ALGEBRAIC NUMBERS, FREE GROUP AUTOMORPHISMS
AND SUBSTITUTIONS ON THE PLANE

PIERRE ARNOUX, MAKI FURUKADO, EDMUND HARRISS, AND SHUNJI ITO

Abstract. There has been much recent work on the geometric representation of Pisot substitutions and the explicit construction of Markov partitions for Pisot toral automorphisms. We give a construction that extends this to the general hyperbolic case.

For the sake of simplicity, we consider a simple example of an automorphism of the free group on 4 generators whose associated matrix has 4 distinct complex eigenvalues, two of them of modulus larger than 1 and the other 2 of modulus smaller than 1 (non-Pisot case). Using this generator, we build substitution polygonal tilings of the contracting plane and the expanding plane of the matrix. We prove that these substitution tilings can be lifted in a unique way to stepped surfaces approximating each of these planes. The vertices of each of these stepped surfaces can be projected to an “atomic surface”, a compact set with fractal boundary contained in the other plane.

We prove that both tilings can be refined to exact self-similar tilings whose tiles have fractal boundaries and can be obtained by iteration or by a “cut and project” method by using the atomic surface as the window.

Using the self-similar tiling, one can build a numeration system associated to a complex $\lambda$-expansion; the natural extension of the $\lambda$-expansion associated with this number system is the linear map obtained by abelianization of the free group automorphism. This gives an explicit Markov partition of this hyperbolic toral automorphism.

The fractal domains can be used to define a pseudo-group of translations which gives transversal dynamics in the sense of Vershik (1994) or numeration systems in the sense of Kamae (2005).

The construction can be extended to a larger class of free group automorphisms, each of which can be used to build substitution rules and dynamical systems.

1. Introduction

Substitution rules for tilings provide a sense of scaling symmetry and are therefore of fundamental interest. In the last few years their study has emerged as an independent area from several branches of mathematics, ranging from dynamical systems [Sie02] to algebraic geometry [FHK02]. Of particular interest are the dynamical systems associated to substitution rules and their underlying spaces. These
appear for example in the study of branched manifolds by Williams [Wil74, Wil01] in analogy to the 1-dimensional case of Thurston’s train tracks [Wil67].

An important part of this study has focused on the 1-dimensional case, where substitutions can be seen as positive endomorphisms of free groups, in particular substitutions of Pisot type (substitutions where the associated matrix is one real eigenvalue greater than 1 and all other eigenvalues have absolute value less than 1). For substitutions of Pisot type, in many cases the tiling space is well understood. It is essentially translations on a compact group. It is conjectured that this applies to all such substitutions (cf. [BK06]). A simple section for the tiling flow is given by the closure of the symbolic shift on the fixed point. In addition, this method allows explicit construction of Markov partitions for the associated toral automorphism.

This paper gives an important step in extending this work to the non-Pisot case. We give the explicit construction of a Markov partition for a hyperbolic toral automorphism on $T^4$ associated to a particular automorphism of the free group on 4 generators, using tools introduced in previous papers ([AI01, AIS01, Ei03]). The construction will also produce stepped surfaces approximating the stable and unstable planes of the toral automorphism, and self-similar tilings of these planes, giving rise to a complex $\beta$-expansion and a complex numeration system. Though we treat a single example, it is possible to generalize these results to a large class of examples. The possibilities of generalization, as well as possible obstructions, are considered in the final section. To say it shortly, the essential requirement is the positivity condition; i.e., a suitable extension of the automorphism (defined below) can be written using only positive terms with a good choice of the basis. This generalizes the property for a group automorphism to be expressible as a substitution for a good choice of the basis, which is essential for the Pisot case. If this positivity condition is realized, the same tools as in the Pisot case can be used. As we remark in the last section, there are some possibilities to obtain similar results in the non-positive case, by using “blocks” instead of elementary faces, in a way that generalizes train tracks for group automorphisms.

The primary interest of this construction is in the generalization of the 1-dimensional theory to higher dimensions. In dimension greater than one even the limits of possible substitution rules are not well understood. This example illustrates a general method to obtain a large class of geometric substitution rules from combinatorial objects (automorphisms of the free group). The dynamical systems constructed here can be considered as a finitely generated pseudo-group which appears as a section of an $R^k$ action on a tiling space. This higher-dimensional structure plays the role of the symbolic shift of the fixed point in the Pisot case. The precise definition of the pseudo-group is more complicated in dimension higher than one; in one dimension, we consider a tiling of the line; there is a natural flow, and there are stopping points for this flow, given by the extremities of the tiles. In two dimensions, we consider a tiling of the plane by parallelograms, with the natural $R^2$-action. The vertices of the parallelograms play the role of the stopping times; the pseudo-group is given by moving from a vertex to another, and it is generated by a finite number of basic moves, along an edge of a parallelogram. These basic moves are not defined everywhere, since there may be vertices which are not adjacent to a given type of tiles and there may be several of them defined in one point.
1.1. Markov partitions for hyperbolic toral automorphisms. It is well known, since the work of Adler and Weiss [AW70], that any hyperbolic automorphism of the 2-torus admits a Markov partition. This partition can be computed explicitly and is made of rectangles whose sides are parallel to the eigendirections.

One can say more in this case: the matrix of such a hyperbolic automorphism can be conjugated to a positive matrix. It is then always possible to find a substitution (positive automorphism of the free group on two letters), whose abelianization is given by the matrix. This substitution admits a fixed point (infinite sequence), and the closure of the orbit of this fixed point along the shift defines a dynamical system which is conjugate to the return map of the flow in the expanding direction to the basis of the Markov partition. Furthermore, any orbit of the flow in the expanding direction has a tiling defined by intersection with the rectangles of the Markov partition. This tiling is made of two different tiles (intervals), and its combinatorics is completely defined by the substitution. This substitution is directly related to the periodic continued fraction expansion of the slope of the expanding direction. A similar result holds for the unstable direction, using the reverse of this continued fraction expansion.

We will just show the simplest example, the Fibonacci map; its matrix is \( \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \).

If we denote by \( \phi \) the golden number \( \frac{1 + \sqrt{5}}{2} \), its expanding eigendirection is the line of slope \( \frac{1}{\phi} \) and its contracting direction the line of slope \( -\phi \) (these two directions are orthogonal, due to the symmetry of the matrix). A Markov partition is described in Figure 1.1. The first return map of the flow in the expanding direction to the basis of the Markov partition is easily shown to be the rotation of angle \( \phi \). The substitution associated to this matrix is \( 0 \mapsto 01, 1 \mapsto 0 \), which admits for a fixed point the Fibonacci infinite word \( 010010100100101 \ldots \), and it is well known that the shift on the closure of the shift orbit of this infinite word is measurably and semi-topologically conjugate to the rotation.

The general theory of hyperbolic systems shows that, in higher dimensions, any hyperbolic toral automorphism of the \( d \)-torus admits a Markov partition; see for example [Sin68, Bow70]. However, the situation is not so simple, since a result of Bowen [Bow78] shows that if \( d > 2 \) and the automorphism is algebraically irreducible, the elements of the Markov partition can never have a smooth boundary: this boundary is a non-trivial fractal set and is difficult to construct explicitly in the general case.

1.2. Pisot substitutions and Rauzy fractals. In the particular case of a positive hyperbolic automorphism of the Pisot type (that is, the expanding space is of dimension 1, and the contracting space of dimension \( d - 1 \)), a remarkable progress was achieved by Rauzy [Rau82]. He showed, in the specific case of the substitution \( \sigma : 1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 1 \), how to build a fundamental domain with fractal boundary \( R \) for the lattice \( \mathbb{Z}^2 \) in the plane, such that a particular translation modulo the lattice, which appears as an exchange of three pieces \( R_1, R_2, R_3 \) of \( R \), is conjugate to the dynamical system associated to \( \sigma \). The domain \( R \) can be constructed in various ways: as the projection of the geometric realization of the fixed point on the contracting plane along the expanding eigendirection, or by iteration of a generalized substitution (see below). It is then not difficult, although it was not immediately recognized at the time, to build an explicit Markov partition for the
Figure 1.1. Markov partition for the Fibonacci automorphism.

\[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{pmatrix}
\]

associated to \( \sigma \) by abelianization. This Markov partition can be seen as a fundamental domain for a lattice tiling of \( \mathbb{R}^3 \) (see Figure 1.2).

This construction has now been generalized to a large class of substitutions and allows us to build explicit Markov partitions for many (and conjecturally for all) Pisot toral automorphisms; see [Sie02, BK06].

1.3. The non-Pisot case. The situation is much more complicated in the non-Pisot case. Symbolic dynamics in this case have been studied from an arithmetic viewpoint in the work of Kenyon-Vershik and of Leborgne (see [KV98, Bor97]), but there are several problems here. The technique that was used only gives a sofic system, and this system is not in general measurably conjugate to the automorphism, but only semi-conjugate; furthermore, it does not seem to give an explicit geometric construction.

In the case of a Pisot toral automorphism, the entropy of the automorphism is the logarithm of the largest eigenvalue. If the matrix is non-negative, this is also the entropy of the subshift of finite type associated to the matrix, so we can expect to find symbolic dynamics in this way. However, in the general case, things are more complicated, since the entropy of the automorphism is the sum of the logarithms
of the moduli of eigenvalues larger than one. One way to recover this number is to consider the exterior power of rank $k$ of the matrix, where $k$ is the dimension of the expanding space. Its largest eigenvalue is the product of all eigenvalues of modulus larger than 1, and if this matrix is positive, the subshift it defines will have the same entropy. This is the case for the automorphism we will study in this paper. Note that there is a strongly related result in a paper by Manning (see [Man02]), where he considers the action by the matrix on the $k$-dimensional homology $H_k(T^n, \mathbb{R})$.

1.4. Self-similar tilings. Since the pioneering work of Penrose, many papers have been devoted to tilings having self-similarity properties, and different generating methods have been proposed: substitution, cut-and-project, etc.

The term self-similar is itself a little confusing in its current usage. Going back to Thurston, but also used by Solomyak, the term has been used to refer to substitution tilings, where the substituted tiles fit exactly over the expanded tiles, as is the case, for example, for the 1-dimensional Fibonacci tiling or for the Penrose-type tiling by two triangles. The case where the tiles do not fit exactly (as in the case of the classical kite-and-dart Penrose tiling) is described as pseudo-self-similar. Of course the term seems to refer simply to tilings with self-similar properties, but these are much more general. The term substitution rule has gained a lot of use as the term for the general case, though there are no clearer, but widely used terms for the distinction between self-similar and pseudo-self-similar.

A simple way to obtain such a self-similar tiling is given by a Markov partition for a toral automorphism. The boxes of the Markov partition intersect the expanding (or contracting) space, giving rise to a tiling of this space by a finite number of tile types; this tiling is obviously self-similar, by the properties of the Markov partition. In many known cases, the re-normalization turns out to be measurably conjugate to such a toral automorphism.
1.5. **Substitution rules on the plane and in higher dimensions.** The substitution rules of the line, found from symbolic substitution rules, can be considered as simply the 1-dimensional case of a more general construction. Substitution rules on the plane were originally developed as a method of showing that an aperiodic set of tiles could indeed tile the plane by Berger [Ber66]. The study of these tilings led to the famous and beautiful example of Penrose [Gar77]. This is an aperiodic set with just two tiles. The tilings that these tiles admit are precisely the tilings generated by the Penrose substitution rule. Pioneering work on such tilings was also done by Thurston [Thu89] and Kenyon [Ken96, Ken92].

To such a tiling one can associate a space of tilings by taking the closure of the orbit of the tiling by translation, and two dynamical systems, a minimal system of entropy 0 corresponding to the translation, and a system of positive entropy corresponding to the substitution or re-normalization.

1.6. **Main results.** We give a brief overview of the results of the paper; precise definitions are given in the following sections.

We consider an automorphism $\sigma$ of the free group on 4 generators, with abelianization matrix $A$. We call $P_c$ (resp. $P_e$) the contracting (resp. expanding) plane and $\pi_c$ (resp. $\pi_e$) the corresponding projections. Our first aim is to build some discrete approximations of the planes $P_c$ and $P_e$, and related tilings, with self-similarity properties given by $\sigma$.

If $e_i, i = 1, \ldots, 4,$ denotes the canonical basis of $\mathbb{R}^4$, and $x \in \mathbb{R}^4$, we denote by $(x, i \wedge j)$ the 2-dimensional square in $\mathbb{R}^4 \{x + se_i + te_j | 0 \leq s, t \leq 1\}$, and we say this 2-face is of type $i \wedge j$. We define the 2-dimensional extension $E^2(\sigma)$ and its dual map $E^2(\sigma)$, which associate to any such integral 2-face (an integer translate of a 2-dimensional face of the unit hypercube) a finite union of 2-faces. These extensions are the main tool of this paper. We first show that there is a stepped surface (surface made of 2-faces) periodic under the action of $E^2(\sigma)$.

**Theorem 1.** There exists a unique stepped surface $S_e$ in good position containing $(-e_3, 3 \wedge 4)$ such that 

$$E^2(\sigma)^3(S_e) = S_e.$$ 

This stepped surface stays within bounded distance of the expanding plane.

There is a related tiling of $P_e$.

**Theorem 2.** There exists a unique tiling containing the tile $\pi_e(-e_3, 3 \wedge 4)$ which is periodic of period 3 under the substitution rule associated to $E^2(\sigma)$. This tiling is the image by $\pi_e$ of the stepped surface $S_e$.

The same theorem holds true for $E^2(\sigma)$:

**Theorem 3.** Define $V = (0, 1 \wedge 2) + (0, 2 \wedge 3) + (0, 3 \wedge 1)$. There exists a unique stepped surface $S_c$ containing $V$ which is fixed by $E^2(\sigma)$. This stepped surface stays within bounded distance of the contracting plane $P_c$. Its projection to the contracting plane defines a polygonal tiling which is invariant under the natural substitution rule associated to $E^2(\sigma)$.

We call $V_c(i \wedge j)$ the set of basic vertices of faces of type $i \wedge j$. By taking a limit to a re-normalization of the faces under $E^2(\sigma)$, we define fractal tiles $Y_c(i \wedge j)$.

---

1 An aperiodic set of tiles is a set of tiles that can tile the plane, but cannot do so periodically.
Figure 1.3. A patch of the substitution tiling of the expanding plane. This gives a glimpse in the 4-dimensional space, since what is really represented is a perspective view of a 2-dimensional stepped surface in a 4-dimensional space, which explains why the figure is somewhat uncomfortable to look at.

These tiles build a self-similar tiling, with the same combinatorics (vertices) as the polygonal tiling; a similar result holds for $E_2(\sigma)$:

**Theorem 4.** The sets $x + Y_e(i \wedge j)$, for $x \in \pi_e(V_e(i \wedge j))$, tile the plane $P_e$, and this tiling is self-similar for the action of the matrix $A^{-1}$ on the plane.

The sets $x + Y_e(i \wedge j)$, for $x \in \pi_e(V_e(i \wedge j))$, tile the plane $P_e$. This tiling is self-similar for the action of the matrix $A^3$ on the expanding plane $P_e$.

These tilings are cut-and-project tilings; the corresponding window is made of tiles that are related in a simple way to the dual tiles:

**Theorem 5.** The tilings generated by the projection of $E_2(\sigma)$ are projection tilings, and the window is a generalized octagon formed of translates of the opposite of the fractal tiles of the dual tiling.

We define a set with fractal boundary $X_e$ composed of 6 tiles $X_e(i \wedge j)$ of the tiling, an expanding map $T$ on $X_e$ related to the matrix $A$, and a digit function $a^*$ on $X_e$, with value in the finite set $\{0, f_1, f_2\}$; we can then define an expansion of any point of $x$. Taking an adequate complex coordinate $\phi_e$ on $P_e$, we get a complex numeration system with a related sequence of digits $\alpha^*$ with value in a finite set of
complex numbers. The admissibility condition is given by the following graph $G_c$:

```
1 ∧ 4 ← 0 1 ∧ 2 0 → 3 ∧ 1
0 0 0 ↓ ↑ 0
3 ∧ 4 ← 2 ∧ 3 4 ∧ 2
```

More precisely:

**Theorem 6.** The dynamical system $(X_c, T)$ is a Markov map with respect to the partition $X_c(i ∧ j)$. Its adjacency matrix is the structure matrix of $E^2(\sigma)$.

The set $\Omega_c$ of admissible digit sequences $(a^*_n)$ are given by the labels of infinite paths in the graph $G_c$ above. Any element $x \in X_c$ can be written in terms of its digit sequence $a^*_n(x)$:

$$x = \sum_{n=1}^{\infty} A^n a^*_n(x).$$

Any complex number $z \in \phi_c(X_c)$ can be written in base $\mu$ as

$$z = \sum_{n=1}^{\infty} \alpha^*_n(z) \mu^n.$$

Furthermore, for any $z \in \mathbb{C}$ we can find $n_0 \in \mathbb{N}$ and an admissible sequence $(\alpha^*_n(z))_{n \geq -n_0}$ such that

$$z = \sum_{n=-n_0}^{\infty} \alpha^*_n(z) \mu^n.$$

A similar complex system exists on the expanding plane, but its definition is more complicated because it is multi-valued.

Finally, we define tiles $X_e(\psi(i ∧ j))$ of the expanding plane related to the window for the tiling of the plane $P_c$, and a compact subset $\hat{X} = \bigcup_{i ∧ j} X_c(i ∧ j) + X_e(\psi(i ∧ j))$, and get the promised Markov partition for $A$:

**Theorem 7.** The set $\hat{X}$ is a fundamental domain for the lattice $\mathbb{Z}^4$. Its canonical decomposition in sets $X_c(i ∧ j) + X_e(\psi(i ∧ j))$ is a Markov partition for the automorphism $A$ of the 4-dimensional torus $\mathbb{R}^4/\mathbb{Z}^4$.

1.7. **Organization of the paper.** In section 2, we will first define the automorphism of the free group of rank 4 and study its basic properties. We will then recall the definitions of the higher-dimensional extensions of free group automorphisms and their duals, as given in [AIS01] for substitutions, later generalized to automorphisms of free groups in [Ei03], as we will use them heavily in the rest of the paper, and state the basic properties which will be used throughout the paper.

In section 3, we will prove Theorem 1 there exists a stepped surface which projects homeomorphically on the expanding plane along the contracting plane and which is periodic of period 3 for the 2-dimensional extension of the automorphism. This stepped surface stays within bounded distance of the expanding plane.
In section 4, we will prove that this surface projects to a polygonal tiling which is defined by a substitution rule (Theorem 2). This polygonal tiling can be defined directly, which gives an independent proof of the existence of the stepped surface.

In section 5, we will prove analogous results (Theorem 3) for the 2-dimensional dual extension, getting a stepped surface which approximates the contracting plane and is invariant by the dual 2-dimensional extension of the automorphism, and a substitution invariant polygonal tiling of the contracting plane.

In section 6, we will show Theorem 4: by a well-known re-normalization procedure, one can get exact substitution tilings whose tiles have fractal boundaries; these tiles are analogous to the Rauzy tiles, but for a non-Pisot case.

In section 7, we will show that each of these tilings can be obtained by a cut-and-project method (Theorem 5). The window for a tiling is obtained by projecting on the complementary plane the vertices of the corresponding stepped surface, and it corresponds to the basic tiles of the other tiling. These tiles also correspond to a graph-directed linear iterated function system.

In section 8, we will define and study a complex numeration system associated to this tiling (Theorem 6).

In section 9, we will prove Theorem 7: we can build a Markov partition for the automorphism by taking the product of corresponding tiles, or equivalently by taking the natural extension of the λ-expansion associated to the complex numeration system.

In section 10, we will show another example and give some additional remarks and open problems.

1.8. Notation.

1.8.1. Linear algebra of σ.

σ automorphism of the free group.

\(\hat{\sigma}^{-1}\) the reversed inverse of \(\sigma\).

\(\{e_1, e_2, e_3, e_4\}\) canonical basis of \(\mathbb{R}^4\).

\(f\) abelianization map.

\(A\) structure matrix of \(\sigma\).

\(\lambda, \bar{\lambda}, \mu, \bar{\mu}\) eigenvalues of \(A\).

\(P_e, P_c\) contracting and expanding planes; \(\pi_e, \pi_c\) corresponding projections.

\(\phi_e, \phi_c\) complex coordinates on \(P_e, P_c\).

\(\Phi\) corresponding map from \(\mathbb{R}^4\) to \(\mathbb{C}^2\).

\(\Lambda\) the lattice of \(\mathbb{C}^2\) corresponding to \(\mathbb{Z}^4\) for the complex structure.

1.8.2. Extensions of substitutions.

\((x, i)\) 1-dimensional face.

\((x, i \land j)\) 2-dimensional face.

\(E_k(\sigma)\) \(k\)-dimensional extension.

\(E^*_k(\sigma)\) dual map of \(E_k(\sigma)\).

\(\phi_k\) Poincaré map in dimension \(k\).

\(F_e\) (resp. \(F_c\)) set of types of positively oriented faces with respect to \(\pi_e\) (resp. \(\pi_c\)).

\(\psi\) duality map on faces, sending a map with a given orientation with respect to \(\pi_e\) to a complementary map with the same orientation with respect to \(\pi_c\). We have \(\psi(F_e) = F_c, \psi(F_c) = F_e\).

\(E^k(\sigma)\) geometric dual map of dimension \(k\) (conjugate of \(E^*_{d-k}(\sigma)\) by \(\phi_{d-k}\)).

\(A^*\) exterior square of \(A\).
1.8.3. Stepped surfaces.
$S_c$ stepped surface approximating $P_c$, periodic under $E_2(\sigma)$ and containing $(-e_3, 3 \wedge 4)$.
$S_e$ stepped surface approximating $P_e$, fixed under $E_2(\sigma)$ and containing $(0, 1 \wedge 2)$.
$V = (0, 1 \wedge 2) + (0, 2 \wedge 3) + (0, 3 \wedge 1)$ starting point of the construction of $S_c$.
$U_c$ basic octagon, seed of the invariant tiling of the contracting plane.
$x_{i \wedge j}$ basic vertex of the face of type $i \wedge j$ in $U_c$.
$V_{c}(i \wedge j)$ set of basic vertices of faces of type $i \wedge j$ in the stepped surface $S_c$.
$V_{e}(i \wedge j)$ set of basic vertices of faces of type $i \wedge j$ in the stepped surface $S_e$.

1.8.4. Fractal tiles.
$Y_{c}(i \wedge j)$ fractal tiles in $P_c$ obtained by re-normalization of $\pi_c(0, i \wedge j)$ by the dual map.
$X_{c}(i \wedge j)$ fractal tiles in $P_c$ defined by $x_{i \wedge j} + Y_{c}(i \wedge j)$; fractal tiles obtained by re-normalization of $\pi_c(x_{i \wedge j}, i \wedge j)$.
$X_{c}$ union of $X_{c}(i \wedge j)$.
$Y_{e}(i \wedge j)$ fractal tiles in $P_e$ obtained by renormalization of $\pi_e(0, i \wedge j)$ by $E_2(\sigma)$.
$Z_{e}(i \wedge j)$ closure of the projection of $V_{c}(\psi(i \wedge j)$ on the expanding plane (atomic surface, or window).
$X_{e}(i \wedge j) = -Z_{e}(i \wedge j)$.

1.8.5. Number system.
$T$ Markov map on $X_c$.
$f_1, f_2$ digits for the numeration system on $X_c$.
$a^*$ digit map on $X_c$.
$\alpha^*$ corresponding complex digit map.

1.8.6. Markov partition.
$\hat{X}$ fundamental domain $\hat{X} = \bigcup_{i \wedge j \in F_c} X_{c}(i \wedge j) + X_{c}(\psi(i \wedge j))$.

2. Definitions and notation

2.1. The example. We consider the free group $F_4$ on 4 generators denoted by 1, 2, 3, 4, and the automorphism $\sigma$ of $F_4$ defined by

$$
\sigma: \begin{align*}
1 & \mapsto 2 \\
2 & \mapsto 3 \\
3 & \mapsto 4 \\
4 & \mapsto 41^{-1}.
\end{align*}
$$

We denote by $f: F_4 \to \mathbb{Z}^4$ the natural abelianization map, and by $e_i, i = 1, \ldots, 4$, the canonical basis of $\mathbb{R}^4$, so that $f(i) = e_i$. The map $f$ conjugates the automorphism $\sigma$ to a linear map of $\mathbb{R}^4$, whose matrix in the canonical basis is the $4 \times 4$ matrix $A$ given by

$$
A = \begin{pmatrix}
0 & 0 & 0 & -1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1
\end{pmatrix}
$$
with the obvious commutation relation:

\[
\begin{array}{c c c c}
F_4 & \sigma & F_4 \\
\downarrow f & \downarrow f & \downarrow f \\
\mathbb{R}^4 & \overset{\lambda}{\longrightarrow} & \mathbb{R}^4
\end{array}
\]

This matrix is in \(SL(4, \mathbb{Z})\), and defines an automorphism of the torus \(T^4\). In section 8 we will give an explicit Markov partition for this automorphism. The matrix \(A\) has characteristic polynomial \(P_A(X) = X^4 - X^3 + 1\) (this is immediately checked, since the matrix is the companion matrix for the polynomial \(P_A\)). This polynomial has no real roots, hence the matrix has two pairs of complex conjugate eigenvalues: one is \(\lambda, \bar{\lambda} \approx 1.01891 \pm 0.602565 i\) of modulus \(\approx 1.18375 > 1\) and the other is \(\mu, \bar{\mu} \approx 0.518913 \pm 0.66661 i\) of modulus \(\approx 0.844772 < 1\).

We will denote by \(P_e\) (resp. \(P_c\)) the expanding plane, corresponding to eigenvalues \(\lambda, \bar{\lambda}\) (resp. the contracting plane), and by \(\pi_e\) (resp. \(\pi_c\)) the projection on \(P_e\) along \(P_c\) (resp. on \(P_c\) along \(P_e\)).

We remark that we can endow \(P_c\) with a complex structure, so that the restriction to the invariant plane \(P_c\) of the action of \(A\) is given by the multiplication by \(\lambda\) and that the vector \(e_1\) projects to 1. This defines a map from \(P_c\) to \(\mathbb{C}\).

**Definition 2.1.** We denote by \(\phi_c\) the map \(P_c \to \mathbb{C}\) uniquely defined by \(\phi_c(A.x) = \lambda \phi_c(x)\) for all \(x \in P_c\), and \(\phi_c(\pi_c(e_1)) = 1\).

Similarly, we denote by \(\phi_e\) the map \(P_e \to \mathbb{C}\) uniquely defined by \(\phi_e(A.x) = \mu \phi_e(x)\) for all \(x \in P_e\), and \(\phi_e(\pi_e(e_1)) = 1\).

We denote by \(\Phi\) the map \(\Phi : \mathbb{R}^4 \to \mathbb{C}^2\) defined by \(\Phi(x) = (\phi_c(\pi_c(x)), \phi_e(\pi_e(x)))\).

We immediately compute \(\phi_c(\pi_c(e_2)) = \lambda\), \(\phi_c(\pi_c(e_3)) = \lambda^2\) and \(\phi_c(\pi_c(e_4)) = \lambda^3\) with corresponding formulas for \(\phi_e(\pi_e(e_i))\), replacing \(\lambda\) by \(\mu\). The map \(\Phi\) conjugates the action of \(A\) to the map \((z, w) \mapsto (\lambda z, \mu w)\). It sends the lattice \(\mathbb{Z}^4\) to a lattice \(\Lambda \subset \mathbb{C}^2\) by the map \(\Phi : \mathbb{Z}^4 \to \Lambda, (a, b, c, d) \mapsto (a + b\lambda + c\lambda^2 + d\lambda^3, a + b\mu + c\mu^2 + d\mu^3)\).

The image by \(\phi_e\) of the projection of the lattice \(\mathbb{Z}^4\) by \(\pi_e\) is the ring \(\mathbb{Z}[\lambda]\). We can see in Figure 2.1 the projection of the canonical basis \((e_1, e_2, e_3, e_4)\) on the two invariant planes.

One of the main aims of this paper is to give a discrete approximation in \(\mathbb{R}^4\) of the contracting and expanding planes by 2-dimensional stepped surfaces with vertices in \(\mathbb{Z}^4\).

**Remark 2.2.** Since both planes are completely irrational, the maps \(\phi_c \circ \pi_c\) and \(\phi_e \circ \pi_e\) from \(\mathbb{Z}^4\) to \(\mathbb{Z}[\lambda]\) or \(\mathbb{Z}[\mu]\) are one-to-one. The reciprocal maps have a nice arithmetic definition: there is a natural conjugacy morphism \(\gamma : \mathbb{Z}[\lambda] \to \mathbb{Z}[\mu]\). We can use it to define the map \(\mathbb{Z}[\lambda] \to \Lambda, z \mapsto (z, \gamma(z))\). In particular, we see that the set \(\mathbb{Z}[\lambda]\) is dense in \(\mathbb{C}\), but its image is a discrete set in \(\mathbb{C}^2\).

2.2. **Extensions of substitutions.** In this paper, we will make use of the 2-dimensional extension of the automorphism \(\sigma\) and its dual map. Here we briefly recall their definition and main properties introduced in [AIS01].

2.2.1. **Geometric representations of words.** It is natural to represent graphically any word in the free monoid on \(d\) letters, or more generally any reduced word representing an element of the free group, as a continuous stepped line in \(\mathbb{R}^d\), starting from the origin and made of segments of unit length parallel to the basis vectors.
We will consider this stepped line as a formal sum of segments \((x, i)\), with \(x \in \mathbb{Z}^d\), and \(i\) an element of the alphabet. In the case of an element of the free group, we must also consider elements such as \((x, i^{-1})\). We take the convention \((x, i^{-1}) = -(x - e_i, i)\), since \((x, i^{-1})\) corresponds to the same segment as \((x - e_i, i)\), but with the opposite orientation.

The consequence of this convention is that the line associated to \(ii^{-1}\) is empty so that it takes care of cancellations. Hence any word and the corresponding reduced word define the same stepped line, which depends only on the element of the free group.

### 2.2.2. 1-dimensional extension.

To any morphism \(\sigma\) on the free monoid we can associate a map on such stepped lines, which sends the stepped line corresponding to \(W\) to the stepped line corresponding to \(\sigma(W)\). One checks that if the stepped line ends at the point \(P \in \mathbb{Z}^d\), its image ends at the point \(AP\), where \(A\) is the abelianization of \(\sigma\). Hence the image of a particular segment of the stepped line does not depend on the rest of the line.

**Definition 2.3.** We denote by \((x, i)\), where \(x \in \mathbb{Z}^d\) and \(i \in \{1, \ldots, d\}\), the unit segment starting at \(x\) and parallel to the basis vector \(e_i\). We can represent any stepped line as a formal sum of such segments.

The previous paragraph allows us to define the 1-dimensional extension \(E_1(\sigma)\) of \(\sigma\); we can give explicit formulas:

**Definition 2.4.** Let \(\sigma\) be an automorphism of the free group on \(d\) generators and let \(A\) be the corresponding matrix.

Let \(\sigma(i) = W^{(i)} = W_1^{(i)}W_2^{(i)} \ldots W_{l_i}^{(i)} = P_k^{(i)}W_k^{(i)}S_k^{(i)}\), where \(W^{(i)}\) is a reduced word in the generators and their inverses, and \(P_k^{(i)}\) (resp. \(S_k^{(i)}\)) is a prefix of length \(k - 1\) (resp. suffix of length \(l_i - k\)) of \(W^{(i)}\).

We define

\[
E_1(\sigma)(x, i) = \sum_{k=1}^{l_i} \left( Ax + f(P_k^{(i)}, W_k^{(i)}) \right),
\]
2.2.3. 2-dimensional extension. Let $\gamma : [0, n] \rightarrow \mathbb{R}^d$ and $\delta : [0, m] \rightarrow \mathbb{R}^d$ be two broken lines, parametrized by their length, corresponding to words $C$ and $D$ of respective length $n$ and $m$. We can define an oriented parametrized surface $\gamma \wedge \delta$ by

$$\gamma \wedge \delta : [0, n] \times [0, m] \rightarrow \mathbb{R}^d (s, t) \mapsto \gamma(s) + \delta(t).$$

It is clear that this surface is made of squares $(x, i) \wedge (y, j)$, and such a square is parallel to the square on the unit vectors $e_i$ and $e_j$ with its lower corner in $x + y$.

**Definition 2.5.** We will denote by $(x + y, i \wedge j)$ the face $(x, i) \wedge (y, j)$ (remark that this face does not depend on $x$ and $y$, but only on the sum $x + y$). The surface $(x, i \wedge i)$ is degenerate, and we will take the convention that $(x, i \wedge i) = 0$ and $(x, i \wedge j) = -(x, j \wedge i)$. For any face $(x, i \wedge j)$, the point $x \in \mathbb{Z}^d$ will be called the basic vertex of the face and $i \wedge j$ will be called the type of the face.

In the sequel, we will consider formal linear combinations of faces; such a combination with all non-zero coefficients equal to 1 will be called a patch of faces. By abuse of language, we will not consider the difference between such a patch of faces and the underlying subset of $\mathbb{R}^d$, since it will not lead to confusion in our setting.

We can then define an action of the automorphism, not only on the segments, but also on the unit squares; this is what we call the 2-dimensional extension $E_2(\sigma)$, acting on formal sums of squares:

**Definition 2.6.** We define the 2-dimensional extension of $\sigma$ by

$$E_2(\sigma)(x + y, i \wedge j) = E_1(\sigma)(x, i) \wedge E_1(\sigma)(y, j).$$

Using the previous notation, we have

$$E_2(\sigma)(x, i \wedge j) = \sum_{m=1}^{l_i} \sum_{n=1}^{l_j} \left( Ax + f(P_m^{(i)}) + f(P_n^{(j)}), W_m^{(i)} \wedge W_n^{(j)} \right).$$

Remark that this formula may appear slightly complicated in general, but in the present case, since all images of letters except one are of length 1, it will be very simple to compute.

We can in the same way define the $k$-dimensional extension $E_k(\sigma)$ of $\sigma$ acting on $k$-dimensional faces.
2.3. Dual maps and their geometric representation. The \( k \)-dimensional maps extend naturally to the vector space of finite weighted sums of \( k \)-dimensional faces, and one can formally define their dual maps; we will first compute it for the 1-dimensional map.

2.3.1. The dual map \( E_1^*(\sigma) \). Using the dual basis \((x, i)^*\) of the canonical basis, we can explicitly compute the dual of the map \( E_1(\sigma) \). The computation turns out to be very easy when the matrix \( A \) is invertible (as an integer matrix, that is, it has determinant 1). Note that this is always the case for a free group automorphism.

**Proposition 2.7.** If the matrix \( A \) is invertible, the dual map is defined, using the previous notation, by

\[
E_1^*(\sigma)(x, i)^* = \sum_{j,k : W^j_k = i} (A^{-1}(x - f(P^{(j)}_k)), j)^*.
\]

**Proof.** By definition of the dual map, we have

\[
\langle E_1^*(\sigma)(x, i)^* | (y, j) \rangle = \langle (x, i)^* | E_1(\sigma)(y, j) \rangle = \langle (x, i)^* \sum_{k=1}^{l_j} (Ay + f(P^{(j)}_k), W^*_k) \rangle.
\]

This product can only take value 0 or 1, and it takes value 1 precisely for the pairs \((y, j)\) such that there exists \( k \) with \( W^j_k = i \), and \( x = Ay + f(P^{(j)}_k) \), or equivalently \( y = A^{-1}(x - f(P^{(j)}_k)) \). This is the formula given in the proposition. \( \square \)

2.3.2. Higher-dimensional dual maps. We can do the same type of computation for higher-dimensional extensions, and recover their dual maps. We will not give the formulas here, as we will see below a much more efficient way to compute these maps for automorphisms of free groups.

2.3.3. Poincaré duality and geometric representation of dual maps. In contrast with the extensions of a substitution, the dual maps are abstract objects, formally defined on the dual basis.

We can however define a kind of Poincaré duality \( \varphi_k \), which sends the dual of a face of dimension \( k \) to a face of dimension \( d - k \) (where \( d \) is the dimension of the space). More formally:

**Definition 2.8.** We define the Poincaré map in dimension \( k \) by

\[
\varphi_k(x, i_1 \wedge \cdots \wedge i_k)^* = (-1)^{i_1 + \cdots + i_k} (x + e_{i_1} + \cdots + e_{i_k}, j_1 \wedge \cdots \wedge j_{d-k}),
\]

where \( \{i_1, \ldots, i_k\} \) and \( \{j_1, \ldots, j_{d-k}\} \) form a partition of \( \{1, 2, \ldots, d\} \), with \( i_1 < \cdots < i_k, j_1 < \cdots < j_{d-k} \).

One of the important properties of this map is that it commutes with the natural boundary and coboundary operators. Using the Poincaré map one can conjugate the dual maps, acting on the dual of faces of dimension \( k \), to maps acting on faces of dimension \( d - k \). This allows us to give a geometric model of these maps.

**Remark 2.9.** The definition of the Poincaré map, and especially the offset for the origin point, is not arbitrary. If we define the Poincaré map in dimension 0 (that is, on points) in the natural way by \( \varphi_0(x, \bullet)^* = (x, 1 \wedge \cdots \wedge n) \), the definition in any dimension follows if we want to get this property of commutation with boundary and coboundary.
Definition 2.10. We denote by $E^{d-k}(\sigma)$ the conjugate, by the map $\varphi_k$, of the dual map $E^*_k(\sigma)$. We call this map the geometric dual map.

A little computation gives a nice formula for the geometric dual map, using the suffix instead of the prefix. We will give it only for the case of the geometric dual map $E^{d-1}$, conjugate of $E^*_1$, and we will for simplicity denote by $(x,i^*)$ the $d-1$-dimensional face in $x$ transverse to the direction $i$:

**Proposition 2.11.** We have

$$E^{d-1}(\sigma)(x,i^*) = \sum_{j,k; W^{(j)}_k = i} \left( A^{-1}(x + f(S^{(j)}_k)), j^* \right).$$

Similar formulas will be recovered below in a simpler way.

2.4. Some remarkable properties of extensions of automorphisms of free groups. In the next sections we will use the following properties:

A first property, proved in [AIS01], is that the extensions of substitutions commute with the boundary and coboundary operators. More precisely, one can define a boundary operator which associates, to any face $f$ of dimension $k$ its boundary $df$, a union of oriented faces of dimension $k-1$. We have the following proposition:

**Proposition 2.12.** $dE^k(\sigma)(f) = E^{k-1}(\sigma)(df)$, and $dE^k(\sigma)(f) = E^{k-1}(\sigma)(df)$: the image of the boundary (of a union of faces) is the boundary of the image (of this union of faces).

A second property is that, in the hyperbolic case (the matrix $A$ has no eigenvalue of modulus one), we can build limit sets on the expanding (resp. contracting) space by iterating the suitable extension (resp. the geometric dual map), projecting to the contracting (resp. expanding) space, and re-normalizing. More precisely:

**Proposition 2.13.** Suppose that the matrix $A$ is hyperbolic, with an expanding space of dimension $k$. Denote by $\pi_e$, as above, the projection on the expanding space, and by $f$ any $k$-dimensional face. Then the sequence of sets $A^{-n}\pi_e(E^k(\sigma)^n(f))$ converges for the Hausdorff metric on compact sets, and similarly for the sequence of sets $A^n\pi_e(E^{d-k}(\sigma)^n(g))$, where $g$ is a $d-k$-dimensional face.

A third important property is the theorem of Ei ([Ei03]):

**Proposition 2.14.** Let $\sigma$ be an automorphism of the free group; let $\sigma^{-1}$ be its inverse and $\tilde{\sigma}^{-1}$ the automorphism obtained from $\sigma^{-1}$ by reversing the order of the words. We have $E^k(\sigma) = E_k(\tilde{\sigma}^{-1})$.

For the case of automorphisms, this greatly reduces the difficulty of computing the dual map.

3. Invariant stepped surface

We come back now to the study of the automorphism $\sigma$ defined above.

3.1. Orientation of 2-dimensional faces. There are 6 elementary 2-dimensional faces, with two possible orientations for each one. As one checks immediately on Figure 21, one can choose an orientation on the expanding plane such that the projection $\pi_e$ on the expanding plane, restricted to faces $(0,i \land j)$ with $i < j$, preserves the orientation.
Definition 3.1. We will say that a face \((x, i \wedge j)\) is positively oriented (with respect to the projection \(\pi_e\)) if \(i < j\). We denote by \(F_e\) the set of types of faces positively oriented with respect to \(\pi_e\).

3.2. Explicit computation of the 2-dimensional extension. To know the 2-dimensional extension of \(\sigma\), \(E_2(\sigma)\), it suffices to compute the images of the unit squares at the origin, \((0, i \wedge j)\). There are 6 of them, and one shows easily:

Proposition 3.2. The map \(E_2(\sigma)\) is defined on the 6 basic faces by

\[
\begin{align*}
(0, 1 \wedge 2) & \mapsto (0, 2 \wedge 3), \\
(0, 1 \wedge 3) & \mapsto (0, 2 \wedge 4), \\
(0, 1 \wedge 4) & \mapsto (0, 2 \wedge 4) + (e_4 - e_1, 1 \wedge 2), \\
(0, 2 \wedge 3) & \mapsto (0, 3 \wedge 4), \\
(0, 2 \wedge 4) & \mapsto (0, 3 \wedge 4) + (e_4 - e_1, 1 \wedge 3), \\
(0, 3 \wedge 4) & \mapsto (e_4 - e_1, 1 \wedge 4).
\end{align*}
\]

We can see the representation of this map in Figure 3.1, or more exactly of its projection to the expanding plane (one cannot represent exactly dimension 4!). It is a remarkable fact that the image of any positively oriented face is a union of positively oriented faces. This property can be generalized:

Definition 3.3. Let \(\tau\) be an automorphism of the free group on \(d\) letters. Suppose that the associated abelianization matrix is hyperbolic, with an expanding space of dimension \(k\) and a contracting space of dimension \(d - k\) (with associated projections \(\pi_e\) and \(\pi_c\)), and choose a positive orientation on the expanding and contracting spaces.

Define a \(k\)-face (resp., a \((d - k)\)-face) to be positively oriented with respect to \(\pi_e\) (resp., \(\pi_c\)) if the projection \(\pi_e\) (resp., \(\pi_c\)) preserves the orientation. (We will just say for the sake of brevity that a face is “positive” when the context is clear.)

We say that \(\tau\) satisfies the positivity condition for the expanding space (resp., contracting space) if the image by \(E_k(\tau)\) of a positive \(k\)-face is a sum of positive faces (resp., the image by \(E^{d-k}(\tau)\) of a positive \((d - k)\)-face is a sum of positive faces).

Remark 3.4. In the Pisot case, the expanding space is 1-dimensional, and 1-faces correspond to basis vectors. It is then easy to check that a Pisot automorphism satisfies the positivity condition for expanding space if and only if it is a substitution (up to a replacement of some generators \(i\) by their inverse \(i^{-1}\)).

The positivity condition can be seen as a generalization of the notion of substitution to the non-Pisot case.

Associated to the operator \(E_2(\sigma)\), we can define in an obvious way a \(6 \times 6\) matrix \(A^*\), indexed by the 6 types of faces, which counts the number of times a type of face occurs in the image of a type of face. This matrix can also be computed in a purely algebraic way, as the exterior square of the matrix \(A\), by computing all \(2 \times 2\) minors. As we have seen above, this matrix is positive. If we order the 6 types by
the lexicographic ordering as above, the matrix is given by

\[
A^* = \begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0
\end{pmatrix}.
\]
Remark 3.5. The concept of the matrix $A^\star$ associated to $E_2(\sigma)$ from the non-Pisot automorphism $\sigma$ was first introduced in [FIR06]. Several properties of $A^\star$ are discussed in it.

Remark 3.6. The eigenvalues of the matrix $A^\star$ are the product of all pairs of different eigenvalues of the matrix $A$. Hence there are two real eigenvalues, $\lambda\bar{\lambda}$, which is the Perron eigenvalue of the positive matrix $A^\star$, and $\mu\bar{\mu}$, and four complex eigenvalues of modulus 1. Remark that $\lambda\bar{\lambda}$ is a Salem number: its minimal polynomial is $X^6 - X^4 - X^3 - X^2 + 1$, and all of its algebraic conjugates, except one, are of modulus 1.

The following proposition will be useful later.

**Proposition 3.7.** The 6-dimensional vector whose coordinates are given by the areas of the projections $\pi_e(0, i \wedge j)$ of the six faces to the expanding plane is a left eigenvector of $A^\star$ associated to the eigenvalue $\lambda\bar{\lambda}$. Similarly, the vector of areas of the projections of the six faces to the contracting plane is a right eigenvector (Perron eigenvector) associated to the same dominant eigenvalue.

**Proof.** The linear map associated to $A$ acts in a natural way on the 6-dimensional space of symplectic forms on $\mathbb{R}^4$, and the matrix of this action in the canonical basis is the transpose of $A^\star$. Consider the map that associates, to a pair of vectors $(u, v)$, the area of the parallelogram generated by $\pi_e(u)$ and $\pi_e(v)$. It is a symplectic form, which is an eigenvector for this action, associated to $|\lambda|^2 = \lambda\bar{\lambda}$. The coordinates in the canonical basis of this form are given, by definition of the canonical basis, by the areas of projection of the canonical faces. A similar proof, replacing $A$ by $A^{-1}$, works for the other vector. $\square$

### 3.3. Stepped surfaces in good position.

**Definition 3.8.** A stepped surface is a union of positively oriented 2-dimensional faces. We will say that a stepped surface is in good position with respect to the expanding plane if the restriction of the projection $\pi_e$ to the stepped surface is a homeomorphism from the stepped surface to the plane $P_e$.

Our first aim is to prove directly that the 2-dimensional extension $E_2(\sigma)$ acts on the stepped surfaces in good position. We will give another, and somewhat shorter, proof of this fact in the next section.

**Proposition 3.9.** The image of a stepped surface in good position (with respect to the expanding plane) is a stepped surface in good position.

**Proof.** We will first define a piecewise linear map from a stepped surface to its image, which sends each face to its image by the 2-dimensional extension.

There is a small difficulty: we want this map to send the boundary of a face to the boundary of its image. This gives some trouble on the face $(x, 3 \wedge 4)$, because here there is a cancellation on the boundary: $343^{-1}4^{-1}$ is sent to $441^{-1}4^{-1}14^{-1}$. If we consider this as a cyclic word, the first letter cancels the last.

On the faces $(x, 1 \wedge 2)$, $(x, 1 \wedge 3)$ and $(x, 2 \wedge 3)$, we take the unique linear map that sends the parallelogram to its image.

On the faces $(x, 1 \wedge 4)$, we divide the parallelogram in two halves, and we send the first half on $(Ax, 2 \wedge 4)$ and the second half on $(Ax + e_4 - e_1, 1 \wedge 2)$, and similarly on $(x, 2 \wedge 4)$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
On the face \((x, 3 \land 4)\), we divide the parallelogram into two triangles and a parallelogram. The first triangle is sent to the segment \([Ax, Ax + e_4]\), the middle parallelogram to the left half of the face \((Ax + e_4 - e_1, 1 \land 4)\), and the last triangle to the other half. Note that this map is degenerate on the first triangle: this takes care of the cancellation on the boundary.

It is easily checked that if two faces have a common boundary, the maps we have just defined agree on the boundary. Hence we can in this way define a continuous piecewise linear map from the stepped surface to its image (this piecewise linear map is degenerate on some triangles, which are sent to segments).

It follows that the image of the stepped surface is connected, as an image of a connected set by a continuous map, and it has no boundary because the preimage of this boundary would be a boundary of the initial stepped surface.

Suppose that the restriction of the projection \(\pi_e\) to the image of the stepped surface is not one-to-one; we can then find two points \(x_1, x_2\) on this image that project to the same point of \(P_e\). They have respective preimages \(y_1, y_2\) on the initial stepped surface. Consider the lift to the stepped surface of the segment joining \(\pi_e(y_1)\) to \(\pi_e(y_2)\). A case study shows that the image of this segment by the piecewise linear map cannot project to the closed curve, which is a contradiction. Hence the restriction of \(\pi_e\) to the image of the stepped surface is one-to-one. It follows that it is a homeomorphism, and the image is in good position. \(\Box\)

We now prove that the iterated image of a face in a stepped surface in good position is connected. We might think that it follows immediately from the previous construction of a continuous map from a stepped surface to its image, but this does not work because the image of \((x, 3 \land 4)\) by the continuous map is not exactly \(E_2(\sigma)(x, 3 \land 4)\): we have added a segment. We need some additional precautions.

Definition 3.10. A finite sequence \((F_n)_{n=1,\ldots,N}\) of faces is said to be \(*\)-connected if each pair of consecutive faces \(F_n, F_{n+1}\) either have a common edge, or one is of the form \((x - e_k, k \land 4)\) and the other is of the form \((x - e_j, i \land j)\), with \(j \geq k\) or \((x, i \land j)\), with \(j < 4\). (Remark that, in the last cases, the two faces share the vertex \(x\).)

A set of faces is said to be \(*\)-connected if every pair of faces in the set can be linked by a \(*\)-connected sequence.

It is clear that a \(*\)-connected set of faces is connected in the usual topological sense.

Proposition 3.11. The image by \(E_2(\sigma)\) of a \(*\)-connected set of faces is \(*\)-connected.

Proof. It is clearly sufficient to prove that the image of a \(*\)-connected sequence is \(*\)-connected. It is obviously true for a face. If two faces share a boundary of type \((x, 1), (x, 2)\) or \((x, 4)\), their images share the image of this boundary. It is also true if they share a boundary of type \((x, 3)\), except in the case where one face is of the type \((x, 3 \land 4)\) and the other one is of the type \((x, i \land 3)\), with \(i < 3\), or \((x - e_4, 3 \land 4)\). In that case, their respective images are \((Ax + e_4 - e_1, 1 \land 4)\) and \((Ax, i + 1 \land 4)\) or \((Ax, 1 \land 4)\); hence it is still \(*\)-connected (the common vertex being \(Ax + e_4\)).

Now suppose that we are in the second case. The image of \((x - e_k, k \land 4)\) contains a face of type \((Ax - e_{k'}, k' \land 4)\), with \(k' = k + 1\) if \(k = 1, 2\) and \(k' = 1\) if \(k = 3\). The
image of \((x - e_j, i \wedge j)\) is \((Ax - e_{j+1}, i + 1 \wedge j + 1)\) if \(i, j < 4\); in that case, it is still *-connected. The image of \((x - e_4, i \wedge 4)\) contains a face of the type \((Ax, 1 \wedge i + 1)\), and it shares with the other face the vertex \(x\) if \(i < 3\) and the boundary segment \((Ax, 4)\) if \(i = 3\). The case of a face of type \((x, i \wedge j)\), with \(j < 4\), is treated in the same way: it shares with the other face the vertex \(x\) or a boundary.

3.4. Periodic stepped surface. We are now in a position to define a stepped surface which is periodic under the action of \(E_2(\sigma)\).

Remark that the image of \((-e_3, 3 \wedge 4)\) is \((-e_1, 1 \wedge 4)\); the image of \((-e_1, 1 \wedge 4)\) contains \((-e_2, 2 \wedge 4)\), whose image contains \((-e_3, 3 \wedge 4)\). Hence \(E_2(\sigma)^3(-e_3, 3 \wedge 4)\) contains \((-e_3, 3 \wedge 4)\). We get the following theorem:

**Theorem 1.** There exists a unique stepped surface \(S_c\) in good position containing \((-e_3, 3 \wedge 4)\) such that

\[
E_2(\sigma)^3(S_c) = S_c.
\]

This stepped surface stays within bounded distance of the expanding plane.

*Proof.* We can find a stepped surface in good position containing \((-e_3, 3 \wedge 4)\) (just consider the plane generated by \(e_3\) and \(e_4\)). By iterating \(E_2(\sigma)^3\) on this surface, we get a sequence of surfaces in good position containing \(E_2(\sigma)^{3n}(-e_3, 3 \wedge 4)\) for all \(n\).

It is immediate that these sets form an increasing sequence: by Proposition 3.11 these sets are connected, hence their union is a connected set in good position. It remains to prove that it has no boundary.

However, we see that, for any segment \(\gamma = (x, i)\), the distance between \(A\gamma\) and \(E_1(\sigma)(\gamma)\) is at most \(\sqrt{7}/2\). This is of course still true when we consider the projection to the expanding plane. The action of \(A\) on the expanding plane is equivalent to the complex multiplication by \(\lambda\). Hence, if the distance of \(\gamma\) to the origin is larger than \(\frac{\sqrt{7}/2}{2(\lambda-1)} \approx 3.84\), its image is strictly farther from the origin. An explicit computation shows that there exists an \(n\) such that the projection of the image of the boundary of \((-e_3, 3 \wedge 4)\) does not intersect the disk of radius 4 (\(n = 21\) works). Hence this image goes to the infinite, and \(\bigcup_{n \in \mathbb{N}} E_2(\sigma)^{3n}(-e_3, 3 \wedge 4)\) is a stepped surface in good position. It is obviously periodic for \(E_2(\sigma)\) of period 3.

Since the action of \(E_2(\sigma)\) is a bounded perturbation of the linear map \(A\), the projection on the contracting plane must stay within bounded distance of 0. Hence the stepped surface stays within bounded distance of the expanding plane. \(\square\)

We see in Figure 3.2 the first iterates of the substitution.

**Remark 3.12.** We have in fact built three periodic surfaces; one can check that these surfaces have some faces in common. From that one can easily prove that they have a common subset of positive density, and it seems highly probable that this set is of asymptotic density 1. This is reminiscent of the condition of strong coincidence for usual substitutions: up to a set of density 0, there is a unique fixed surface.

**Remark 3.13.** We have only considered stepped surfaces with integer vertices. However, our definitions extend readily to stepped surfaces with real vertices, and we could imitate the methods of [BK06] and build an attractor for \(E_2(\sigma)\) with a natural \(\mathbb{R}^2\)-action. We will come back to this point in the last section.
Figure 3.2. The first iterates of $E_2(\sigma)$; remark the periodicity of order 3.
4. **Polygonal substitution tiling**

In this section we consider how morphisms can be used to construct substitution rules and tilings of the plane (the techniques outlined here should apply in higher dimensions as well, but we will not consider them explicitly here). This technique was originally used by Kenyon [Ken96]. We will explain how to define from $E_2(\sigma)$ a substitution rule on the expanding plane, and deduce from that another proof of the fact that $E_2(\sigma)$ acts on the surfaces in good position.

We start with a simple remark:

**Lemma 4.1.** Under the projection $\pi_e$ the paths $\sigma(i ji^{-1}j^{-1})$, for each choice of $i$ and $j$, do not self-intersect, except for $\sigma(34^{-1}1^{-1}4^{-1}) = 441^{-1}4^{-1}14^{-1}$ which can be cyclically reduced.

**Proof.** By computation. □

We also remark that, by definition of the matrix $A$, the path corresponding to $\sigma(i)$ and the image by $A$ of the path corresponding to $i$ have the same extremities.

We now consider tilings of the plane $P_e$ by the six tiles $\pi_e(i \wedge j)$; such tilings obviously exist (each tile is a parallelogram, hence it can tile the plane). We can define a substitution rule on these tilings:

**Proposition 4.2.** The projection of $E_2(\sigma)$ to the expanding eigenspace $P_e$ of $A$ defines a substitution rule on $P_e$ with tiles $\pi_e(i \wedge j)$.

**Proof.** We consider such a tiling of the plane. We first act by the matrix $A$ and obtain a new tiling with edges of type $A.i$; we now replace each such edge with a path $\sigma(i)$. Because of the previous lemma, these paths do not intersect in their
interior (except in the case of the image of $3 \land 4$), and they determine polygons which are made of one or two tiles.

In fact, we have just replaced a tile of the form $\pi_e(x, \iota \land j)$ by $\pi_e(E_2(\sigma)(x, \iota \land j))$. The construction ensures that the tiles in the image have disjoint interior and cover the plane, hence the substitution rule acts on the space of tilings (Figure 3.1 in fact shows exactly this substitution rule).

This gives us an alternate proof of Proposition 3.9:

**Proof.** The projection of the surface defines a tiling with vertices in $\mathbb{Z}[\lambda]$. Its image by the substitution rule is also a tiling with vertices in $\mathbb{Z}[\lambda]$. However, since the projection from $\mathbb{Z}^4$ to $\mathbb{Z}[\lambda]$ is one-to-one, this tiling lifts in a unique way to a surface, which is in good position by construction. An immediate computation shows that this surface is the image of the initial one by $E_2(\sigma)$.

**Proposition 4.3.** By iterating any non-empty patch, we can cover disks of arbitrarily large radius.

**Proof.** We consider the number of tiles which could be on the boundary of the patches. The number of tiles in the $n$-th patch starting from the tile $e_1 \land e_2$ is given by $|A^n(1, 0, 0, 0, 0)^T|$. The ratio of the number of tiles from $n$ to $n + 1$ will therefore tend towards the leading eigenvalue of $A^*$ as $n$ goes to infinity. The number of edges is at most $|A^n(2, 2, 0, 0)^T|$, where $A'$ is the matrix for $\sigma$ that counts every instance of a letter or its inverse. Again, the ratio from $n$ to $n + 1$ will tend towards the leading eigenvalue of $A'$ as $n$ goes towards infinity. The number of tiles on the edge of a patch will also be strictly less than this. The leading eigenvalue of $A^*$ is approximately 1.40126, and the leading eigenvalue of $A'$ is approximately 1.38028. As 1.38028 is less than 1.40126, the ratio of the number of tiles to the length of the boundary goes to infinity with $n$. This is only possible if the successive patches contain arbitrarily large disks: if any tile was at bounded distance from the boundary, the ratio of the number of tiles to the length of the boundary would be bounded.

Direct computation using the structure matrix of $E_2(\sigma)$ shows that the tile $e_1 \land e_2$ is contained in the iterate of order eight of any patch. Hence the preceding proof applies to any non-empty patch.

Using the previous section, we see that the tile $\pi_e(-e_3, 3 \land 4)$ belongs to its third image by the substitution rule. By iterating on this face, and using the previous proposition, we can prove:

**Theorem 2.** There is a unique tiling containing the tile $\pi_e(-e_3, 3 \land 4)$ which is periodic of period 3 under the substitution rule associated to $E_2(\sigma)$. This tiling is the image by $\pi_e$ of the stepped surface $S_e$ defined above.

Figure 1.3 shows a patch of this tiling.

5. **Dual extension**

We would like to build a discrete approximation to the contracting plane. This is possible by using the dual extension defined above. We will need some notation for this dual extension.
5.1. **Orientation of dual faces.** Projection on the expanding and contracting planes do not define the same orientation on elementary faces, and we need to use a different set of faces for the dual extension. We define a set of faces with the same orientation with respect to the projection \( \pi_c \) on the contracting plane (see Figure 5.1).

**Definition 5.1.** We denote by \( F_c \) the set of positively oriented faces for the projection \( \pi_c \) and a particular orientation on \( P_c \), chosen so that we have
\[
F_c = \{ (0, 1 \wedge 2), (0, 3 \wedge 1), (0, 1 \wedge 4), (0, 2 \wedge 3), (0, 4 \wedge 2), (0, 3 \wedge 4) \}.
\]

**Remark 5.2.** There is a simple rule to relate the orientation of faces with respect to the expanding or contracting planes: let \((0, i \wedge j)\) be a face with positive orientation with respect to \( \pi_e \). Let \((0, k \wedge l)\) be the complementary face, with \( i, j, k, l \) all distinct. The face \((0, k \wedge l)\) is positively oriented with respect to \( \pi_c \) if and only if the permutation \((i, j, k, l)\) is an even permutation.

**Definition 5.3.** We denote by \( \psi \) the map \( \psi : F_c \rightarrow F_e \) defined by \( \psi(i \wedge j) = k \wedge l \), where \((i, j, k, l)\) is an even permutation.

An immediate computation shows that the map \( \psi \), extended to all possible pairs \( i \wedge j \), is an involution. Hence it makes sense to also denote by \( \psi \) the inverse map \( F_e \rightarrow F_c \). The map \( \psi \) will be useful in sections 8 and 9.

5.2. **Dual extension.** We can compute the geometric dual map by using Proposition 2.14. The inverse of the automorphism is given by
\[
\sigma^{-1} : \begin{array}{rcl}
1 & \mapsto & 4^{-1}3 \\
2 & \mapsto & 1 \\
3 & \mapsto & 2 \\
4 & \mapsto & 3.
\end{array}
\]

From this, we deduce the geometric dual map \( E^2(\sigma) \):

**Proposition 5.4.** The map \( E^2(\sigma) \) is defined on the basic faces by
\[
\begin{align*}
(0, 1 \wedge 2) & \mapsto (0, 3 \wedge 1) + (e_3 - e_4, 1 \wedge 4), \\
(0, 3 \wedge 1) & \mapsto (0, 2 \wedge 3) + (e_3 - e_4, 4 \wedge 2), \\
(0, 1 \wedge 4) & \mapsto (e_3 - e_4, 3 \wedge 4), \\
(0, 2 \wedge 3) & \mapsto (0, 1 \wedge 2), \\
(0, 4 \wedge 2) & \mapsto (0, 3 \wedge 1), \\
(0, 3 \wedge 4) & \mapsto (0, 2 \wedge 3).
\end{align*}
\]

It also satisfies the positivity condition.

We check immediately that \( E^2(\sigma)^3(0, 1 \wedge 2) \) contains \((0, 1 \wedge 2)\). In fact, we have an even stronger property: we get a fixed surface.

**Lemma 5.5.** Let \( V = (0, 1 \wedge 2) + (0, 2 \wedge 3) + (0, 3 \wedge 1) \). Then \( V \) is contained in \( E^2(\sigma)(V) \).
We deduce immediately:

**Theorem 3.** There exists a unique stepped surface $S_c$ containing $V$ which is fixed by $E^2(\sigma)$. This stepped surface stays within bounded distance of the contracting plane $P_c$. Its projection to the contracting plane defines a polygonal tiling which is invariant under the natural substitution rule associated to $E^2(\sigma)$. 

Figure 5.1. The dual map $E^2(\sigma)$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Consider the sequence containing each of the six basic tiles.

**Proof.** Consider the sequence $\mathcal{V}_n$ defined by $\mathcal{V}_0 = \mathcal{V}$, and $\mathcal{V}_{n+1} = E^2(\sigma)(\mathcal{V}_n)$. From the previous lemma, it is clear that this sequence is increasing. Using the same techniques as above, we can prove that these sets are connected subsets of a surface in good position and that they contain an arbitrary large disk centered at the origin.

It will prove useful in the sequel to consider a generating patch for this tiling containing each of the six basic tiles.
Definition 5.6. We define the patch \( U_c \) by
\[
U_c := (0, 1 \wedge 2) + (0, 3 \wedge 1) + (e_3 - e_4, 1 \wedge 4) \\
+ (0, 2 \wedge 3) + (e_3 - e_4, 4 \wedge 2) + (e_2 - e_4, 3 \wedge 4).
\]
For all \( i \wedge j \in F_c \), we denote by \( x_{i \wedge j} \) the basic vertex of the file of type \( i \wedge j \) of \( U_c \), so that \( x_{i \wedge j} = 0 \) if \( i \wedge j = 1 \wedge 2, 2 \wedge 3, 3 \wedge 1 \), \( x_{i \wedge j} = e_3 - e_4 \) if \( i \wedge j = 1 \wedge 4, 4 \wedge 2 \), and \( x_{i \wedge j} = e_2 - e_4 \) if \( i \wedge j = 3 \wedge 4 \).

A direct computation shows that \( U_c \) is included in its image by \( E^2(\sigma) \) (the image contains 2 more tiles; see Figure 5.2 to understand the action of \( E^2(\sigma) \) on \( U_c \)).

5.3. Vertices of the invariant stepped surface. The invariant stepped surface is a union of faces \( (x, i \wedge j) \). We will be interested at a later time in the vertices of this stepped surface.

Definition 5.7. We denote by \( V_c(i \wedge j) \) the subset of \( Z^4 \) of basic vertices of faces of type \( i \wedge j \in F_c \) belonging to the stepped surface \( S_c \).

The sets \( V_c(i \wedge j) \) satisfy a remarkable set equation:

Lemma 5.8. We have:
\[
\begin{align*}
V_c(1 \wedge 2) & = A^{-1}V_c(2 \wedge 3), \\
V_c(3 \wedge 1) & = A^{-1}V_c(4 \wedge 2) \cup A^{-1}V_c(1 \wedge 2), \\
V_c(1 \wedge 4) & = A^{-1}V_c(1 \wedge 2) + e_3 - e_4, \\
V_c(2 \wedge 3) & = A^{-1}V_c(3 \wedge 1) \cup A^{-1}V_c(3 \wedge 4), \\
V_c(4 \wedge 2) & = A^{-1}V_c(3 \wedge 1) + e_3 - e_4, \\
V_c(3 \wedge 4) & = A^{-1}V_c(1 \wedge 4) + e_3 - e_4.
\end{align*}
\]

Proof. This is an immediate consequence of the substitution rule. For example, a face of type \( 1 \wedge 2 \) has to be the image by the generalized substitution of a face of type \( 2 \wedge 3 \); hence the first equality \( V_c(1 \wedge 2) = A^{-1}V_c(2 \wedge 3) \). The other relations are obtained in the same way, paying due care to the possible offset.

We can rephrase this set equation in a form that will be useful later by using the map \( \psi \) and multiplying by \( A \) on both sides:
\[
\begin{align*}
AV_c(\psi(1 \wedge 2)) & = V_c(\psi(2 \wedge 3)) + e_1, \\
AV_c(\psi(1 \wedge 3)) & = V_c(\psi(2 \wedge 4)) + e_1, \\
AV_c(\psi(1 \wedge 4)) & = V_c(\psi(2 \wedge 4)) \cup V_c(\psi(1 \wedge 2)), \\
AV_c(\psi(2 \wedge 3)) & = V_c(\psi(3 \wedge 4)) + e_1, \\
AV_c(\psi(2 \wedge 4)) & = V_c(\psi(1 \wedge 3)) \cup V_c(\psi(3 \wedge 4)), \\
AV_c(\psi(3 \wedge 4)) & = V_c(\psi(1 \wedge 4)).
\end{align*}
\]

6. Self-similar fractal tilings

There is a well-known re-normalization procedure to obtain an exact self-similar tiling from a substitution tiling with expanding substitution rule: start with a basic tile, apply the substitution, re-normalize, and iterate the procedure (for more details, see [PS01]). In this section, we are going to apply explicitly this procedure for the tiling defined in the previous section.
We prefer to deal with the tiling of the contracting plane defined by the dual map because it is fixed by the substitution rule, while the tiling on the expanding plane is only periodic for its substitution rule, which makes things more complicated to state.

6.1. Definition of the fractal tile. We have the following easy proposition from [AIS01]:

**Proposition 6.1.** For any tile \((0, i \wedge j)\), the sequence of sets

\[ A^n \pi_c \left( E^2(\sigma)^n(0, i \wedge j) \right) \]

converges in the Hausdorff metric.

**Proof.** It is easy to check that, for any patch \(T\) in the stepped surface, the Hausdorff distance between \(\pi_c(T)\) and \(A \pi_c \left( E^2(\sigma)(T) \right)\) is uniformly bounded, because it is true for any face \((x, i \wedge j)\).

From this it follows that the Hausdorff distance between two successive sets in the sequence \(A^n \pi_c \left( E^2(\sigma)^n(0, i \wedge j) \right)\) decreases geometrically fast; hence the sequence converges to a unique compact set (see [AIS01, AI01] for more details). \(\square\)

**Definition 6.2.** We will denote by \(Y_c(i \wedge j) \in P_c\) the limit of the sequence \(A^n \pi_c \left( E^2(\sigma)^n(0, i \wedge j) \right)\).

Since the action of \(E^2(\sigma)\) is a bounded distortion of \(A^{-1}\), a simple computation shows the following:

**Lemma 6.3.** For any \(x \in \mathbb{Z}^4\), the sequence \(A^n \pi_c \left( E^2(\sigma)^n(x, i \wedge j) \right)\) converges to \(\pi_c(x) + Y_c(i \wedge j)\).

We will be especially interested in the re-normalization of the basic tiles \((x_i \wedge j, i \wedge j)\) of \(U_c\) defined in the previous section.

**Definition 6.4.** We denote by \(X_c(i \wedge j)\) the sets \(\pi_c(x_i \wedge j) + Y_c(i \wedge j)\) obtained by the re-normalization of the tiles of \(U_c\), and by \(X_c\) the union \(\bigcup_{i \wedge j \in F_c} X_c(i \wedge j)\).

Figure 6.1 shows the limit tile \(X_c\). Note that, due to cancellations on the boundary, the topological arrangement of the fractal tiles is not exactly the same as that of the parallelogram tiles, a common boundary has disappeared, and two contact points have been enlarged to a curve boundary.

6.2. Properties of the fractal tiles. A first important remark is that the action of the substitution on the boundaries of the faces is given, as explained in Proposition 2.14 by the automorphism \(\tilde{\sigma}^{-1}\).

**Lemma 6.5.** The automorphism \(\tilde{\sigma}^{-1}\) has no cancellation on iterated images of letters.

**Proof.** \(\tilde{\sigma}^{-1}\) is defined by

\[
\begin{align*}
\tilde{\sigma}^{-1} : & \quad 1 \mapsto 34^{-1} \\
& \quad 2 \mapsto 1 \\
& \quad 3 \mapsto 2 \\
& \quad 4 \mapsto 3.
\end{align*}
\]
Hence cancellations can only come from the word $4^{-1}1$ or its inverse, since all letters have images with different initial and final letter, except 1 and 4. Consider the following set of words of length 2:

$$E = \{12^{-1}, 21^{-1}, 14^{-1}1^{-1}, 23^{-1}, 32^{-1}, 34^{-1}, 1^{-1}2, 2^{-1}1,$$

$$1^{-1}3, 3^{-1}1, 2^{-1}3, 3^{-1}2, 3^{-1}4, 4^{-1}3\}.$$

We say that a word is $E$-admissible if any of its factors of length 2 belong to $E$ (that is, the word belongs to the Markov system defined by $E$). A direct computation shows that if a word is $E$-admissible, its image by $\tilde{\sigma}^{-1}$ is also $E$-admissible. Since 1 has image $34^{-1}$, and all the other letters eventually have 1 as its image, it follows that all iterated images of letters are $E$-admissible. Since the word $4^{-1}1$ and its inverse are not in $E$, there can be no cancellations on the images of the letters. □

Remark 6.6. In the language of Bestvina-Handel (see [BH92]), this means that the natural action of the automorphism $\tilde{\sigma}^{-1}$ on the rose with four petals (the graph with one vertex and four edges) is a train-track map.

Proposition 6.7. The Hausdorff dimension of the boundary of the fractal tile has dimension less than 2; hence this boundary has measure 0.

Proof. If a sequence of compact sets $X_n$ converges to $X$ in the Hausdorff topology, the boundary of $X$ is included in the Hausdorff limit of the boundaries of $X_n$, when this limit exists. But the previous lemma implies that, for any tile $(0, i \wedge j)$, the number of segments in the boundary of the image $E^2(\sigma)^n(0, i \wedge j)$ grows exponentially with ratio $\alpha \approx 1.22074$, where $\alpha$ is the largest eigenvalue of the matrix $B$ that counts the number of occurrences of each letter in the image of each letter ($B$ is the absolute value of the abelianization matrix of $\tilde{\sigma}^{-1}$; see Proposition 4.3). We can then compute the Hausdorff dimension of the limit of the boundary, which is $\log(\alpha)/\log(|\lambda|) \approx 1.18242$. 

\[\text{Figure 6.1. } U_c \text{ and } X_c.\]
Hence the Hausdorff dimension of the boundary is less than 2, and the boundary has measure 0. □

From this proposition, we get several interesting consequences. We first give a simple lemma on Hausdorff limits.

**Lemma 6.8.** Let \((X_n)\) (resp. \((Y_n)\)) be a sequence of compact sets that are the closure of their interior, and such that, for all \(n\), \(X_n\) and \(Y_n\) have disjoint interior. Suppose that the sequence \(X_n\) (resp. \(Y_n\)) converges in the Hausdorff metric to \(X\) (resp. \(Y\)) and that their boundaries also have a limit. Any point in the intersection \(X \cap Y\) is contained in the intersection of the limit of the boundaries.

**Proof.** If \(x\) is in the intersection, it means that we can find a sequence \((x_n) \in X_n\) (resp. \((y_n) \in Y_n\)) that converges to \(x\). The segment \([x_n, y_n]\) must contain a point \(x'_n\) (resp. \(y'_n\)) in the boundary of \(X_n\) (resp. \(Y_n\)), and these two sequences also converge to \(x\). □

We recall a proposition from [AI01, Prop.9]:

**Lemma 6.9.** If a sequence \(X_n\) converges to \(X\) in the Hausdorff metric, the measure of \(X\) is at least the upper limit of the measures of the \(X_n\).

From this, we get:

**Proposition 6.10.** Let \((x, i \wedge j)\) and \((y, k \wedge l)\) be two faces of the stepped surface. Then the sets \(\pi_c x + Y_c(i \wedge j)\) and \(\pi_c y + Y_c(k \wedge l)\) have disjoint interior. Furthermore, the sets \((0, i \wedge j)\) and \(Y_c(i \wedge j)\) have the same measure.

**Proof.** Since \(E^2(\sigma)\) acts on the stepped surface in such a way that any face is in the image of exactly one face of the stepped surface, it is clear that the sets \(A^n \pi_c (E^2(\sigma)^n(x, i \wedge j))\) and \(A^n \pi_c (E^2(\sigma)^n(y, k \wedge l))\) have disjoint interior. Hence any point in the intersection of the limits should belong to the limit of the boundary. However, this set has measure 0, and hence an empty interior; hence the interiors of \(\pi_c x + Y_c(i \wedge j)\) and \(\pi_c y + Y_c(k \wedge l)\) cannot intersect.

By Proposition 3.7, we know that all the sets \(A^n \pi_c (E^2(\sigma)^n(x, i \wedge j))\) have the same measure, because the vector of measures of the six faces is an eigenvector for the linear map associated to \(E^2(\sigma)\). Hence the measure of \(Y_c(i \wedge j)\) is at least that of \((0, i \wedge j)\). But if we take a large ball in \(P_c\) and cover it by projections of disjoint faces, we can also find in this ball the same number of disjoint limit sets, which have at least the same measure. This implies that these limit sets have exactly the same measure as the faces, otherwise they should have an intersection of non-zero measure. □

The fact that the vector of measures of the limit sets is an eigenvector for the map associated to \(E^2(\sigma)\) is also a consequence of the following, which gives another characterization of the sets \(X_c(i \wedge j)\), as an iterated function system.
Figure 6.2. $X_c(i \land j), \bigcup_{i \land j \in F_c} X_c(i \land j)$, and $\bigcup_{i \land j \in F_c} A^{-1}X_c(i \land j)$. 
Proposition 6.11. The six sets $X_c(i \wedge j)$ are disjoint and satisfy the following set equation:

\begin{align*}
A^{-1}X_c(1 \wedge 2) &= X_c(3 \wedge 1) \cup X_c(1 \wedge 4), \\
A^{-1}X_c(3 \wedge 1) &= X_c(2 \wedge 3) \cup X_c(4 \wedge 2), \\
A^{-1}X_c(1 \wedge 4) &= X_c(3 \wedge 4), \\
A^{-1}X_c(2 \wedge 3) &= X_c(1 \wedge 2), \\
A^{-1}X_c(4 \wedge 2) &= X_c(3 \wedge 1) + \pi_c(e_2 - e_3), \\
A^{-1}X_c(3 \wedge 4) &= X_c(2 \wedge 3) + \pi_c(e_1 - e_3).
\end{align*}

This proposition is nothing other than the action of $E^2(\sigma)$ taken to the limit; it will prove essential in section 8. We can see this set equation in Figure 6.2.

6.3. Self-similar fractal tiling of the contracting plane. An immediate consequence of the previous results is the following theorem:

Theorem 4. The sets $x + Y_c(i \wedge j)$, for $x \in \pi_c(V_c(i \wedge j))$, tile the plane $P_c$, and this tiling is self-similar for the action of the matrix $A^{-1}$ on the plane.

Proof. It is an easy consequence of the previous results: since by definition two faces of the invariant stepped surface project to parallelograms in $P_c$ with disjoint interior, the corresponding fractal tiles also have disjoint interior. The rest is left to the reader. \hfill \Box

Figure 6.3 shows a patch of this fractal tiling.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fractal_tiling.png}
\caption{A patch of the self-similar fractal tiling of the contracting plane.}
\end{figure}

6.4. Self-similar fractal tiling of the expanding plane. A similar construction applies to the expanding plane. It is slightly more complicated: first because we start with a periodic tiling, so we have to work with the third power of the substitution rule, and we get several fixed points; second because the automorphism used to compute the boundary is not in train-track form. There are cancellations, and it is not easy to compute exactly the Hausdorff measure of the boundary.
Figure 6.4. The tiles $Y_e(i \wedge j)$. 
It is however not difficult to prove that it is less than 2, so everything works as in the contracting plane. More computation shows that when we apply the automorphism $\sigma$ to a letter, the sequence of words we obtain increases exponentially, with a ratio $\beta \approx 1.20322$, where $\beta$ is the unique real root of the polynomial $X^7 - X^2 - X - 1$. Hence the Hausdorff dimension of the boundary is $\log(\beta)/\log(|\lambda|) \approx 1.0967$. This dimension is smaller than that for the tiles in the contracting plane because the growth of the boundary is slower, due to the cancellations. We can see in Figure 6.4 that the boundary indeed looks smoother for the tiles on the expanding plane.

We first need some notation.

**Definition 6.12.** For any $i \land j \in F_e$, the sequence of sets $A^{-n}\pi_e(E_2(\sigma)(0, i \land j))$ converges in the Hausdorff metric. We denote this limit by $Y_e(i \land j)$.

**Remark 6.13.** These sets are not disjoint, as can be seen in the figures (all of them contain the origin as an inner point). This should not come as a surprise: this is already the case for the similar construction in the trivial example of the Fibonacci automorphism, where the projections of the two basis vectors on the expanding line intersect.

**Proposition 6.14.** The sets $Y_e(i \land j)$ satisfy the following equations:

- $AY_e(1 \land 2) = Y_e(2 \land 3)$,
- $AY_e(1 \land 3) = Y_e(2 \land 4)$,
- $AY_e(1 \land 4) = Y_e(2 \land 4) \cup (Y_e(1 \land 2) + \pi_e(e_4 - e_1))$,
- $AY_e(2 \land 3) = Y_e(3 \land 4)$,
- $AY_e(2 \land 4) = Y_e(3 \land 4) \cup (Y_e(1 \land 3) + \pi_e(e_4 - e_1))$,
- $AY_e(3 \land 4) = Y_e(1 \land 4) + \pi_e(e_4 - e_1)$.

Using the vertices of the faces of the stepped surface $S_e$ defined in section 3, we get a dual tiling of the expanding plane. We need a definition to state the theorem.
We denote by $V_e(i \wedge j)$ the set of basic vertices of the faces of type $i \wedge j$ of the stepped surface $S_e$.

**Theorem 4 (Dual version).** The sets $\pi_e(x) + Y_e(i \wedge j)$, for $x \in V_e(i \wedge j)$, tile the plane. This tiling is self-similar for the action of the matrix $A^3$ on the expanding plane $P_e$.

We can see in Figure 6.5 a patch of this self-similar tiling; the curves in the figure are traces of the cancellations on the boundary.

## 7. Windows and Atomic Surfaces

In this section we show that the tilings generated by the substitution rules described above can be considered as windowed projection tilings (also called cut-and-project tilings). Their windows are made up of the fractal tiles of the other tiling.

More precisely, we want to prove that the vertices belonging to the stepped surface approximating one of the invariant planes are selected by a window in the other plane. Remark first that the basic vertex of any face of $S_e$ is a vertex of the stepped surface but that not all vertices appear in this way; some vertices may appear several times as a common basic vertex for different faces. We want to find a rule to enumerate the vertices. There are many ways to do this, but, as our tiles are parallelograms, we may use the following convention.

**Definition 7.1 (Vertex-tile map).** Let $T$ be a tiling of the plane by parallelograms. Choose a direction $d$ in the plane that is not parallel to any edge of the tiling. The vertex-tile map is the map $V$ from vertices of the tiling to tiles, where $V(v)$ is the tile in the direction $d$ of the vertex $v$.

**Lemma 7.2.** The vertex-tile map is a bijection.

**Proof.** Consider a parallelogram tile $T$ in $T$. As $T$ is a parallelogram we may put its four corners together to divide a circle into four. The direction $d$ will lie in precisely one of these four angles (as it is not parallel to the edges of $T$). Thus the vertex-tile map will map precisely one vertex to $T$. This holds for all tiles in $T$, thus the vertex-tile map is a bijection.

In the following proof we will need to use the construction of a canonical projection tiling. This is a general method used to construct quasiperiodic tilings of the plane and has been studied in great detail; for example see [Sen95]. The construction was originally used by de Bruijn who showed that it gave an alternative construction of the Penrose tiling [BS]. The general result of when such tilings have a substitution rule is given in [Har03, HL04]. For this example we need only define two canonical projection tilings:

**Definition 7.3 (Canonical projection tilings).** Consider the incidence matrix $A$ of the morphism $\sigma$, with the two eigenspaces $P_e$ and $P_c$. Let $A$ be the set of points $\mathbb{Z}^4 \cap (P_e + \mathcal{H})$, where $\mathcal{H}$ is the unit hypercube in dimension 4. Let $\mathcal{F}$ be the collection of 2-dimensional faces of the cubic lattice in $\mathbb{R}^4$ whose four vertices belong to $A$.

We define the canonical projection tiling $C_e$ to be $\pi_e(\mathcal{F})$. Note that $P_e$ and $P_c$ are so positioned that this projection gives a tiling; in other words, it is the projection of a surface in good position.
Figure 7.1. The structure of the window for the canonical projection tiling $C_e$. In (a) we obtain the window region for the tiles $\pi_e(2 \wedge 3)$. The vertices associated to these tiles under the vertex-tile bijection are at the ends of edges $\pi_e(e_2)$ and $\pi_e(e_3)$. Therefore, we consider the regions of the window from which a translation by $\pi_e(e_2)$ and $\pi_e(e_3)$ will stay in the window. The intersection of these two regions contains all the points associated to all vertices of the tiling which are at the ends of edges $\pi_e(e_2)$ and $\pi_e(e_3)$. Thus this region can be associated to the tile $\pi_e(2 \wedge 3)$. In (b) the window regions for all six tiles are shown.

We define the tiling $C_c$ in an analogous manner with the roles of $P_e$ and $P_c$ reversed.

We now wish to use the window to understand the tiling. To do this we need to show how points in the window correspond to points in the tiling. In particular we can divide the window into regions which contain all the points corresponding to a given tile type under the vertex-tile bijection. This is best described in a picture; see Figure 7.1.

The shape of the window and its regions can now be summed up in a lemma.
The projection of $E_2(\sigma)$

First iteration on the window

Second iteration on the window

$n$-th iteration on the window

Figure 7.2. Building the window of the substitution rule, from the canonical projection window.

Lemma 7.4. The tilings $C_e$ and $C_c$ have an octagonal window that can be divided into six regions, each corresponding to the vertices of a particular tile in the vertex-tile bijection.

Proof. The window is $\pi_c(\mathcal{H})$, for $C_e$, and $\pi_e(\mathcal{H})$, for $C_c$. Thus it is an octagon in both cases. The six regions are described in Figure 7.1. □

Theorem 5. The tilings generated by the projection of $E_2(\sigma)$ are projection tilings, and the window is a generalized octagon formed of translates of the fractal tiles of the dual tiling.
Proof. Consider the tiling $C_e$ and its associated window $\Omega_e$. This is the projection of a surface in good position. We may therefore apply the substitution rule $E_2(\sigma)$ to obtain a new surface in good position by Proposition 3.9. Iterating the substitution rule we obtain a sequence of tilings. The limit of the sequence of tilings is a tiling generated by the substitution rule, as every patch will have infinitely many predecessors. To obtain the window for the limit tiling we therefore consider how $E_2(\sigma)$ effects the window and how this behaves under the limit.

This is very similar to what we did in the previous section. The first stages of constructing the window are shown in Figure 7.2. Consider the effect of $E_2(\sigma)$ on the window. The action of $E_2(\sigma)$ first applies the matrix $A$ that acts as a linear map on the window (as the window space $P_e$ is an eigenspace of $A$). We now have the tiling $A C_e$ with window $A \Omega_e$, shown in the second column. The window is divided into six regions by Lemma 7.4, each corresponding to a single tile type. Each tile type is translated by a set of vectors in $Z^4$ under $E_2(\sigma)$. Each window region is therefore translated by the corresponding projections to $P_e$, shown in the third column. As the new surface is in good position none of these translated regions overlap. We have therefore obtained the window for the new tiling. Finally, in the fourth column we can consider the new window, with the regions coloured according to their associated tiles. This process can be iterated, yielding a sequence of windows. The action is given explicitly by the set equation on the vertices of the tiling (see Lemma 5.8). Call $Z_e(i \wedge j)$ the closure of the projection by $\pi_e$ of the set $V_e(\psi(i \wedge j))$. From the equations, we get

\[
\begin{align*}
AZ_e(1 \wedge 2) &= Z_e(2 \wedge 3) + \pi_e(e_1), \\
AZ_e(1 \wedge 3) &= Z_e(2 \wedge 4) + \pi_e(e_1), \\
AZ_e(1 \wedge 4) &= Z_e(2 \wedge 4) \cup Z_e(1 \wedge 2), \\
AZ_e(2 \wedge 3) &= Z_e(3 \wedge 4) + \pi_e(e_1), \\
AZ_e(2 \wedge 4) &= Z_e(3 \wedge 4) \cup Z_e(1 \wedge 3), \\
AZ_e(3 \wedge 4) &= Z_e(1 \wedge 4).
\end{align*}
\]

Now consider the boundary of the windows. The six regions of the canonical projection’s window are bounded by four lines. After applying the substitution rule, the new regions are bounded by four stepped lines. At each further iteration we obtain four new stepped lines. The new stepped lines are made up of the previous set of stepped lines under the contraction. We may therefore obtain the morphism $\tilde{\sigma}^{-1}$ which generates the boundary of the infinite limit. From this we may observe that the construction of the window is the dual substitution and that the edges of the window at each stage are replaced under the morphism $\tilde{\sigma}^{-1}$.

To construct a section of the boundary, therefore, we iterate the morphism $\tilde{\sigma}^{-1}$ on the edges, re-normalizing by a contracting root of $A$ for each iterate. As was proved in the previous section, the limit is a curve of Hausdorff dimension smaller than 2.

We may show the corresponding result for the dual tiling by following an analogous proof starting with $C_e$ and using the dual substitution. \qed
The set equation given above is very similar to the set equation satisfied by $Y_c(i \land j)$. Indeed, the sets $Z_c$ are congruent by a translation to the opposite of $Y_c$, and we can show:

**Proposition 7.5.** We have $Z_c(i \land j) = \pi_c(v_{i\land j}) - Y_c(i \land j)$, where the vectors $v_{i\land j}$ are defined by

\[
\begin{align*}
v_{1\land 2} &= e_1 + e_2 - e_4, \\
v_{1\land 3} &= e_1 + e_3 - e_4, \\
v_{1\land 4} &= e_1, \\
v_{2\land 3} &= e_1 + e_3 - e_4, \\
v_{2\land 4} &= e_2, \\
v_{3\land 4} &= e_3.
\end{align*}
\]

**Proof.** If we replace $Z_c(i \land j)$ by $v_{i\land j} - K(i \land j)$ in the system above, a direct computation shows that we recover the set equation for $Y_c$. However, such a system can have only one solution, hence $K(i \land j) = Y_c(i \land j)$.

In section 8 we will need the following definition and lemma:

**Definition 7.6.** We define the sets $X_c(i \land j)$ by $X_c(i \land j) = -Z_c(i \land j)$.

**Lemma 7.7.** For any $x \in V_c(i \land j)$, the set $\pi_c(x) + X_c(\psi(i \land j))$ contains 0.

**Proof.** By definition, $Z_c(\psi(i \land j))$ contains the projection of $x \in V_c(i \land j)$; hence $X_c(\psi(i \land j))$ contains $-\pi_c(x)$ and $\pi_c(x) + X_c(\psi(i \land j))$ contains 0.

8. **Complex numeration system associated to the tilings**

We now want to define a numeration system on the set $X_c$ and, using the map $\phi_c$ defined in section 2, a related numeration system in base $\mu$ on the complex plane. We will do so by defining a Markov map on $X_c$ and symbolic dynamics for this Markov map. We first give notation.

**Definition 8.1.** We will denote by $f_1$ the vector $f_1 = \pi_c(e_2 - e_3)$ and by $f_2$ the vector $f_2 = \pi_c(e_1 - e_3)$.

**Definition 8.2.** We define the transformation $T : X_c = \bigcup_{i\land j} X_c(i \land j) \to X_c$ by

\[
T(x) := \begin{cases}
A^{-1}x & \text{if } x \in X_c(i \land j), i \land j \neq 4 \land 2, 3 \land 4, \\
A^{-1}x - f_1 & \text{if } x \in X_c(4 \land 2), \\
A^{-1}x - f_2 & \text{if } x \in X_c(3 \land 4).
\end{cases}
\]

**Lemma 8.3.** The transformation $T$ is well defined, except perhaps on a set of measure 0.

**Proof.** This is an immediate consequence of Proposition 8.1. The image of $X_c(i \land j)$ by $A^{-1}$ is included in $X_c$ except for $X_c(4 \land 2)$ and $X_c(3 \land 4)$, in which case it is taken back to $X_c$ by the respective subtraction of $f_1$ or $f_2$.

The only problem that could occur is with the points that belong to two sets $X_c(i \land j)$ at the same time, and hence to their boundary, because the transformation is not well defined in this case. However, this set is of measure 0.

**Remark 8.4.** From now on, we will systematically neglect sets of measure 0. A more detailed analysis is needed to take care of them, but we do not want to concentrate on the matter in this paper.
Definition 8.5. We define the digit function $a^* : X_c \to \{0, f_1, f_2\}$ by

$$a^*(x) := \begin{cases} 0 & \text{if } x \in X_c(i \land j), \ i \land j \neq 4 \land 2, \ 3 \land 4, \\ f_1 & \text{if } x \in X_c(4 \land 2), \\ f_2 & \text{if } x \in X_c(3 \land 4). \end{cases}$$

For any $x \in X_c$, we can hence define a sequence of digits by $a^*_n(x) := a^*(T^{n-1}x)$, $n \in \mathbb{N}$. We want to characterize admissible sequences of digits corresponding to points of $X_c$. We define the graph $G_c$ below.

![Graph](graph.png)

Then, we obtain the following theorem.

Theorem 6. The dynamical system $(X_c, T)$ is a Markov map with respect to the partition $X_c(i \land j)$. Its adjacency matrix is the structure matrix of $E^2(\sigma)$.

The set $\Omega_c$ of admissible digit sequences $(a^*_n)$ is given by the labels of infinite paths in the graph $G_c$ above. Any element $x \in X_c$ can be written in terms of its digit sequence $a^*_n(x)$:

$$x = \sum_{n=1}^{\infty} A^na^*_n(x).$$

Proof. Proposition 8.11 shows not only that $T$ is well-defined, but also that the partition $X_c(i \land j)$ is a Markov partition. The graph $G_c$ follows immediately from the Markov condition.

It is now clear by recurrence that one can write $x = A^NT^N(x) + \sum_{n=1}^{N} A^na^*_n(x)$. However, since $T^N(x)$ is bounded, $A^NT^N(x)$ goes exponentially fast to 0, which proves the required equality. \hfill $\square$

It is not difficult from this point to prove the following proposition:

Proposition 8.6. The set $\Omega_c$ is shift invariant; the dynamical system $(\Omega_c, S)$ (where $S$ is the shift map) is a sofic system, with a unique measure of maximal entropy. The dynamical system $(X_c, T)$ also has a unique measure of maximal entropy and is measurably conjugate to $(\Omega_c, S)$.

We remark that it is easy to determine the partition from the digit sequence: $a^*_1(x) = f_1$ if and only if $x \in X_c(4 \land 2)$; $a^*_1(x) = f_2$ if and only if $x \in X_c(3 \land 4)$; $a^*_1(x) = 0$ and $a^*_2(x) = f_2$ if and only if $x \in X_c(1 \land 4)$. For the remaining three sets, the condition is slightly more complicated: $x$ is in $X_c(3 \land 1)$ if and only if the first non-zero digit is of the form $a^*_{3n+2} = f_1$ or $a^*_{3n+5} = f_1$, and similar conditions for the 2 remaining sets. There are of course points admitting several expansions, and it is possible to characterize them (by a rather large automaton).
Using the map $\phi_c$ defined in section 2, we can rephrase the previous theorem to obtain a complex expansion in base $\mu$, where $\mu$ is the small eigenvalue of $A$ defined in section 2.

**Definition 8.7.** For a complex number $z \in \phi_c(X_c)$, we define a sequence of complex digits, with value in the set $\{0, \mu - \mu^2, 1 - \mu^2\}$, by

$$\alpha^*_n(z) = \phi_c(a^*_n(T^{n-1}(\phi_c^{-1}(z)))).$$ 

Since the origin 0 is contained in the interior of $X_c$, for any $x \in P_c$, we can find a power $A^n$ of $A$ such that $A^n x \in X_c$. This gives us an expansion for all complex numbers in base $\mu$. Hence we get a complex version of the previous theorem:

**Theorem 6 (Complex version).** Any complex number $z \in \phi_c(X_c)$ can be written in base $\mu$ as

$$z = \sum_{n=1}^{\infty} \alpha^*_n(z)\mu^n.$$ 

Furthermore, for any $z \in \mathbb{C}$, we can find $n_0 \in \mathbb{N}$ and an admissible sequence $(\alpha^*_n(z))_{n \geq -n_0}$ such that

$$z = \sum_{n=-n_0}^{\infty} \alpha^*_n(z)\mu^n.$$ 

**Remark 8.8.** It is easy to prove that this expansion is unique for almost all complex numbers. The “integral part” of the expansion, for $n \leq 0$, defines the tile of the self-similar tiling of the complex plane to which the number belongs, which gives an arithmetic interpretation of the tiling.

**Remark 8.9.** There is a unique infinite path in $G_c$ with all labels 0, which alternates between $1 \land 2$, $3 \land 1$, $2 \land 3$. This proves that the point 0 belongs to all three sets, and hence to their boundary.

We can define a natural extension of the one-sided shift on $\Omega_c$; a simple way is to consider bi-infinite admissible sequences. Another way is to define the “dual graph” $G_e$ for $G_c$, as below:

Please note that the dual graph $G_e$ is given not only by reversing the direction of each edge of $G_c$, but also by changing the name of vertices by the map $\psi$ given in Definition 5.3. We can then define admissible sequences $(\beta_n)_{n \in \mathbb{N}}$ for $G_e$ and define an expansion on the expanding plane. There is a complex version, where numbers

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
can be written in an almost unique way

\[ z = \sum_{n=-n_0}^{\infty} \beta_n^*(z) \lambda^{-n}. \]

We must however be careful, because the sets we define by this expansion (in a similar way to the previous expansion) do not have disjoint interiors; the corresponding map must be defined on an indexed union. We will make this more precise in the next section.

9. Markov Partition for the Toral Automorphism

We consider the set

\[ \hat{X} = \bigcup_{i \land j \in F_c} \hat{X}(i \land j) = \bigcup_{i \land j \in F_c} X_c(i \land j) + X_c(\psi(i \land j)). \]

**Lemma 9.1.** The Lebesgue measure of \( \hat{X} \) is 1.

**Proof.** Remark first that the sets \( \hat{X}(i \land j) \) are disjoint up to a set of measure 0 (because the same property is true for the sets \( X_c(i \land j) \) in the plane \( P_c \)). So it is sufficient to prove that the sum of the measures of \( \hat{X}(i \land j) \) is 1.

However, we have proved above that the Lebesgue measure of \( X_c(i \land j) \) is the same as the measure of the projected parallelogram \( \pi_c(0, i \land j) \), and the same property holds for \( X_c(0, i \land j) \).

Hence we are reduced to proving that the sum of the volumes of the six parallelopetes \( \pi_c(0, i \land j) + \pi_c(0, \psi(i \land j)) \) is equal to 1.

But this turns out to be just the formula for the computation of the determinant of the canonical basis of \( \mathbb{Z}^4 \), written in a basis whose first two vectors are in \( P_c \) and last two vectors are in \( P_c \): consider the 6 possible matrices of order 2 which can be built from the first two rows, and the 6 complementary matrices from the last two rows.

**Lemma 9.2.** For any \( x \in V_c(i \land j) \), the set \( x + X_c(i \land j) + X_c(\psi(i \land j)) \) intersects \( P_c \) in \( \pi_c(x) + X_c(i \land j) \).

**Proof.** We can write \( x + X_c(i \land j) + X_c(\psi(i \land j)) = \pi_c(x) + X_c(i \land j) + \pi_c(x) + X_c(\psi(i \land j)) \). But, from Lemma 9.1 we know that \( 0 \in \pi_c(x) + X_c(\psi(i \land j)) \); hence this set intersects \( P_c \) in \( \pi_c(x) + X_c(i \land j) \). □

We deduce the immediate corollary:

**Corollary 9.3.** The set \( \bigcup_{i \land j \in F_c} \bigcup_{x \in V_c(i \land j)} x + X_c(i \land j) + X_c(\psi(i \land j)) \) contains the plane \( P_c \).

**Proof.** The intersection of \( P_c \) with the set under consideration is, by the previous lemma, equal to \( \bigcup_{i \land j \in F_c} \bigcup_{x \in V_c(i \land j)} \pi_c(x) + X_c(i \land j) \). However, this is exactly the invariant tiling, and it covers the plane. □

**Theorem 7.** The set \( \hat{X} \) is a fundamental domain for the lattice \( \mathbb{Z}^4 \). Its canonical decomposition in sets \( X_c(i \land j) + X_c(\psi(i \land j)) \) is a Markov partition for the automorphism \( A \) of the 4-dimensional torus \( \mathbb{R}^4 / \mathbb{Z}^4 \).
Proof: Since $\hat{X}$ has volume one, to prove that it is a fundamental domain, it is enough to prove that $\mathbb{Z}^4 + \hat{X}$ covers all of $\mathbb{R}^4$. But the previous theorem proves that $\bigcup_{(i,j) \in F_c} X_e(i \wedge j) + \mathbb{Z}^4 \subset \mathbb{R}^4$. It is then clear that, for any $y \in \mathbb{Z}^4$, $\bigcup_{(i,j) \in F_c} \left( X_e(i \wedge j) + y \right) \subset \mathbb{Z}^4 + \hat{X}$ contains the plane $P_c$. Hence the union of translates of $\hat{X}$ contains a dense subset of $\mathbb{R}^4$. Since this set is closed, as a locally finite union of closed sets, it contains all of $\mathbb{R}^4$.

The Markov partition is an immediate consequence of the set equations given above. □

Remark 9.4. We could use the complex numeration system defined above in this setting. The map $\Phi : \mathbb{R}^4 \to \mathbb{C}^2$ conjugates the automorphism to a map which is locally of the form $(z, w) \mapsto (\lambda z + c_1, \mu w + c_2)$, where $c_1, c_2$ are locally constant; $\lambda$ is the digit of the numeration system. To each point in $\hat{X}$, we can associate symbolic dynamics and a bi-infinite sequence of digits which is the natural extension of the Markov map $T$ defining the complex numeration system.

Remark 9.5. The set $\hat{X}$ tiles the space $\mathbb{R}^4$ periodically. The real meaning of Lemma 9.2 is that the invariant tiling of the plane $P_c$ is obtained by intersecting the plane with the translates of $\hat{X}$. We can in the same way consider the tiling obtained by intersecting the plane $P_c$ with the translates of $X$. The problem here is that the origin belongs to three sets $X_e(i \wedge j)$, as can be seen in Figure 5.2 (and as is evident also from the numeration system, since the zero sequence of digits has three possible initial states). Hence there are three possible tilings for the expanding plane, corresponding to the three stepped surfaces seen above.

It seems very probable that these three stepped surfaces coincide out of a set of asymptotic density 0. This is reminiscent of the Coincidence Condition for Pisot substitutions, which implies that if a substitution has several fixed points, then these coincide except on a set of density 0.

10. ADDITIONAL REMARKS AND EXAMPLES

10.1. Other examples. We consider the following automorphism:

$$
\tau : \begin{align*}
1 & \mapsto 12 \\
2 & \mapsto 32 \\
3 & \mapsto 4 \\
4 & \mapsto 1^{-1}.
\end{align*}
$$

This example also satisfies the positivity condition. We can again get two dual tilings, stepped surfaces, and a Markov partition for the associated toral automorphism.

Similarly for the following example:

$$
\rho : \begin{align*}
1 & \mapsto 213 \\
2 & \mapsto 21 \\
3 & \mapsto 2134 \\
4 & \mapsto 134.
\end{align*}
$$
In that case, the contracting eigenvalues are complex and the expanding eigenvalues are real. We can see in Figure 10.1 the fractal tile in the contracting plane and in Figure 10.2 the corresponding fractal tile in the expanding plane. (The iterated function system associated to the fractal tiles on the expanding plane is given by an affine map, so we call this tile self-affine instead of self-similar; the figure shows a privileged direction corresponding to the largest eigenvector.) This case proves that there are substitutions that satisfy the positivity condition.

Finally, we remark that there is, as we will show elsewhere, an infinite family of automorphisms satisfying the positivity condition. Indeed, let $a \in \mathbb{Z}$, and $b$ satisfying $b = -a - 1, -a, -a + 1$, and define the automorphism $\phi_{a,b}$ by

$$
\phi_{a,b} : 
\begin{align*}
1 & \mapsto 2 \\
2 & \mapsto 3 \\
3 & \mapsto 4 \\
4 & \mapsto 2^a 4^b 1^{-1}.
\end{align*}
$$

One can prove that, for any value of $a$ and $b$ as above, $\phi_{a,b}$ is a hyperbolic automorphism satisfying the positivity condition. All possible combinations of real and complex eigenvalues occur in this family.
The automorphism $\phi_{3,3}$ is a totally real example; we can see the corresponding tiles in Figures 10.3 and 10.4. They seem to have rather bad topological properties, with a high dimension for the boundary.

![Image](image.png)

**Figure 10.3.** The self-affine fractal tile of the contracting plane for $\phi_{3,3}$.

10.2. **Open problems.**

10.2.1. **Transversal dynamics.** There is a natural $\mathbb{R}^2$ action on the plane $P_e$ and on the space of tilings generated by the substitution rule, and another one on the plane $P_c$. These generalize the stable and unstable flows for an automorphism of the torus $\mathbb{T}^2$ and give rise to some kind of transversal dynamics in the sense of Vershik.

For the 2-dimensional torus, these stable and unstable flows have a nice symbolic dynamic. By taking a section, we can consider the first return map and obtain a substitution dynamical system.

The analogue in the present case is not so clear; one of the problems is that the definition of a section for an $\mathbb{R}^2$-action is not simple. However, we can get a pseudo-group of partially defined exchange of pieces by considering the 8 different ways the faces of a hypercube can project to a plane (as compared with the 2 ways the faces of a 3-dimensional cube can project, giving rise to the “flip” movement). We can find an interesting geometric realization of this pseudo-group in terms of tiling: if we consider the space of all polygonal tilings generated by the substitution rule and the subsets of tilings which have a vertex at the origin, we can define elementary moves by translating the tiling along an edge adjacent to the origin. The pseudo-group is generated by these elementary moves (which are not everywhere defined),
and we can approximate in this way any translation up to the half-diameter of the tiles. The pseudo-group is a discrete approximation of the $\mathbb{R}^2$-action, and this is probably the best generalization of the 1-dimensional case that we can expect.

10.2.2. Faces with real vertices. In fact, the formulas for $E_2(\sigma)$ and $E^2(\sigma)$ extend immediately to faces with arbitrary real vertices, instead of integer vertices. We can in this way define real surfaces in good position with respect to $\pi_e$ (or $\pi_c$) and consider the action of $E_2(\sigma)$ on this set.

This is the basic idea of [BK06]. It should not be difficult to prove that the limit invariant set for this action is the space of surfaces generated by the translates of the invariant surface we have found along the invariant plane. This space admits a natural $\mathbb{R}^2$ action by translation, which is the transversal dynamics considered above, and an invariant measure; it is certainly measurably isomorphic to the torus $\mathbb{T}^4$. 

**Figure 10.4.** The self-affine fractal tile of the expanding plane for $\phi_{3,3}$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
This construction can certainly be generalized to a large class of automorphisms, in any dimension (although pictures will be more difficult to get in dimension higher than 4).

10.2.3. Positivity condition. We want to generalize the result. The positivity condition, which is essential to our work, generalizes the condition that a Pisot automorphism has a representative which is a substitution.

Is it possible to generalize our results to all hyperbolic free group automorphisms satisfying the positivity condition? Recall that, already, in the Pisot case, this is not trivial: it is still not proved that all Pisot substitutions have discrete spectrum. This is the well-known discrete spectrum conjecture which states, in its most restricted form, that the symbolic dynamical system associated with a unimodular Pisot substitution has discrete spectrum. This conjecture has several equivalent formulations in terms of combinatorics or geometry. One of these is that the limit tile obtained by renormalization has the same measure as the initial polygonal tile. This is what we proved in section 6, by showing that the limit of the renormalized boundary has measure 0.

It would be interesting to characterize by simple conditions the automorphisms that satisfy the positivity condition; for information on this problem, see [Man02, Rob96].

10.2.4. Non-positive automorphisms. There has been some progress in the Pisot case for automorphisms which are not substitutions. It is possible in that case to define in a more complicated way a symbolic dynamical system and a geometric representation.

Can we obtain results of this type in the hyperbolic non-Pisot case? Some of our results still work in this framework. We can obtain fractal tiles and stepped surfaces which stay within bounded distance of the invariant planes, but these stepped surface are generally not in good position. This suggests that it would be useful to have a theory of “algebraic” tilings with overlapping tiles. The state of the question is very unclear at the moment (see [Fur06]).

We would like to thank the anonymous referee for his detailed comments and corrections on the paper.

References


Some properties of invertible substitutions of rank d, and higher dimensional substitutions, Osaka Journal of Mathematics 40 (2003), 543–562. MR1988704


The dynamical properties of Penrose tilings, Transactions of the American Mathematical Society 348 (1996), 4447–4464. MR1355301 (97a:52041)

Quasicrystals and geometry, Cambridge University Press, 1995. MR1340198 (96c:52038)


Markov partitions and U-diffeomorphisms, Funkts. Anal. Prilozhen. 2 (1968), 64–89. MR0230385 (38:1361)

Groups, tilings and finite state automata, AMS Colloquium Lectures (1989), Lectures notes distributed in conjunction with the Colloquium Series.


IML UPR-CNRS 9016, Campus de Luminy case 907, 13288 Marseille Cedex 9, France
E-mail address: arnoux@iml.univ-mrs.fr

Faculty of Business Administration, Yokohama University, 79-4, Tokiwadai, Hodogaya-ku, Yokohama, 240-8501, Japan
E-mail address: furukado@ynu.ac.jp

Department of Mathematics, University of Arkansas, Fayetteville, Arkansas 72701
E-mail address: edmund.harriss@mathematicians.org.uk

Graduate School of Natural Science and Technology of Kanazawa University, Kakuma-machi, Kanazawa, 920-1192, Japan
E-mail address: ito@t.kanazawa-u.ac.jp