{V} = \{V_j\}_{j \in Z}$ is a family of closed subplaces of $L^2(R^n)$ and using both the above definition and the properties of the set Q it is easy to see that this family satisfies the following properties:

- $\mathscr{V}$ is an increasing family, namely, $V_j \subset V_{j+1}$ for all integers j,
- $\cap_{j \in Z} V_j = \{0\}$,
- $V_j = U_A^{-j} V_0$ for each integer j,
- V_0 is invariant under τ_k for each k in Z^n,
- $\{\tau_k \chi_Q : k \in Z^n\}$ is a complete orthonormal basis for V_0.

In view of this, to see that $\mathscr{V}$ is a multiresolution analysis associated with (Z^n, A) with scaling function χ_Q, it suffices to show that $\cup_{j \in Z} V_j$ is dense in $L^2(R^n)$.

To this end, let $P_j f$ be the orthogonal projection of f onto V_j, let $\phi_j(x) = |\det A|^j \chi_Q(-A^j x)$, and observe that

$$P_j f(x) = \phi_j * f(A^{-j} k), \quad \text{whenever } x - A^{-j} k \in A^{-j} Q, \quad (17)$$

where

$$\phi_j * f(x) = \int_{R^n} \phi_j(x - y) f(y)\, dy$$

is the convolution of ϕ_j and f. Now, it is easy to see that

$$\phi_j * f - f \to 0 \quad \text{in} \quad L^2(R^n) \text{ as } j \to \infty, \quad (18)$$

for all f in $L^2(R^n)$. In view of (17), the difference between $P_j f$ and $\phi_j * f$ may be expressed as

$$E_j(x) = \phi_j * f(x) - P_j f(x)$$
$$= \int_{R_n} \Big\{ f(x - y) - \sum_{k \in Z^n} \chi_{A^{-j}Q}(x - A^{-j} k) f$$
$$\cdot (A^{-j} k - y) \Big\} \phi_j(y)\, dy. \quad (19)$$

If we call the expression in braces $F_j(x, y)$ and take f to be continuous with compact support and $|y| \leq 1$ then for any positive ϵ we may write

$$\int_{R^n} |F_j(x, y)|^2\, dx < \epsilon^2, \quad (20)$$

for sufficiently large j. Hence, in this case, if we take the $L^2(R^n)$ norm of E^j and apply the integral variant of Minkowski's inequality to the right-hand side of (19), inequality (20) implies that

$$\| E_j \|_{L^2(R^n)} < \epsilon,$$

for sufficiently large j. In other words,

$$\phi_j * f - P_j f \to 0 \quad \text{in} \quad L^2(R^n) \quad \text{as } j \to \infty, \quad (21)$$

whenever f is continuous with compact support. Since such f are dense in $L^2(R^n)$, (21) together with (18) imply that $\cup_{j \in Z} V_j$ is dense in $L^2(R^n)$. □

Suppose

$$\phi = c \chi_Q \quad (22)$$

is a scaling function for a multiresolution analysis associated with (Z^n, A). In view of this, we know that $|Q| = c = 1$ and

$$\phi(x) = \sum_{k \in \mathscr{K}} \phi(Ax - k), \quad (23)$$

where $\phi = \chi_Q$ and $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a collection of representatives of distinct cosets in Z^n / AZ^n. Thus, if we are interested in constructing a multiresolution analysis whose scaling function ϕ is of the form (22), a reasonable approach seems to be the following: find an appropriate collection of lattice points $\mathscr{K}$ and solve the functional equation (23).

Since the solution of (23) is a fixed point of the mapping

$$\psi(x) \to \sum_{k \in \mathscr{K}} \psi(Ax - k),$$

it is quite natural to apply fixed point iteration to solve for ϕ. Namely, start with an initial function ϕ_0 and define the sequence $\phi_1, \phi_2, \phi_3, \cdots$ via

$$\phi_{N+1}(x) = \sum_{k \in \mathscr{K}} \phi_N(x) \tag{24}$$

and hope that the sequence converges to ϕ. Since the desired solution is the characteristic function of a set Q whose Z^n translates tile, it is reasonable to begin the iteration with $\phi_0 = \chi_{Q_0}$ where Q_0 has the same properties.

Suppose Q_0 satisfies (10) and (12) and $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a collection of representatives of distinct cosets of Z^n/AZ^n. If $\phi_0 = \chi_{Q_0}$ and ϕ_1 is related to ϕ_0 via (24) then $\phi_1 = \chi_{Q_1}$ where

$$Q_1 = \bigcup_{k \in \mathscr{K}} A^{-1}(Q_0 + k)$$

also satisfies (10) and (12); that Q_1, satisfies (12) follows from the fact that $\mathscr{K}$ is a full collection of distinct representatives of Z^n/AZ^n and that it satisfies (10) follows from $|Q_1| = 1$. By induction we may conclude that ϕ_{N+1}, $N = 0, 1, 2, \cdots$, is the characteristic function of the set Q_{N+1} defined by

$$Q_{N+1} = \bigcup_{k \in \mathscr{K}} A^{-1}(Q_N + k). \tag{25}$$

Observe that (25) looks like an iterated function system in the sense of Barnsley [1]. Convergence of schemes like (25) is usually considered in terms of the following metric defined on the space of subsets of R^n:

$$\rho(P, Q) = \max\{r(P, Q), r(Q, P)\} \tag{26}$$

where

$$r(P, Q) = \sup_{x \in P} \inf_{y \in Q} |x - y|.$$

It is well known that when equipped with the metric ρ the class of compact subsets of R^n is a complete metric space. If the mapping $x \to A^{-1}x$ is a contraction and Q_0 is compact then the iteration (25) converges to a compact set Q; for example, see [1].

Unfortunately, in our considerations the mapping $x \to A^{-1}x$ is not necessarily a contraction, see Example 1) in Section V-C. Nevertheless since all the eigenvalues of A^{-1} are less than one in absolute value it follows that

$$\|A^{-j}x\| \le C\lambda^j j^s \|x\|, \tag{27}$$

for all $x \in R^n$, where C, s, λ are positive constants and $\lambda < 1$. This is easily seen by writing A in its Jordan form. Inequality (27) allows us to state the following.

Lemma 3: Suppose $\mathscr{K} = \{k_1, \cdots, k_m\}$ is a finite collection of lattice points in Z^n and Q is the compact set defined by

$$Q = \left\{x \in R^n : x = \sum_{j=1}^{\infty} A^{-j}\epsilon_j,\ \epsilon_j \in \mathscr{K}\right\}. \tag{28}$$

If Q_0 is any compact set then the sequence of sets Q_1, $Q_2, \cdots$, defined by

$$Q_{N+1} = \bigcup_{i=1}^{m} A^{-1}(k_i + Q_N) \tag{29}$$

converges in the metric ρ to the set Q.

Remarks:

- Note that Q is well defined and bounded by virtue of (27).
- It follows immediately from the definition that the set Q satisfies the self-similarity relation

$$Q = \bigcup_{i=1}^{m} A^{-1}(k_i + Q).$$

Proof: Using (25) repeatedly, we see that

$$\begin{aligned} Q_N &= \bigcup_{i=1}^{m} (A^{-1}k_i + A^{-1}Q_{N-1}) \\ &= \bigcup_{i=1}^{m} \bigcup_{j=1}^{m} (A^{-1}k_i + A^{-2}k_j + A^{-2}Q_{N-2}) \\ &= \cdots = \bigcup_{(\epsilon_1, \epsilon_2, \cdots, \epsilon_N) \in \mathscr{K}^N} \left(\sum_{j=1}^{N} A^{-j}\epsilon_j + A^{-N}Q_0\right), \end{aligned} \tag{30}$$

where $\mathscr{K}^N$ is the collection of all N tuples $(\epsilon_1, \epsilon_2, \cdots, \epsilon_N)$ whose components are in $\mathscr{K}$. Hence, if x is any element in Q_N then there is a point x_0 in Q_0 such that $x = \sum_{j=1}^{N} A^{-j}\epsilon_j + A^{-N}x_0$. So by choosing an element y in Q of the form $y = \sum_{j=1}^{N} A^{-j}\epsilon_j + A^{-N}y_0$ where $y_0 \in Q$ we may write

$$\begin{aligned} \inf_{y \in Q} |x - y| &\le \inf_{y_0 \in Q} |A^{-N}(x_0 - y_0)| \\ &\le C\lambda^N N^s \inf_{y_0 \in Q} |x_0 - y_0| \\ &\le C\lambda^N N^s r(Q_0, Q), \end{aligned} \tag{31}$$

where the inequalities follow by virtue of (27) and the definition of $r(Q_0, Q)$. Taking supremum over $x \in Q_N$ shows that $r(Q_N, Q)$ can be made arbitrarily small by choosing N sufficiently large. Similar reasoning shows that $r(Q, Q_N)$ can also be made arbitrarily small by choosing N sufficiently large. It now follows that the sequence of compact sets $\{Q_N\}$ converges to Q in the sense of the metric ρ and, as a consequence, Q is compact. □

In what follows, we will often be considering collections $\mathscr{K} = \{k_1, \cdots, k_m\}$ of lattice points in conjunction with the scaling relation (13) or the self-similarity (11). In this context it is a minor inconvenience if 0 is not in $\mathscr{K}$. For example, the set defined by (28) does not contain finite sums of the specified form. In this particular case this inconvenience can be remedied by re-expressing the elements x as follows:

$$x = (A - I)^{-1}k_1 + \sum_{j=1}^{\infty} A^{-j}(\epsilon_j + k_1),$$

where the ϵ_j's are in $\mathscr{K}$. In the arguments used below there is no loss of generality in assuming that $\mathscr{K}$ contains 0 in view of the following easily verifiable lemma.

Lemma 4: Suppose $\mathscr{K}_1$ is a collection of lattice points and $\mathscr{K}_2 = x_0 + \mathscr{K}_1$ for some point x_0 in R^n. If $\phi(x)$ satisfies

$$\phi(x) = \sum_{k \in \mathscr{K}_1} \phi(Ax - k),$$

then $\psi(x) = \phi(x - (A - I)^{-1}x_0)$ satisfies

$$\psi(x) = \sum_{k \in \mathscr{K}_2} \psi(Ax - k).$$

Similarly, if Q satisfies (11) where the union is taken over k_i in $\mathscr{K}_1$ then $Q + (A - I)^{-1}x_0$ satisfies (11) where the union is taken over k_i in $\mathscr{K}_2$.

Next we define two sequences of useful measures. If $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a collection of lattice points in Z^n that contains 0 then the sequences of measures $\{\mu_N\}$ and $\{\nu_N\}$, $N = 1, 2, \cdots$, are defined as follows:

$$\mu_N(x) = \frac{1}{q} \sum_{k \in \mathscr{K}} \delta(x - A^{-N}k), \tag{32}$$

where $\delta(x)$ is the unit Dirac measure at the origin and

$$\nu_{N+1} = \mu_{N+1} * \nu_N, \tag{33}$$

where $\nu_1 = \mu_1$. Note that the support of ν_N is the finite collection of points

$$\mathscr{Q}_N = \left\{ x \in R^N : x = \sum_{j=1}^{N} A^{-j}\epsilon_j,\ \epsilon_j \in \mathscr{K} \right\}. \tag{34}$$

Note that $\{\mathscr{Q}_N\}$, $N = 0, 1, 2, \cdots$, is the sequence of sets generated via (25) with $Q_0 = \{0\}$.

Suppose $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a full collection of representatives of distinct cosets of Z^n/AZ^n. To avoid needless repetition of these phrases we say that such a collection is a *full collection of digits*. Of course we refer to elements of such a set $\mathscr{K}$ as *digits*.

Lemma 5: Suppose $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a collection of q distinct lattice points in Z^n. Then any integrable solution of

$$\phi(x) = \sum_{k \in \mathscr{K}} \phi(Ax - k) \tag{35}$$

is unique up to multiplication by a constant and has support in the compact set

$$Q = \left\{ x \in R^n : x = \sum_{j=1}^{\infty} A^{-j}\epsilon_j,\ \epsilon_j \in \mathscr{K} \right\}.$$

Proof: Without loss of generality, assume $\mathscr{K}$ contains zero. Suppose ϕ is an integrable solution of (35). Note that (35) may be re-expressed as

$$\phi(x) = \mu_1 * \phi_1(x) = \int \phi_1(x - y)\, d\mu_1(y), \tag{36}$$

where $\phi_N(x) = q^N \phi(A^N x)$ and μ_1 is defined by (32). By induction it follows that for $N = 1, 2, \cdots$,

$$\phi(x) = \nu_N * \phi_N(x), \tag{37}$$

where ν_N is defined by (33). Now, if ψ is in $L^p(R^n)$, $1 \le p < \infty$, $\phi_N * \psi$ converges to $c\psi$ in $L^p(R^n)$ as $N \to \infty$ where $c = \int_{R^n} \phi(x)dx$. Since for each N ν_N has total variation one it follows that $\nu_N * \phi_N * \psi - c\nu_N * \psi$ converges to zero in $L^p(R^n)$ as $N \to \infty$. Since for any $N = 1, 2, \cdots, \phi = \nu_N * \phi_N = c\nu_N + (\nu_N * \phi_N + c\nu_N)$ and $\nu_N * \phi_N - c\nu_N$ converges weakly to zero we may conclude that $c\nu_N$ converges weakly to ϕ as $N \to \infty$. Thus the support of ϕ must be in Q. Since the sequence μ_N is uniquely defined by the functional equation (35) and it converges weakly to a constant multiple of ϕ it follows that ϕ is unique up to multiplication by constants. □

Theorem 2: Suppose $\mathscr{K} = \{k_1, \cdots, k_q\}$ is a full collection of digits and suppose the compact set Q is defined by

$$Q = \left\{ x \in R^N : x = \sum_{j=1}^{\infty} A^{-j}\epsilon_j,\ \epsilon_j \in \mathscr{K} \right\}. \tag{38}$$

Then the set Q has the following properties.

1) If Q_0 is any compact set then the sequence of sets $Q_1, Q_2, \cdots$, defined by (25) converges in the metric ρ to Q.
2) $AQ \simeq \cup_{i=1}^{q}(k_i + Q)$.
3) $\cup_{k \in Z^n}(Q + k) \simeq R^n$.
4) $(Q + k_i) \cap (Q + k_j) \simeq \emptyset$ wherever both k_i and k_j are in $\mathscr{K}$ and $k_i \ne k_j$.
5) $\phi = |Q|^{-1/2}\chi_Q$ is the unique solution, in the $L^1(R^n)$ sense, of

$$\phi(x) = \sum_{k \in \mathscr{K}} \phi(Ax - k),$$

which has $L^2(R^n)$ norm 1.
6) The sequence of measures $\{\nu_N\}$, $N = 0, 1, 2, \cdots$, defined by (33) converges weakly to $|Q|^{-1}\chi_Q$. In other words,

$$\lim_{N \to \infty} \int_{R^n} \psi(x)\, d\nu_N(x) = \frac{1}{|Q|} \int_{R^n} \psi(x)\chi_Q(x)\, dx,$$

for all functions ψ that are continuous and bounded on R^n.
7) If $\{Q_N\}$, $N = 0,1,2,\cdots$, is the sequence of sets generated via (25) with $Q_0 = [-1/2, 1/2]^n$, then the corresponding sequence of characteristic functions $\{\chi_{Q_N}\}$ converges weakly to $|Q|^{-1}\chi_Q$. In other words,

$$\lim_{N \to \infty} \int_{R^n} \psi(x)\chi_{Q_n}(x) d(x) = \frac{1}{|Q|} \int_{R^n} \psi(x)\chi_Q(x)\, dx,$$

for all functions ψ that are continuous and bounded on R^n.

Remarks:

- note that Property 3) implies that $|Q| \geq 1$;
- as will be clear from the proof, Property 7) holds if $Q_0 = [-1/2, 1/2]^n$ is replaced by any compact set Q_0 that satisfies (10) and (12).

Proof: The first two assertions are simply implied by Lemma 3.

To see item 3) let $Q_0 = [-1/2, 1/2]^n$ and observe that for each N, $N = 0, 1, 2, \cdots, Q_N$ satisfies (12) by virtue of the fact that $\mathscr{K}$ is a full collection of representatives of distinct cosets of Z^n/AZ^n, namely, $Z^n = \cup_{k \in \mathscr{K}}(k + AZ^n)$. This implies that for every x in R^n there is a sequence of lattice points $\{m_N\}$ such that $x - m_N$ is in Q_N. Since the Q_N's are all contained in a fixed ball, $\{m_N\}$ is a bounded sequence as well, and thus, contains a constant subsequence $m_{N_j} = m$, $j = 1, 2, \cdots$. Finally, since $x - m \in Q_{N_j}$ for infinitely many indexes, it follows that $x - m$ is in Q by virtue of the fact that $\{Q_N\}$ converges to Q in the ρ metric.

Since 3) implies that $|Q| \geq 1$, item 4) follows from the identity

$$\left| \bigcup_{k \in \mathscr{K}} (k + Q) \right| = q|Q|,$$

which is implied by 2).

That χ_Q satisfies the identity in item 5) follows immediately from 2) and 4). The fact concerning uniqueness is the assertion of Lemma 5.

To see item 6) let ψ be any bounded continuous function on R^n. Write

$$\begin{aligned} \int_{R^n} \psi(x)\, d\nu_N(x) &= \frac{1}{q^N} \sum_{x \in \mathscr{Q}_N} \psi(x) \\ &= \frac{1}{|Q|} \left\{ |A^{-N}Q| \sum_{x \in \mathscr{Q}_N} \psi(x) \right\} \\ &= \frac{1}{|Q|} \int_{R^n} \psi_N(x)\, dx, \end{aligned}$$

where

$$\psi_N(x) = \sum_{y \in \mathscr{Q}_N} \psi(y) \chi_Q(A^N(x - y))$$

is a simple function which converges to $\psi(x)\chi_Q(x)$ almost everywhere and is dominated by a constant multiple of χ_Q. The dominated convergence theorem now implies the desired result.

Item 7) follows from an argument analogous to the one used to show 6). □

It should now be clear how to construct scaling functions ϕ for multiresolution analyses associated with (Z^n, A), which have characteristic functions as scaling functions.

- Start with compact set Q_0 and a full collection $\mathscr{K} = (k_1, \cdots, k_q\}$ of representatives of distinct cosets of Z^n/AZ^n.
- Find Q as the limit of the iteration (25).
- If $|Q| = 1$ the algorithm is successful and $\phi = \chi_Q$ is the scaling function of a multiresolution analysis associated with (Z^n, A). Otherwise the algorithm fails.

The reason one must check the condition $|Q| = 1$ is that the requirement that $\mathscr{K}$ be a full collection of representatives of distinct cosets of Z^n is a necessary but not sufficient condition on this set of indices. Indeed, examples show that Q need not satisfy this condition and the algorithm may fail.

The following theorem gives various equivalent conditions that guarantee that this algorithm be successful.

Theorem 3: Suppose $\mathscr{K} = (k_1, \cdots, k_q\}$ is a collection of representatives of distinct cosets of Z^n/AZ^n and the compact set Q is defined by (38). Then the following statements are equivalent:

1) χ_Q is a scaling function for a multiresolution analysis associated with (Z^n, A).
2) $|Q| = 1$.
3) $Q \cap (k + Q) \simeq \emptyset$ for all k in Z^n which are different from 0.
4) The sequence of measures $\{\nu_N\}$, $N = 0, 1, 2, \cdots$, defined by (33) converges weakly to χ_Q. In other words,

$$\lim_{N \to \infty} \int_{R^n} \psi(x)\, d\nu_N(x) = \int_{R^n} \psi(x)\chi_Q(x)\, dx,$$

for all functions ψ that are continuous and bounded on R^n.

5) If $\{Q_N\}$, $N = 0, 1, 2, \cdots$, is the sequence of sets generated via (25) with $Q_0 = [0, 1]^n$, then the corresponding sequence of characteristic functions $\{\chi_{Q_N}\}$ converges to χ_Q in measure.
6) (Cohen's condition) Let $\hat{\mu}$ be the Fourier transform of the measure $\mu = \mu_1$ where μ_1 is defined by (32), namely

$$\hat{\mu}(\xi) = \frac{1}{q} \sum_{k \in \mathscr{K}} e^{-i\langle k, \xi \rangle}. \tag{39}$$

There exists a compact set K that contains a neighborhood of the origin and which satisfies

- $\cup_{k \in Z^n}(2\pi k + K) = R^n$,
- $K \cap (2\pi k + K) \simeq \emptyset$, whenever $k \neq 0$,

such that if $B = A^*$ then

$$|\hat{\mu}(B^{-j}\xi)| > 0 \tag{40}$$

holds for all $\xi \in K$ and $j \geq 1$.

Proof: That items 1), 2), and 3) are equivalent is essentially the content of Lemma 1 and Theorem 1.

That items 2) and 4) are equivalent follows immediately from item 6) of Theorem 2.

To see that 2) and 5) are equivalent observe that $|Q_n| = 1$, $n = 1, 2, \cdots$, and by virtue of dominated convergence

$$\int_{R^n} \chi_{Q_N}(x)\,dx = \int_{R^n} \chi_Q(x)\,dx,$$

so it is quite clear that 5) implies 2).

To see that 2) implies 5), let ϵ be any positive number and observe that the regularity of Lebesgue measure and the compactness of Q imply that there is a positive δ such that $|Q + B_\delta| < \epsilon$, where $B_\delta = \{x \in R^n: |x| < \delta\}$. Since $\{Q_N\}$ converges to Q in the ρ metric it follows that for N sufficiently large $Q_N \subseteq Q + B_\delta$. In this case $Q_N \cup Q \subseteq Q + B_\delta$ so

$$|Q_N \cup Q| \le |Q + B_\delta| < 1 + \epsilon.$$

This and the fact that $|Q_N| = |Q| = 1$ gives

$$\begin{aligned}|Q_N \cup Q| &= |Q_N| + |Q| - |Q_N \cap Q| \\ &= 2 - |Q_N \cap Q| < 1 + \epsilon\end{aligned}$$

so $|Q_N \cap Q| > 1 - \epsilon$. Thus, we may conclude that

$$|Q_N \Delta Q| = |Q_N \cup Q| - |Q_N \cap Q| < 2\epsilon,$$

whenever N is sufficiently large. Since $|\chi_Q - \chi_{Q_N}| = \chi_{Q_N \Delta Q}$, the last inequality implies the desired result.

To see how $\hat{\mu}$ fits in let $\{Q_N\}$, $N = 0, 1, 2, \cdots$, be the sequence of sets generated by (25) with $Q_0 = [1/2, 1/2]^n$ and let $\phi_N = \chi_{Q_n}$, $N = 0, 1, 2, \cdots$, and $\phi = |Q|^{-1}\chi_Q$. Recall that the Fourier transform $\phi \to \hat{\phi}$ is defined by

$$\hat{\phi}(\xi) = \int_{R^n} \phi(x) e^{-i(\xi, x)}\,dx.$$

Let ν_N be the measure defined by (33) and observe that

$$\hat{\nu}_N(\xi) = \prod_{j=1}^{N} \hat{\mu}(B^{-j}\xi) \quad \text{and}$$

$$\hat{\phi}_N(\xi) = \hat{\nu}_N(\xi)\hat{\phi}_0(B^{-N}\xi).$$

In view of items 6) and 7) of Theorem 2 we see that

$$\lim_{N\to\infty} \hat{\nu}_N(\xi) = \hat{\phi}(\xi) \quad \text{and} \quad \lim_{N\to\infty} \hat{\phi}_N(\xi) = \hat{\phi}(\xi),$$

for all ξ in R^n and

$$\hat{\nu}_N(\xi)\hat{\phi}(B^{-N}\xi) = \hat{\phi}(\xi), \tag{41}$$

for $N = 1, 2, \cdots$.

Now, to see that item 1) implies 6) let $\phi = \chi_Q$ and recall that the fact that the family $\{\phi(x - k)\}_{k \in Z^n}$ is orthonormal implies that

$$\hat{\phi}(0) = 1 \quad \text{and} \quad \sum_{k \in Z^n} |\hat{\phi}(\xi - 2\pi k)|^2 = 1,$$

for almost all ξ in R^n. From this and the smoothness properties of $\hat{\phi}$ it follows that there is a finite subset $\mathscr{L}$ of Z^n such that

$$\sum_{k \in \mathscr{L}} |\hat{\phi}(\xi - 2\pi k)|^2 > 1/2,$$

for all ξ in the cube $\Omega = [-\pi, \pi]^n$. The last inequality implies that if N is the number of elements in $\mathscr{L}$ then for each ξ in Ω there is an open cube Ω_ξ and a lattice point k_ξ in $\mathscr{L}$ such that

$$|\hat{\phi}(\eta - 2\pi k_\xi)|^2 > \frac{1}{2N},$$

for all η in Ω_ξ. Let $\Omega_{\xi_0}, \Omega_{\xi_1}, \cdots, \Omega_{\epsilon_m}$ be any finite subcollection of these cubes which covers Ω and such that $\Omega_{\xi_0} = \Omega_0$ is a cube centered at 0. Finally, let K_0 be the closure of $\Omega_{\xi_0} \cap \Omega$ and let K_j be the closure of $(\Omega_{\xi_j} \cap \Omega) \setminus (\cup_{l=0}^{j-1} K_l)$. It is clear that $K = \cup_{j=0}^{m}(k_{\xi_j} + K_j)$ does the desired job by virtue of (41).

To complete the proof we show that item 6) implies 2).

Let $\phi = |Q|^{-1}\chi_Q$. Since $\|\phi\|_{L^2(\mathbf{R}^n)} = |Q|^{-1/2}$, to see the desired result it suffices to show that $\|\phi\|_{L^2(\mathbf{R}^n)} = 1$. Since $\|\phi_N\|_{L^2(R^n)} = 1$, $N = 1, 2, \cdots$, Plancherel's formula will imply the desired result if

$$\lim_{N\to\infty} \int_{R^n} |\hat{\phi}_N(\xi)|^2\,d\xi = \int_{R^n} |\hat{\phi}(\xi)|^2\,d\xi. \tag{42}$$

This is where the set K comes in. First observe that

$$\sum_{k \in Z^n} |\hat{\phi}_0(\xi - 2\pi k)|^2 = 1,$$

so, by virtue of the fact that $\hat{\nu}_N$ is $2\pi B^N Z^n$ periodic, setting $\Omega = [-\pi, \pi]^n$ we may write

$$\begin{aligned}\int_{R^n} &|\hat{\phi}_N(\xi)|^2\,d\xi \\ &= \int_{R^n} |\hat{\nu}_N(\xi)\phi_0(B^{-N}\xi)|^2\,d\xi \\ &= \sum_{k \in Z^n} \int_{B^N\Omega} |\hat{\nu}_N(\xi)|^2 |\phi_0(B^{-N}\xi - 2\pi k)|^2\,d\xi \\ &= \int_{B^N\Omega} |\hat{\nu}_N(\xi)|^2\,d\xi = \int_{B^N K} |\hat{\nu}_N(\xi)|^2\,d\xi \\ &= \int_{R^n} |\hat{\nu}_N(\xi)\chi_{B^N K}(\xi)|^2\,d\xi\end{aligned}$$

or, more briefly,

$$\int_{R^n} |\hat{\phi}_N(\xi)|^2\,d\xi = \int_{R^n} |\hat{\nu}_N(\xi)\chi_{B^N K}(\xi)|^2\,d\xi. \tag{43}$$

Next observe that there is a positive constant C such that

$$|\hat{\phi}(\xi)| > C, \qquad \text{for all } \xi \text{ in } K. \tag{44}$$

(If this were not the case then $\hat{\phi}(\xi) = 0$ for some ξ in K. The hypothesis and (41) then imply that $\hat{\phi}(B^{-N}\xi) = 0$ for $N = 1, 2, \cdots$, contradicting the fact that $\hat{\phi}(0) = 1$.) Inequality (44) implies that $|\hat{\phi}(B^{-N}\xi)| > C$ for all ξ in $B^N K$ so by virtue of (41), we may write

$$|\hat{\nu}_N(\xi)\chi_{B^N k}(\xi)| \le C^{-1}|\phi(\xi)|. \tag{45}$$

Finally, since K contains a neighborhood of 0, it follows that

$$\lim_{N\to\infty} \hat{\nu}_N(\xi)\chi_{B^N k}(\xi) = \hat{\phi}(\xi)$$

for all ξ in R^n so that (45) implies that

$$\lim_{N\to\infty} \int_{R^n} |\hat{\nu}_N(\xi)\chi_{B^{N}K}(\xi)|^2\, d\xi = \int_{R^n} |\hat{\phi}(\xi)|^2\, d\xi$$

by virtue of dominated convergence. In view of (43), this gives the desired result. □

IV. Wavelet Bases of Haar Type

To construct a piecewise constant wavelet basis associated with (Z^n, A) we use the results of Section III and follow Meyer's recipe outlined in Section II.

First let Q be any set satisfying items 1)–4) of Theorem 1. Then χ_Q is a scaling function for a multiresolution analysis $\mathscr{V} = \{\cdots, V_0, V_1, \cdots\}$ associated with (Z^n, A). Next we need to identify the subspace W_0, the orthogonal complement of V_0 in V_1. Since V_0 is the collection of all functions of the form

$$f(x) = \sum_{k\in Z^n} a_k \chi_Q(x-k),$$

where $\{a_k\}$ is in $l^2(Z^n)$, it is not difficult to see the following.

Lemma 6: W_0 is the collection of all functions of the form

$$f(x) = \sum_{k\in Z^n} c_k q^{1/2} \chi_Q(Ax-k), \tag{46}$$

where the sequence of coefficients $\{c_k\}$ is in $l^2(R^n)$ and satisfies

$$\sum_{k\in\mathscr{K}} c_{k+al} = 0, \qquad \text{for all } l \text{ in } Z^n. \tag{47}$$

Here $\mathscr{K} = \{k_1, \cdots, k_q\}$ is the collection of coset representatives appearing in the self-similarity relation (11) for Q.

From (46) and (47) it is not difficult to construct a collection of $q-1$ basic wavelets whose existence is guaranteed by Meyer's result. Indeed one can easily verify the following.

Lemma 7: Suppose $U = (u_{ij})$ is a unitary matrix $q\times q$ matrix whose first row is constant, namely $u_{1j} = q^{-1/2}$, $j = 1,\cdots,q$. Let $\mathscr{K} = \{k_1,\cdots,k_q\}$ be the collection of coset representatives as in Lemma 6. Then the collection of functions $\{\psi_1,\cdots,\psi_{q-1}\}$ defined by

$$\psi_{i-1}(x) = \sum_{j=1}^{q} u_{ij} q^{1/2} \chi_Q(Ax - k_j) \qquad i = 2,\cdots,q \tag{48}$$

is a collection of elementary basic wavelets corresponding to the simple multiresolution analysis associated with (Z^n, A) whose scaling function is χ_Q. In other words, the support of ψ_i is contained in Q, $i = 1,\cdots,q-1$, and the collection $\{\tau_k\psi_i,\ k\in Z^n,\ i = 1,\cdots,q-1\}$ is a complete orthonormal system for W_0.

Conversely, every collection of elementary basic wavelets that arises from the multiscale analysis associated with (Z^n, A) whose scaling function is χ_Q is of form (48).

In the case $q = 2$, there is essentially only one matrix $U = (u_{ij})$ that satisfies the property described in the lemma. Namely, $u_{11} = u_{12} = 1/\sqrt{2}$ and $u_{21} = -u_{22} = \pm 1/\sqrt{2}$. When $q \ge 3$ there are many such examples. Specifically, we mention the case

$$u_{ij} = \frac{1}{q^{1/2}} \cos\frac{(i-1)(j-1)2\pi}{q},$$

$i = 1,\cdots,q$ and $j = 1,\cdots,q$.

As a corollary we state the following.

Theorem 4: Suppose $\psi_1,\cdots,\psi_{q-1}$ is the collection of functions defined in Lemma 7. Then the system

$$U_A^{-j}\tau_k\psi_i \qquad j\in Z,\ k\in Z^n,\ i = 1,\cdots,q-1$$

is a complete orthonormal basis for $L^2(R^n)$.

V. Examples

A. Generalities

Numerical experiments lead to various observations. These include the following.

- Fixing the dilation matrix A but varying the choice of digits can result in wildly varying Q's. Some cases appear totally unrelated while others appear to be some sort of dilates of each other.
- Certain choices of dilation matrix A can give rise to Q's that are simple parallelepipeds when an appropriate choice of digits $\mathscr{K}$ is used. Other choices of A never give rise to such simple Q's; the corresponding Q's are always "fractals."

The following proposition sheds some light on the first item.

Lemma 8: Suppose $\mathscr{K}_1$ and $\mathscr{K}_2$ are two full sets of digits for A. Let Q_1 and Q_2 be the self-similar sets satisfying (11), with $\mathscr{K} = \mathscr{K}_1$ and $\mathscr{K}_2$, respectively. If there is a linear transformation B that commutes with A and such that $\mathscr{K}_2 = B\mathscr{K}_1$ then $Q_2 = BQ_1$.

Remark: Examples show that the hypothesis that B commutes with A is essential for the conclusion.

Proof: Write

$$ABQ_1 = BAQ_1 = B\Big\{\bigcup_{k\in\mathscr{K}_1}(k+Q_1)\Big\} = \bigcup_{k\in\mathscr{K}_1}(Bk + BQ_1) = \bigcup_{k\in\mathscr{K}_2}(k + BQ_1).$$

Since the solution of (11) is unique the last string of equalities implies the desired result. □

Concerning the second item, here is a characterization of dilation matrices A that can give rise to self-similar sets Q, which are simple parallelepipeds with the appropriate choice of digits $\mathscr{K}$.

Lemma 9: The self-similar tile Q resulting as the limit of the iteration (25) can be a parallelepiped, if and only if the dilation factorizes as $A = CDPC^{-1}$. Here C is an integer-valued, invertible matrix with determinant ± 1, P a permutation matrix, $(Px)_i = x_{\pi(i)}$ for some permutation $\pi\in S_n$, and D is a diagonal matrix with entries $d_i\in Z$ along the diagonal such that on each cycle of π $|d_i| > 1$ for at least one i; in other words, such that $|d_i d_{\pi(i)} d_{\pi^2(i)} \cdots d_{\pi^{n-1}(i)}| > 1$ for each $i = 1, 2, \cdots, n$.

Proof: Assume that A is of the described form, then A is indeed a dilation. Since C, P, D are integer-valued, A leaves Z^n invariant. The conditions on D, P guarantee that DP, hence, $A = CDPC^{-1}$ have all eigenvalues $|\lambda_i| > 1$. To see this, let $D^{(k)}$ be the diagonal matrix with the permuted entries $d_i^{(k)} = d_{\pi^k(i)}$, and check that $P^k D = D^{(k)} P^k$. After rearranging, $(DP)^n = DD'D'' \cdots D^{(n-1)}P^n = \Pi_{k=1}^n D^{(k)}$, one obtains a diagonal matrix $\tilde{D}$ with elements $d_i = d_i d_{\pi(i)} d_{\pi^2(i)} \cdots d_{\pi^{n-1}(i)}$. The eigenvalues of DP are nth roots of D_0, and its eigenvalues $|\lambda_i| > 1$, if and only if in the cycle determined by $i \mid d_{\pi^k(i)}| > 1$ for at least one k.

P just interchanges the coordinate axes and thus $PQ_0 = Q_0$. Therefore, $DPQ_0 = DQ_0$ is a parallelepiped with edges parallel to the coordinate axes and side-lengths $|d_i|$. Obviously $DPQ_0 = \cup_{i=1}^q (k_i + Q_0)$ for an appropriate set $\{k_i, i = 1, \cdots, q\} \subseteq Z^n$ of the form $\{1 \in Z^n : 0 \le l_i < d_i$ or $d_i \le l_i < 0\}$. In the general case, $C \ne Id$, set $Q = CQ_0$ and $k_i' = Ck_i$, then $CDPC^{-1}Q = \cup_{i=1}^q (Ck_i + CQ_0) = \cup_{i=1}^q (k_i' + Q)$.

Conversely, assume that Q is a parallelepiped $Q = \{x : x = \sum_{i=1}^n x_i e_i, a_i \le x_i \le b_i, i = 1, 2, \cdots, n\}$ for some numbers a_i, b_i and n linearly independents vectors $e_i \in R^n$. Upon writing $a = \sum a_i e_i$ and C for the invertible matrix with columns $(b_i - a_i)e_i$, one obtains $Q = CQ_0 + a$.

The self-similarity of Q (11) becomes $C^{-1}A(CQ_0 + a) = \cup_{i=1}^q C^{-1}(k_i + a + CQ_0)$ or

$$C^{-1}ACQ_0 = \bigcup_{i=1}^{q} \left(C^{-1}(k_i + a - Aa) + Q_0\right). \quad (49)$$

In other words, the parallelepiped $C^{-1}ACQ_0$ is a union of unit cubes. This is only possible if the edges of $C^{-1}ACQ_0$ are parallel to the coordinate axes and $C^{-1}(k_i + a - Aa) \in \{l \in Z^n : 0 \le l_i < d_i$ or $-d_i \le l_i < 0\}$ for some $d_i \in Z$. Therefore $A' = C^{-1}AC$ maps the coordinate axes onto themselves, $A'\delta_i = d_i \delta_{\pi(i)}$, where $d_i \in Z$, $\pi \in S_n$ is a permutation, and δ_i is the ith unit vector. Thus $A' = DP$ and $A = C^{-1}DPC$. Finally, $R^n = C^{-1}R^n = C^{-1}(\cup_{k \in Z^n}(k + Q)) = C^{-1}(\cup_{k \in Z^n}(k + a + CQ_0)) = \cup_{k \in Z^n}(C^{-1}a + C^{-1}k + Q_0)) = \cup_{k \in Z^n}(C^{-1}k + Q_0))$ implies that $C^{-1}Z^n = Z^n$. A symmetric argument yields $CZ^n = Z^n$, from which $C \in SL(n, Z)$ follows. □

Remarks:

- As we have just seen, the choice $\mathscr{K} = \{l \in Z^n : 0 \le l_i < d_i$ or $-d_i \le l_i < 0\}$ produces a parallelepiped as the self-similar tile when A is of the appropriate form. It is, however, not the only possible choice of digits in this case. As we will see (Example 3 in Section V-C), other sets of digits can be used and yield more interesting tilings.
- Since the set Q depends on the choice of digits, it should be clear that there may be many different multiresolution analyses associated with (Z^n, A), which consists of simple functions. In certain cases, Lemmas 4 and 8 are useful in relating different multiresolution analyses that arise in this way.

The following examples represent a very small sampling of a very large smorgasbord.

B. Univariate Examples

The case $A = 2$ is relatively uninteresting in this context. Since all sets of possible digits are related via shifts and multiplications by integers, the fact that $(A - I)^{-1} = I$ together with Lemmas 4 and 8 implies that the only Q's that lead to scaling functions are integer translates of the interval [0, 1]. Thus, we may conclude that *there is only one multiresolution analysis associated with (Z, 2) that consists of simple functions; its scaling function is the characteristic function of the interval [0, 1] and the corresponding elementary wavelet basis is the classical Haar system.* For a similar result obtained from a different point of view see [3].

The case $A = 3$ is already more interesting. The choice of digits $\{0, 1, 2\}$ leads to $Q = [0, 1]$. Since $(A - I)^{-1} = 1/2$, the choice of digits $\{1, 2, 3\}$ leads to $Q = [1/2, 3/2]$. These Q's lead to two different multiresolution analyses associated with $(Z, 3)$. The choice of digits $\{0, 1, 5\}$ leads to a disconnected set Q that can be shown to have measure one by using item 6) of Theorem 3; the corresponding characteristic function is the scaling function of yet another multiresolution analysis associated with $(Z, 3)$. Using Lemma 7, one can construct many different elementary wavelet bases corresponding to each of these examples.

These examples should give the flavor of what happens in the general univariate case.

C. Two-Dimensional Examples

1) The dilation

$$A = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$$

is of the type described in Lemma 9. Choosing $k_1 = (0, -1)$, $k_2 = (1, -1)$ one obtains $Q = [0, 1]^2$ as the basic tile, since $AQ = [0, 2] \times [-1, 0] = (0, -1) + Q \cup (1, -1) + Q$. The basic wavelet ψ is defined by $\psi(x) = 1$ for $x \in [0, 1] \times [0, 1/2)$, $\psi(x) = -1$ for $x \in [0, 1] \times [1/2, 1]$ and $\psi(x) = 0$, elsewhere. The corresponding elementary wavelet basis for $L^2(R^2)$ is $2^{j/2}\psi(A^j x - k)$, $j \in Z$, $k \in Z^2$; it is generated by one function only.

In this case one can easily determine all the simple multiresolution analyses associated with (Z^2, A). Indeed we have the following:

There are exactly three different simple multiresolution analyses associated with (Z^2, A). Their scaling functions are the characteristic functions of the squares Q, $Q + x_0$ and $Q - x_0$ where $Q = [0, 1]^2$ and $x_0 = (-1/3, 1/3)$.

To see this observe that any acceptable pair of digits must be Z^2 translates of one of the pairs $\{(0, 0), (l, m)\}$ where $l = 2j + 1$, and j, $m \in Z$. Now, the pair $\{(0, 0), (l, m)\}$ is the image of the pair $\{(0, 0), (1, 0)]$ under the linear transformation whose matrix is given by

$$B = \begin{pmatrix} l & -2m \\ m & l \end{pmatrix}$$

and which commutes with A. In view of Lemma 8 and the fact that the determinant of B is $l^2 + 2m^2$, it follows that the only such pairs that can give rise to a Q of measure one are $\{(0,0), (1,0)\}$ and $\{(0,0), (-1,0)\}$. Since these pairs are Z^2 translates of each other we may conclude by virtue of Lemma 4 that all the acceptable sets Q are of the form $Q_0 + (A - I)^{-1}k$, $k \in Z^2$. Since

$$(A - I)^{-1} = -\frac{1}{3}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix},$$

it is easy to check that each such square is a Z^2 translate of one of the three squares previously given.

2) The matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

rotates a vector by $\pi/4$ and stretches it by a factor of $\sqrt{2}$. Choosing $k_1 = (0,0)$, $k_2 = (1,0)$, the algorithm (25) produces the set Q shown in Fig. 1, which is known as the twin dragon set in the flowery language of fractals [1].

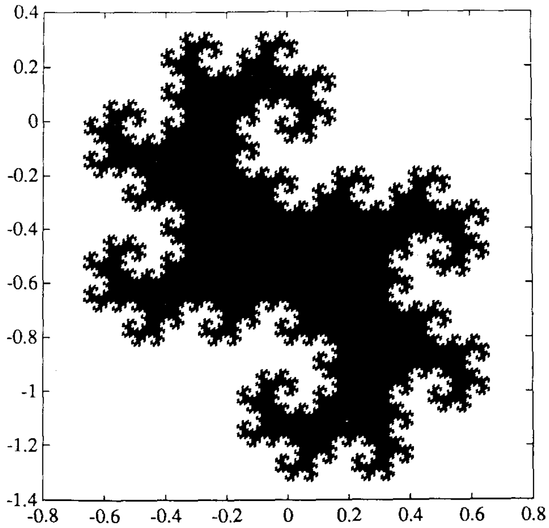

Fig. 1. Tile described in Example 2) in Section V-C.

In order to check that the characteristic function of Q is a scaling function for a multiresolution analysis we verify Cohen's condition. In this case $\hat{\mu}(\xi_1, \xi_2) = (1 + e^{i\xi_1})/2$ has zeroes at $\xi_1 = (2n+1)\pi$, ξ_2 arbitrary. The natural guess $K = \Omega = [-\pi, \pi]^2$ does not satisfy (40), because $\hat{\mu}(\pm A^{-1}(\pi, \pi)) = \hat{\mu}(\pm\pi, 0) = 0$. This can be avoided by moving a neighborhood of the critical points $\pm(\pi, \pi)$ by $\mp 2\pi$ as follows: Let $U_{\pm 1} = \Omega \cap B_{\pm 1}$ where $B_{\pm 1} = \{x : |x \mp (\pi, \pi)| < \delta\}$. Then $K = (\Omega \setminus (U_1 \cup U_{-1})) \cup (U_1 + (-2\pi, 0)) \cup (U_{-1} + (2\pi, 0))$ does the job.

The function ψ defined by $\psi(x) = 1$ for $x \in A^{-1}Q$, $\psi(x) = -1$ for $x \in A^{-1}Q + (1/2, -1/2)$, $\psi(x) = 0$ elsewhere is the basic elementary wavelet in this case, see Fig. 2. The collection $2^{j/2}\psi(A^j x - k)$, $j \in Z$, $k \in Z^2$, which is generated by the one basic wavelet is a complete orthonormal basis for $L^2(R^2)$.

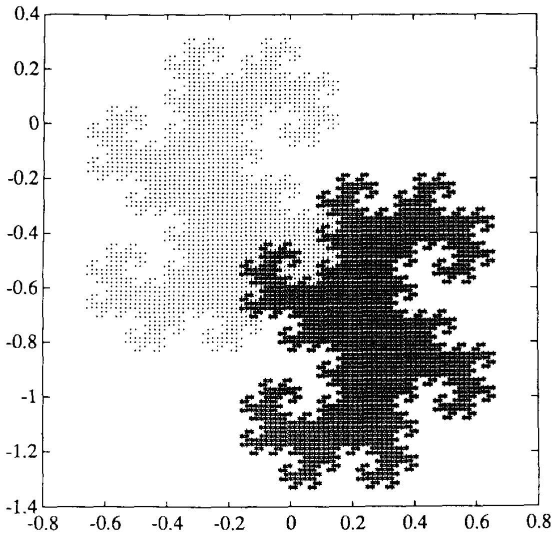

Fig. 2. Wavelet corresponding to tile in Fig. 1.

By using reasoning similar to that used in the previous example we can easily determine all the simple multiresolution analyses associated with (Z^2, A). Indeed, we have the following.

There are exactly two different simple multiresolution analyses associated with (Z^2, A). Their scaling functions are the characteristic functions of the twin dragon set Q described above and BQ where B is a rotation about the origin by the angle $\pi/2$.

3) At first glance the usual homogeneous dilation $A = 2I$, as in the univariate case, does not seem very interesting. If $k_1 = (0,0)$, $k_2 = (1,0)$, $k_3 = (0,1)$, $k_4 = (1,1)$, then clearly $Q = [0,1]^2$ and χQ is the scaling function for multiscale analysis $\mathscr{V}_1$ associated with (Z^2, A), which is the obvious generalization of the univariate dyadic multiscale analysis considered in Section V-B. The corresponding elementary wavelet basis is generated by three basic wavelets that are easily constructed using the recipe given in Section IV.

Choosing as the digits $k_1 = (0,0)$, $k_2 = (1,1)$, $k_3 = (0,1)$, $k_4 = (1,2)$, one obtains the parallelogram with corners k_i, $i = 1, 2, 3, 4$ as the self-similar set Q. The characteristic function of Q is the scaling function for a multiscale analysis $\mathscr{V}_2$ associated with (Z^2, A), which is different from $\mathscr{V}_1$ previously mentioned.

However, if the digits are chosen $k_1 = (0,0)$, $k_2 = (1,0)$, $k_3 = (0,1)$, $k_4 = (-1,-1)$, then the algorithm (25) converges to a Cantor-like set Q, which is shown in Fig. 3. To see that χ_Q is a scaling function for a multiresolution analysis associated with (Z^n, A), we check Cohen's condition: The zeros of $\hat{\mu}(\xi_1, \xi_2) = (1 + e^{i\xi_1} + e^{i\xi_2} + e^{-i(\xi_1+\xi_2)})/4$ are at the points $((2k+1)\pi, l\pi)$ or $(l\pi, (2k+1)\pi)$, $l, k \in Z$. Then $K = [-\pi, \pi]^2$ does the job since $\hat{\mu}$ does not vanish $2^{-j}K \subseteq \frac{1}{2}K = [-\pi/2, \pi/2]^2$. It should be clear that the corresponding multiresolution analysis is different from $\mathscr{V}_1$

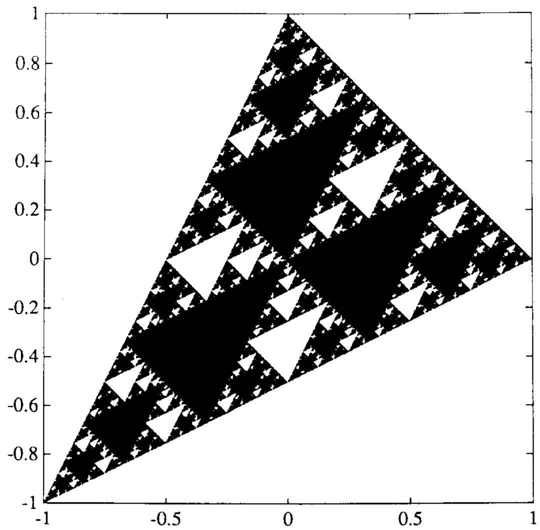

Fig. 3. Tile described in last paragraph of Example 3) in Section V-C.

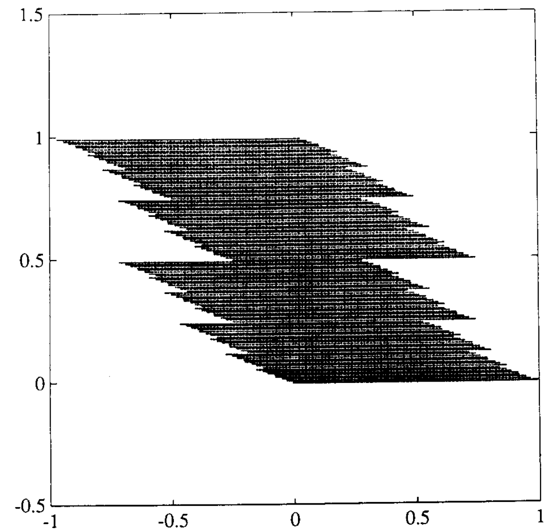

Fig. 5. One of the tiles described in Example 4) in Section V-C.

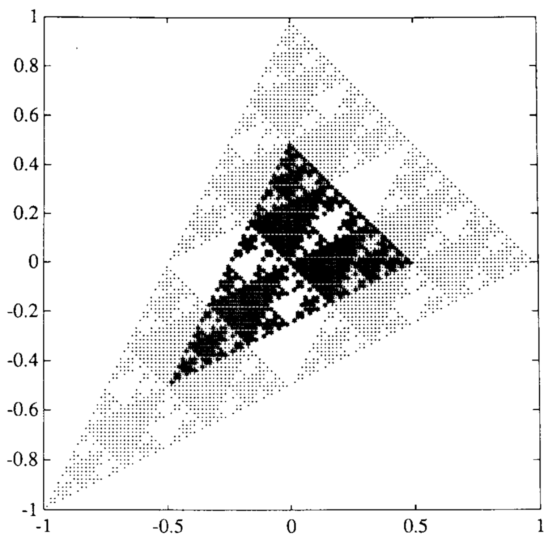

Fig. 4. Wavelet corresponding to tile in Fig. 2.

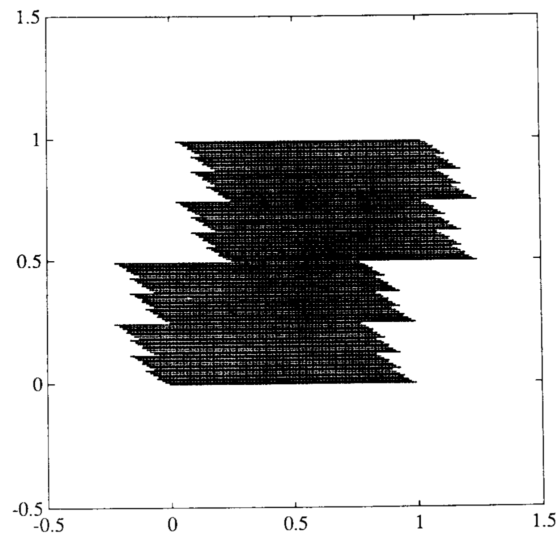

Fig. 6. Second tile described in Example 4) in Section V-C.

and $\mathcal{V}_2$ previously mentioned. The corresponding elementary wavelet bases are generated by three basic wavelets that are easily constructed using the recipe given in Section IV, see Fig. 4.

4) Figs. 5 and 6 show the tiles obtained from the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},$$

with the digits $k_1 = (0,0)$, $k_2 = (1,0)$, $k_3 = (0,1)$, $k_4 = (1,1)$ and the digits $k_1 = (0,0)$, $k_2 = (1,0)$, $k_3 = (1,1)$, $k_4 = (2,1)$, respectively. It should be clear from the picture that the first set has measure one; this can also be checked by verifying Cohen's condition with $K = [-\pi, \pi]^2$. The set in Fig. 6 is the image of the first under the linear map that sends $(1,0)$ to $(1,0)$ and $(0,1)$ to $(1,1)$; this transformation maps the first set of digits into the second, commutes with A, and has determinant one.

5) Figs. 7 and 8 show the tiles obtained from the matrices

$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & -1 \\ 1 & -2 \end{pmatrix}$$

with the digits $\{(0,0), (0,1), (0,2), (1,0), (1,2), (2,0), (2,1), (2,2), (4,4)\}$ and $\{(0,0), (1,0), (0,-1)\}$, respectively.

VI. Miscellaneous Remarks

For more details and background concerning dyadic multiresolution analysis and wavelet bases, including the classi-

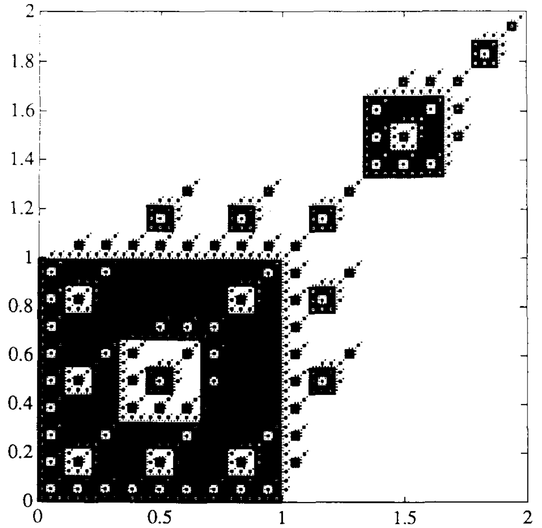

Fig. 7. Tile corresponding to $A = 3I$ described in Example 5) in Section V-C.

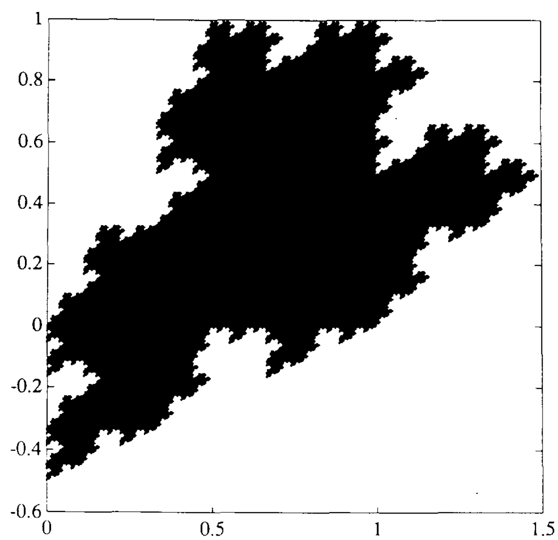

Fig. 8. Second tile described in Example 5) in Section V-C.

cal Haar system, see [3], [6], [7], [9] and the references cited there; for the more general case see [8].

The fixed point iteration (24) is called the *cascade algorithm* in [3] in the case considered there. Tilings of R^n that are not necessarily self-similar arise naturally in other contexts also; for example, see [10] and [4]. The proof of the equivalence of item 6) to the other items in Theorem 3 is an easy modification of the argument found in [2], we include it for the sake of completeness.

In the case $|\det A| = 2$, the martingale version of a classical theorem of Littlewood and Paley, see [5, Theorem 5.3.8], implies that the elementary wavelet bases constructed here are also unconditional bases for $L^p(R^n)$, $1 < p < \infty$.

We emphasize that the specific examples considered here represent a small selection of a wealth of interesting examples. The numerical experiments were done with Matlab software on a Sun 3 workstation.

Acknowledgment

The authors wish to acknowledge Cohen who showed us his condition and made [2] available to us. After this paper was submitted for publication, W. Lawton kindly brought to our attention the related work [11]. These authors propose an interesting framework for multiresolution analyses associated with (Γ, A) whose scaling sequences are finite and, in that spirit, obtain some of the results found here.

References

[1] M. Barnsley, *Fractals Everywhere.* New York: Academic Press, 1988.
[2] A. Cohen, "Ondelettes, analyses multirésolutions et filtres miroirs en quadrature," *Ann. Inst. Poincaré*, vol. 7, pp. 439-459, 1990.
[3] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. in Pure and Applied Math.*, vol. 41, pp. 909-996, 1988.
[4] C. deBoor, C. K. Hollig, and S. Riemenschneider, "Convergence of cardinal series," *Proc. Amer. Math. Soc.*, vol. 98, pp. 457-460, 1986.
[5] R. E. Edwards and G. I. Gaudry, *Littlewood-Paley and Multiplier Theory.* New York: Springer, 1977.
[6] P. G. Lemarie and Y. Meyer, "Ondelette et bases hilbertiennes," *Revista Matematica Iberoamericana*, vol. 2, pp. 1-18, 1986.
[7] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, pp. 69-88, 1989.
[8] Y. Meyer, "Ondelettes, fonctions splines et analyses graduées," Rapport CEREMADE n.8703, 1987.
[9] —, *Ondelettes et Opérateurs.* Paris: Hermann, 1990.
[10] W. P. Thurston, "Groups, Tilings, and Finite State Automata," in *Lecture Notes, Summer Meetings Amer. Math. Soc.*, Bolder, CO 1989.
[11] W. M. Lawton and H. L. Resnikoff, "Multidimensional wavelet bases," Aware Inc. tech. rep., submitted to *SIAM J. Math. Anal.*