Remarks on Self-Affine Tilings

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

CONTENTS

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 $n \le 3$, is a lattice-tile.

INTRODUCTION

We study certain tilings of $\mathbb{R}^n$ defined by "generalized decimal expansions" in which the base ten and the digits $0, \dots, 9$ 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

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

where $N \in \mathbb{Z}$ and all the v_i belong to $\mathcal{D}$. To guarantee convergence, A is assumed to be expanding, that is, all its eigenvalues have absolute value greater than 1. Since A is an integer matrix, 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 [Lagarias and Wang a], $\mathcal{D}$ 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 $k \ge 0$ and $v_i \in \mathcal{D}$, and Q to be the set of all infinite sums $A^{-1}v_{-1} + A^{-2}v_{-2} + \cdots$, for $v_i \in \mathcal{D}$. In the usual decimal expansion, Q is the unit interval and $I = \{0, 1, 2, \dots\}$. The translates of Q by elements of I tile some subset of $\mathbb{R}^n$. By enlarging I we can obtain a set G' such that the translates of Q by elements of G' tile all of $\mathbb{R}^n$ (see Proposition 1.5). In fact, $G' \subseteq G = I - I$, the set of differences of elements of I.

All authors receive support from CNPq and MCT, Brazil.

1058-6458/96 $0.50 per page

FIGURE 1. Lattice-tile for $A = \left(\begin{smallmatrix} 0 & 1 \\ -3 & 2 \end{smallmatrix}\right)$ and $\mathcal{D} = \{(0,0),(1,0),(-1,5)\}$.

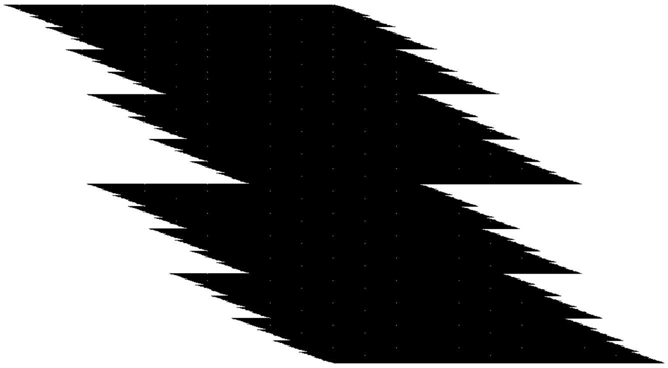

FIGURE 2. Lattice-tile for $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ and $\mathcal{D} = \{(0,0),(1,0),(0,1),(1,1)\}$.

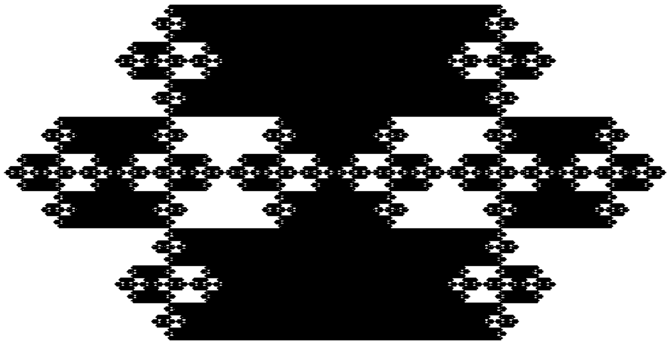

FIGURE 3. Lattice-tile for $A = \left(\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}\right)$ and $\mathcal{D} = \{(-1,-1),(0,-1),(1,-1),(-2,0),(0,0),(2,0),(-1,1),(0,1),(1,1)\}$.

The following simple example 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, and in this case $G' = G$. If G is a lattice, we call Q a *lattice-tile*. It is shown in [Gröchenig 1994; Gröchenig and Haas] that G is always a lattice for $n = 1$. On the other hand, [Lagarias and Wang a, Example 2.3] shows that G is not always a lattice for $n > 1$.

The tile Q itself can be intriguingly complex. Figure 1 shows Q for $A = \left(\begin{smallmatrix} 0 & 1 \\ -3 & 2 \end{smallmatrix}\right)$ and $\mathcal{D} = \{(0,0), (1,0),(-1,5)\}$. One can show that Q has infinitely many connected components, each with infinitely many holes.

By contrast, Figure 2 shows Q for $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ and $\mathcal{D} = \{(0,0),(1,0),(0,1),(1,1)\}$. This tile is connected and simply connected; there are density points of Q on the boundary.

As a last example, Figure 3 shows a case where Q is connected but not simply connected.

The rest of this article is structured as follows. 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 [Bandt 1991, Gröchenig and Haas, Gröchenig and Madych 1992, Kenyon 1992, Lagarias and Wang a–c, Vince 1993]. In Section 2 we investigate certain aspects of the case $q = 2$: we prove that Q is always connected, and that G is a lattice for $n \leq 3$. Finally, in Section 3, we derive an algorithm for checking if G is a lattice given A and $\mathcal{D}$ (see [Vince 1993] for another algorithm).

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 [Lagarias and Wang a]. Many 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, that is, one whose eigenvalues have absolute value greater than 1. We write $\det A = \pm q$, where q is a positive integer. Reducing A to Jordan canonical form, we see that, for any bounded set $B \subset \mathbb{R}^n$, 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 $0 \in \mathcal{D}$ and $\mathcal{D} - \mathcal{D} \cap A\mathbb{Z}^n = 0$; in other words, the elements of $\mathcal{D}$ are distinct modulo $A\mathbb{Z}^n$. It follows that $\mathbb{Z}^n$ is the disjoint union of the q cosets $r + A\mathbb{Z}^n$, for $r \in \mathcal{D}$.

(Throughout this paper $X + Y$ will denote the set of all sums $x + y$ with $x \in X$, $y \in Y$, and likewise $X - Y$. Infinite sums are defined if all sets contain 0: 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 $\mathbb{Z}^n = \mathcal{D} + A\mathbb{Z}^n$, we have $\mathbb{Z}^n = \mathcal{D} + A\mathcal{D} + A^2\mathbb{Z}^n$ and so on. We write

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

This gives a unique representation for each element of I, because if $r_1 + \cdots + A^k r_k = s_1 + \cdots + A^k s_k$ we get $r_1 = s_1$ modulo $A\mathbb{Z}^n$, so $r_1 = s_1$ and $r_i = s_i$ for all i by induction. As already remarked, for $n = 1$, $A = 3$ and $\mathcal{D} = \{0, 4, 11\}$ we have $I = \{0, 4, 11, 12, 16, \ldots\}$.

Now set $\tau Z = A^{-1}(\mathcal{D} + Z)$, for any $Z \subseteq \mathbb{R}^n$. 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 $AQ_1 \subset A^2 Q_2 \subset \cdots \subset I$. The compact set

$$\overline{\bigcup_k Q_k}$$

is invariant under τ. Following [Hutchinson 1981] (compare [Falconer 1985] and [Lagarias and Wang a]), we examine how τ acts on the compact subsets of $\mathbb{R}^n$ in order to characterize this invariant set.

Recall that, for any metric space X, the *Hausdorff metric* on the space of compact nonempty subsets of X is defined by

$$d(K, L) = \inf\{\varepsilon \mid K \subseteq N_\varepsilon(L) \text{ and } L \subseteq N_\varepsilon(K)\},$$

where $N_\varepsilon(K)$ is the open ε-neighborhood of K. Let $H(X)$ be this space of subsets with the Hausdorff metric. It is well known that, if X is complete, so is $H(X)$.

Clearly, τ maps $H(\mathbb{R}^n)$ into itself. We know that some power τ^N of τ is a contraction; let Q be its unique fixed point. 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. Furthermore, if we apply τ repeatedly to any $K \in H(\mathbb{R}^n)$, the iterates tend to Q in the Hausdorff metric (consider the subsequences of the form $k = jN + k_0$, for k_0 fixed and $j \to \infty$). Therefore:

Lemma 1.1. *The map $\tau : H(\mathbb{R}^n) \to H(\mathbb{R}^n)$ has a unique fixed point Q, and $Q = \lim_{k\to\infty} \tau^k K$ for any $K \in H(\mathbb{R}^n)$.* □

In particular, taking $K = \{0\}$, we get $\tau^k K = Q_k$, so $Q = \overline{\bigcup_k Q_k}$ 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 multiplication by A takes $\mathbb{Z}^n$ to itself, it induces a map of $\mathbb{T}^n$ to itself, also denoted A. Now A^{-1} acts on subsets of $\mathbb{T}^n$ taking B to $A^{-1}B$, so it induces a map $A^{-1} : H(\mathbb{T}^n) \to H(\mathbb{T}^n)$. Also π induces a map $\pi : H(\mathbb{R}^n) \to H(\mathbb{T}^n)$, which is continuous; indeed, π decreases distances. Since the digits form a complete set of representatives modulo $A\mathbb{Z}^n$, we see that $\pi\tau = A^{-1}\pi$. If we take a compact set $K \subset \mathbb{R}^n$ that projects surjectively to $\mathbb{T}^n$ and apply τ repeatedly, we get $\pi Q = \pi \lim_k \tau^k K = \lim_k \pi\tau^k K = \lim_k A^{-k}\pi K = \mathbb{T}^n$, showing that $\mathbb{Z}^n + Q = \mathbb{R}^n$ as claimed.

By Baire's theorem, this also implies the following result (compare [Lagarias and Wang a, Theorem 1.1]):

Proposition 1.2. *Q has nonempty interior.* □

We next look at the self-similarity properties of Q (compare [Falconer 1985]). Let $|X|$ be Lebesgue measure of X; all sets considered will be measurable. We say that two sets K and L *overlap* if their intersection has positive measure.

Lemma 1.3. *The translates of $A^{-k}Q$ by elements of Q_k do not overlap. The translates of Q by elements of I do 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 first assertion follows. For the second, note that if $u, v \in I$, there is some k such that $u, v \in A^k Q_k$. If u and v are distinct, $A^{-k}u + A^{-k}Q$ and $A^{-k}v + A^{-k}Q$ don't overlap, and therefore neither do $u + Q$ and $v + Q$. □

Definition 1.4. Let $X, \mathcal{H}, Y \subset \mathbb{R}^n$ be measurable sets, where X is bounded with nonempty interior. We say that $\mathcal{H} + X$ is a *tiling* of Y if $Y = \mathcal{H} + X$ and the translates of X by elements of $\mathcal{H}$ don't overlap.

We now define G as the set $I - I$ of differences of I. From Proposition 1.2 and the self-similarity of Q (Lemma 1.3), we can recover the following tiling property [Gröchenig and Haas].

Proposition 1.5. *There exists a subset G' of G such that the translates of Q by elements of G' tile $\mathbb{R}^n$. Furthermore, $G' - G' \subseteq G$.*

Proof. We know that Q has nonempty interior. Given $r > 0$, we use the expansiveness of A to find k such that a ball of radius $r + \operatorname{diam} Q$ fits inside $A^k Q$. But $A^k Q = A^k Q_k + Q$, and the translates of Q by elements of $A^k Q_k$ are nonoverlapping, by Lemma 1.3. Thus, for some element v_r of $A^k Q_k$, the translates of Q under $G_r = A^k Q_k - v_r \subset G$ tile (a superset of) a ball of radius r around the origin.

We now have tilings of arbitrarily large regions around the origin; we will use them to 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 $G_1^1, G_2^1, \dots$ of $G_1, G_2, \dots$, all meeting B_1 in the same set. For $s = 2$, we take a subsequence $G_1^2, G_2^2, \dots$ of $G_1^1, G_2^1, \dots$, all meeting B_2 in the same set. We repeat the process for increasing s. Define G' to be the set of elements of G belonging to all but finitely many of the sets $G_1^1, G_1^2, G_1^3, \dots$: by construction, $G' + Q$ is a tiling of $\mathbb{R}^n$.

To prove that $G' - G' \subseteq G$, we observe that $G_r - G_r = Q_{k_r} - Q_{k_r} \subseteq I - I = G$. The result follows since any finite subset of G' is contained in a translate of some G_r. □

Next we consider how τ acts on the measurable subsets of Q. If K and L are measurable sets, we write $K \equiv L$ if the symmetric difference $(K \cup L) \setminus (K \cap L)$ has measure zero. We say that K is *τ-invariant* if $K \equiv \tau K$.

We first observe that, $Z \subseteq Q$ implies $|\tau Z| = |Z|$. For, setting $Y = Q \setminus Z$, we have $|\tau Z| \le |Z|$ and $|\tau Y| \le |Y|$ by the definition of τ. But $\tau Z \cup \tau Y = \tau Q = Q = Z \cup Y$, so Z and τZ have the same measure. Thus, if $Z \subseteq \tau Z$ or if $\tau Z \subseteq Z$, the set Z is τ-invariant.

At the same time, we have the following ergodicity result (see the end of this section):

Proposition 1.6. *Any τ-invariant subset of Q has measure 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 such that

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

Thus, given $\delta > 0$, we have

$$\frac{|Z \cap N_\varepsilon(x)|}{|Q \cap N_\varepsilon(x)|} \ge 1 - \tfrac{1}{2}\delta$$

for $\varepsilon > 0$ sufficiently small. 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 Lemma 1.3, the tiny sets $v + A^{-k}Q$, for v in some subset of Q_k, cover $N_\varepsilon(x)$, except for a narrow margin around the boundary of $N_\varepsilon(x)$ (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 $1 - \frac{1}{2}\delta$, a contradiction.

By hypothesis, $Z \equiv \tau^k Z$. Hence $Z \equiv Q_k + A^{-k}Z$. Again by Lemma 1.3, we have

$$(Q_k + A^{-k}Z) \cap (v + A^{-k}Q) \equiv v + A^{-k}Z.$$

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 1.7. *Except on a set of measure zero,* $\pi : Q \to \mathbb{T}^n$ *is* f*-to-one for some fixed integer* f.

Proof. Let $Q^{(k)}$ be the set of $x \in Q$ such that $x + \mathbb{Z}^n$ 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, take $x, y \in Q$ such that $x - y \in \mathbb{Z}^n$. Since $Q = \tau Q$, we have $x = A^{-1}(x' + r)$ and $y = A^{-1}(y' + s)$ 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 Proposition 1.6, $|Q^{(k)}|$ is 0 or $|Q|$. Now take f 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 [Lagarias and Wang a] for more general digit sets.

Two other consequences are worth noticing: the Lebesgue measure of Q is always an integer, and the boundary of Q has measure zero (since $\mathring{Q}$ has positive measure and maps inside itself under τ.)

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

Proposition 1.8. G *is a lattice if and only if the translates of* Q *by elements of* G *tile* $\mathbb{R}^n$.

Proof. Assume G is a lattice, and take v, w in G. By hypothesis, $v - w = x - y$ where $x, y \in I$. If v and w are distinct, $x + Q$ and $y + Q$ do not overlap, by Lemma 1.3, so neither do $v + Q$ and $w + Q$.

Conversely, if the G-translates of Q don't overlap, $G = G'$ since $G' + Q$ is a tiling. Thus, by Proposition 1.5, $G - G \subseteq G$ and G is a lattice. □

From Propositions 1.5 and 1.8 we get:

Theorem 1.9. *If* G *is a lattice,* $G + Q$ *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 1.10. $G \cap Q = \mathbb{Z}^n \cap Q$ *implies* $G = \mathbb{Z}^n$.

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 our earlier example, 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 $[0, \frac{11}{2}]$. 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 Lemma 1.10, $G = \mathbb{Z}$. From Theorem 1.9, $\mathbb{Z} + Q$ is a tiling and $|Q| = 1$.

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

Proposition 1.11. *G is a lattice if and only if $G \cap Q = \mathcal{G} \cap Q$.*

Proof. G a lattice implies $G \cap Q = \mathcal{G} \cap Q$ trivially. Conversely, assume $G \cap Q = \mathcal{G} \cap Q$: since Q is bounded and $G' + Q = \mathbb{R}^n$, the set 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 $\mathcal{G}$ is contained in $\mathbb{Z}^n$ and spans $\mathbb{R}^n$ since G' does; thus $\mathcal{G}$ 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 Lemma 1.10 is well-defined. Now follow the proof of Lemma 1.10 with $\mathcal{G}$ instead of $\mathbb{Z}^n$. $\square$

The question whether or not G is a lattice has been settled in certain cases.

Theorem 1.12 [Gröchenig 1994]. *G is always a lattice if $n = 1$.*

Example 1.13 [Lagarias and Wang a]. G is not a lattice for

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

For the reader's convenience, we include a proof of Theorem 1.12, based on arguments in [Gröchenig 1994; Gröchenig and Haas]. We first present a series of auxiliary definitions and results. Consider the n-dimensional case for a moment. We denote by $\chi_k(g)$ the coefficient of $z^k = z_1^{k_1} \cdots z_n^{k_n}$, for $k \in \mathbb{Z}^n$, in a Laurent series $g \in \mathbb{R}[z_1, z_1^{-1}, \ldots, z_n, z_n^{-1}]$.

The key to the proof is the introduction of the *tiling polynomial* T, defined by $\chi_k(T) = |Q \cap (k + Q)|$. Clearly, T is constant if and only if the translates of Q under elements of $\mathbb{Z}^n$ do not overlap.

We set $D = \sum_{d \in \mathcal{D}} z^d$, and $\bar{D} = \sum_{d \in \mathcal{D}} z^{-d}$. (The notation comes from the fact that $\bar{D}(z) = \overline{D(z)}$ when $|z| = 1$.)

We first show that

$$\chi_k(qT) = \chi_{Ak}(D\bar{D}T) \tag{1.1}$$

for all $k \in \mathbb{Z}^n$, where, as we recall, $q = |\det A|$. By Lemma 1.3 and the fact that $Q = \tau Q$, we have

$$\begin{aligned} \chi_k(qT) &= q\,|Q \cap k{+}Q| = |AQ \cap Ak{+}AQ| \\ &= |A\tau Q \cap Ak{+}A\tau Q| = |\mathcal{D}{+}Q \cap Ak{+}\mathcal{D}{+}Q| \\ &= \sum_{d,d' \in \mathcal{D}} |d{+}Q \cap Ak{+}d'{+}Q| \\ &= \sum_{d,d' \in \mathcal{D}} |Q \cap Ak{+}d' - d{+}Q| = \chi_{Ak}(D\bar{D}T), \end{aligned}$$

by the definition of D and T.

Given g, we define $\hat{g}$ by $\chi_k(\hat{g}) = \chi_{Ak}(D\bar{D}g)$. Thus (1.1) may be rewritten as $\hat{T} = qT$. Notice also that $\hat{1} = q$, since $\mathcal{D}$ is a complete set of residues modulo $A\mathbb{Z}^n$.

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

Lemma 1.14. *For $n = 1$ and $|z| = 1$ we have*

$$q\hat{g}(z) = \sum_{w^q = z} g(w)|D(w)|^2.$$

Proof. Write $\tilde{g} = D\bar{D}g$. Then

$$\tilde{g}(z) = \hat{g}(z^q) + zg_1(z^q) + \cdots + z^{q-1}g_{q-1}(z^q)$$

for appropriate $g_1, \ldots, g_{q-1}$. Setting $w^q = z$, we get $\tilde{g}(w) = \hat{g}(z) + wg_1(z) + \cdots + w^{q-1}g_{q-1}(z)$. But $\sum_{w^q=z} w^j = 0$ for $0 < j < q$. It follows that $q\hat{g}(z) = \sum_{w^q=z} \tilde{g}(w)$. $\square$

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

$$\sum_{w^q=z} |D(w)|^2 = q^2 \tag{1.2}$$

and

$$\sum_{w^q=z} |D(w)|^2 \big(T(w) - T(z)\big) = 0. \tag{1.3}$$

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 1.15. *If T is nonconstant*, $\gcd(\mathcal{D}) > 1$.

Proof. The set E of extrema of T in $\mathbb{S}^1$ is finite and has at least two points. Also, if $z \in E$, the difference $T(z) - T(y)$ has the same sign for all $y \in \mathbb{S}^1$. Take $z \in E$ distinct from 1. By (1.2) and (1.3), there is some w_0 with $w_0^q = z$, 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 (1.3), $|D(w)| = 0$ for such w and, from (1.2), $|D(w_0)| = q$ 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 $w_0^q = z \neq 1$. $\square$

Proof of Theorem 1.12. By the previous lemma, if $\gcd(\mathcal{D}) = 1$ the translates of Q by $\mathbb{Z}$ do not overlap. Since $\mathbb{Z} + Q = \mathbb{R}$, the translates of Q by $\mathbb{Z}$ tile $\mathbb{R}$. 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. $\square$

We close this section by describing briefly how the results of this section are related to the study of expanding toral epimorphisms [Katznelson 1971; Mañé 1987]. Katznelson determined which toral epimorphisms are Bernoulli in terms of the eigenvalues of A, where A is the integral matrix representing the epimorphism $\bar{A}$. 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 $(1/q, \dots, 1/q)$, where $q = |\det A|$. In fact, in many cases the generalized decimal expansion (base A) provides an equivalence between $\bar{A}$ on $\mathbb{T}^n$ and a one-sided Bernoulli shift. Because $Q = A^{-1}\mathcal{D} + A^{-1}Q$ and the translates are nonoverlapping, 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 $\pi \circ S = \bar{A} \circ \pi$ whenever S is defined. By Theorem 1.9, $\pi : Q \to \mathbb{R}^n/G$ is an equivalence between (Q, S) and $(\mathbb{R}^n/G, \bar{A})$, provided G is a lattice. In particular, when G is a lattice, Proposition 1.6 can be deduced from the known fact [Mañé 1987] that $\bar{A}$ 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. The case $q = 2$ has certain special features that we now explore. We begin by showing that Q is connected by constructing a space-filling curve in Q. We then prove two theorems (2.5 and 2.6), which guarantee that in many cases G is a lattice.

From Lemma 1.3, we have $Q = Q_k + A^{-k}Q$. Thus Q is the union of 2^k *k-pieces*, each of the form $w + A^{-k}Q$, for $w \in Q_k$. We start by remarking that the two 1-pieces $A^{-1}Q$ and $A^{-1}v + A^{-1}Q$ have nonempty intersection. For, if they were disjoint, so would be the four 2-pieces, and, inductively, the 2^k k-pieces. Now Q has nonempty interior and therefore contains a ball, which is covered by k-pieces whose diameter can be taken smaller than that of the ball. This would contradict the fact that the ball is connected.

We now 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. Let

$$J_k = [0,1] \cap \mathbb{Z}/6^k = \{0, 1/6^k, \dots, 1\}.$$

We say that a function $\gamma : J_k \to Q$ is *admissible* if there is at least one point of $\gamma(J_k)$ in the interior of each k-piece, and for any two consecutive points r, s of J_k there is some k-piece containing both $\gamma(r)$ and $\gamma(s)$.

Any admissible $\gamma : J_k \to Q$ extends to an admissible $\tilde{\gamma} : J_{k+1} \to Q$. To see this, 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 assumption, $\gamma(a_0)$ 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. Arbitrarily choose $\tilde{\gamma}(a_1), \tilde{\gamma}(a_3), \tilde{\gamma}(a_5) \in P_0 \cap P_1$ (which is nonempty since the two 1-pieces of Q intersect), and further choose $\tilde{\gamma}(a_2) \in \mathring{P}_0$ and $\tilde{\gamma}(a_4) \in \mathring{P}_1$. This defines $\tilde{\gamma}$, which is easily seen to be admissible.

Theorem 2.1. *Q is path-connected if $q = 2$. 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 preceding observation 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: it contains points in the interior of each k-piece for all k. $\square$

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 $n \le 3$.

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, the conditions $g(A)v = 0$ and $g(A) = 0$ are equivalent. 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 absolute value 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 in the form $f(A)$, with f reduced. The set G corresponds to the set of all polynomials (reduced or not) with coefficients 0, 1 or -1. This gives the following result:

Proposition 2.2. *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 $\mathcal{D}$, it 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

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

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 $2I = Aq_A(A)$, we see that, if f is reduced, $f(A) \in A\mathbb{Z}[A]$ if and only if $f(0)$ is even. We now define a *carrying operation* $C : \mathbb{Z}[A] \to \mathbb{Z}[A]$ (compare Lemma 1.10). If f is reduced there is a unique $\varepsilon \in \{0, 1, -1\}$ such that

$$f = xg + 2c + \varepsilon \quad \text{with } g \in \mathbb{Z}[x] \text{ and } c \in \mathbb{Z} \tag{2.1}$$

and $|2c + \varepsilon| = |2c| + |\varepsilon|$. We define $C(f)$ as the (reduced) polynomial $g + cq_A$.

Lemma 2.3. *If $\|q_A\| \le 2$ then $\|C(f)\| \le \|f\|$, and equality implies $f(A) = Ah(A)$ for some h.*

Proof. Clearly $\deg(g) \le n-2$ in (2.1). Thus $\|f\| = \|xg\| + |2c+\varepsilon| = \|g\| + |2c| + |\varepsilon|$ and $\|C(f)\| = \|g + cq_A\| \le \|g\| + |2c|$. Thus $\|C(f)\| \le \|f\| - |\varepsilon|$. It follows that $\|C(f)\| = \|f\|$ implies $\varepsilon = 0$ and $f(M) = Ag(A) + cAq_A(A)$. $\square$

Lemma 2.4. *If $g(A)$ may be written as $A^k g_k(A)$ for all k, then $g(A) = 0$.*

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

Theorem 2.5. *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 $\|q_A\| \le 2$ then G is a lattice for all digit sets $\mathcal{D}$.*

Proof. For any f, the sequence $\|C^k(f)\|$ is eventually constant by Lemma 2.3 (since $\|q_A\| \le 2$). By Lemmas 2.3 and 2.4, $\|C^k(f)\|$ is eventually zero. But if $h = C(g)$ belongs to G then so does g since, by definition, $g(A) = Ah(A) + \varepsilon I$. Since $0 \in G$, we conclude that $f(A) \in G$. Thus $\mathcal{G} = G$ and G is a lattice. $\square$

Now consider polynomials of the form $p(x) = \delta x^l + \varepsilon x^k - 2$, where $l > k > 0$, $\delta = \pm 1$ and $\varepsilon = \pm 1$. We obtain a criterion for such a polynomial to be expanding. (An example that is *not* expanding is $x^2 + x - 2$.) If p is not expanding, let α be a root with $|\alpha| \leq 1$. Then $|\delta\alpha^l| \leq 1$ and $|\varepsilon\alpha^k| \leq 1$. Hence $\delta\alpha^l = 1$ and $\varepsilon\alpha^k = 1$. Thus, if the equations $\delta x^l = 1$ and $\varepsilon x^k = 1$ have no common solution over the complex numbers, p is expanding.

By the same token, the polynomial

$$\frac{x^l + x^k - 2}{x^c - 1},$$

where $c = \gcd(l, k)$, is always expanding. For one easily sees that the numerator does not have any multiple roots. If α is a root of the numerator with $|\alpha| \leq 1$, then $\alpha^l = \alpha^k = 1$. Hence $\alpha^c = 1$, and α is not a root of the quotient polynomial.

Theorem 2.6. *Let A have characteristic polynomial*

$$p(x) = \frac{x^l + x^k - 2}{x^c - 1},$$

where $c = \gcd(l, k)$. Then G is a lattice. The same is true if we replace A by $-A$.

Proof. Since $x^l + x^k - 2$ has no repeated roots, $x^c - 1$ and p are prime. Consider 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 $M = \left(\begin{smallmatrix} A & 0 \\ 0 & Z \end{smallmatrix}\right)$, which we write as $A \oplus Z$. For $f(M) \in \mathbb{Z}[M]$ define $C(f(M))$ as before, namely

$$\begin{aligned} C(f(M)) &= M^{-1}(f(M) - \varepsilon I) \\ &= A^{-1}(f(A) - \varepsilon I) \oplus Z^{-1}(f(Z) - \varepsilon I). \end{aligned}$$

(The ε we get from $\mathbb{Z}[M]$ need not be the ε we would get for C on $\mathbb{Z}[A]$.) Lemma 2.3 continues to hold since no assumption concerning eigenvalues was made there. As for Lemma 2.4, 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. $\square$

As an application, we have:

Theorem 2.7. *G is a lattice if $q = 2$ and $n \leq 3$.*

Proof. For $q = 2$ and $n = 2$ there are six possibilities, with characteristic polynomials $X^2 - 2$, $X^2 + 2$, $X^2 - X + 2$, $X^2 + X + 2$, $X^2 - 2X + 2$, $X^2 + 2X + 2$. All but the last two are covered by Theorem 2.5. The case $X^2 + 2X + 2$ follows from Theorem 2.6 (with $c = 1$), and $X^2 - 2X + 2$ follows by replacing A by $-A$.

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 terms of even degree. The only cases not covered by Theorems 2.5 and 2.6 are $X^3 - X^2 - X + 2$, $X^3 + X^2 - X - 2$, $X^3 - 2X + 2$ and $X^3 - 2X - 2$, which are easily checked by the algorithm of the next section. $\square$

3. DETERMINING IF G IS A LATTICE

Let A be a fixed $n \times n$ integer expanding matrix, and $\mathcal{D} \ni 0$ a set of coset representatives of $A\mathbb{Z}^n$. We describe an algorithm to determine if in this situation G is a lattice.

Let $\mathcal{G}'$ be the lattice generated by $\mathcal{D}, A\mathcal{D}, \ldots, A^{n-1}\mathcal{D}$, and let $\mathcal{G}$ be the smallest lattice containing $\mathcal{D}$ and satisfying $A\mathcal{G} \subseteq \mathcal{G}$. We claim that $\mathcal{G} = \mathcal{G}'$, 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, $u \in \mathcal{G}'$ implies $Au \in \mathcal{G}'$ (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 $\mathcal{D}, A\mathcal{D}, \dots, A^{n-1}\mathcal{D}$ 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 $\tau X \subseteq X$—equivalently, that $Au + v - v' \in X$ implies $u \in X$. Such sets clearly exist (for example, $X = Q$) 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.

Given A and a bound on the size of the elements of $\mathcal{D}$, we will find numbers $0 < N_1 < N_2$ with the following properties: $Q \subseteq [-N_1, N_1]^n$, and if $u \notin [-N_2, N_2]^n$ the forward orbit of u by $w \mapsto Aw + v - v'$, for $v, v' \in \mathcal{D}$, never enters $[-N_1, N_1]^n$.

Let k_m and k_M be two positive numbers such that $1 < k_m < k_M$ and $k_m < |\lambda| < k_M$ for any eigenvalue λ of A; since A is an expansion, it is clearly possible to choose such numbers. We can now choose an invertible matrix M with

$$k_m|u| \le |MAM^{-1}u| \le k_M|u|$$

for all u. Defining $\|u\| = |Mu|$, this becomes $k_m\|u\| \le \|Au\| \le k_M\|u\|$. Thus, if $r = \max_{v\in\mathcal{D}} \|v\|$, we have $\|u\| \le r/(k_m - 1)$ 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 $\|u\| \le r/(k_m - 1)$. Once N_1 is fixed, take any N_2 such that all u with $\|u\| \le \max_{w\in[-N_1,N_1]^k} \|w\|$ belong to the cube $[-N_2, N_2]^k$.

After N_1 and N_2 have been chosen, reserve one 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 $Au + d_1 - d_2$, for $d_1, d_2 \in \mathcal{D}$. 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 choice of N_1 and N_2 guarantees 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 if and only if

$$G_* \cap [-N_1, N_1]^n = \mathbb{Z}^n \cap [-N_1, N_1]^n.$$

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 $n \le 3$ and $q = 2$, G is a lattice (Theorem 2.7).

ACKNOWLEDGEMENT

We take this opportunity to thank the referee for drawing our attention to a number of references.

REFERENCES

[Bandt 1991] C. Bandt, "Self-similar sets, 5: Integer matrices and fractal tilings of $\mathbb{R}^n$", *Proc. Amer. Math. Soc.* **112** (1991), 549–562.

[Falconer 1985] K. J. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge, 1985.

[Gröchenig 1994] K. Gröchenig, "Orthogonality criteria for compactly supported scaling functions", *Appl. Comp. Harmonic Anal.* **1** (1994), 242–245.

[Gröchenig and Haas] K. Gröchenig and A. Hass, "Self-similar lattice tilings", to appear in *J. Fourier Anal.*

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

[Hutchinson 1981] J. E. Hutchinson, "Fractals and self-similarity", *Indiana Univ. Math. J.* **30** (1981), 713–747.

[Katznelson 1971] Y. Katznelson, "Ergodic automorphisms of $\mathbb{T}^n$ are Bernoulli shifts", *Israel J. Math.* **10** (1971), 186–195.

[Kenyon 1992] R. Kenyon, "Self-replicating tilings", pp. 239–264 in *Symbolic dynamics and its applications* (edited by P. Walters), Contemporary Mathematics **135**, American Mathematical Society, Providence, RI, 1992.

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

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

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

[Mañé 1987] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer, Berlin, 1987.

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

Derek Hacon, Departamento de Matemática, PUC-Rio, Rua Marquês de São Vicente, 225, Rio de Janeiro, RJ 22453-900, Brasil (derek@mat.puc-rio.br)

Nicolau C. Saldanha, Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina, 110, Rio de Janeiro, RJ 22460-320, Brasil (nicolau@impa.br)

J. J. P. Veerman, Departamento de Matemática, Universidade Federal de Pernambuco, Cidade Universitária, Rua Prof. Luis Freire, Recife, PE 50740-540 (veerman@dmat.ufpe.br)

Received May 24, 1994; accepted in revised form February 5, 1995