Some remarks on self-affine tilings

Derek Hacon, Nicolau C. Saldanha, J. J. P. Veerman

Abstract: We study self-affine tilings of $\mathbb{R}^n$ with special emphasis on the two-digit case. We prove that in this case the tile is connected and, if $n \leq 3$, is a lattice-tile.

0. Introduction

The present article is devoted to the study of certain tilings of $\mathbb{R}^n$ defined by "generalized decimal expansions" in which the base 10 and the digits $0, \ldots, 9$ are replaced by an integer $n \times n$ matrix A and a finite subset $\mathcal{D}$ of $\mathbb{Z}^n$. The expansions in question are of the form

$$\sum_{-\infty < i \leq N} A^i v_i$$

where $N \in \mathbb{Z}$ and all the v_i belong to $\mathcal{D}$. To guarantee convergence, A is assumed to be expanding (i.e. all its eigenvalues have modulus greater than 1). Since A is an integer matrix, it follows that its determinant is $\pm q$ where q is a positive integer. $\mathcal{D}$ is required to consist of exactly q elements, one for each coset of $A\mathbb{Z}^n$ in $\mathbb{Z}^n$ (in the terminology of [LW1], $\mathcal{D}$ is a *standard digit set*). We assume for simplicity that 0 belongs to $\mathcal{D}$.

Corresponding to the integer and fractional parts in the usual decimal expansion, define I to consist of all sums of the form $v_0 + \cdots + A^k v_k$ (where $k \geq 0$ and $v_i \in \mathcal{D}$) and Q to be the set of all infinite sums $A^{-1}v_{-1} + A^{-2}v_{-2} + \cdots \quad (v_i \in \mathcal{D})$. In the usual decimal expansion, Q is the unit interval and $I = \{0, 1, 2, \ldots\}$. The set $I + Q$ tiles some subset of $\mathbb{R}^n$ and we may enlarge I to obtain a set G' so that $G' + Q$ is a tiling of $\mathbb{R}^n$ (1.13). In fact, $G' \subseteq G = I - I$, the set of differences of elements of I. The following example, although simple, illustrates several features of the general case: if $n = 1$, $A = 3$ and $\mathcal{D} = \{0, 4, 11\}$ then $I = \{0, 4, 11, 12, 16, \ldots\}$ and $\mathbb{R}$ is tiled by the translates of Q by $\mathbb{Z}$ (see the end of section 1). It turns out that G' is a lattice if $G = I - I$ is one also; in this case, $G' = G$. If G is a lattice, we call Q a *lattice-tile*. Gröchenig and Haas ([Gr], [GrH]) have shown that for $n = 1$, G is always a lattice. On the other hand, an example due to Lagarias and Wang ([LW1], Example 2.3) shows that, for $n > 1$, G is not always a lattice.

Section 1 consists mostly of a review of basic results on self-affine tilings of $\mathbb{R}^n$. This section overlaps substantially with results of other authors (see [B], [GrM], [Ke], [V] and the pre-prints [GrH], [LW1], [LW2], [LW3]); we take this opportunity of thanking the referee for drawing our attention to a number of these references. In section 2 we investigate certain aspects of the case $q = 2$: we prove that Q is always connected and that, for $n \leq 3$, G is a lattice. Finally, in section 3, we derive an algorithm for checking if G is a lattice given A and $\mathcal{D}$ (see [V] for another algorithm).

All three authors receive support from CNPq and MCT, Brazil.

1. Basic Results

In this section we establish some basic results on tilings of $\mathbb{R}^n$ which are self-affine in the terminology of [LW1]. Some of these results are due independently to other authors; we refer the reader to the references given in the introduction.

Let A be an expanding integer $n \times n$ matrix (i.e. all of whose eigenvalues have modulus greater than 1). Then, of course, $\det A = \pm q$ where q is a positive integer. Reducing A to Jordan canonical form, one has:

Lemma 1.1: *For any bounded set B the diameter of $A^{-k}B$ tends to 0 as $k \to \infty$.* ∎

We also suppose given (or choose) a set $\mathcal{D}$ of q elements of $\mathbb{Z}^n$ such that:

(1) 0 belongs to $\mathcal{D}$.

(2) The elements of $\mathcal{D}$ are distinct modulo $A\mathbb{Z}^n$, in the sense that if $r, s \in \mathcal{D}$ and $r - s \in A\mathbb{Z}^n$ then $r = s$.

It follows that $\mathbb{Z}^n$ is the disjoint union of the q cosets $r + A\mathbb{Z}^n$ $(r \in \mathcal{D})$. Here we denote by $X + Y$ the set of all sums $x + y$ where $x \in X$, $y \in Y$. Infinite sums are defined if all sets contain 0. Thus, if $0 \in X_k$ for all k, $X_1 + X_2 + \cdots$ is the increasing union of the $X_1 + \cdots + X_k$ for all k.

Since $\mathbb{Z}^n = \mathcal{D} + A\mathbb{Z}^n$, one has $\mathbb{Z}^n = \mathcal{D} + A\mathcal{D} + A^2\mathbb{Z}^n$ and so on. We write

$$I = \mathcal{D} + A\mathcal{D} + A^2\mathcal{D} + \cdots.$$

This representation of I is unique, for if $r_1 + \cdots + A^k r_k = s_1 + \cdots + A^k s_k$ then $r_1 = s_1$ modulo $A\mathbb{Z}^n$, so $r_1 = s_1$. Similarly $r_2 = s_2, \ldots, r_k = s_k$.

Example 1.2: *For $n = 1$, $A = 3$ and $\mathcal{D} = \{0, 4, 11\}$ we have $I = \{0, 4, 11, 12, 16, \ldots\}$.*

Definition 1.3: *If $Z \subseteq \mathbb{R}^n$ define $\tau Z = A^{-1}(\mathcal{D} + Z)$.*

Then $\tau^k Z = Q_k + A^{-k}Z$ where $Q_k = \tau^k(\{0\}) = A^{-1}\mathcal{D} + \cdots + A^{-k}\mathcal{D}$. Clearly $Q_1 \subset Q_2 \subset \cdots$ and $AQ_1 \subset A^2 Q_2 \subset \cdots \subset I$. Notice that the compact set $\overline{\bigcup_k Q_k}$ is invariant under τ. Following [H] (compare [F2] and [LW1]), we examine how τ acts on the compact subsets of $\mathbb{R}^n$ in order to characterize this set.

Definition 1.4: *$H(X)$ is the space of compact non-empty subsets of a metric space X, equipped with the Hausdorff metric:*

$$d(K, L) = \inf\{\epsilon \mid K \subseteq N_\epsilon(L), L \subseteq N_\epsilon(K)\}$$

where $N_\epsilon(K)$ is the open ϵ-neighbourhood of K.

It is well known that if X is complete so is $H(X)$. Clearly, τ maps $H(\mathbb{R}^n)$ into itself. By 1.1, some power τ^N of τ is a contraction mapping. Let Q be the unique fixed point of τ^N. Since $\tau^N \tau Q = \tau Q$, one has $\tau Q = Q$, by uniqueness. If $K \in H(\mathbb{R}^n)$ and $\tau K = K$ then $\tau^N K = K$ so that $K = Q$, again by uniqueness. We have proved:

Lemma 1.5: *$\tau : H(\mathbb{R}^n) \to H(\mathbb{R}^n)$ has a unique fixed point Q.* ∎

Furthermore, for any $K \in H(\mathbb{R}^n)$, $Q = \lim_{k\to\infty} \tau^k K$, the limit being in the Hausdorff metric: consider the subsequences of the form $k = jN + k_0$, k_0 fixed, $j \to \infty$. For example, taking $K = \{0\}$, $Q = \overline{\bigcup_k Q_k}$ and the finite sets Q_k are approximations to Q (in the Hausdorff metric).

We next ask how $\mathbb{R}^n$ may be represented in terms of I and Q. We first show that $\mathbb{R}^n = \mathbb{Z}^n + Q$, in other words, that $\pi Q = \mathbb{T}^n$ where $\pi : \mathbb{R}^n \to \mathbb{T}^n$ is the quotient map onto the n-torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Since $A\mathbb{Z}^n \subseteq \mathbb{Z}^n$, A induces a map $\overline{A} : \mathbb{T}^n \to \mathbb{T}^n$. Now $\overline{A}^{-1}$ acts on subsets of $\mathbb{T}^n$ in the obvious way, taking B to $\overline{A}^{-1}B$. Clearly we obtain a map $W : H(\mathbb{T}^n) \to H(\mathbb{T}^n)$. Also π induces $\pi : H(\mathbb{R}^n) \to H(\mathbb{T}^n)$, which is continuous; indeed, π decreases distances. Since the digits form a complete set of representatives modulo $A\mathbb{Z}^n$, $\pi\tau = W\pi$.

Lemma 1.6: $\mathbb{Z}^n + Q = \mathbb{R}^n$.

Proof: Let $K \in H(\mathbb{R}^n)$ be such that $\pi K = T$. Then $\pi Q = \pi \lim_k \tau^k K = \lim_k \pi\tau^k K = \lim_k W^k \pi K = T$. ■

The above lemma implies the following result (compare Theorem 1.1 in [LW1]).

Corollary 1.7: *Q has non-empty interior.*

Proof: Follows from 1.6 by Baire's theorem. ■

We next look at the self-similarity properties of Q (compare [F2]). Let $|X|$ be Lebesgue measure of X; all sets considered will be measurable.

Definition 1.8: *Z_1 and Z_2 overlap iff $|Z_1 \cap Z_2| > 0$. A sum $X + Y$ does not overlap iff the sets $x + Y$ do not overlap for $x \in X$.*

Lemma 1.9: *$Q_k + A^{-k}Q$ does not overlap.*

Proof: Q_k contains at most q^k points and $|A^{-k}Z| = q^{-k}|Z|$ for any Z. Since $Q = \tau^k Q$, the lemma follows. ■

Corollary 1.10: *$I + Q$ does not overlap.*

Proof: If $u, v \in I$ and $u \neq v$ then, for some k, where $u, v \in A^k Q_k$. Since, by 1.9, $A^{-k}u + A^{-k}Q$ and $A^{-k}v + A^{-k}Q$ do not overlap neither do $u + Q$ and $v + Q$. ■

Definition 1.11: *If $X, \mathcal{H}, Y \subset \mathbb{R}^n$ are measurable sets, where X is bounded with non-empty interior, we shall say that $\mathcal{H} + X$ is a tiling of Y if $Y = \mathcal{H} + X$ and $\mathcal{H} + X$ does not overlap.*

Definition 1.12: *$G = I - I$.*

From 1.7, and the self-similarity of Q (1.9), we can recover the following tiling property (see [GrH]).

Proposition 1.13: *There exists a subset G' of G such that $G' + Q$ tiles $\mathbb{R}^n$. Furthermore, $G' - G' \subseteq G$.*

Proof: From 1.7, Q has non-empty interior. Since A is expanding, given r there exists k_r such that a ball of radius $r + \operatorname{diam}(Q)$ fits inside $A^{k_r}Q$. But $A^{k_r}Q = A^{k_r}Q_{k_r} + Q$, and this sum is non-overlapping by 1.9. Thus, for some element v_r of $A^{k_r}Q_{k_r}$, $G_r + Q$ tiles (a superset of) a ball of radius r around the origin, where $G_r = A^{k_r} Q_{k_r} - v_r \subset G$. Consider such G_r for all positive integers r.

We now have tilings of arbitrarily large regions around the origin; using these, we assemble a tiling of $\mathbb{R}^n$. Given any positive integer s, the intersections of G_r with B_s, the ball of radius s around the origin, can only produce finitely many different sets. Thus, there is an infinite subsequence of values of r for which these intersections are all equal. Starting with $s = 1$, we take a subsequence $G^1_1, G^1_2, \ldots$ of $G_1, G_2, \ldots$, all meeting B_1 in the same set. For $s = 2$, we take a subsequence $G^2_1, G^2_2, \ldots$ of $G^1_1, G^1_2, \ldots$, all meeting B_2 in the same set. We repeat the process for $s = 3, 4, 5, \ldots$. Define G' to be the set of elements of G belonging to all but finitely many of the sets $G^1_1, G^2_1, G^3_1, \ldots$: by construction, $G' + Q$ is a tiling of $\mathbb{R}^n$.

To prove the last assertion, we observe that $G_r - G_r = Q_{k_r} - Q_{k_r} \subseteq I - I = G$ and therefore $G' - G' \subseteq G$, since any finite subset of G' is contained in some G_r. ∎

Next consider how τ acts on the (measurable) subsets of Q.

Definition 1.14: *$Z_1 \equiv Z_2$ means $|Z_1 \cup Z_2| = |Z_1 \cap Z_2|$.*

Definition 1.15: *A subset Z of Q is τ-invariant iff $Z \equiv \tau Z$.*

Lemma 1.16: *If $Z \subseteq Q$ then $|\tau Z| = |Z|$.*

Proof: Let $Y = Q \backslash Z$. Then, by definition of τ, $|\tau Z| \leq |Z|$ and $|\tau Y| \leq |Y|$. But $\tau Z \cup \tau Y = \tau Q = Q = Z \cup Y$ so that $|\tau Z| = |Z|$. ∎

Thus, if $Z \subseteq \tau Z$ or if $\tau Z \subseteq Z$ then Z is τ-invariant. We now have the following ergodicity result (see the end of this section).

Proposition 1.17: *The measure of any τ-invariant subset of Q is either 0 or $|Q|$.*

Proof: Suppose that Z is τ-invariant and $|Z| > 0$. It follows directly from the Lebesgue density theorem that we may choose a point x at which Z is dense, i.e., such that

$$\lim_{\epsilon \searrow 0} \frac{|Z \cap N_\epsilon(x)|}{|N_\epsilon(x)|} = 1.$$

Thus, given $\delta > 0$, we have, for $\epsilon > 0$ sufficiently small,

$$\frac{|Z \cap N_\epsilon(x)|}{|Q \cap N_\epsilon(x)|} \geq 1 - \delta/2.$$

For k large enough, the diameter of $A^{-k}Q$ will be small enough, in comparison with ϵ, to ensure the existence of $v \in Q_k$ such that

$$\frac{|Z \cap (v + A^{-k}Q)|}{|Q \cap (v + A^{-k}Q)|} \geq 1 - \delta.$$

Indeed, by 1.9, the tiny sets $v + A^{-k}Q$, for v in some subset of Q_k, cover $N_\epsilon(x)$, except for a narrow margin around the boundary of $N_\epsilon(x)$ (of negligible size), without overlapping;

if the above ratio were smaller than $1-\delta$ for all such v the ratio in the previous inequality would be smaller than $1-\delta/2$, a contradiction.

By hypothesis, $Z \equiv \tau^k Z$. Hence $Z \equiv Q_k + A^{-k}Z$. By 1.9, $(Q_k + A^{-k}Z) \cap (v + A^{-k}Q) \equiv v + A^{-k}Z$. Hence

$$\frac{|Z|}{|Q|} = \frac{|v + A^{-k}Z|}{|v + A^{-k}Q|} \geq 1 - \delta.$$

Since $\delta > 0$ is arbitrary, we are done. ■

Corollary 1.18: *Except on a set of measure zero, $\pi : Q \to \mathbb{T}^n$ is ℓ to 1 for some fixed integer ℓ.*

Proof: Let $Q^{(k)} = \{x \in Q \mid x + \mathbb{Z}^n$ meets Q in at least k points (including x)$\}$. Clearly, $Q = Q^{(1)} \supseteq Q^{(2)} \supseteq \cdots$, and $Q^{(k)}$ is empty for large k (since Q is bounded) and the $Q^{(k)}$ are all measurable.

To verify that $Q^{(k)} \subseteq \tau Q^{(k)}$ for all k, let $x, y \in Q$ such that $x - y \in \mathbb{Z}^n$. Since $Q = \tau Q$, we have $x = A^{-1}(x' + r), y = A^{-1}(y' + s)$ for some $x', y' \in Q$ and $r, s \in \mathcal{D}$. Clearly, $x' - y' \in \mathbb{Z}^n$ and if $x' = y'$ then $r - s \in A\mathbb{Z}^n$ so that $r = s$ and, therefore, $x = y$. This proves that $Q^{(k)} \subseteq \tau Q^{(k)}$ and that $Q^{(k)}$ is τ-invariant for all k. By 1.17, $|Q^{(k)}|$ is 0 or $|Q|$; take ℓ to be the largest value of k for which $|Q^{(k)}| = |Q|$. ■

Notice that in this proof we only use the fact that $(\mathcal{D} - \mathcal{D}) \cap A\mathbb{Z}^n = \{0\}$; see [LW1] for more general digit sets.

Corollary 1.19: *The Lebesgue measure of Q is always an integer.* ■

Corollary 1.20: *The boundary of Q has measure zero.*

Proof: By 1.7 $|\text{Int}Q| > 0$. Also $\tau(\text{Int}Q) \subseteq \text{Int}Q$ by definition of τ. Now apply 1.17. ■

We now show that if G is a lattice, then $G' = G$.

Proposition 1.21: *G is a lattice if and only if $G + Q$ tiles $\mathbb{R}^n$.*

Proof: First assume G to be a lattice. Let v and w be distinct points of G. By hypothesis, $v - w = x - y$ where $x, y \in I$. By 1.10, $x + Q$ and $y + Q$ do not overlap so neither do $v + Q$ and $w + Q$.

Conversely, if $G + Q$ does not overlap then $G = G'$ since $G' + Q$ is a tiling. Thus, by 1.13, $G - G \subseteq G$ and G is a lattice. ■

From 1.13 and 1.21 we have:

Theorem 1.22: *If G is a lattice then $G + Q$ is a tiling and the Lebesgue measure of Q is equal to the index of G in $\mathbb{Z}^n$.* ■

We may deduce a criterion for G to be a lattice.

Lemma 1.23: *If $G \cap Q = \mathbb{Z}^n \cap Q$ then $G = \mathbb{Z}^n$.*

Proof: Define $C : \mathbb{Z}^n \to \mathbb{Z}^n$ by $C(v) = A^{-1}(v + r)$ where r is the unique element of $\mathcal{D}$ such that $v + r \in A\mathbb{Z}^n$. Clearly, $C^{-1}G \subseteq G$. For any $v \in \mathbb{Z}^n$ and any $k \geq 1$ we have, by definition, $C^k(v) \in \tau^k(\{v\})$. But $Q = \lim_k \tau^k(\{v\})$ and $\mathbb{Z}^n$ is discrete, so eventually $C^k(v) \in Q$. By hypothesis, $C^k(v) \in G$, i.e., $v \in C^{-k}G$. Hence $v \in G$, as required. ■

Returning to Example 1.2, we show that $G = \mathbb{Z}$ for $n = 1$, $A = 3$ and $\mathcal{D} = \{0, 4, 11\}$. It is easy to see from the definition that Q is contained in $[0, 11/2]$. We generate all integers in this interval:

$$0 \xrightarrow{4} 4 \xrightarrow{-11} 1 \xrightarrow{0} 3 \xrightarrow{-7} 2$$

and

$$0 \xrightarrow{4} 4 \xrightarrow{-7} 5.$$

Here $a \xrightarrow{b} c$ means $c = 3a + b$. For example, $2 = 4 \cdot 3^3 - 11 \cdot 3^2 + 0 \cdot 3^1 - 7 \cdot 3^0$. We thus have $G \cap Q = \{0, 1, 2, 3, 4, 5\} = \mathbb{Z} \cap Q$ and, by 1.23, $G = \mathbb{Z}$. From Theorem 1.22, $\mathbb{Z} + Q$ is a tiling and $|Q| = 1$.

Let $\mathcal{G}$ be the lattice generated by G; thus, G is a lattice iff $G = \mathcal{G}$. As we shall see in Section 3, $\mathcal{G}$ is easily computable.

Proposition 1.24: *G is a lattice if and only if $G \cap Q = \mathcal{G} \cap Q$.*

Proof: Trivially, if G is a lattice then $G \cap Q = \mathcal{G} \cap Q$. Conversely, assume $G \cap Q = \mathcal{G} \cap Q$: since Q is bounded and $G' + Q = \mathbb{R}^n$, G' is not contained in a proper vector subspace of $\mathbb{R}^n$. By definition, $AG \subseteq G$ so $A\mathcal{G} \subseteq \mathcal{G}$. Now $\mathcal{G} \subseteq \mathbb{Z}^n$ and spans $\mathbb{R}^n$ since G' does hence $\mathcal{G}$ is isomorphic to $\mathbb{Z}^n$. Also $\mathcal{D} \subset \mathcal{G}$ (since $\mathcal{D} \subset G$) and the elements of $\mathcal{D}$ are distinct modulo $A\mathbb{Z}^n$ and hence modulo $A\mathcal{G}$ (since $\mathcal{G} \subseteq \mathbb{Z}^n$). Therefore, $\mathcal{D}$ contains precisely one element in each coset of $A\mathcal{G}$ in $\mathcal{G}$ and $C : \mathcal{G} \to \mathcal{G}$ as in the proof of 1.23 is well defined. Now follow the proof of 1.23 with $\mathcal{G}$ instead of $\mathbb{Z}^n$. ■

The question whether or not G is a lattice has been settled in various cases. In dimension 1, Gröchenig and Haas prove:

Theorem 1.25: ([Gr]) *If $n = 1$ then G is always a lattice.*

On the other hand, one has the following example due to Lagarias and Wang:

Example 1.26: ([LW1]) *G is* not *a lattice in the case*

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \qquad \mathcal{D} = \{(0,0), (3,0), (0,1), (3,1)\}.$$

For the reader's convenience, we include a proof of 1.25 which is based on arguments in [Gr] and [GrH]. We first present a series of auxiliary definitions and results. Consider the n dimensional case for a moment. We denote the coefficient of $z^k = z_1^{k_1} \cdots z_n^{k_n}$, $k \in \mathbb{Z}^n$, in a Laurent series $g \in \mathbb{R}[z_1, z_1^{-1}, \ldots, z_n, z_n^{-1}]$ by $\chi_k(g)$. The key to the proof is the introduction of the polynomial T defined by $\chi_k(T) = |Q \cap (k + Q)|$. Clearly, T is constant iff $\mathbb{Z}^n + Q$ does not overlap.

Definition 1.27: *The tiling polynomial T is defined by $\chi_k(T) = |Q \cap (k + Q)|$.*

Definition 1.28: *If $g = \sum a_k z^k$ then $\overline{g} = \sum a_k z^{-k}$.*

The notation comes from $\overline{g(z)} = \overline{g}(z)$ when $|z| = 1$.

Definition 1.29: $D = \sum_{d \in \mathcal{D}} z^d$.

Lemma 1.30: *$\chi_k(qT) = \chi_{Ak}(D\overline{D}T)$ for all $k \in \mathbb{Z}^n$.*

Proof: We recall that $q = |\det A|$. Since $Q = \tau Q$ (1.5) and by 1.9, we have $\chi_k(qT) = q|Q \cap k + Q| = |AQ \cap Ak + AQ| = |A\tau Q \cap Ak + A\tau Q| = |\mathcal{D} + Q \cap Ak + \mathcal{D} + Q| = \sum_{d,d' \in \mathcal{D}} |d + Q \cap Ak + d' + Q| = \sum_{d,d' \in \mathcal{D}} |Q \cap Ak + d' - d + Q| = \chi_{Ak}(D\overline{D}T)$, by definition of D and T. ■

Definition 1.31: *$\hat{g}$ is given by $\chi_k(\hat{g}) = \chi_{Ak}(D\overline{D}g)$.*

Thus, the lemma above may be rewritten as $\hat{T} = qT$. Notice that, since $\mathcal{D}$ is a complete set of residues modulo $A\mathbb{Z}^n$, we have $\hat{1} = q$.

Returning to the 1-dimensional case, we take $A = q$.

Lemma 1.32: *For $|z| = 1$, $q\hat{g}(z) = \sum_{w^q = z} g(w)|D(w)|^2$.*

Proof: Write $\tilde{g} = D\overline{D}g$. Then $\tilde{g}(z)$ may be written in the form $\hat{g}(z^q) + zg_1(z^q) + \cdots + z^{q-1}g_{q-1}(z^q)$ for appropriate $g_1, \ldots, g_{q-1}$. So if $w^q = z$ then $\tilde{g}(w) = \hat{g}(z) + wg_1(z) + \cdots + w^{q-1}g_{q-1}(z)$. But $\sum_{w^q=z} w^j = 0$ for $0 < j < q$. It follows that $q\hat{g}(z) = \sum_{w^q=z} \tilde{g}(w)$. ■

Taking $g = 1$ and $g = T$, we obtain, for $|z| = 1$,

$$\sum_{w^q=z} |D(w)|^2 = q^2 \tag{1}$$

and

$$\sum_{w^q=z} |D(w)|^2 \big(T(w) - T(z)\big) = 0. \tag{2}$$

Since $\chi_{-k}(T) = \chi_k(T)$, the restriction of T to the unit circle $\mathbb{S}^1 = \{z \mid |z| = 1\}$ is real valued.

Lemma 1.33: *If T is non-constant then* $\gcd(\mathcal{D}) > 1$.

Proof: The set E of extrema of T in $\mathbb{S}^1$ is finite and consists of at least two points. Also, if $z \in E$, then, for all $y \in \mathbb{S}^1$, $T(z) - T(y)$ has the same sign. Let $z \neq 1$, $z \in E$. Then, by (1) and (2), there is some w_0, $w_0^q = z$, for which $T(w_0) = T(z)$ and therefore $w_0 \in E$. For each z there is at least one such w_0 and different values of z correspond to different w_0; finiteness of E then guarantees that, given z, there is exactly one such w_0 and therefore $T(w) \neq T(z)$ if $w \neq w_0$ and $w^q = z$. Again from (2), $|D(w)| = 0$ for such w and, from (1), $|D(w_0)| = q$ and therefore $w_0^d = 1$ for all $d \in \mathcal{D}$ (since $0 \in \mathcal{D}$). It follows that $w_0^{\gcd(\mathcal{D})} = 1$ implying $\gcd(\mathcal{D}) > 1$ since $w_0^q = z \neq 1$. ■

Proof of Theorem 1.25: By the previous lemma, if $\gcd(\mathcal{D}) = 1$ then $\mathbb{Z} + Q$ does not overlap. Since $\mathbb{Z} + Q = \mathbb{R}$ (1.6), $\mathbb{Z} + Q$ is a tiling. But $\mathcal{G} \subseteq \mathbb{Z}$ and $\mathcal{G} + Q$ is a tiling; it follows that $\mathcal{G} = \mathbb{Z}$ and $G = \mathbb{Z}$ is a lattice. ■

We close this section by describing briefly how the results of this section are related to the study of expanding toral epimorphisms ([Ka], [M]). Katznelson determined which toral epimorphisms are Bernoulli in terms of the eigenvalues of A, where A is the integral matrix representing the epimorphism $\overline{A}$. From Katznelson's theorem and the classification of Bernoulli shifts by their entropy it follows that any expanding toral epimorphism is equivalent to the shift of type $(1/q, \ldots, 1/q)$, where $q = |\det(A)|$. In fact, the generalized decimal expansion (base A) provides an equivalence between $\overline{A}$ on $\mathbb{T}^n$ and a one-sided

Bernoulli shift. From the fact that $Q = A^{-1}\mathcal{D} + A^{-1}Q$ is non-overlapping it follows that one may define (almost everywhere on Q) the shift S by $S(x) = Ax - r$ where r is such that $x \in A^{-1}r + A^{-1}Q$. It is easily verified that $\pi \circ S = \overline{A} \circ \pi$ whenever S is defined. By 1.22, $\pi : Q \to \mathbb{R}^n/G$ is an equivalence between (Q, S) and $(\mathbb{R}^n/G, \overline{A})$ provided G is a lattice. In particular, when G is a lattice, Proposition 1.17 can be deduced from the known fact ([M]) that $\overline{A}$ is ergodic.

2. The two digit case

Throughout this section we assume A to be an expanding $n \times n$ integer matrix with $q = |\det(A)| = 2$ so that $\mathcal{D}$ consists of two digits, 0 and v (say). The case $q = 2$ has certain special features which we explore in this section. We begin by showing that Q is connected by constructing a space-filling curve in Q. We then prove two theorems (2.10 and 2.12) which guarantee that in many cases G is a lattice.

From 1.9, we have $Q = Q_k + A^{-k}Q$. Thus Q is the union of 2^k *k-pieces*, each of the form $w + A^{-k}Q$ (where $w \in Q_k$).

Lemma 2.1: *The intersection of the two 1-pieces of Q is non-empty.*

Proof: Assume for a contradiction that the two 1-pieces $A^{-1}Q$ and $A^{-1}v + A^{-1}Q$ are disjoint. Then the four 2-pieces of Q are also disjoint and, in general, the 2^k k-pieces are all disjoint. We know that Q has non-empty interior and therefore contains a ball, which is covered by k-pieces whose diameter can be taken to be smaller than that of the ball. This contradicts the fact that the ball is connected. ■

We construct a surjective continuous function $\gamma : [0, 1] \to Q$ by first defining it on the 6-adic numbers in $[0, 1]$ and then passing to the limit.

Definition 2.2: $J_k = [0, 1] \cap \mathbb{Z}/6^k = \{0, 1/6^k, \ldots, 1\}$.

Definition 2.3: $\gamma : J_k \to Q$ *is* admissible *iff:*

(1) There is at least one point of $\gamma(J_k)$ in the interior of each k-piece.

(2) For any two consecutive points r, s of J_k there is some k-piece containing both $\gamma(r)$ and $\gamma(s)$.

Lemma 2.4: *Any admissible $\gamma : J_k \to Q$ extends to an admissible $\tilde{\gamma} : J_{k+1} \to Q$.*

Proof: Let $a_0, a_1, a_2, a_3, a_4, a_5, a_6$ be consecutive points of J_{k+1} with $a_0, a_6 \in J_k$. By (2), $\gamma(a_0)$ and $\gamma(a_6)$ both lie in some k-piece. This k-piece is the union of two $(k+1)$-pieces, P_0 and P_1, say. Arbitrarily choose $\tilde{\gamma}(a_1), \tilde{\gamma}(a_3), \tilde{\gamma}(a_5) \in P_0 \cap P_1$ (which is non-empty by Lemma 2.1), $\tilde{\gamma}(a_2) \in \mathrm{Int}(P_0)$ and $\tilde{\gamma}(a_4) \in \mathrm{Int}(P_1)$. This defines $\tilde{\gamma}$, which is easily seen to be admissible. ■

Theorem 2.5: *If $q = 2$ then Q is path-connected. Moreover, there exists a continuous surjective map γ from $[0, 1]$ to Q.*

Proof: Start by defining $\gamma(0)$ and $\gamma(1)$ arbitrarily. Use the previous lemma to define γ on the union of all J_k. The function γ is uniformly continuous because steps of size 6^{-k} correspond to arcs contained in a k-piece and the diameter of a k-piece tends to zero.

Thus, γ can be continuously extended over $[0,1]$. Furthermore, γ is surjective because its image is dense, since it contains points in the interior of each k-piece for all k. ■

Another interesting feature of the case $q = 2$ is that it allows us to produce many examples of expanding matrices A such that, for all digit sets, G is a lattice. In particular, this is true for $n \leq 3$.

Let A be an expanding matrix with $q = |\det(A)| = 2$. The lattice $\mathcal{G}$ consists of all vectors of the form $g(A)v$ where $g \in \mathbb{Z}[x]$. Also, since $\mathcal{G}$ has rank n, $g(A)v = 0$ if and only if $g(A) = 0$. We thus identify $\mathcal{G}$ with $\mathbb{Z}[A]$ which is, by definition, the ring of all matrices of the form $g(A)$, where $g \in \mathbb{Z}[x]$. The characteristic polynomial p_A of A is irreducible in $\mathbb{Z}[x]$: if it could be factored, one of the factors would have constant term 1 and its roots could not have modulus greater than 1. Therefore, p_A is also the minimal polynomial of A and $g(A) = 0$ if and only if g is a multiple of p_A. A polynomial f of degree less than or equal to $n-1$ is said to be *reduced*. Every element of $\mathbb{Z}[A]$ may be written uniquely as $f(A)$, where f is reduced. The set G corresponds to the set of all polynomials (reduced or not) with coefficients 0, 1 or -1. We have therefore proved the following result:

Proposition 2.6: *G is a lattice if and only if every reduced polynomial f can be written as $g + p_A h$, where the coefficients of g are 0, 1 or -1. In particular, if G is a lattice for some choice of $\mathcal{D}$, then G is a lattice for all $\mathcal{D}$.*

We call a polynomial *expanding* if all its roots lie outside the unit circle. For a given degree n, there exist only a finite number of expanding polynomials with integer coefficients and constant term ± 2 since the other coefficients of the polynomial are bounded, being functions of the roots. Thus, up to conjugation by an integer invertible matrix, there exist only a finite number of $n \times n$ expanding matrices with $q = 2$.

Define the reduced polynomial q_A by

$$p_A(x) = 2 - xq_A(x);$$

thus, the relation $p_A(A) = 0$ becomes $q_A(A) = 2A^{-1}$. We give $\mathbb{Z}[A]$ the (Manhattan) norm $\|f(A)\| = |a_{n-1}| + \cdots + |a_0|$ where $f(x) = a_{n-1}x^{n-1} + \cdots + a_0$ is reduced. Using the relation $2I = Aq_A(A)$, it is clear that, for reduced f, $f(A) \in A\mathbb{Z}[A]$ if and only if $f(0)$ is even. We now define a *carrying operation* $C : \mathbb{Z}[A] \to \mathbb{Z}[A]$ (compare 1.23). If f is reduced then there is a unique $\epsilon = 0$, 1 or -1 such that

(i) $f = xg + 2c + \epsilon$ where $g \in \mathbb{Z}[x]$ and $c \in \mathbb{Z}$,

(ii) $|2c + \epsilon| = |2c| + |\epsilon|$.

Definition 2.7: *With the above notation, $C(f)$ is the (reduced) polynomial $g + cq_A$.*

Lemma 2.8: *If $\|q_A\| \leq 2$ then $\|C(f)\| \leq \|f\|$ and equality implies $f(A) = Ah(A)$ for some h.*

Proof: Clearly, in *(i)* above, $\deg(g) \leq n-2$. Thus $\|f\| = \|xg\| + |2c+\epsilon| = \|g\| + |2c| + |\epsilon|$ and $\|C(f)\| = \|g + cq_A\| \leq \|g\| + |2c|$. Thus $\|C(f)\| \leq \|f\| - |\epsilon|$ and so, if $\|C(f)\| = \|f\|$ then $\epsilon = 0$ and $f(M) = Ag(A) + cAq_A(A)$. ■

Lemma 2.9: *If $g(A)$ may be written as $A^k g_k(A)$ for all k, then $g(A) = 0$.*

Proof: We have $g_k(A) = A^{-k}g(A)$. By 1.1 $g_k(A) \to 0$. Since $\mathbb{Z}[A]$ is a lattice, eventually $g_k(A) = 0$. Hence $g(A) = 0$. ■

Theorem 2.10: *Let A be an expanding matrix with $q = |\det(A)| = 2$. Let q_A be defined by $p_A(x) = 2 - xq_A(x)$ where p_A is the characteristic polynomial of A. If $\|q_A\| \leq 2$ then G is a lattice for all digit sets $\mathcal{D}$.*

Proof: For any f, $\|C^k(f)\|$ is eventually constant by 2.8 (since $\|q_A\| \leq 2$). By 2.8 and 2.9, $\|C^k(f)\|$ is eventually zero. But if $h = C(g)$ belongs to G then so does g since, by definition, $g(A) = Ah(A) + \epsilon I$. Since $0 \in G$, we conclude that $f(A) \in G$. Thus $\mathcal{G} = G$ and G is a lattice. ■

We now obtain a criterion for a polynomial of the form $\pm x^\ell \pm x^k - 2$ to be expanding. Not all such are expanding (for example, $x^2 + x - 2$ is not) but we do have the following simple test.

Lemma 2.11: *Let $p = \delta x^\ell + \epsilon x^k - 2$ where $\ell > k > 0$, $\delta = \pm 1$ and $\epsilon = \pm 1$. If the equations $\delta x^\ell = 1$ and $\epsilon x^k = 1$ have no common solution in $\mathbb{C}$ then p is expanding.*

Furthermore, $\frac{x^\ell + x^k - 2}{x^c - 1}$ is expanding where $c = \gcd(\ell, k)$.

Proof: Suppose $p(\alpha) = 0$ and $|\alpha| \leq 1$. Then $|\delta\alpha^\ell| \leq 1$ and $|\epsilon\alpha^k| \leq 1$. Hence $\delta\alpha^\ell = 1$ and $\epsilon\alpha^k = 1$, as required.

Let now $g = x^\ell + x^k - 2$. It is easily seen that g does not have any multiple roots. If $g(\alpha) = 0$ and $|\alpha| \leq 1$ then $\alpha^\ell = \alpha^k = 1$. Hence $\alpha^c = 1$. Thus $\frac{x^\ell + x^k - 2}{x^c - 1}$ is expanding. ■

Theorem 2.12: *If A has characteristic polynomial $\frac{x^\ell + x^k - 2}{x^c - 1}$ where $c = \gcd(\ell, k)$ then G is a lattice. The same is true if we replace A by $-A$.*

Proof: Since $x^\ell + x^k - 2$ has no repeated roots, $x^c - 1$ and $p = \frac{x^\ell + x^k - 2}{x^c - 1}$ are prime. Let Z be the $c \times c$ matrix

$$Z = \begin{pmatrix} 0 & 0 & \cdots & & 1 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{pmatrix}.$$

The minimum polynomial of Z is $x^c - 1$. Let

$$M = \begin{pmatrix} A & 0 \\ 0 & Z \end{pmatrix},$$

which we write as $A \oplus Z$. For $f(M) \in \mathbb{Z}[M]$ define $C(f(M))$ as before, namely as $M^{-1}(f(M) - \epsilon I)$, that is $A^{-1}(f(A) - \epsilon I) \oplus Z^{-1}(f(Z) - \epsilon I)$. Observe that the ϵ we get from $\mathbb{Z}[M]$ need not be the ϵ we would get for C on $\mathbb{Z}[A]$. Lemma 2.8 continues to hold since no assumption concerning eigenvalues was made there. As for 2.9, suppose $g(M) = M^i g_i(M)$ for all i. Since $g(M) = g(A) \oplus g(Z)$ we have $g(A) = A^i g_i(A)$ for all i and, therefore, $g(A) = 0$ since A is expanding. Thus any $f(A)$ belongs to G and $G = \mathbb{Z}[A]$ is a lattice. ■

As an application, we have:

Theorem 2.13: *If $q = 2$ and $n \leq 3$ then G is a lattice.*

Proof: For $q = 2$, $n = 2$ there are six possibilities; their characteristic polynomials are: $X^2-2, X^2+2, X^2-X+2, X^2+X+2, X^2-2X+2, X^2+2X+2$. All but the last two are covered by Theorem 2.10; X^2+2X+2 follows from 2.12 (for $c = 1$) and X^2-2X+2 by replacing A by $-A$. Similarly, for $n = 3$ we have fourteen possibilities: $X^3+2, X^3-X+2, X^3+X^2+2, X^3-X^2-X+2, X^3+X^2+X+2, X^3-2X+2, X^3+2X^2+2X+2$ and seven more obtained by reversing the signs of the even degree terms. The only cases not covered by Theorem 2.10 or 2.12 are $X^3-X^2-X+2, X^3+X^2-X-2, X^3-2X+2$ and X^3-2X-2 and these four cases are easily checked by the algorithm of the next section.

■

3. Computations

Let A be a fixed $n \times n$ integer expanding matrix and $\mathcal{D}$ be a set of $q = |\det(A)|$ elements of $\mathbb{Z}^n$, including 0, such that the difference of two distinct elements of $\mathcal{D}$ is never in $A\mathbb{Z}^n$. We next describe an algorithm to determine if in this situation G is a lattice.

Let $\mathcal{G}'$ be the lattice generated by $\mathcal{D}, A\mathcal{D}, \ldots, A^{n-1}\mathcal{D}$; we claim that $\mathcal{G} = \mathcal{G}'$ where $\mathcal{G}$ is the smallest lattice containing $\mathcal{D}$ with $A\mathcal{G} \subseteq \mathcal{G}$. Indeed, since $\mathcal{D}, A\mathcal{D}, \ldots, A^{n-1}\mathcal{D} \subseteq G$, it follows that $\mathcal{G}' \subseteq \mathcal{G}$. On the other hand, if $u \in \mathcal{G}'$ it follows that $Au \in \mathcal{G}'$ (since the minimum polynomial of A has degree at most n) and therefore $Au + v - v' \in \mathcal{G}'$ for any $v, v' \in \mathcal{D}$. Thus, $A\mathcal{G}' \subseteq \mathcal{G}'$ and $\mathcal{G} \subseteq \mathcal{G}'$.

It is now easy to check whether $\mathcal{G} = \mathbb{Z}^n$: start with the vectors $\mathcal{D}, A\mathcal{D}, \ldots, A^{n-1}\mathcal{D}$ and try to get the canonical basis by linear combinations. By a linear change of coordinates, we can assume $\mathcal{G} = \mathbb{Z}^n$.

We would now like to consider a bounded set X with the property that if $Au+v-v' \in X$ then $u \in X$, or, equivalently, that $\tau X \subseteq X$. Such sets clearly exist (e.g., $X = Q$) but they are not always easy to obtain. In particular, for certain matrices A, X may not be taken as a round ball or cube around the origin, however large. We could work with somewhat more complicated bounded sets but we prefer to work instead with *two* sets. Let therefore $0 < N_1 < N_2$ be such that:

i. $Q \subseteq [-N_1, N_1]^n$,

ii. if $u \notin [-N_2, N_2]^n$ then the forward orbit of u by $w \mapsto Aw + v - v'$, $v, v' \in \mathcal{D}$, never enters $[-N_1, N_1]^n$.

It is easy to see that N_1 and N_2 as above exist and we now show, given A and a bound on the size of the elements of $\mathcal{D}$, how to obtain such numbers. Let ℓ_m and ℓ_M be two positive numbers with the following properties: $1 < \ell_m < \ell_M$ and for any eigenvalue λ of A, $\ell_m < |\lambda| < \ell_M$; since A is an expansion, it is clearly possible to choose such numbers. We can now choose an invertible matrix M with $\ell_m|u| \leq |MAM^{-1}u| \leq \ell_M|u|$ for all u. Defining $\|u\| = |Mu|$, this becomes $\ell_m\|u\| \leq \|Au\| \leq \ell_M\|u\|$. Thus, if $r = \max_{v\in\mathcal{D}} \|v\|$, we have $\|u\| \leq r/(\ell_m - 1)$ for all $u \in Q$ and we can take any N_1 such that the cube $[-N_1, N_1]^k$ contains all points u with $\|u\| \leq r/(\ell_m - 1)$. Once N_1 is fixed, take any N_2 such that all u with $\|u\| \leq \max_{w\in[-N_1,N_1]^k} \|w\|$ belong to the cube $[-N_2, N_2]^k$.

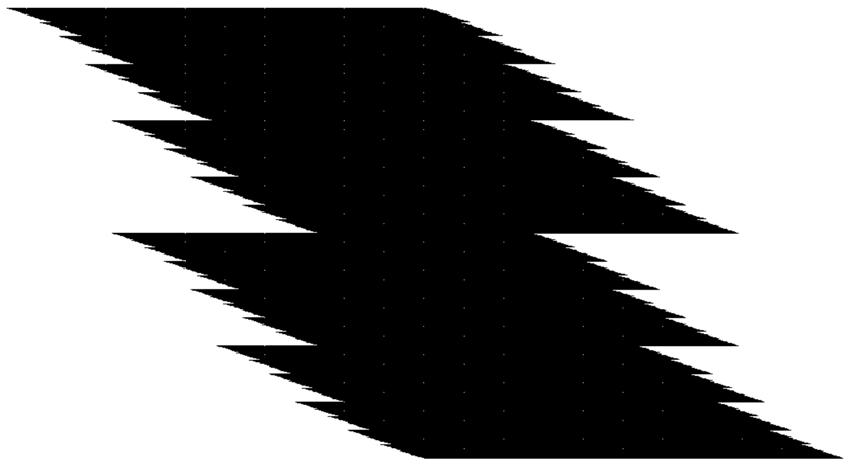

Figure 1

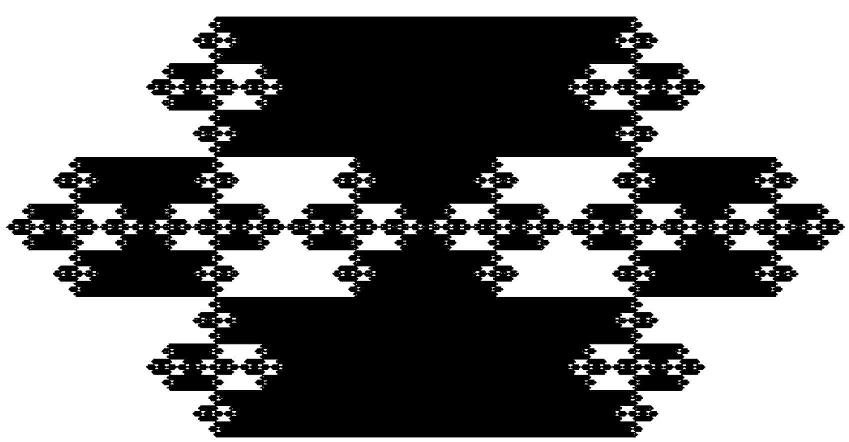

Figure 2

After N_1 and N_2 have been chosen, reserve a bit of memory for every integral element of the cube $[-N_2, N_2]^n$ to indicate whether that element is known to be in G. Start with only the bit for the zero vector turned on. Perform then the following process: for each vector u whose associated bit is on, turn on all vectors of the form $Au + d_1 - d_2$, $d_1, d_2 \in \mathcal{D}$. A second bit associated to each vector indicates whether this process has already been carried out for it. The process stops when no vector has only one of the two associated bits turned on; let G_* be the set of vectors marked at the end. The properties of N_1 and N_2 guarantee, however, that $G_* \cap [-N_1, N_1]^n = G \cap [-N_1, N_1]^n$ (notice, however, that we usually do not have $G_* \cap [-N_2, N_2]^n = G \cap [-N_2, N_2]^n$). Since $Q \subseteq [-N_1, N_1]^n$ and we assume $\mathcal{G} = \mathbb{Z}^n$, G is a lattice iff $G_* \cap [-N_1, N_1]^n = \mathbb{Z}^n \cap [-N_1, N_1]^n$.

This algorithm was applied to various random matrices and digit sets and G always turned out to be a lattice. This suggests that the examples of Lagarias and Wang (where G is not a lattice) must be relatively rare. Also, one example from each conjugacy class of 3×3 expanding integer matrices A with $q = |\det(A)| = 2$ was tested, thus completing the proof that, for $n \leq 3$ and $q = 2$, G is a lattice.

Figure 3

We present pictures of various tiles in order to illustrate some of their properties. Figure 1 shows Q for

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \qquad \mathcal{D} = \{(0,0), (1,0), (0,1), (1,1)\};$$

notice that Q is connected and simply connected; there are density points of Q on the boundary. Figure 2 shows Q for

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix},$$
$$\mathcal{D} = \{(-1,-1), (0,-1), (1,-1), (-2,0), (0,0), (2,0), (-1,1), (0,1), (1,1)\};$$

notice that Q is connected but not simply connected. Figure 3 shows Q for

$$A = \begin{pmatrix} 0 & 1 \\ -3 & 2 \end{pmatrix}, \qquad \mathcal{D} = \{(0,0),(1,0),(-1,5)\};$$

it is easy to show that Q has infinitely many connected components, each with infinitely many holes.

Bibliography

[B] Bandt, C., *Self-similar sets 5 — Integer matrices and fractal tilings of* $\mathbb{R}^n$, Proc. Amer. Math. Soc., 112, 549-562.

[F1] Falconer, K. J., *Fractal geometry*, John Wiley and sons, Chichester, 1990.

[F2] Falconer, K. J., *The geometry of fractal sets*, Cambridge University Press, Cambridge, 1985.

[Gi] Gilbert, W. J., *Geometry of Radix Representations*, in *The geometric vein: the Coxeter festschrift*, edited by Davis, C., Grünbaum, B. and Sherk, F. A., Springer-Verlag, New York, 1981.

[Gr] Gröchenig, K., *Orthogonality criteria for compactly supported scaling functions*, Applied and Computational Harmonic Analysis, 1, 242-245 (1994).

[GrH] Gröchenig, K. and Hass, A., *Self-similar lattice tilings*, Journal of Fourier Analysis, to appear.

[GrM] Gröchenig, K. and Madych, W., *Multiresolution analysis, Haar basis and self-similar tilings*, IEEE Trans. Info. Th., IT-38, No. 2, part II, 556-568 (1992).

[H] Hutchinson, J. E., *Fractals and self-similarity*, Indiana Univ. Math. J., 30 (1981), no. 5, 713-747.

[Ka] Katznelson, Y., *Ergodic automorphisms of* $\mathbb{T}^n$ *are Bernoulli shifts*, Israel Journal of Math., vol. 10 (1971) p. 186-195.

[Ke] Kenyon, R., *Self-replicating tilings*, in *Symbolic dynamics and its applications*, (P. Walters Ed.), Contemporary Math., vol. 135, 239-264 (1992).

[Kn] Knuth, D. E., *The Art of computer programming, vol. 2: seminumerical algorithms*, Addison-Wesley, Reading, Mass. 1969.

[LW1] Lagarias, J. C. and Wang, Y., *Self-affine tiles in* $\mathbb{R}^n$, Advances in Math., to appear.

[LW2] Lagarias, J. C. and Wang, Y., *Integral self-affine tiles in* $\mathbb{R}^n$ *I. Standard and non-standard digit sets*, J. London Math. Soc., to appear.

[LW3] Lagarias, J. C. and Wang, Y., *Integral self-affine tiles in* $\mathbb{R}^n$ *II. Lattice tilings*, preprint.

[M] Mañe, R., *Ergodic theory and differentiable dynamics*, Springer-Verlag, Berlin – New York, 1987.

[O] Odlyzko, A. M., *Non-negative digit sets in positional number systems*, Proc. London Math. Soc., 3rd series, 37, 213-229 (1978).

[S] Strichartz, R. S., *Wavelets and self-affine tilings*, Constructive Approx., to appear.

[V] Vince, A., *Replicating tesselations*, SIAM J. Discr. Math., 6, 501-521 (1993).

Derek Hacon, Depto. de Matemática, PUC-Rio
e-mail: derek@mat.puc-rio.br

Nicolau C. Saldanha, IMPA and PUC-Rio
e-mail: nicolau@impa.br; http://www.impa.br/~nicolau/

J. J. P. Veerman, Depto. de Matemática, UFPe
e-mail: veerman@dmat.ufpe.br

PUC-Rio, Depto. de Matemática,
Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ 22453-900, BRAZIL

IMPA,
Estr. Dona Castorina, 110, Rio de Janeiro, RJ 22460-320, BRAZIL

UFPE, Depto. de Matemática,
Cidade Universitaria, Rua Prof. Luis Freire, Recife, PE 50740-540,BRAZIL