7. Spectral theory and geometric representation of substitutions

From geometry to symbolic dynamics. As explained in Chap. 5, symbolic dynamical systems were first introduced to better understand the dynamics of geometric maps. Indeed, by coding the orbits of a dynamical system with respect to a cleverly chosen finite partition indexed by the alphabet $\mathcal{A}$, one can replace the initial dynamical system, which may be difficult to understand, by a simpler dynamical system, that is, the shift map on a subset of $\mathcal{A}^{\mathbb{N}}$.

This old idea was used intensively, up to these days, particularly to study dynamical systems for which past and future are disjoint, such as toral automorphisms or pseudo-Anosov diffeomorphisms of surfaces. These systems with no memory, whose entropy is strictly positive, are coded by subshifts of finite type, defined by a finite number of forbidden words. Some very important literature has been devoted to their many properties (see [265]). The partitions which provide a good description for a topological dynamical system, leading to a subshift of finite type, are called Markov partitions (a precise definition will be given in Sec. 7.1).

Self-Similar dynamics: where substitutions naturally appear. For a dynamical system, it is a usual problem to try to understand the local structure of its orbits. A classical method to study this problem is to consider the first return map (Poincaré map) over an appropriate neighborhood of a given point. For some systems such as toral quadratic rotations or some interval exchanges with parameters living in a quadratic extension, the system defined by the first return map on some subset is topologically conjugated to the original system. One can say that the original dynamical system has a *self-similar structure*. A basic idea is that, in general, as soon as self-similarity appears, a substitution is hidden behind the original dynamical system.

A significant example is the addition φ of the golden ratio α on the one-dimensional torus $\mathbb{T}$: the first return map of this map on the interval $[1-\alpha, 1[$ is the addition of α over $\mathbb{T}$, up to a reversal of the orientation and a renormalization (see Chap. 6).

Indeed, let us choose the natural partition of $\mathbb{T}$ by intervals of continuity $\mathbb{T} = I_2 \cup I_1$, with $I_2 = [0, 1-\alpha[$ and $I_1 = [1-\alpha, 1[$. The first return map

[1] This chapter has been written by A. Siegel

over I_1, denoted by φ_1, is conjugate to the initial map. The conjugacy maps the initial partition onto a partition $J_1 \cup J_2$ of I_1. An easy computation gives $J_1 = [1-\alpha, 2-2\alpha[$ and $J_2 = [2-2\alpha, 1[$. It is immediate to check that J_1 is included in I_1, its image is I_2 and its second image is included in I_1, whereas J_2 is included in I_1 and is immediately mapped into the return subset I_1. The Fibonacci substitution $1 \mapsto 12$ and $2 \mapsto 1$ naturally appears here.

More precisely, for any point $x \in J_1$, if we denote by w the symbolic sequence given by the orbit of x with respect to the partition $I_1 \cup I_2$, and by v the symbolic sequence for the induced map with respect to the partition $J_1 \cup J_2$, it is clear that w is the image of v by the Fibonacci substitution. In particular, the symbolic sequence of the unique fixed point of the conjugacy is the fixed point of the Fibonacci substitution. It is then easy to prove that all the symbolic sequences associated with points in $\mathbb{T}$ belong to the symbolic dynamical system generated by the substitution (see also Chap. 6).

This situation is quite general: in case of self-similarity, under suitable hypotheses, the trajectories (with respect to the partition) of points in the return subset, before they come back into the subset, define a substitution. In that case, the codings of the trajectories of points of the full system belong to the symbolic system associated with this substitution (see Chaps. 1 and 5).

Is this a good representation? We have defined a coding map from the geometric system onto the substitutive system. The question is: how far is this map from being a bijection?

For the example of the toral addition of the golden ratio, we can define an inverse map, from the symbolic system onto the torus. It is proved that this map is continuous, 2-to-1, and 1-to-1 except on a countable set (see Chap. 6); this is the best possible result, given the fact that one of the sets is connected and the other one a Cantor set.

For other examples, the question can be much more difficult.

It is natural then to focus on the reverse question: given a substitution, which self-similar actions are coded by this substitution?

For the Morse substitution, it was proved that the symbolic dynamical system associated with this substitution is a two-point extension of the dyadic odometer, that is, the group $\mathbb{Z}_2$ of 2-adic integers (see Chap. 5).

The three-letter equivalent of the Fibonacci substitution is the Tribonacci substitution $1 \mapsto 12$, $2 \mapsto 13$, $3 \mapsto 1$. G. Rauzy, with methods taken from number theory, proved in 1981 that the symbolic dynamical system associated with this substitution is measure-theoretically isomorphic, by a continuous map, to a domain exchange on a self-similar compact subset of $\mathbb{R}^2$ called the *Rauzy fractal* [350]. Tiling properties of the Rauzy fractal bring an isomorphism between the substitutive system and a translation on the two-dimensional torus.

Considering these examples, we see the connection between the search for a geometric interpretation of symbolic dynamical systems and understanding whether substitutive dynamical systems are isomorphic to already known dynamical systems or if they are new. Since substitutive dynamical systems are deterministic, i.e., of zero entropy, they are very different from a subshift of finite type. Hence, the following question is natural: which substitutive dynamical systems are isomorphic to a rotation on a compact group? More generally, what is their maximal equicontinuous factor ?

Some history. A precise answer was obtained for substitutions of constant length during the seventies [229, 230, 281, 135]. We know that the maximal equicontinuous factor of a substitutive system of constant length l is a translation on the direct product of the adic group $\mathbb{Z}_l$ and a finite group. There exists a measure-theoretic isomorphism between such a substitutive system and its maximal equicontinuous factor, if and only if the substitution satisfies a combinatorial condition called the *coincidence condition*.

It is natural but more difficult to study substitutions of nonconstant length. G. Rauzy first tackled this question with a complete study of the system associated with the Tribonacci substitution [350]. Then, B. Host made a significant advance by proving that all eigenfunctions of primitive substitutive dynamical systems are continuous [215]. Thus, the two main dynamical classifications (up to measure-theoretic isomorphism and topological conjugacy) are equivalent for primitive substitutive systems. In the continuation of this, it was proved that the spectrum of a substitutive system can be divided into two parts. The first part has an arithmetic origin, and depends only on the incidence matrix of the substitution. The second part has a combinatorial origin, and is related to the return words associated with the fixed point of the substitution [174].

Many papers deal with conditions for a substitutive dynamical system to have a purely discrete spectrum [215, 216, 266, 267, 268, 439, 292, 411, 412, 413, 206]. Some are necessary conditions, others are sufficient conditions. In particular, P. Michel, B. Host and independently A. N. Livshits, whose work was restated and generalized by M. Hollander, defined a combinatorial condition on two-letter substitutions called *coincidence condition*, which generalizes the condition of F. M. Dekking for constant length substitutions, for two-letter primitive substitutions. This condition is equivalent with the purely discrete spectrum property.

Except for the Tribonacci substitution [350] and some examples developed by B. Solomyak [411, 412], these works do not give an explicit realization of the maximal equicontinuous factor. For every unimodular substitution of Pisot type on d letters, P. Arnoux and S. Ito build explicitly a self-similar compact subset of $\mathbb{R}^{d-1}$ called generalized *Rauzy fractal* [43]. They generalize the coincidence condition to all nonconstant length substitutions, and prove that this condition is sufficient for the system first to be semi-topologically conjugate to a domain exchange on the Rauzy fractal of the substitution, second to admit as a topological factor a minimal translation on the $(d-1)$-dimensional torus. The techniques they use do not provide results for a measure-theoretic isomorphism between the substitutive system and its toral topological translation factor, as is the case for a two-letter alphabet.

An alternative construction of the Rauzy fractal associated with a substitution of Pisot type can be obtained by generalizing the techniques used by G. Rauzy [350, 353], as developed by V. Canterini and A. Siegel [102, 103]. This construction is based on the use of formal power series, and provides new proofs for the results of P. Arnoux and S. Ito in [43]. Moreover, this point of view allows one to give a combinatorial necessary and sufficient condition for a substitutive unimodular system of Pisot type to be measure-theoretically isomorphic to its toral topological translation factor [398]. This has consequences for the construction of explicit Markov partitions for toral automorphisms, the main eigenvalue of which is a Pisot number [397]. An interesting feature of those Markov partitions is that their topological and geometrical properties can be studied, such as connectedness [101], fractal boundary [289, 291, 290, 44, 208] or simple connectedness [224].

Description of the chapter. The aim of this chapter is to provide a more precise description of the above results. We have no claim to be exhaustive in our exposition since each of the notions we refer to requires a lot of related material to be exposed in detail.

The aim of Sec. 7.1 is to provide a brief introduction to subshifts of finite type, related to Markov partitions and self-similarity. The aim of Sec. 7.2 is to illustrate the deep relationship between substitutive dynamical systems and shifts of finite type, via the tool of adic transformations.

The entire Sec. 7.3 is devoted to the spectral theory of substitutive dynamical systems. Section 7.3.2 presents an overview of the general spectral theory of substitutive systems of nonconstant length. In Sec. 7.3.3 the main attention is devoted to the spectral properties of systems associated with the extensively studied class of substitutions of Pisot type.

In Sec. 7.4 we expose in details the construction and the properties of the Rauzy fractal defined by G. Rauzy to get a geometric realization of the system associated with the Tribonacci substitution. As a fifth part, in Sec. 7.5 we explore in details the construction and properties of the Rauzy fractal for any substitution of Pisot type. A special attention is devoted to their tiling properties. Finally, the aim of Sec. 7.6 is to explain how a geometrical representation of substitutions of Pisot type provides explicit Markov partitions for some Pisot toral automorphisms.

7.1 Shifts of finite type: introduction

Our investigation of the properties of substitutive dynamical systems will need the use of the most well known class of symbolic dynamical systems, that is, shifts of finite type. These shifts are introduced in Sec. 7.1.1. Geometrically, shifts of finite type are related to Markov partitions. Section 7.1.2 is devoted to them. These Markov partitions can be seen as a generalization of self-similar tilings, defined in Sec. 7.1.3.

7.1.1 Shifts of finite type

While they have been around implicitly for a long time, the systematic study of shifts of finite type began with Parry [323], Smale [410] and R. Williams [460]. A complete presentation can be found in [265].

Definition. Let $\mathcal{A}$ be a finite set. Endowed with the shift map, a subset X of $\mathcal{A}^{\mathbb{Z}}$ is a *bilateral shift of finite type* if and only if there exists a finite set $\mathcal{F} \subset \mathcal{A}^{\star}$ of finite words over $\mathcal{A}$ such that X is the set of biinfinite words in $\mathcal{A}^{\mathbb{Z}}$ which have no factor in $\mathcal{F}$. The subshift X is also denoted $X_{\mathcal{F}}$.

Equipped with the topology on $\mathcal{A}^{\mathbb{Z}}$, a shift of finite type is a symbolic dynamical system, since it is compact and invariant under the shift. One defines similarly the notion of *unilateral shift of finite type.*

Example 7.1.1. The bilateral fullshift $\{1, \ldots, d\}^{\mathbb{Z}}$ over a d-letter alphabet is a shift of finite type with no forbidden word.

Example 7.1.2. The *golden mean shift of finite type* is the set of all binary sequences with no consecutive 1's. The associated set of forbidden words is $\mathcal{F} = \{11\}$.

Example 7.1.3. The *even shift*, that is, the set of all binary sequences (that is sequences over $\{0, 1\}$) so that between any two 1's there are an even number of 0's, is not a shift of finite type: it cannot be described by a finite number of constraints.

Higher presentation. Let (X, S) be a symbolic dynamical system over the alphabet $\mathcal{A}$. A basic construction defines a new system by looking at the blocks of consecutive letters in the language of X, and by considering them as letters from a new alphabet.

Indeed, let n be an integer. The new alphabet is $\mathcal{A}^{[n]} = \mathcal{L}_n(X)$, that is, the set of blocks of length n that appear in at least one element of X. Define the mapping $\beta_n : X \to \left(\mathcal{A}^{[n]}\right)^{\mathbb{Z}}$ by $\beta_n(w)_i = w_i \ldots w_{i+n-1}$ for all $w \in X$. Let $X^{[n]} = \beta_n(X)$ and let $S^{[n]}$ denote the shift map over $\left(\mathcal{A}^{[n]}\right)^{\mathbb{Z}}$.

Then $(X^{[n]}, S^{[n]})$ is a symbolic dynamical system, called the *n-th higher presentation of* (X, S). If (X, S) is a shift of finite type, $(X^{[n]}, S^{[n]})$ is also a shift of finite type.

Example 7.1.4. The 2-th higher presentation of the golden mean shift of finite type (Example 7.1.2) is the shift of finite type over the alphabet $\mathcal{A}^{[2]} = \{a = 10,\, b = 01,\, c = 00\}$, associated with the following set of forbidden words: $\mathcal{F}^{[2]} = \{aa, bb, bc, ca\}$.

Graphs. A *finite graph* $\mathcal{G}$ consists of a finite set of *vertices* or *states* $\mathcal{V}$ together with a finite set of *edges* $\mathcal{E}$. If e is an edge, $i(e) \in \mathcal{V}$ is the *initial vertex* and $t(e) \in \mathcal{V}$ is the *terminal vertex*. Two vertices can be connected by more than one edge.

The *adjacency matrix* of a finite graph $\mathcal{G}$ has size the cardinality of $\mathcal{V}$. If $(i,t) \in \mathcal{V}^2$, its coefficient of index (i,t) is equal to the number of edges in $\mathcal{E}$ with initial state i and terminal state t. A finite graph is thus uniquely determined by its adjacency matrix (up to graph isomorphism).

Shift associated with a graph. Let $\mathbf{A}$ be a matrix with coefficients in $\{0,1\}$; the *bilateral shift associated with the finite graph* $\mathcal{G}$ of adjacency matrix $\mathbf{A}$ is defined as the following set of biinfinite paths in the graph $\mathcal{G}$, endowed with the shift map:

$$X_{\mathcal{G}} = X_{\mathbf{A}} = \{((e_n)_{n\in\mathbb{Z}} \in \mathcal{E}^{\mathbb{Z}};\ \forall n \in \mathbb{Z}, t(e_n) = i(e_{n+1})\}.$$

One can easily check that this set is closed and hence compact, and invariant under the shift map in $\mathcal{V}^{\mathbb{Z}}$.

The *unilateral shift* is defined as the set of right paths in $\mathcal{G}$, that is, as the set of paths indexed by $\mathbb{N}$.

Example 7.1.5. The shift associated with the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ is the subset of $\{a,b,c\}^{\mathbb{N}}$ that contains all the labels of walks in the graph shown in Fig. 7.1.

Example 7.1.6. The bilateral full-shift $\{1,\ldots,d\}^{\mathbb{Z}}$ over a d-letter alphabet is the shift associated with the matrix $\mathbf{A} = [d]$ of size 1 associated with the graph shown in Fig. 7.2.

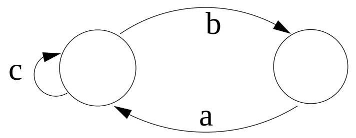


Fig. 7.1. Graph for Example 7.1.5.

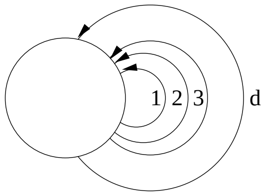


Fig. 7.2. Graph for the fullshift over d letters.

Relationship between the two categories of subshifts. One should remark that any shift associated with a graph is a shift of finite type: the set of forbidden words simply consists of pairs of edges ef such that $t(e) \neq i(f)$.

On the contrary, not every shift of finite type is the shift associated with a graph: for instance, the golden mean shift (Example 7.1.2) cannot be described by a graph.

However, any shift of finite type can be recoded, using a higher block presentation, to be the shift associated with a graph:

Theorem 7.1.7 (see [265]). *Let $(X_{\mathcal{F}}, S)$ be a shift of finite type. Let M denotes the maximum of the lengths in $\mathcal{F}$. Then there exists a graph $\mathcal{G}$ such that the shift associated with $\mathcal{G}$ is the M-th higher presentation of (X, S).*

Example 7.1.8. Up to a recoding of the labels ($a = 10$, $b = 01$, $c = 00$), the 2-th higher presentation of the golden mean shift described in Example 7.1.4, is the shift associated with the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, given in Example 7.1.5.

This results means that shifts of finite type are essentially the same as shifts associated with graphs (up to a recoding). From now on, we identify shifts of finite type and shifts associated with a graph.

Labelled graphs. An *automaton* $(\mathcal{G}, t)$ (also called *labelled graph* by the set $\mathcal{A}$) consists of a finite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ together with a labeling map of the edges $t : \mathcal{E} \to \mathcal{A}$.

Sofic systems. Endowed with the shift map, a subset of $\mathcal{A}^{\mathbb{Z}}$ or $\mathcal{A}^{\mathbb{N}}$ is called a (unilateral or bilateral) *sofic system* if it is defined as the labeling of infinite (unilateral or bilateral) paths of an automaton. Endowed with the topology over $\mathcal{A}^{\mathbb{Z}}$ or $\mathcal{A}^{\mathbb{N}}$, such a system is still a symbolic dynamical system.

Example 7.1.9. The set of all binary sequences so that between any two 1's there is an even number of 0's is a sofic system, associated with the automaton shown in Fig. 7.3. We recall that it is not a shift of finite type.

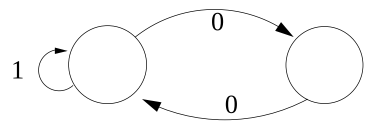


Fig. 7.3. The graph of Example 7.1.9.

Connection between shifts of finite type and sofic systems. A shift of finite type is nothing else that a sofic system associated with an automaton whose all edges have different labels. If so, the automaton is said to be *deterministic*.

These two categories of symbolic dynamical systems are fundamentally connected:

Theorem 7.1.10 (see [265]). *Any sofic system is a topological factor of a shift of finite type, and conversely.*

7.1.2 Markov partitions

As we already explained, one of the main sources of interest in symbolic dynamics is its use in representing other dynamical systems. Indeed, to describe the orbits of an invertible map f on a space $\mathcal{E}$, one may divide the space X into a finite number d of pieces $\mathcal{E}_i$, and then track the orbit of a point $z \in \mathcal{E}$ by keeping a record of which of these pieces $f^n(z)$ lands in. This defines a symbolic sequence $x \in \{1 \ldots d\}^{\mathbb{Z}}$ defined by $x_n = i$, if $f^n(z) \in \mathcal{E}_i$. In short, the pieces are said to be a *Markov partition of* $\mathcal{E}$ if, first the set of all symbolic sequences associated with the points of $\mathcal{E}$ is a shift of finite type, and, secondly, the mapping from $\mathcal{E}$ onto the subshift which maps any point on its coding is precise enough. We refer the reader to [265, 242, 243, 395] for more details on the concepts introduced in this section.

Precise definitions. A *topological partition* of a metric space $\mathcal{E}$ is a finite collection $\Lambda = \{\mathcal{E}_1, \ldots, \mathcal{E}_d\}$ of disjoint open sets whose closure covers $\mathcal{E}$, that is, $\mathcal{E} = \cup \overline{\mathcal{E}_i}$.

Let $(\mathcal{E}, f)$ be a dynamical system. If f is invertible, the *bilateral symbolic dynamical system* associated with a topological partition Λ is the set $X_{\mathcal{E}}$ endowed with the shift map:

$$X_{\mathcal{E}} = \{(x_n)_{n\in\mathbb{Z}} \in \{1, \ldots, d\}^{\mathbb{Z}};\ \exists z \in \mathcal{E}, \forall n \in \mathbb{Z}, f^n(z) \in \mathcal{E}_{x_n}\}.$$

If f is non-invertible, the *one-sided symbolic dynamical system* associated with a topological partition Λ consists of sequences indexed by $\mathbb{N}$, corresponding to the nonnegative orbits.

Definition 7.1.11. *Let $(\mathcal{E}, f)$ be an invertible (respectively non-invertible) dynamical system. A topological partition $\Lambda = \{\mathcal{E}_1, \ldots, \mathcal{E}_d\}$ of $\mathcal{E}$ is said to be a* Markov partition *of $\mathcal{E}$ if*

- *the bilateral (respectively one-sided) symbolic dynamical system $(X_{\mathcal{E}}, S)$ associated with Λ is a shift of finite type;*
- *for every symbolic sequence $x \in X_{\mathcal{E}}$, the intersection*

$$\bigcap_{-\infty}^{+\infty} \overline{f^{-k}(\mathcal{E}_{x_k})} \left(\textit{respectively } \bigcap_{k=0}^{+\infty} \overline{f^{-k}(\mathcal{E}_{x_k})} \right)$$

  consists of exactly one point.

Remarks.

- By definition of the dynamical system $X_{\mathcal{E}}$, the intersection above is never empty. The fact that it consists of exactly one point means that the coding mapping from $\mathcal{E}$ onto $X_{\mathcal{E}}$ is one-to-one.
- There are number of variants on the definition of Markov partition in the literature, many involving geometry of the pieces $\mathcal{E}_i$. This one is simple to state but is somewhat weaker than other variants.

Proposition 7.1.12 (see [265]). *Let $\Lambda = \{\mathcal{E}_1, \dots, \mathcal{E}_d\}$ be a Markov partition of the invertible (respectively non-invertible) dynamical system $(\mathcal{E}, f)$. Let $X_{\mathcal{E}}$ be the bilateral (respectively one-sided) symbolic dynamical system associated with Λ.*

Then the inverse map $\varphi : X_{\mathcal{E}} \to \mathcal{E}$ of the coding mapping is well-defined as follows:

$$\forall x \in X_{\mathcal{E}},\ \{\varphi(x)\} = \bigcap_{-\infty}^{+\infty} \overline{f^{-k}(\mathcal{E}_{x_k})} \ \left(\textit{respectively } \bigcap_{k=0}^{+\infty} \overline{f^{-k}(\mathcal{E}_{x_k})} \right).$$

The map φ is continuous, onto, and realizes a commutative diagram between the shift map on $X_{\mathcal{E}}$ and f on $\mathcal{E}$, that is, $\varphi \circ S = f \circ \varphi$.

Example. The most simple example of a Markov partition is the doubling map on the circle: the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is provided with the non-invertible action $f(z) = 2z \bmod 1$. Subdivide $\mathbb{T}$ into two equal subintervals $I_1 = [0, 1/2)$ and $I_2 = [1/2, 1)$. The alphabet here shall be $\mathcal{A} = \{1, 2\}$. For every $z \in \mathbb{T}$ we define $x_n \in \{1, 2\}$ by $f^n(z) \in I_{x_n}$. The sequence $x = (x_n)_{n \geq 0}$ is an element of the full-shift $\{1, 2\}^{\mathbb{N}}$. Notice that x is nothing else that the sequence of digits of z in its binary expansion. The action of f corresponds to the shift map: the action of multiplying by 2 an element of $\mathbb{T}$, and taking the rest modulo 1 simply shifts the binary digits to the left and delete the first digit. Hence, $\mathbb{T} = I_1 \cup I_2$ realizes a Markov partition for f.

More generally, Markov partitions can be seen as generalizations of the binary system (see [244]).

General interest of Markov partitions. Markov partitions allow a good combinatorial understanding of symbolic dynamical systems. They are partitions of a space $\mathcal{E}$ which describe by means of a shift of finite type, the dynamics of a map acting on $\mathcal{E}$, this description preserving properties such as density of periodic points, transitivity or mixing. The existence of Markov partitions allows one to deduce from the properties of the shift of finite type a "quasi-immediate" description of certain dynamical properties of the system.

For instance, Adler and Weiss show in [7] that topological entropy is a complete invariant for the measure-theoretic isomorphism of continuous ergodic automorphisms of the torus $\mathbb{T}^2$. Their proof is based on the construction of an explicit Markov partition of the torus $\mathbb{T}^2$ with respect to the automorphism.

The intersection property. More precisely, Adler and Weiss define a partition of $\mathbb{T}^2$ which is natural relatively to a linear automorphism f . The key to prove that this partition is a Markov partition is the following *intersection property*:

- the image under f of the intersection of a cylinder $\mathcal{E}_i$ with an affine subspace directed with the stable direction of f is included in a intersection of the same type,

- the image under f^{-1} of the intersection of a cylinder $\mathcal{E}_i$ with an affine subspace directed with the unstable direction of f is included in a intersection of the same type.

In other variants on the definition of Markov partition, this property appears to be a hypothesis in the definition.

Existence of Markov partitions. R. Bowen in [87] shows that for every hyperbolic diffeomorphism of a manifold $\mathcal{E}$, there exists a Markov partition of $\mathcal{E}$. His proof is not constructive.

One should remark that, except for the one and two dimensional-cases, Markov partitions for toral automorphisms are made of pieces that cannot be simple parallelograms. Indeed, R. Bowen proved in [89] that for hyperbolic automorphisms of $\mathbb{T}^3$, the boundary of the elements of the Markov partition are typically fractals: their Haussdorf dimension is not an integer.

Explicit Markov partitions. Markov partitions remained abstract objects for a long time: except for linear automorphisms of the two-dimensional torus and some automorphisms of $\mathbb{T}^3$ [36, 61, 289, 291, 290], explicit constructions of Markov partitions for toral hyperbolic automorphisms were obtained only recently. For precise details about results on this subject, see [244].

Making use of β-developments defined by A. Bertrand-Mathis (see [180] for a general reference), B. Praggastis [336] builds tilings which induce Markov partitions for toral hyperbolic automorphisms whose dominant eigenvalue is a unimodular Pisot number. These constructions are particularly simple and explicit for the automorphisms associated with the companion matrix of a polynomial of the type $x^n - d_1x^{n-1} - \cdots - d_n$, with $d_1 \geq d_2 \geq \cdots \geq d_n = 1$. Indeed, the β-development associated with the Pisot roots of such polynomials admits very special properties [181] that are intensively used by B. Praggastis.

More generally, on the one hand, S. Leborgne [262, 261], on the other hand, R. Kenyon and A. Vershik [245], study the case of hyperbolic automorphisms of $\mathbb{T}^n$, using algebraic properties of the eigenvalues of the automorphism to code its action. Just as with B. Praggastis, those codings are based on a numeration system with a non-integral basis.

Thus, with an equivalent of β-developments, S. Leborgne represents hyperbolic automorphisms by sofic systems. Independently, choosing particular digits in expansions with a non-integral basis, R. Kenyon et A. Vershik construct a representation by a shift of finite type. This representation is not totally explicit, since the digits are chosen at random.

Relation with substitutions. We will see in Sec. 7.6 how a geometrical representation of substitutive systems of Pisot type provides a Markov partition for the incidence matrix of the substitution, the Markov partition being described by the combinatorial representation in terms of subshifts exposed in Sec. 7.2.2.

7.1.3 Self-similarity

Markov partitions usually generate tilings with a specific property: self-similarity. Let us set up precise definitions about this notion.

Tilings. A collection Λ of compact subsets of $\mathbb{R}^n$ is a *tiling* of a set $\mathcal{E}$ which is the closure of its interior if and only if:

- $\mathcal{E} = \cup_{C\in\Lambda} C$,
- every element of Λ is the closure of its interior,
- every compact subset of $\mathcal{E}$ intersects a finite number of elements of Λ,
- elements of Λ have disjoint interior.

Periodic tilings. The tiling is said to be a *periodic tiling* of $\mathbb{R}^n$ modulo $\mathcal{L}$, where $\mathcal{L}$ is a lattice, if and only if

- $\mathbb{R}^n = \cup_{C\in\Lambda,\, \mathbf{z}\in\mathcal{L}} (C+\mathbf{z})$,
- for all $\mathbf{z} \in \mathcal{L}$, $(C+\mathbf{z}) \cap \operatorname{Int} C' \neq \emptyset$ implies $\mathbf{z} = \mathbf{0}$ and $C = C'$.

Self-similar sets. A tiling is said to be *self-similar* if it is invariant under the action of a given expanding diagonal map in $\mathbb{R}^n$, the image of each tile under the mapping being an exact union of original tiles. More precise and general definitions can be found in [242, 243, 431].

A self-similar tiling can be seen as a Markov partition for the expanding map associated with the tiling, and can be seen as a generalization of the decimal system, through the theory of the representation of numbers in non-standard bases (for more details, see [244]).

A compact subset $\mathcal{E}$ of $\mathbb{R}^n$ is said to be *self-similar* if there exists a finite partition of $\mathcal{E}$ whose translates under a given lattice generate a self-similar and periodic tiling of $\mathbb{R}^n$.

7.2 Substitutive dynamical systems and shifts of finite type

The spectral theory of substitutive systems of constant length was developed during the seventies (see Sec. 7.3). One of the main problems raised by the generalization of this theory to substitutions of nonconstant length was to find a general way to decide positions in which any factor of the fixed point appears, that is, to "desubstitute" the substitutive system. The important notion of recognizability deals with this problem. As it will be explained in Sec. 7.2.1, it took quite a long time to be correctly defined and to get general results about it.

The problem of recognizability being solved, an important part of the results on spectral theory of nonconstant length substitutive systems was obtained by using a deep relationship between substitutive dynamical systems and an other important class of symbolic dynamical systems, that is, shifts of finite type. In Sec. 7.2.2 we focus on this relationship.

In Sec. 7.2.3, we expose how the notion of recognizability allows one to make explicit the results of Sec. 7.2.2, thanks to the useful tool of prefix-suffix expansion.

7.2.1 Recognizability

Let σ be a primitive substitution over the alphabet $\mathcal{A}$, and let u be a biinfinite fixed point of σ. We thus have $u = \sigma(u)$. Let E_1 be the following set of lengths

$$E_1 = \{0\} \cup \{|\sigma(u_0 \ldots u_{p-1})|, |\sigma(u_{-p} \ldots u_{-1})|;\, p > 0\}.$$

Since $u = \sigma(u)$, for every factor $W = u_i \ldots u_{i+|W|-1}$ of u, there exists a rank j, a length l, a suffix S of $\sigma(u_j)$ and a prefix P of $\sigma(u_{j+l+1})$ such that

$$W = S\sigma(u_{j+1}) \ldots \sigma(u_{j+l})P,$$

and such that $E_1 \cap \{i, \ldots, i + |W| - 1\} = (i - k) + E_1 \cap \{k, \ldots, k + |W| - 1\}$, with $k = |\sigma(u_0 \ldots u_j)| - |S|$, which means that it is equivalent to cut u with respect to σ at the ranks i and k.

We say that $[S, \sigma(u_{j+1}), \ldots, \sigma(u_{j+l}), P]$ is the *1-cutting at the rank i* of W, and that W comes from the word $u_j \ldots u_{j+l+1}$, which will be called the *ancestor word* of W.

By minimality, the factor W appears infinitely many times in the sequence u. The question of the unicity of the 1-cutting of W is thus natural. Let us briefly survey the existing literature on this question.

In [282], J. C. Martin calls *rank one determined* substitutions for which every long enough factor of a fixed point admits a 1-cutting and an ancestor word independent from the rank of apparition of the factor. The author claims that any substitution on a two-letter alphabet which is not shift-periodic is rank one determined. His proof is not convincing.

Later, B. Host [215] and M. Queffélec [340] introduce the notion of *recognizable substitution*, that is, substitutions for which the 1-cutting of any long enough factor is independent of the rank of occurrence of the factor, except maybe for a suffix of the word, the length of this suffix being bounded. More precisely, a substitution σ is said to be *unilaterally recognizable* if there exists $L > 0$ such that $u_i \ldots u_{i+L-1} = u_j \ldots u_{j+L-1}$, with $i \in E_1$ implies $j \in E_1$.

B. Host proves in [215] that this property is equivalent to the fact that the image $\sigma(X_\sigma)$ under σ of the substitutive dynamical system X_σ associated with σ is an open set. G. Rauzy announced that he had a short proof of the fact that a constant length substitution which is one-to-one on the set of letters and which is not shift-periodic is unilaterally recognizable [340]. Nobody could check this proof. However, a new short proof in this case was recently obtained in [35]. Concerning substitutions of nonconstant length, the question remained unsettled.

In [306], B. Mossé studies the question of unilateral recognizability. She proves that a substitution is not necessarily unilaterally recognizable: a sufficient condition for a substitution not to be unilaterally recognizable is that for every couple of distinct letters (a, b), $\sigma(a)$ is a strict suffix of $\sigma(b)$, or conversely.

Example 7.2.1. The substitution $1 \mapsto 1112$ and $2 \mapsto 12$ is not unilaterally recognizable.

To get a general result of recognizability, B. Mossé introduces a new notion of recognizable substitution, for which the 1-cutting of any long enough factor is independent from the rank of apparition of the factor, except maybe for a suffix and a prefix of the factor, the lengths of those prefixes and suffixes being bounded. More precisely, a substitution σ is said to be *bilaterally recognizable* if there exists $L > 0$ such that $u_{i-L} \dots u_{i+L} = u_{j-L} \dots u_{j+L}$, with $i \in E_1$ implies $j \in E_1$.

The following theorem shows that this notion of recognizability is the right one:

Theorem 7.2.2 (Mossé [306, 307]). *Let σ be a primitive substitution with a non-periodic fixed point u. Then the substitution σ is bilaterally recognizable, and the ancestor word of every factor of u is unique except on the end of the factor: there exists $L > 0$ such that if $u_{i-L} \dots u_{j+L} = u_{i'-L} \dots u_{j'+L}$, then $u_i \dots u_j$ and $u_{i'} \dots u_{j'}$ have the same 1-cutting and the same ancestor word at ranks i and i'.*

Thus, if a substitution is primitive and not shift-periodic, then one can always *desubstitute* any factor of the fixed point, except possibly on the ends of the factor.

Remark. There exist nontrivial substitutions which are shift-periodic, for instance $1 \mapsto 121$, $2 \mapsto 21212$. It is not so easy to recognize if a substitution is not shift-periodic. However, this problem is solved thanks to an algorithm (see [199, 321]).

If we just suppose from now on that σ is a primitive substitution which is not shift-periodic, σ has not necessarily a fixed point but has at least a periodic point. Consequently, a power of σ is bilaterally recognizable. Since the dynamical systems associated with a substitution or with any of its iterations are identical, we deduce from Theorem 7.2.2 that any word of the bilateral symbolic dynamical system X_σ associated with the substitution σ can be desubstituted in a unique biinfinite word:

Corollary 7.2.3. *Let σ be a primitive substitution which is not shift-periodic. Let X_σ be the substitutive dynamical system associated with σ. Then we have*

$$\forall w \in X_\sigma,\ \exists v \in X_\sigma,\ v \ \textit{unique};\ w = S^k\sigma(v),\ \ \textit{and}\ \ 0 \le k < |\sigma(v_0)| \tag{7.1}$$

This corollary means that any word w in X_σ can by cut or desubstituted in a unique way in the following form:

$$w = \ldots \mid \underbrace{\ldots}_{\sigma(v_{-1})} \mid \underbrace{w_{-k} \ldots w_{-1}.w_0 \ldots w_l}_{\sigma(v_0)} \mid \underbrace{\ldots}_{\sigma(v_1)} \mid \underbrace{\ldots}_{\sigma(v_2)} \mid \ldots \tag{7.2}$$

where the word $v = \ldots v_{-n} \ldots v_{-1} v_0 v_1 \ldots v_n \ldots$ belongs to X_σ.

Note that this property is satisfied only by two-sided substitutive dynamical systems, and is not satisfied by a one-sided substitutive dynamical system. A counter-example is again given by the substitution defined by $1 \mapsto 1112$ and $2 \mapsto 12$.

This notion of recognizability is an important hypothesis used by most of the authors to get results about the spectrum of substitutive systems of nonconstant length. Before B. Mossé's work, most authors had to suppose that the substitutions they considered were recognizable.

7.2.2 Markov compactum and adic transformation

The object of this section is to focus on a deep relationship between shifts of finite type and substitutive dynamical systems, which allows one to transpose some of the properties of shifts of finite type to substitutive dynamical systems. The results of this section will be explicitly illustrated in Sec. 7.2.3, by using the notion of recognizability exposed in Sec. 7.2.1.

Markov compactum. At the beginning of the eighties, A. M. Vershik defined a new type of dynamical systems, called *adic systems*, their support being called Markov compacta [438]. In short, a *Markov compactum* is the set of paths in an infinite labeled graph. One can define arbitrary a partial order between the labels of the edges of the graph, and deduce a partial order between the elements of the Markov compactum. The choice of the successor with respect to this partial order is a continuous map called *adic transformation*.

A motivation for the introduction of this new family of systems is the question of the approximation of ergodic systems. Indeed, every automorphism of a Lebesgue space with an ergodic invariant measure is measure-theoretically isomorphic to an adic transformation. One can refer to [438] for more details about these systems.

Stationary Markov compactum. As a particular case of Markov compactum one can define the notion of *stationary Markov compactum.* A Markov compactum is said to be stationary if its graph is a tree and each level of the tree has the same structure. This implies that a stationary Markov compactum can be defined as a shift of finite type. More precisely, let $\mathbf{M}$ be a $d \times d$ matrix, with entries 0 or 1. The stationary Markov compactum associated with $\mathbf{M}$ is the set of the infinite paths in the infinite graph with levels indexed by $0, 1, 2 \ldots$, and having d vertices on each level. The edges connect vertices of the i-th level with vertices of the $i+1$-th level according to the adjacency matrix $\mathbf{M}$. Thus, the Markov compactum is the following set:

$$X_{\mathbf{M}} = \{(x_i)_{i\geq 0} \in \{1 \ldots d\}^{\mathbb{N}};\ \ \forall i \geq 0,\ X_{x_i, x_{i+1}} = 1\}.$$

There exists a canonical map which acts on this compactum: the one-sided shift map defined by $T((x_i)_{i\geq 0}) = (x_i)_{i\geq 1}$. This map is a dynamical system with positive entropy, the properties of which are quite well known.

Partial ordering on the Markov compactum. The edges issued from a vertex i are totally ordered by the number of the exit vertex of the edge. From this partial order on vertices can be deduced the following partial order on the paths $X_{\mathbf{M}}$:

$$(x_i)_{i\geq 0} \prec (y_i)_{i\geq 0} \text{ if there exists } i_0 \text{ such that } \begin{cases} x_i = y_i, & \forall\, i > i_0, \\ x_n < y_n. \end{cases}$$

In other words, a path $x = (x_i)_{i\geq 0}$ is said to precede a path $y = (y_i)_{i\geq 0}$ if they differ in finitely many terms and in the last such term the edge of x precedes the edges of y.

The Markov compactum is said to be *proper* if $X_{\mathbf{M}}$ contains exactly one maximal point and one minimal point with respect to the partial ordering.

Adic transformation. One may check that any non-maximal path $x \in X_{\mathbf{M}}$ admits an immediate successor $\tau(x)$, that is, the minimum of y in $X_{\mathbf{M}}$ preceded by x. This defines the *adic transformation* τ on $X_{\mathbf{M}}$, except on the set of the maximal elements which is finite when $\mathbf{M}$ is primitive.

Theorem 7.2.4 ([438, 439]). *The adic transformation associated with a primitive matrix is minimal and uniquely ergodic. The unique invariant measure is equivalent to the probability measure on $X_{\mathbf{M}}$ which is of maximal entropy for the shift map.*

Note that the probability measure which is of maximal entropy for the shift map is well-known: by primitivity, the matrix $\mathbf{M}$ admits a dominant eigenvalue α, and right and left eigenvectors $\mathbf{u} = (u_i)_i$, and $\mathbf{v} = (v_i)_i$ respectively, with positive coefficients.

One can normalize these vectors so that $< \mathbf{v}, \mathbf{u} >= 1$. One sets $p_i = v_i\, u_i$ and $p_{i,j} = a_{i,j}u_j/(\alpha u_i)$. The matrix $\mathbf{P}$ defined by $\mathbf{P} = (p_{i,j})$ is thus *stochastic*, and (p_i) is a right eigenvector associated with the eigenvalue 1.

The probability measure μ defined as follows over the cylinders of the shift of finite type is easily seen to be shift invariant, and is of *maximal entropy* (see for instance [445]):

$$\mu([a_0, a_1, \ldots, a_n]) = p_{a_0} p_{a_0,a_1} p_{a_1,a_2} \cdots p_{a_{n-1},a_n}.$$

Remark. Though the unique invariant measure for the adic transformation on $X_{\mathbf{M}}$ is equivalent to the probability measure on $X_{\mathbf{M}}$ which is of maximal entropy for the shift map, these maps do not have the same entropy: the maximal entropy of the shift map is strictly positive, though the entropy of the adic transformation is equal to zero.

Stationary Markov compactum and adic transformations. A. N. Livshits proves in [267] that the stationary adic transformation associated with a primitive $d \times d$ matrix $\mathbf{M}$ is measure-theoretically isomorphic to the dynamical system associated with the substitution defined on $\{1, \ldots, d\}$ by $i \mapsto 1^{\mathbf{M}_{1,i}} 2^{\mathbf{M}_{2,i}} \ldots d^{\mathbf{M}_{d,i}}$, as soon as the substitution σ is recognizable. More generally, a proof of the fact that every primitive substitutive dynamical system is isomorphic to an adic transformation is sketched in [439], the definition of adic transformation being adapted to a slightly more general context than stationary Markov compactum.

These results were completed by A. Forrest [175], who proved the following result in terms of Bratteli diagrams, which is a slightly different version of Markov compactum: he proves that the measure-theoretic isomorphism of Theorem 7.2.4 turns to be a topological conjugacy.

Theorem 7.2.5 (Forrest [175, 156]). *Each adic transformation on a stationary proper Markov compactum is topologically conjugate either to a primitive substitutive dynamical system or to an odometer. Conversely, each primitive substitutive dynamical system or odometer is topologically conjugate to an adic transformation on a stationary proper Markov compactum.*

The proof of A. Forrest is not constructive. It was completed and restated in the above terms by F. Durand, B. Host and C. F. Skau in [156]. In these papers, the authors give a new proof using substitutions, return words, and more precisely results about recognizability stated in Sec. 7.2.1 (for a definition of return words, see Chap. 3 and Sec. 7.3.2). Their approach provides an algorithm which computes for any substitutive dynamical system an adic transformation on a Markov compactum which is conjugate to a substitutive system. One can also refer to [217] for more details.

7.2.3 Desubstitution and prefix-suffix expansion

The computation of an adic transformation conjugated to a given substitutive dynamical system in [156] is not very natural and not very easy. Using the properties of recognizability, G. Rauzy initiated in [350] and [353] the idea of representing a substitutive dynamical system directly thanks to (7.1) and its interpretation given in (7.2). Let us illustrate this approach in more detail.

Desubstitution. It will be recalled that a consequence of the works of B. Mossé is that every word in X_σ can be expanded in the following form:

$$w = \dots \mid \underbrace{\dots}_{\sigma(y_{-1})} \mid \underbrace{w_{-k}\dots w_{-1}.w_0\dots w_l}_{\sigma(y_0)} \mid \underbrace{\dots}_{\sigma(y_1)} \mid \underbrace{\dots}_{\sigma(y_2)} \mid \dots$$

with $\dots y_{-n}\dots y_{-1}y_0y_1\dots y_n\dots$ in X_σ.

Let $p = w_{-k}\dots w_{-1}$ denote the prefix of $\sigma(y_0)$ of length k and $s = w_1\dots w_l$ the suffix of length l. The word w is completely determined by the word y and the decomposition of $\sigma(y_0)$ as $\sigma(y_0) = pw_0s$. Let $\mathcal{P}$ be the finite set of all such decompositions:

$$\mathcal{P} = \{(p,a,s) \in \mathcal{A}^\star \times \mathcal{A} \times \mathcal{A}^\star;\ \exists b \in \mathcal{A} \text{ and } \sigma(b) = pas\}\,.$$

Example 7.2.6. For the Fibonacci substitution $1 \mapsto 12$ and $2 \mapsto 1$, one gets:

$$\mathcal{P} = \{(\varepsilon,1,2),(1,2,\varepsilon),(\varepsilon,1,\varepsilon)\}.$$

For the substitution $1 \mapsto 1112$ and $2 \mapsto 12$, one gets:

$$\mathcal{P} = \{(\varepsilon,1,112),(1,1,12),(11,1,2),(111,2,\varepsilon),(\varepsilon,1,2),(1,2,\varepsilon)\}.$$

Thus, Theorem 7.2.2 implies that we can define, on the one hand, a continuous *desubstitution map* on X_σ (which sends w to y), on the other hand, a partition corresponding to the decomposition of $\sigma(y_0)$.

Theorem 7.2.7. *For any primitive substitution σ the following* desubstitution map θ *and the* prefix-suffix coding map γ *are well defined and continuous on X_σ:*

$$\theta : \begin{cases} X_\sigma \to X_\sigma \\ w \ \mapsto y \end{cases} \qquad \textit{such that } w = S^k\sigma(y) \textit{ and } 0 \le k < |\sigma(y_0)|,$$
$$\gamma : \begin{cases} X_\sigma \to \mathcal{P} \\ w \ \mapsto (p,w_0,s) \end{cases} \qquad \textit{such that } \sigma(y_0) = pw_0s \textit{ and } |p| = k\,.$$

Prefix-suffix expansion. In [102] and independently in [211], the itineraries of the points of X_σ under the desubstitution according to the partition defined by γ are studied. More precisely, the *prefix-suffix expansion* is the map Γ defined on X_σ by:

$$\forall w \in X_\sigma, \quad \Gamma(w) = (\gamma(\theta^i w))_{i\ge 0} = (p_i,a_i,s_i)_{i\ge 0} \in \mathcal{P}^{\mathbb{N}}\,.$$

Exercise 7.2.8. Let $w \in X_\sigma$ and $\Gamma(w) = (p_i,a_i,s_i)_{i\ge 0}$ be its prefix-suffix development. Suppose that nor all the prefixes p_i neither all the suffixes s_i are empty from a certain rank. Prove that w and $\Gamma(w)$ satisfy the following relationship:

$$w = \dots\sigma^n(p_n)\dots\sigma(p_1)p_0\dot{a}_0s_0\sigma(s_1)\dots\sigma^n(s_n)\dots$$

Important questions are

- the identification of the image of Γ,
- the injectivity of Γ,
- the identification of an action on $\Gamma(X_\sigma)$ such that Γ is a measure-theoretic isomorphism between that action and the shift map S on X_σ.

The image and the injectivity of Γ: prefix-suffix shift of finite type. The definition of Γ leads us to define the shift of finite type described by the automaton, called *prefix-suffix automaton* of the substitution σ, which has $\mathcal{A}$ as a set of vertices and $\mathcal{P}$ as a set of label edges: there is an edge labeled by $e = (p, a, s)$ from a to b if and only if $pas = \sigma(b)$.

The *prefix-suffix shift of finite type* is the set $\mathcal{D}$ of labels of infinite walks in this automaton. By definition, it is the support of a shift of finite type.

Example 7.2.9. The prefix-suffix automaton of the Fibonacci substitution $1 \mapsto 12$, $2 \mapsto 1$ is shown in Fig. 7.4.

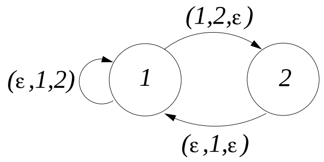


Fig. 7.4. Prefix-suffix automaton for the Fibonacci substitution.

The prefix-suffix of finite type for the Fibonacci substitution is the subset of $\{(\varepsilon, 1, 2), (1, 2, \varepsilon), (\varepsilon, 1, \varepsilon)\}^{\mathbb{N}}$ which consists of the labels of paths of the automaton.

It is proved in [102] and [211] that Γ is onto $\mathcal{D}$ and almost everywhere one-to-one on X_σ.

Prefix-suffix partial order. We explained before that a stationary Markov compactum is nothing else than the support of a shift of finite type, provided with a partial order deduced from a partial order on the labels of the edges of the subshift.

The preceding construction associates with a substitution the support of a shift of finite type, that is, the prefix-suffix shift of finite type $\mathcal{D}$. Thus, to define a Markov compactum on this set, we only need a partial order on the set $\mathcal{P}$ of labels of $\mathcal{D}$. However, the substitution σ provides a very natural partial order: two labels (p, a, s) and (q, b, r) are comparable for this order if they issue from the same letter, i.e., $pas = qbr = \sigma(c)$ for a letter c. The order of comparison comes from the length of the prefix p. More precisely, let us define

$$(p, a, s) \prec (q, b, r) \text{ if there exists } c \in \mathcal{A} \text{ such that } \begin{cases} pas = qbr = \sigma(c), \\ |p| < |q|. \end{cases} \tag{7.3}$$

As explained is Sec. 7.2.2, this order on the set of labels $\mathcal{P}$ induces an order on the support $\mathcal{D}$ of a shift of finite type, obtained as a lexicographical order from the right: a path $e = (e_i)_{i\geq 0} \in \mathcal{D}$ is said to precede a path $f = (f_i)_{i\geq 0} \in \mathcal{D}$ if they differ in finitely many terms and in the last such term i_0, e_{i_0} precedes f_{i_0} for the partial order defined at (7.3):

$$(e_i)_{i\geq 0} \prec (f_i)_{i\geq 0} \text{ if there exists a rank } i_0 \text{ such that } \begin{cases} \forall\, i > i_0,\, e_i = f_i \\ e_{i_0} \prec f_{i_0} \end{cases}. \tag{7.4}$$

Remark. Denote $e_{i_0} = (p_{i_0}, a_{i_0}, s_{i_0})$ and $f_{i_0} = (q_{i_0}, b_{i_0}, r_{i_0})$. As $e_{i_0+1} = f_{i_0+1}$, we have $p_{i_0}a_{i_0}s_{i_0} = q_{i_0}b_{i_0}r_{i_0}$, and e_{i_0} is comparable with f_{i_0} for the prefix-suffix partial order. The condition $e_{i_0} \prec f_{i_0}$ means that the length of the prefix p_{i_0} is less than the length of the prefix q_{i_0}.

Prefix-suffix adic transformation. Following the construction described in Sec. 7.2.2, the prefix-suffix partial order on the support $\mathcal{D}$ of a shift of finite type provides an adic transformation, called *prefix-suffix adic transformation* and denoted by τ. This map is an immediate successor transformation on the set of non-maximal points of $\mathcal{D}$.

Exercise 7.2.10. 1. Prove that a point $e = (p_i, a_i, s_i)_{i\geq 0} \in \mathcal{D}$ is maximal for the prefix-suffix partial order if and only if all the suffixes s_i are empty.

2. Let $e = (p_i, a_i, s_i)_{i\geq 0} \in \mathcal{D}$ be a non-maximal point in $\mathcal{D}$. Let i_0 be the smallest integer such that $s_{i_0} \neq \varepsilon$. Prove that $\tau(e)$ is the unique path $f = (q_i, b_i, t_i)_{i\geq 0} \in \mathcal{D}$ defined by:

$$\begin{array}{ll} \forall i > i_0,\, (q_i, b_i, t_i) = (p_i, a_i, s_i), & \\ i = i_0 & q_ib_it_i = p_ia_is_i \text{ and } |q_i| = |p_i| + 1, \\ i < i_0 & q_i = \varepsilon \text{ and } b_it_i = \sigma(b_{i+1}). \end{array}$$

Exercise 7.2.11. Consider the substitution $1 \mapsto 21$, $2 \mapsto 13$, $3 \mapsto 1$. Let u be the periodic point defined by

$$u = \lim_{n\to\infty} \sigma^{2n}(1.1) = \ldots 1321212111321.1321212111\ldots.$$

1. Prove that the prefix-suffix expansion of u is:

$$\Gamma(u) = m = (\varepsilon, 2, 1)(\varepsilon, 1, 3)(\varepsilon, 2, 1)(\varepsilon, 1, 3)\ldots.$$

2. Prove that

$$\begin{aligned} \tau(m) &= (2, 1, \varepsilon)(\varepsilon, 1, 3)(\varepsilon, 2, 1)(\varepsilon, 1, 3)\ldots \\ \tau^2(m) &= (\varepsilon, 1, \varepsilon)(1, 3, \varepsilon) \quad - \quad\quad - \quad \ldots \\ \tau^3(m) &= (\varepsilon, 1, 3)(\varepsilon, 2, 1)(2, 1, \varepsilon) \quad - \quad \ldots. \end{aligned}$$

3. Prove that

$$\tau^7(m) = (\varepsilon, 1, 3)(\varepsilon, 2, 1)(\varepsilon, 1, \varepsilon)(1, 3, \varepsilon)[(\varepsilon, 2, 1)(\varepsilon, 1, 3)]^\infty.$$

The prefix-suffix adic transformation on the set $\mathcal{D}$ provides an example of Markov compactum as defined in Sec. 7.2.2. The following theorem means that this explicit Markov compactum realizes explicitly the main result in [439] which states that any primitive substitutive dynamical system is measure-theoretically isomorphic to an adic transformation on a Markov compactum.

The conjugacy map is not a topological conjugacy map but a semi-conjugacy, contrary to the one obtained in [156]. But this conjugacy map can be deduced in a simple, natural and explicit way from the substitution.

Theorem 7.2.12 ([102, 211]). *Let σ be a primitive substitution which is not shift-periodic and (X_σ, S) the dynamical system generated by σ. The prefix-suffix expansion is a continuous mapping onto the shift of finite type $\mathcal{D}$. This map is one-to-one except on the orbit of periodic points of σ.*

This prefix-suffix expansion is a semi-topological conjugacy between the shift map S on X_σ and an adic transformation on $\mathcal{D}$, considered as a Markov compactum when the set $\mathcal{P}$ of edges is provided with the natural prefix-suffix partial ordering coming from the substitution.

It can be noted that the prefix-suffix automaton which has allowed us to define the Markov compactum $\mathcal{D}$ is the most complete form of a large class of automata appearing in literature.

It was introduced in [102] in order to extend and to make more precise the prefix automaton (sometimes also called "automate à la provençale"), and its dual suffix automaton, defined by G. Rauzy in the seminal papers [350, 353] with the final goal to represent substitutive dynamical systems of Pisot type by domain exchanges in the Euclidean space. The difference between the prefix automaton and the prefix-suffix automaton is that the subshift generated by the second one is of finite type, while the one generated by the first automaton is only sofic. This has important consequences for the injectivity of geometric representations of substitutive dynamical systems defined thanks to the subshifts (see [350, 353, 103] and Sec. 7.4).

Moreover, the prefix-suffix automaton can be found in a coded form among the number theory works which followed [350]: first in a work by J.-M. Dumont and A. Thomas [151] about numeration scales associated with certain substitutions; then among their development by V. Sirvent [404] who proves some properties of the Rauzy fractal (see also Sec. 7.4). We can also mention the definition by T. Kamae [231] of colored tilings associated with a weighted substitution, the layout of the tiles following rules being close to the transition rules of the prefix automaton, but in a non-explicit way.

To finish this brief survey, the prefix-suffix automaton also appears in the context of language theory: the prefix-suffix automaton projects, on the one hand, onto the automaton defined by A. Maes [276] in his works about decidability of arithmetical theory, on the other hand, onto the automaton defined by P. Narbel [311] to study the boundary of a language using the description with trees of the set of paths in his automaton.

7.3 Spectral theory of substitutive dynamical systems

Substitutive dynamical systems were introduced by W. H. Gottschalk [191], as examples of symbolic dynamical systems, the study of which was initialized in [303, 304]. The attention first focused on systems associated with a substitution of *constant length.* Let us recall that a substitution is said to be of *constant length* if the image of any letter of the alphabet under the substitution contains the same number of letters. We give in Sec. 7.3.1 some results about these substitutions.

At the end of the seventies, the spectral theory of constant length substitutive dynamical systems was quite well known, and people started to focus on spectral theory of nonconstant length substitutive dynamical systems. We will give in the next section the main results about this theory. Thanks to recognizability and Markov compactum, they got the results on the spectral theory of substitutive systems presented in Sec. 7.3.2. A special attention was devoted to the class of substitutions of Pisot type. This will be the object of Sec. 7.3.3.

7.3.1 Constant length substitutions

First topological and ergodic properties. They were obtained at the beginning of the seventies: T. Kamae proves in [229] that the condition of primitivity is sufficient for the minimality of the system. Then he gives a sufficient condition for the system to have a purely discrete spectrum [230]. B. G. Klein shows that primitive substitutive systems of constant length are uniquely ergodic and that their topological entropy equals zero [249].

Partial results about the maximal equicontinuous factor. In [125], E. M. Coven and M. S. Keane focus on a class of constant length substitutions over a two-letter alphabet which generate a minimal system, related to Toeplitz substitutions. They prove that the maximal equicontinuous factor of these systems is the n-adic group $\mathbb{Z}_n$, where n is the length of the substitution. If n is not prime, $\mathbb{Z}_n$ is defined to be the product of the p-adic groups $\mathbb{Z}_p$, for all primes p that divide n .

J. C. Martin generalizes this to substitutions of constant length n over a finite alphabet [281]: if the substitution is one-to-one on the set of letters, the maximal equicontinuous factor of the substitutive dynamical system is the addition of $(1, 1)$ on $\mathbb{Z}_n \times \mathbb{Z}/m\mathbb{Z}$, where m in an integer depending on the substitution. J. C. Martin then gives some conditions for the factor to be isomorphic to the system: some are necessary conditions, the others are sufficient.

Determining the maximal equicontinuous factor. Finally, F. M. Dekking generalizes these results for the constant length case, by introducing new methods:

Theorem 7.3.1 (Dekking [135]). *Let σ be a non-periodic substitution of constant length n. Let u be a periodic point for σ. We call* height *of the substitution the greatest integer m which is coprime with n and divides all the strictly positive ranks of occurrence of the letter u_0 in u. The height is less that the cardinality of the alphabet.*

The maximal equicontinuous factor of the substitutive dynamical system associated with σ is the addition of $(1,1)$ on the abelian group $\mathbb{Z}_n \times \mathbb{Z}/m\mathbb{Z}$.

F. M. Dekking gives an algorithmic method for computing the height of a substitution.

Exercise 7.3.2. Show that the Morse substitution has height 1. Deduce the maximal equicontinuous factor of the system associated with this substitution.

Exercise 7.3.3.
1. Give the height of the substitution $1 \mapsto 121$, $2 \mapsto 312$, $3 \mapsto 213$.
2. Prove that the spectrum (*defined in Chap. 1*) of the system associated with this substitution is

$$\left\{ e^{2in\pi/3^m + 2ik\pi/2};\ n, k \in \mathbb{Z}, m \in \mathbb{N}, \right\}.$$

The question of isomorphism for substitutions of height 1. F. M. Dekking also introduces the combinatorial condition which later will be called *coincidence condition*: a constant length substitution σ satisfies this condition if there exist two integers k, n such that the images of any letter of the alphabet under σ^k has the same n-th letter. Generalizing [230], F. M. Dekking proves that a substitution of height 1 satisfies this condition if and only if the associated substitutive dynamical system is measure-theoretically isomorphic to its maximal equicontinuous factor, that is, if the system has a purely discrete spectrum:

Theorem 7.3.4 (Dekking [135]). *Let σ be a substitution of constant length and of height 1. The substitutive dynamical system associated with σ has a purely discrete spectrum if and only if the substitution σ satisfies the condition of coincidence.*

Exercise 7.3.5. Prove that the substitutive dynamical system associated with the Morse substitution is not isomorphic to a translation on $\mathbb{Z}_2$.

Note that this result was already shown in Chap. 5.

Pure base of a substitution. If the height h of a primitive substitution σ of constant length n is different from 1, the *pure base* of σ is the substitution η defined as follows:

- The alphabet $\mathcal{A}_1$ of η is the collection of all blocks of length h appearing at position kh in a periodic point u for σ.

- For every block $W \in \mathcal{A}_1$ of length h, $\eta(W) = W_1 W_2 \dots W_n$, with $W_i \in \mathcal{A}_1$ and $W_1 W_2 \dots W_n = \sigma(W)$.

One should notice that η is primitive, of constant length n, and that its height is 1. Up to this tool, the result of Dekking can be generalized to all substitutions of constant length.

Theorem 7.3.6 (Dekking [135]). *Let σ be a substitution of constant length. The substitutive dynamical system associated with σ has a purely discrete spectrum if and only if its pure base satisfies the condition of coincidence.*

Exercise 7.3.7. Prove that the pure base of $1 \mapsto 121$, $2 \mapsto 312$, $3 \mapsto 213$ is given by the substitution $a \mapsto aab$, $b \mapsto aba$. Has this substitution a purely discrete spectrum?

7.3.2 Nonconstant length substitutions

The first partial results about the spectrum of substitutions of nonconstant length were obtained by studying a class of nonconstant length substitutive systems which are measure-theoretically isomorphic to constant length substitutive systems [125, 135]. Later, F. M. Dekking et M. S. Keane proved that strongly mixing substitutive dynamical systems do not exist, but weak-mixing is possible [139]. P. Michel proved in [292] that the dynamical system associated with a primitive substitution of nonconstant length is uniquely ergodic, see also 1. M. Queffélec studies in detail the spectral type of substitutive systems in [340, 342]. Most of the results cited above appear in this last reference.

These general topological and ergodic properties having been studied, authors focused on the explicit description of the spectrum of a substitutive dynamical system. They also took interest in obtaining conditions for a substitutive system to have a purely discrete spectrum, partially continuous spectrum, or continuous spectrum. The description of the known results is the aim of this section.

Rational eigenvalues and the characteristic polynomial of the incidence matrix. J. C. Martin obtains a partial result about the spectrum of nonconstant length substitutions on a two-letter alphabet in [282]. His work suffers from a too weak definition of the property of recognizability (see Sec. 7.2.1). If we take into account the results about recognizability proved later, his results become:

Theorem 7.3.8 (Martin [282]). *Let σ be a not shift-periodic and primitive substitution on a two-letter alphabet. Let $X^2 - tX + d$ be the characteristic polynomial of the incidence matrix of σ and let $r \in \mathbb{Q}$. Let i_0 be the greatest positive integer such that:*

- *the prime factors of* i_0 *divide both* $|\sigma(1)|$ *and* $|\sigma(2)|$,
- i_0 *is prime with* r,
- i_0 *divides* $|\sigma(1)| - |\sigma(2)|$.

Then $\exp(2\pi i r)$ *is an eigenvalue of the dynamical system associated with* σ *if and only if there exist integers* k, m, n *such that* $r = k/i_0 + m/(pgcd(d,t))^n$.

A consequence of this theorem is the characterization of p-adic measure-theoretic factors of substitutive systems over a two-letter alphabet:

Exercise 7.3.9. 1. Let σ be a not shift-periodic and primitive substitution over a two-letter alphabet. Let p be a prime number. Show that the addition of 1 on $\mathbb{Z}_p$ is a measure-theoretic factor of the substitutive dynamical system associated with σ if and only if p divides the determinant and the trace of the characteristic polynomial of the incidence matrix of σ.
2. Find the p-adic factors of the systems defined by the substitutions $1 \mapsto 1112$, $2 \mapsto 12$, and $1 \mapsto 11222$, $2 \mapsto 1222$.

In particular, there exist non-unimodular substitutive systems which do not admit any p-adic factor. Such a behavior is totally different from the case of constant length substitutions. This result has been generalized by F. Durand:

Theorem 7.3.10 (Durand [155]). *Let* σ *be a primitive substitution, the incidence matrix* $\mathbf{M}_\sigma$ *of which has an irreducible characteristic polynomial* $\chi_{\mathbf{M}_\sigma}$. *Let* p *be a prime number. The addition of 1 over* $\mathbb{Z}_p$ *is a measure-theoretic factor of the substitutive dynamical system associated with* σ *if and only if* p *divides all the coefficients of* $\chi_{\mathbf{M}_\sigma}$ *except the dominant one, that is, if and only if* $\mathbf{M}_\sigma$ *is nilpotent modulo* p.

A new proof of this result is given in [396] by studying the ramifications of the determinant of the incidence matrix of σ in the Galois extension of its dominant eigenvalue.

Coboundaries: a full description of the spectrum. In [215], B. Host introduces a good notion of recognizability and thanks to this notion proves the following important result.

Theorem 7.3.11 (Host [215]). *Let* σ *be a not shift-periodic and primitive substitution. Each eigenfunction of the substitutive dynamical system associated with* σ *is continuous.*

To be more precise, since we deal with functions in $\mathcal{L}^2$, Theorem 7.3.11 means that any class of eigenfunctions contains a continuous eigenfunction.

Exercise 7.3.12. Let σ be a not shift-periodic and primitive substitution. Show that a minimal toral rotation or the addition of 1 in the p-adic integer set $\mathbb{Z}_p$ is a topological factor of the substitutive dynamical system associated with σ if and only if it is a measure-theoretic factor of this system.

B. Host deduces from Theorem 7.3.11 a complete description of the spectrum of the nonconstant length substitutions, which are recognizable, and one-to-one on the letters. This description is a generalization of [135] and [282]. Here, $\mathbb{U} \subset \mathbb{C}$ denotes the unit circle.

Definition 7.3.13. *A* coboundary *of a substitution σ is defined as a map $h : \mathcal{A} \to \mathbb{U}$ such that there exists a map $f : \mathcal{A} \to \mathbb{U}$ with $f(b) = f(a)h(a)$, for every word ab of length 2 which belongs to the language of the substitution.*

In the most simplest cases the only coboundary is the trivial one, that is, the constant function equal to 1. However, there exist some substitutions with nontrivial coboundaries:

Exercise 7.3.14. 1. Prove that the substitution $1 \mapsto 12$, $2 \mapsto 13$, $3 \mapsto 1$ has no nontrivial coboundary. *(Hint: note that the language of this substitution contains* 12, 21 *and* 11.*)*
2. Prove that any primitive and not shift-periodic substitution over two letters has no nontrivial coboundary.
3. Find a nontrivial coboundary for the substitution $1 \mapsto 1231$, $2 \mapsto 232$, $3 \mapsto 3123$.

In Sec. 7.5.3 we give a method to construct substitutions with nontrivial coboundaries.

Theorem 7.3.15 (Host [215]). *Let σ be a not shift-periodic and primitive substitution over the alphabet $\mathcal{A}$. A complex number $\lambda \subset \mathbb{U}$ is an eigenvalue of (X_σ, S) if and only if there exists $p > 0$ such that for every $a \in \mathcal{A}$, the limit $h(a) = \lim_{n\to\infty} \lambda^{|\sigma^{pn}(a)|}$ is well defined, and h is a coboundary of σ.*

Exercise 7.3.16. Let σ be the substitution defined by $1 \mapsto 12121$ and $2 \mapsto 122$.

1. Prove that $|\sigma^{n+2}(a)| = 5|\sigma^{n+1}(a)| + 4|\sigma^n(a)|$, for $a = 1, 2$.
2. Deduce that $2|\sigma^n(2)| - |\sigma^n(1)| = 1$, for every nonnegative n.
3. Conclude that the system associated with σ is weakly mixing (that is, 1 is its only eigenvalue; see Chap. 1).

Since the constant function equal to one 1 is always a coboundary, a sufficient condition for λ to be an eigenvalue is the following:

Corollary 7.3.17. *Let σ be a not shift-periodic and primitive substitution. If there exists p such that $\lambda \in \mathbb{C}$ satisfies $\lim \lambda^{|\sigma^{pn}(a)|} = 1$ for every letter a of the alphabet, then λ is an eigenvalue of the substitutive dynamical system associated with σ. We say that λ is an* eigenvalue associated with the trivial coboundary.

Let **1** be the column vector whose all coordinates are equal to 1.

Exercise 7.3.18. Let $\lambda = \exp(2i\pi\beta)$ of modulus 1. Prove that $\lim \lambda^{|\sigma^n(a)|} = 1$ for every letter a, if and only if $\beta {}^t\mathbf{M}_\sigma{}^n \mathbf{1}$ tends to zero modulo $\mathbb{Z}^d$.

B. Host provides a method to identify the set of eigenvalues associated with the trivial coboundary.

Proposition 7.3.19 (Host [215]). *Let σ be a primitive d-letter substitution. Let X_σ be the substitutive system associated with σ and $\mathbf{M}_\sigma$ be its incidence matrix.*

The complex $\exp(2i\pi\beta)$ is an eigenvalue for X_σ associated with the trivial coboundary if and only if $\beta\mathbf{1} = \mathbf{x}_1 + \mathbf{x}_2$, where there exist two integers p and q such that ${}^t\mathbf{M}_\sigma{}^{pn}\mathbf{x}_1$ tends to zero in $\mathbb{R}^d$ and $\mathbf{x}_2 \in {}^t\mathbf{M}_\sigma^{-q}\mathbb{Z}^d$.

An important example of application can be found among substitutions of Pisot type, which will be treated in Sec. 7.3.3.

In [174], S. Ferenczi, C. Mauduit et A. Nogueira give a new simplified version of the proofs of Theorems 7.3.11 and 7.3.15. They deduce from these new proofs a more explicit characterization of the spectrum of primitive substitutive dynamical systems, in terms of polynomials (see Theorem 7.3.27).

Combinatorial conditions: a mix of coincidences and return words. The parallel between substitutive systems and adic systems presented in the preceding section, and the ideas developed by B. Host in [215], led many authors to give some necessary and sufficient conditions for a substitutive dynamical system to have a purely discrete spectrum, a partially continuous spectrum or to be weakly mixing (see the survey in [439]).

For instance, A. N. Livshits stated a relationship between the fact that a substitutive system has a purely discrete spectrum and the existence of a code of the system made of modified return words on one side, coincidences on letters on the other side. To precise this, let us introduce the following notions:

Definition 7.3.20. *Let σ be a primitive and not shift-periodic substitution defined over the alphabet $\mathcal{A}$.*

For any letter a, the a-blocks *are the elements of the set Δ_a:*

$$\Delta_a = \left\{ (a_1 \dots a_{r-1}, a_2 \dots a_r) \text{ with } \left\{ \begin{array}{l} a_1 \dots a_r \in \mathcal{L}(X_\sigma) \\ i = 1, r, \, \forall\, n > 0, \, |\sigma^n(a_i)| = |\sigma^n(a)| \\ i \neq 1, r, \, \exists\, n > 0, \, |\sigma^n(a_i)| \neq |\sigma^n(a)| \end{array} \right. \right\}$$

Note that if the substitution is of constant length, then $\Delta_a = \{(b,b),\, bb \in \mathcal{L}(X_\sigma)\}$. Moreover, the set Δ_a contains the *return words* of the substitution, as introduced by F. Durand in [153, 152].

Definition 7.3.21. *Let σ be a primitive and not shift-periodic substitution. A* return word *is a word $W = a_1 \dots a_{k-1} \in \mathcal{L}(X_\sigma)$ such that there exists a letter a_k with*

- $a_1 \dots a_{k-1} a_k \in \mathcal{L}(X_\sigma)$,
- *for every large enough* n, $\begin{cases} \sigma^n(a_k) = \sigma^n(a_1), \\ \forall\, j \neq 1, k,\ \sigma^n(a_j) \neq \sigma^n(a_1). \end{cases}$

The associated return time is defined to be $r_n(W) = |\sigma^n(a_1 \dots a_{k-1})|$.

A. N. Livshits gives the following characterization of substitutive systems with a purely discrete spectrum. This characterization was announced in [266] and [439], and proved in [268]. As defined in Chap. 1, $\mathbf{l} : \mathcal{A} \to \mathbb{N}^d$ denotes the abelianization map.

Theorem 7.3.22 (Livshits [266]). *Let σ be a primitive and not shift-periodic substitution defined over the alphabet $\mathcal{A}$.*

Suppose that there exists a letter a and a finite set Δ of pairs (W_{i_k}, W_{j_k}) with $\mathbf{l}(W_{i_k}) = \mathbf{l}(W_{j_k})$, which satisfy

$$\forall\,(V_1, V_2) \in \Delta \cup \Delta_a,\ \exists\, m > 0,$$

$$\text{such that } \begin{cases} \sigma^m(V_1) = W_{l_1} \dots W_{l_s} \\ \sigma^m(V_2) = W_{j_1} \dots W_{j_s} \end{cases} \text{ with } \begin{cases} \forall\, r,\ (W_{l_r}, W_{j_r}) \in \Delta \cup \Delta_a \\ \exists\, r,\ W_{l_r} = W_{j_r} \end{cases}.$$

Then the substitutive dynamical system associated with σ has a purely discrete spectrum.

As before, examples of application of this theorem can be found among substitutions of Pisot type, which will be treated in detail in the next section (see in particular Theorem 7.3.35).

The converse of Theorem 7.3.22 can be expressed under the following form.

Theorem 7.3.23 (Livshits [266]). *Let σ be a primitive and not shift-periodic substitution. Let $c \neq d$ and a be letters such that $ac, ad \in \mathcal{L}(X_\sigma)$.*

Suppose that there exist a finite set of words $\{W_i\}$ and an increasing sequence of integers m_i satisfying for all i:

$$\begin{cases} \sigma^{m_i}(c) = W_{l_1} \dots W_{l_s} A \\ \sigma^{m_i}(d) = W_{k_1} \dots W_{k_s} B \end{cases} \text{ with } \begin{cases} A = \varepsilon \text{ or } B = \varepsilon, \\ \forall\, r,\ \mathbf{l}(W_{l_t}) = \mathbf{l}(W_{k_t}), \\ \forall\, r,\ W_{l_t} \neq W_{k_t}. \end{cases}$$

Then the substitutive dynamical system associated with σ has a partially continuous spectrum.

For instance the system associated with the substitution $1 \mapsto 23$, $2 \mapsto 12$ and $3 \mapsto 23$ admits a continuous spectral component. However, the substitution is of constant length, so that the system admits $\mathbb{Z}_2$ as a factor (Theorem 7.3.1) and is not weakly mixing.

Another consequence of this theorem is that the Morse substitution system has a partially continuous spectrum, which has been studied: the maximal spectral type is known to be a Riesz product (see the bibliography of

[340]). Finally, the substitutive system associated with the Morse substitution is measure-theoretically isomorphic to a two-point extension of its maximal equicontinuous factor, that is, the addition of 1 on $\mathbb{Z}_2$ (see also Chap. 5).

Finally, A. N. Livshits gives a sufficient condition for a substitutive dynamical system to be weakly mixing, that is, to have no nontrivial eigenvalue.

Theorem 7.3.24 (Livshits [267]). *Let σ be a not shift-periodic and primitive substitution.*

If every eigenvalue of the incidence matrix of σ has a modulus greater than or equal to 1, then the substitutive dynamical system associated with σ has no irrational eigenvalue.

If there exists an element of the language of σ, denoted $W = a_1 \dots a_k = a_1 V$ with $a_1 = a_k$, such that for all integers n, the lengths $|\sigma^n(V)|$ are coprime to each other for $n > N$, then the substitutive dynamical system associated with σ has no rational eigenvalue (except 1).

Note that in the previous theorem, W is a return word as soon as a_1 does not appear is V before the last rank. As an application, the substitution $1 \mapsto 112$ and $2 \mapsto 12222$ introduced in [139] is weakly mixing [266].

Polynomial conditions. B. Solomyak makes the above theorem somewhat more precise by giving a characterization of irrational eigenvalues of substitutive systems in a more explicit way than Host's one.

Theorem 7.3.25 (B. Solomyak [411]). *Let σ be a primitive substitution such that its incidence matrix $\mathbf{M}_\sigma$ admits an irreducible characteristic polynomial. Let α_k, $k = 1 \dots m$, be the eigenvalues of $\mathbf{M}_\sigma$ such that $|\alpha_k| \geq 1$.*

The substitutive system associated with σ admits at least one irrational eigenvalue if and only if there exists a polynomial $P \in \mathbb{Z}[X]$ such that $P(\alpha_j) = P(\alpha_k)$ for every $j, k \leq m$.

If so, $\exp(2\pi i P(\alpha_1))$ and $\exp(2\pi i P(\alpha_1)^n)$ are eigenvalues of the substitutive system.

Remark.

- Note that the numbers $\exp(2i\pi n P(\alpha_1))$ are eigenvalues as soon as the number $\exp(2\pi i P(\alpha_1))$ is an eigenvalue.
- Theorem 7.3.25 does not state that any irrational eigenvalue is of the form $\exp(2\pi i P(\alpha_1))$.

For instance, the spectrum of the system associated with the substitution $1 \mapsto 1244$, $2 \mapsto 23$, $3 \mapsto 4$, $4 \mapsto 1$ contains the set $\exp(2\pi i \mathbb{Z}\sqrt{2})$.

Theorem 7.3.25 produces other sufficient conditions for a substitutive system to have nontrivial rational spectrum, in particular in the Pisot case (see Sec. 7.3.3).

Exercise 7.3.26 ([411]).

- Let s denotes the smallest prime divisor of the cardinality d of the alphabet of a primitive substitution σ. Suppose that the number of eigenvalues of $\mathbf{M}_\sigma$ greater than 1 is strictly greater that d/s. Show that the substitutive dynamical system associated with σ has no irrational eigenvalue.
- For any prime p, any substitutive dynamical system with irrational eigenvalues over a p-letter alphabet is of Pisot type.

In [174], S. Ferenczi, C. Mauduit and A. Nogueira got interested in the relationship between coboundaries and return words.

Proposition 7.3.27 (Ferenczi, Mauduit, Nogueira [174]). *Let σ be a not shift-periodic and primitive substitution and (X_σ, S) be the substitutive dynamical system associated with σ.*

A complex number λ with modulus 1 is an eigenvalue of the substitutive dynamical system (X_σ, S) if and only if $\lim_{n\to\infty} \lambda^{r_n(W)} = 1$ for every return word W.

This proposition is illustrated for the Chacon sequence in Chap. 5.

Using techniques from the study of substitutive normal sets [283], the authors give a constructive version of the proofs of B. Host [215]. They get a characterization of eigenvalues which generalizes (no more hypothesis of irreducibility) and gives a partial converse to Theorem 7.3.25. This characterization is too long to be cited here. A consequence of it is the following:

Theorem 7.3.28 (Ferenczi, Mauduit, Nogueira [174]). *Let σ be a primitive and not shift-periodic substitution and α_k, $k = 1 \dots r$, be the eigenvalues of $\mathbf{M}_\sigma$ such that $|\alpha_k| \geq 1$. Let β be irrational.*

- *If $\lambda = \exp(2\pi i\beta)$ is an eigenvalue of X_σ associated with the trivial coboundary – that is, $\lim \lambda^{|\sigma^n(a)|} = 1$ for every letter a – then there exists a polynomial $P \in \mathbb{Z}[X]$ such that $P(\alpha_i) = \beta$ for every i.*
- *If $P \in \mathbb{Z}[X]$ is a polynomial with $P(\alpha_i) = \beta$ for every i, then there exists an integer k such that $\exp(2\pi ikP(\alpha_i))$ is an eigenvalue of the substitutive system.*

Thus, an integer polynomial which is constant over the set of eigenvalues of $\mathbf{M}_\sigma$ of modulus greater than one, does not provide directly an eigenvalue for the system: the constant may have to be multiplied by an integer to become the argument of an eigenvalue.

As an application, the spectrum of the system associated with the substitution considered above $1 \mapsto 1244$, $2 \mapsto 23$, $3 \mapsto 4$, $4 \mapsto 1$, is exactly the set $\exp(2\pi i\mathbb{Z}\sqrt{2})$: Theorem 7.3.25 implied that the spectrum contains this set. The converse is stated by Theorem 7.3.28.

Conclusion. The spectrum of a primitive dynamical system can be divided into two parts. The first part is the set of complex numbers associated with the trivial coboundary; according to Theorem 7.3.28, these eigenvalues depend only on the incidence matrix of the substitution, and, more precisely, for the irrational eigenvalues, on the matrix eigenvalues which are greater than 1. In this sense, this class of eigenvalues is quite natural.

In a more exotic and surprising way, the substitutive system can have some eigenvalues which are associated with a nontrivial coboundary, and not associated with the trivial coboundary. Such eigenvalues are not described by Theorem 7.3.28. These eigenvalues depend heavily on the return times and more generally on the combinatorics of the substitution, which in other words gives them a non-commutative aspect.

There exist some unpublished partial results about a substitution with rational noncommutative eigenvalues. We do not know any example of substitution with irrational noncommutative eigenvalue.

7.3.3 Substitutions of Pisot type

Some of the general results of Sec. 7.3.2 have special consequences for the spectral theory of Pisot type substitutive systems.

Eigenvalues associated with the trivial coboundary. As claimed by B. Host in [216], Theorem 7.3.11 and Theorem 7.3.15 allow one to determine the spectrum associated with the trivial coboundary, in the unimodular case.

Proposition 7.3.29 ([216]). *Let σ be a unimodular substitution of Pisot type. The group of eigenvalues of X_σ associated with the trivial coboundary is generated by the frequencies of letters in any word of the system, that is, by the coordinates of a right normalized eigenvector associated with the dominant eigenvalue of the incidence matrix of the substitution.*

Hints to deduce this result from Theorem 7.3.15 and Lemma 7.3.19 are the following:

Exercise 7.3.30. Let σ be a unimodular substitution of Pisot type over d letters. Let $\mathbf{M}_\sigma$ be the incidence matrix of σ and $\mathbf{1}$ be the column vector whose all coordinates are equal to 1. Let $\mathbf{u}_1$ be an expanding eigenvector of $\mathbf{M}_\sigma$, normalized so that the sum of its coordinates is equal to one.

1. Let $\mathbf{x} \in \mathbb{R}^d$. Show that ${}^t\mathbf{M}_\sigma{}^n\mathbf{x}$ tends to zero in $\mathbb{R}^d$ if and only if $\mathbf{x}$ is orthogonal to $\mathbf{u}_1$.
2. Show that ${}^t\mathbf{M}_\sigma{}^{-1}\mathbb{Z}^d = \mathbb{Z}^d$.
3. Deduce that $\exp(2i\pi\beta)$ is an eigenvalue of X_σ associated with the trivial coboundary if and only β is an integer combination of the coordinates of $\mathbf{u}_1$.
4. Note that the coordinates of $\mathbf{u}_1$ are the frequencies of letters in any point in the system and deduce Proposition 7.3.29.

Some basic number theory allows one to precise Proposition 7.3.29. We recall that the characteristic polynomial of a Pisot type substitution is always irreducible over $\mathbb{Q}$ (see Chap. 1).

Exercise 7.3.31. Let σ be a substitution of Pisot type and α be the dominant eigenvalue of its incidence matrix.

1. Prove that the coordinates of a dominant right eigenvector of $\mathbf{M}_\sigma$, normalized so that the sum of its coordinates is equal to one, belong to $\mathbb{Z}[\alpha]$ *(Hint: these coordinates are described by a system of equations with rational coefficients.)*
2. Prove that these coordinates are rationally independent *(Hint: the eigenvalues of the incidence matrix are simple and algebraic conjugates; the eigenvectors are linearly independent.)*
3. If σ is unimodular, show that the group of eigenvalues associated with the trivial coboundary is $\exp(2\pi i\mathbb{Z}[\alpha])$.

The result of Exercise 7.3.31 can be obtained independently and generalized (out of the unimodular context) as a consequence of Theorem 7.3.25:

Proposition 7.3.32. *Let α denote the dominant eigenvalue of the incidence matrix of a substitution of Pisot type. The spectrum of the substitutive dynamical system associated with σ contains the set* $\exp(2\pi i\mathbb{Z}[\alpha])$.

In particular, substitutive dynamical systems of Pisot type are never weakly mixing.

Proposition 7.3.32 means that Pisot type substitutions are quite close to constant length substitutions with respect to the fact that both are never weakly mixing: constant length substitutive dynamical systems always have nontrivial rational eigenvalues whereas Pisot type substitutive dynamical systems always have irrational eigenvalues. But if a substitution is neither of Pisot type, nor of constant length, then completely different behaviors can occur, such as the absence of spectrum for which we gave sufficient conditions in Theorem 7.3.24. An example of such a substitution is $1 \mapsto 112$, $2 \mapsto 12222$.

Since $\mathbb{Z}[\alpha]$ is of rank $d-1$, we deduce from Proposition 7.3.32 and Exercise 7.3.12 that:

Corollary 7.3.33. *Any Pisot type substitutive dynamical system over d letters admits as a topological factor a minimal translation on the torus $\mathbb{T}^{d-1}$.*

Rational spectrum. In [413], M. Solomyak focuses on the rational part of the spectrum of unimodular Pisot type substitutive dynamical systems:

Theorem 7.3.34 (M. Solomyak [413]). *The rational spectrum of a unimodular substitutive dynamical system of Pisot type is finite.*

Note that this is not true in general for non-unimodular substitutions, as stated in Theorem 7.3.10.

A specific class of examples: β-numeration. A special attention was devoted by B. Solomyak in [412, 411] to substitutions of Pisot type whose incidence matrix has a specific form, such as to be a companion matrix.

For instance, the dynamical system associated with the unimodular substitution of Pisot type $1 \mapsto 123\ldots d$, $2 \mapsto 1$, $3 \mapsto 2, \ldots, d \mapsto d-1$ has a purely discrete spectrum, and is measure-theoretically isomorphic to the rotation, over $(d-1)$-dimensional torus, by a normalized right dominant eigenvector for the incidence matrix.

From a similar point of view, substitutions defined by $1 \mapsto 1^{a_1}2$, $2 \mapsto 1^{a_2}3$, $\ldots, d-1 \mapsto 1^{a_d-1}d$ and $d \mapsto 1^{a_d}$ have a very specific matrix. The numeration systems associated with the dominant eigenvalue of such a matrix has been studied in the context of β-numeration. From the properties of β-expansion, B. Solomyak deduces in [412] that substitutions of Pisot type which are of the above form have a purely discrete spectrum.

Two-letter substitutions: the coincidence condition. M. Hollander and B. Solomyak restate in [206] the proof of the results of A. N. Livshits and apply them to the case of Pisot type substitutions over a two-letter alphabet.

He proves that the systems associated with such substitutions have a purely discrete spectrum if the substitution satisfies a combinatorial condition which generalizes the coincidence condition introduced by F. M. Dekking for substitutions of constant length (see Sec. 7.3.1). Indeed, a substitution over the alphabet $\{1, 2\}$ is said to satisfy the *coincidence condition* if there exist two integers k, n such that $\sigma^n(1)$ and $\sigma^n(2)$ have the same k-th letter, and the prefixes of length $k-1$ of $\sigma^n(1)$ and $\sigma^n(2)$ have the same image under the abelianization map.

This definition appeared in literature before [206] (see [340]). However, the following results, although conjectured, were not proved before.

Theorem 7.3.35 (Hollander, Solomyak [206]). *Let σ be a substitution of Pisot type over a two-letter alphabet which satisfies the coincidence condition. Then the substitutive dynamical system associated with σ has a purely discrete spectrum.*

Remark. The condition of coincidence and the condition of no nontrivial coboundary are distinct: we already stated that two-letter substitutions have no nontrivial coboundary, whereas there exist two-letters substitutions with no coincidence, for instance the Morse sequence. ■

Up to recently, it was not known whether there existed a Pisot type two-letter substitution which did not satisfy the coincidence condition. A preprint from M. Barge and B. Diamond answers that question, by proving the following:

Theorem 7.3.36 (Barge, Diamond [54]). *Any substitution of Pisot type over a two-letter alphabet satisfies the coincidence condition.*

The proof is a very nice mix of combinatorics and geometry.

7.4 The Rauzy fractal

The spectral theory of substitutive dynamical systems implies a better understanding of these systems. In particular, under some conditions such as unimodularity or Pisot type, these symbolic dynamical systems are not weakly mixing, and their spectrum is the spectrum of a toral rotation. An important remaining question is whether these systems have a purely discrete spectrum. To answer that question, G. Rauzy developed the idea of looking for a geometric representation of substitutive dynamical systems, as a rotation on a suitable space. This representation being explicit, the problem of purely discrete spectrum may be explored through the question of the representation's injectivity.

Following that point of view, the remainder of this chapter is devoted to the reverse of the question that initiated the interest for substitutions: indeed, as symbolic dynamical systems were first introduced to understand better the dynamics of a given geometric map, one can ask for the reverse, that is, which geometric actions are coded by a given substitution?

7.4.1 Where symbolic dynamics codes geometry...

Geometric representation. M. Queffélec defines in [340] (pp. 140) the notion of geometric representation:

Definition 7.4.1. *A* geometric representation *of a symbolic dynamical system* (X, S) *is a continuous map* φ *from* X onto *a geometric dynamical system* $(\mathcal{X}, T)$*, on which there exists a partition indexed by the alphabet* $\mathcal{A}$*, such that every word in* X_u *is the itinerary of a point of* $(\mathcal{X}, T)$ *with respect to the partition.*

It is thus natural to ask which systems admit such a geometric representation, and, if so, how precise the representation is, that is, what is the degree of injectivity of the representation. The representation is considered to be precise if the map φ is a semi-topological conjugacy.

Nontrivial examples of representations. There have been many partial answers given to Queffélec's question: in [49], G. Rauzy and P. Arnoux introduce the so-called *Arnoux-Rauzy sequences*, introduced in Chap. 1; they show that each such sequence codes the orbit of a point under an exchange of six intervals on the circle.

In [85], M. Boshernitzan and I. Kornfeld give explicitly a conjugacy map between the dynamics of the restriction to the limit of a piecewise translation and the dynamics of a substitutive system.

In [166], S. Ferenczi obtains a geometric realization of the Chacon sequence as exduction of a triadic rotation (see also Chap. 5).

Recently, J. Cassaigne, S. Ferenczi and L. Q. Zamboni have exhibited an example of a non-substitutive Arnoux-Rauzy sequence whose properties are very far from those of substitutive Arnoux-Rauzy sequences: in particular, their example cannot be geometrically realized as a natural coding of a toral translation [110] (we will come back to this topic in Chap. 12).

In [407], V. F. Sirvent gives a geometrical realization of the Tribonacci substitution as a dynamical system defined on a geodesic lamination on the hyperbolic disc. This result is generalized in the same paper, for all Arnoux-Rauzy sequences.

Representations by a translation on a compact abelian group. The spectral theory of substitutive dynamical systems provides some information about the geometric realizations by a translation on a compact group that one may get: if a substitutive dynamical system is measure-theoretically isomorphic to a translation on a compact group, this translation must have in its spectrum all the eigenvalues of the dynamical system. As a consequence, the compact group must be a n-dimensional torus if the spectrum of the system is irrational and of rank n. The compact group must admit as a factor, either finite groups, if the spectrum contains rational numbers, or p-adic groups, if there are enough rational eigenvalues (as is the case for constant length substitutions). In a general case, the compact group may be a solenoid.

According to the spectral theory of Pisot type substitutive dynamical systems presented in Sec. 7.3.3, and under this hypothesis of no non-trivial coboundary, unimodular Pisot type substitutive systems are good candidates to have a purely discrete spectrum, that is, to be measure-theoretically isomorphic to a toral translation, this translation being obtained with the frequencies of occurrences of letters in any word of the system (see Proposition 7.3.29). A method to test this hypothesis is to try to explicit a conjugacy between the shift map on the substitutive system and the toral translation formerly described. Then, the problem will be to study the injectivity of the conjugacy.

The idea of representing geometrically substitutive systems of Pisot type was developed first by G. Rauzy [350], within the framework of discrepancy for some real sequences. The aim of this section is to expose Rauzy's work about the Tribonacci substitution and the development given by many authors to this original and fundamental example.

7.4.2 The Tribonacci substitution

In 1981, G. Rauzy generalized in [350] the dynamical properties of the Fibonacci substitution to a three-letter alphabet substitution, called Tribonacci substitution or Rauzy substitution, and defined by

$$\sigma(1) = 12 \qquad \sigma(2) = 13 \qquad \sigma(3) = 1.$$

The incidence matrix of this substitution is $\mathbf{M}_\sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Its characteristic polynomial is $X^3 - X^2 - X - 1$, and its set of roots consists of a real number $\beta > 1$ and two complex conjugates α and $\overline{\alpha}$. Since it satisfies the equation $\alpha^3 = \alpha^2 + \alpha + 1$, the complex number α is called a *Tribonacci number*, in reference to the Fibonacci equation $X^2 = X + 1$.

In particular, the incidence matrix $\mathbf{M}$ admits as eigenspaces in $\mathbb{R}^3$ an expanding one-dimensional direction and a contracting plane.

7.4.3 Geometric construction of the Rauzy fractal

Let u denote a biinfinite word which is a periodic point for σ, for instance $u = \sigma^{3^\infty}(1) \cdot \sigma^\infty(1)$. This infinite word u is embedded as a broken line in $\mathbb{R}^3$ by replacing each letter in the periodic point by the corresponding vector in the canonical basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ in $\mathbb{R}^3$.

An interesting property of this broken line is that it remains at a bounded distance of the expanding direction of $\mathbf{M}_\sigma$, turning around this line. When one projects the vertices of the broken line on the contracting plane , parallel to the expanding directing, then one obtains a bounded set in a two-dimensional vector space. The closure of this set of points is a compact set denoted by $\mathcal{R}$ and called the *Rauzy fractal* (see Fig. 7.5).

To be more precise, denote by π the linear projection in $\mathbb{R}^3$, parallel to the expanding directing, on the contracting plane, identified with the complex plane $\mathbb{C}$. If $u = (u_i)_{i \in \mathbb{Z}}$ is the periodic point of the substitution, then the Rauzy fractal is

$$\mathcal{R} = \overline{\left\{ \pi \left(\sum_{i=0}^{n} \mathbf{e}_{u_i} \right) \; ; \; n \in \mathbb{Z} \right\}}.$$

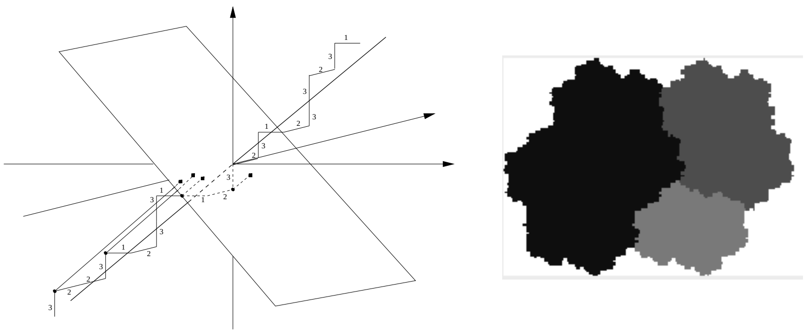


Fig. 7.5. The projection method to get the Rauzy fractal for the Tribonnacci substitution.

Three subsets of the Rauzy fractal can be distinguished. Indeed, for each letter $j = 1, 2, 3$, the *cylinder* $\mathcal{R}_j$ is defined to be the closure of the set of origins of any segment on the broken line which is parallel to the canonical vector $\mathbf{e}_j$:

$$\mathcal{R}_j = \overline{\left\{ \pi\left(\sum_{i=0}^{n} \mathbf{e}_{u_i}\right) ;\ n \in \mathbb{Z},\ u_n = j \right\}}.$$

The union of these three cylinders covers the compact $\mathcal{R}$, and G. Rauzy proved in [350] that their intersection has measure zero:

Proposition 7.4.2 (Rauzy [350]). *The cylinders constitute a measurable partition of the Rauzy fractal:*

$$\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3, \quad \textit{the union being disjoint in measure.}$$

7.4.4 Dynamics over the Rauzy fractal

One can note that it is possible to move on the broken line, from a vertex to the following one, thanks to a translation by one of the three vectors of the canonical basis. In the contracting plane, this means that each cylinder $\mathcal{R}_i$ can be translated by a given vector, i.e., $\pi(\mathbf{e}_i)$, without going out of the Rauzy fractal:

Exercise 7.4.3. Prove that for every letter $i = 1, 2, 3$, the following inclusion is true:

$$\mathcal{R}_i - \pi(\mathbf{e}_i) \subset \mathcal{R}.$$

Dynamics. Thus, the following map φ, called a *domain exchange* (see Fig. 7.6) is defined for any point of the Rauzy fractal which belongs to a unique cylinder. According to Proposition 7.4.2, the cylinders intersect on a set of measure zero, so that this map is defined almost everywhere on the Rauzy fractal:

$$\forall x \in \mathcal{R}, \quad \varphi(x) = x - \pi(\mathbf{e}_i), \quad \text{if } x \in \mathcal{R}_i.$$

Denote by $\mathcal{R}'_i$ the image of the cylinder $\mathcal{R}_i$ through the domain exchange:

$$\mathcal{R}'_i = \mathcal{R}_i - \pi(\mathbf{e}_i).$$

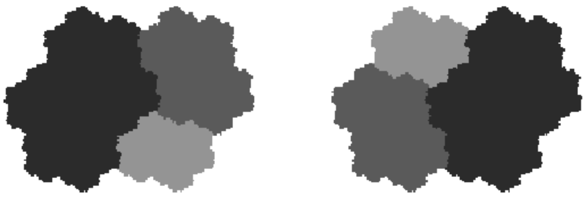


Fig. 7.6. Domain exchange over the Rauzy fractal.

Symbolic dynamics. It is natural to code, up to the partition defined by the 3 cylinders, the action of the domain exchange φ over the Rauzy fractal $\mathcal{R}$. G. Rauzy proved in [350] that the coding map, from $\mathcal{R}$ into the three-letter alphabet full shift $\{1,2,3\}^{\mathbb{Z}}$ is almost everywhere one-to-one. Moreover, this coding map is onto the substitutive system associated with the Tribonacci substitution. Thus we have the following result:

Theorem 7.4.4 (Rauzy, [350]). *The domain exchange φ defined on the Rauzy fractal $\mathcal{R}$ is semi-topologically conjugate to the shift map on the symbolic dynamical system associated with the Tribonacci substitution.*

7.4.5 Geometric interpretation of the substitution: self-similarity

Number theory. Note that G. Rauzy obtains these results by proving that the Rauzy fractal can be defined numerically as the complex set of power series in α whose digits consist in all the sequences of 0 and 1 with no three consecutive ones:

$$\mathcal{R} = \left\{ \sum_{i \geq 0} \varepsilon_i \, \alpha^i ; \;\; \varepsilon_i \in \{0,1\}; \, \varepsilon_i \varepsilon_{i+1} \varepsilon_{i+2} = 0 \right\} \subset \mathbb{C}. \tag{7.5}$$

The condition $\varepsilon_0 = 0$ produces the cylinder $\mathcal{R}_1$, whereas the cylinder $\mathcal{R}_2$ is defined by the condition $\varepsilon_0 \varepsilon_1 = 10$, and $\mathcal{R}_3$ by $\varepsilon_0 \varepsilon_1 = 11$.

The Tribonacci substitutive dynamical system X_σ has a special property: since the letter 1 appears only at the beginning of $\sigma(1)$, $\sigma(2)$ and $\sigma(3)$, we have

$$[1] = \sigma\left(X_\sigma\right), \quad [2] = S\left(\sigma^2(X_\sigma)\right), \quad [3] = S\left(\sigma[2]\right).$$

The abelianization of the shift map S on the broken line is the translation to the following vertex, which projects on the contracting plane as the domain exchange φ. In the same point of view, the abelianization of the substitution σ is the matrix $\mathbf{M}$, which projects on the contracting plane as a complex multiplication by the Tribonacci number α. So that we get:

$$\mathcal{R}_1 = \alpha \mathcal{R}, \quad \mathcal{R}_2 = \alpha^2 \mathcal{R} + 1 \quad \mathcal{R}_3 = \alpha^3 \mathcal{R} + \alpha + 1. \tag{7.6}$$

This means that the Rauzy fractal $\mathcal{R}$ is partitioned by contractions of itself. We say that $\mathcal{R}$ has a *self-similar structure.*

Exercise 7.4.5. Deduce (7.6) from (7.5).

7.4.6 Tilings : toral translation as a factor

The domain exchange φ is defined only almost everywhere. This problem is solved when one quotients the contracting plane by the lattice $\mathcal{L} = \mathbb{Z}\,\pi(\mathbf{e}_1 -$

$\mathbf{e}_3) + \mathbb{Z}\,\pi(\mathbf{e}_2 - \mathbf{e}_3)$. Indeed, this quotient maps the contracting plane onto a two-dimensional torus; the 3 vectors $\pi(\mathbf{e}_1)$, $\pi(\mathbf{e}_2)$ and $\pi(\mathbf{e}_3)$ map onto the same vector on the torus. Thus, the projection of the domain exchange φ on the quotient is a toral translation.

G. Rauzy proved in [350] that the restriction of the quotient map to the Rauzy fractal is onto and almost everywhere one-to-one. Finally, we get that the domain exchange on the Rauzy fractal, which is known to be semi-topologically conjugate to the Tribonacci substitutive dynamical system, is also measure-theoretically isomorphic to a minimal translation on the two-dimensional torus.

More geometrically, the fact that the quotient map is one-to-one means that any translate of $\mathcal{R}$ by a vector of the lattice $\mathcal{L}$ intersects $\mathcal{R}$ on a set of measure zero. The fact that the restriction of the quotient map to the Rauzy fractal is onto means that the union of these translated Rauzy fractal recovers the plane. In other words, the Rauzy fractal generates a periodic tiling of the plane (see Fig. 7.7).

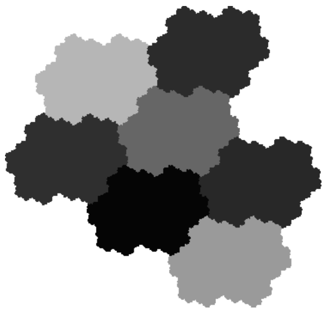


Fig. 7.7. The periodic and autosimilar tiling generated by the Rauzy fractal.

The self-similarity of $\mathcal{R}$ implies that the induced tiling is quasi-periodic. This is analogous to the well understood connection between Sturmian sequences, tilings of the line and quasi-crystals (see [39]): through methods similar to those used by G. Rauzy in [350], E. Bombieri and J. Taylor use substitutions to exhibit a connection between quasi-crystals and number theory in [82] (see also [52, 244, 386, 431] and [440] for a study of the Rauzy fractal as a quasicrystal).

Finally, by mixing dynamics, self-similarity and number theory, we get the three following equivalent results:

Theorem 7.4.6 (Rauzy, [350]).

- Geometry. *The Rauzy fractal generates a self-similar periodic tiling of* $\mathbb{C}$.
- Dynamics. *The symbolic dynamical system generated by the Tribonacci substitution is measure-theoretically isomorphic to a toral translation.*

- Spectral theory. *The Tribonacci substitutive dynamical system has a purely discrete spectrum.*

Additional geometric properties. Using methods introduced by F. M. Dekking in [137], S. Ito and M. Kimura obtain in [222] an alternative construction of the Rauzy fractal and use it to compute the Hausdorff dimension of the boundary of the Rauzy fractal, which is strictly larger than 1 (see Chap. 8). V. Sirvent obtains equivalent results with a different method in [401]. By using numeration systems, A. Messaoudi obtains in [289, 291, 290] additional properties of the fractal $\mathcal{R}$ including an explicit boundary parametrization.

7.5 Geometric realization of substitutions of Pisot type

The case of the Tribonacci substitution being quite well understood, the question of its generalization to a larger class of substitutions is thus natural. Which substitutions generate self-similar and periodic tilings? Which ones code the action of a toral translation? From a spectral point of view, a general formulation for these questions is: which substitutions generate a dynamical system with a purely discrete spectrum?

7.5.1 Generalized Rauzy fractals

For the Tribonacci substitution, the two main properties that allow us the construction of the Rauzy fractal are the following:

- The incidence matrix of the substitution has a one-dimensional expanding direction.
- The broken line corresponding to any periodic point of the substitution remains at a bounded distance from the expanding direction.

Generalized Rauzy fractals. These properties are satisfied as soon as the modulus of all but one of the eigenvalues of the substitution incidence matrix is strictly less than 1. This means that the dominant eigenvalue of the matrix is a *Pisot number* and the substitution is *of Pisot type.*

Definition 7.5.1. *Let σ be a unimodular Pisot type substitution over d letters. Let $\alpha_1 > 1, \ldots, \alpha_r$ be the real eigenvalues of the incidence matrix of σ, $\alpha_{r+1}, \overline{\alpha_{r+1}}, \ldots, \alpha_{r+s}, \overline{\alpha_{r+s}}$ be its complex eigenvalues. Let $\mathbf{v}_1, \ldots \mathbf{v}_{r+s} \in \mathbb{C}^d$ be right eigenvectors of the incidence matrix associated with each of the eigenvalues α_i, $1 \leq i \leq r+s$.*

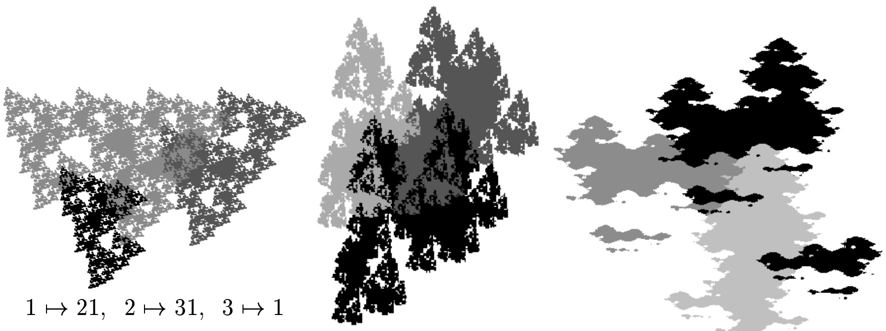
For any finite word W, we define the following vector in $\mathbb{R}^{r-1} \times \mathbb{C}^s$:

$$\delta(W) = \sum_{i=1}^{|W|} \left((\mathbf{v}_2)_{W_i}, \ldots, (\mathbf{v}_{r+s})_{W_i} \right) \in \mathbb{R}^{r-1} \times \mathbb{C}^s .$$

Let $u = (u_i)_{i\in\mathbb{Z}}$ *be a periodic point for* σ. *Then the* generalized Rauzy fractal *or* Rauzy fractal *associated with the substitution* σ *is the following compact set:*

$$\mathcal{R} = \overline{\{\delta\,(u_0 \dots u_i)\,;\; i \in \mathbb{N}\}} \subset \mathbb{R}^{r-1} \times \mathbb{C}^s .$$

Examples of generalized Rauzy fractals are shown in Fig. 7.8.



$1 \mapsto 21,\ \ 2 \mapsto 31,\ \ 3 \mapsto 1$

$1 \mapsto 132,$
$2 \mapsto 21323,$
$3 \mapsto 321323132$

$1 \mapsto 11223,\ \ 2 \mapsto 123,\ \ 3 \mapsto 2$

Fig. 7.8. Examples of generalized Rauzy fractals.

Since $r + 2s = d$, $\mathbb{R}^{r-1} \times \mathbb{C}^s$ is a $\mathbb{R}$-vector space of dimension $d-1$, so that the Rauzy fractal associated with a unimodular substitution of Pisot type can be considered as a part of $\mathbb{R}^{d-1}$. Let us illustrate this point for the Tribonacci substitution, by proving that the Rauzy fractal defined in Sec. 7.4.3 and the one defined above are identical:

Exercise 7.5.2. Let σ be the Tribonacci substitution and α be the Tribonacci number, with $|\alpha| < 1$.

1. Prove that the vector $\mathbf{u}_\alpha = (a_1, a_2, a_3) = (1/\alpha, 1/\alpha^2, 1/\alpha^3)$ is a normalized right eigenvector of $\mathbf{M}_\sigma$ associated with α.
   Then δ is a mapping from $\{1,2,3\}^*$ on $\mathbb{C}$: $\delta(W_1 \dots W_n) = \sum_{1\le i\le n} a_{W_i}$.
2. Determine a left eigenvector $\mathbf{v}_\alpha$ of $\mathbf{M}_\sigma$ associated with α and which satisfies ${}^t\mathbf{v}_\alpha \mathbf{u}_\alpha = 1$.
3. Let $\mathbf{w}_1$, $\mathbf{w}_2$ be its real and imaginary parts: $\mathbf{v}_\alpha = \mathbf{w}_1 + i\mathbf{w}_2$. Show that the family $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a real basis of the contracting plane of $\mathcal{H}$ of $\mathbf{M}_\sigma$.
4. Let $q_{\mathcal{H}} : \mathbb{C} \to \mathcal{H}$ be following $\mathbb{R}$-linear embedding:

   $$\forall z = x + iy \in \mathbb{C},\ q_{\mathcal{H}}(z) = 2x\mathbf{w}_1 - 2y\mathbf{w}_2.$$

   Prove that up to this identification, the map δ coincides with the linear projection $\pi : \mathbb{R}^3 \to \mathcal{H}$ parallel to the expanding direction of $\mathbf{M}_\sigma$:

   $$\forall\, W \in \{1,2,3\}^*,\ q_{\mathcal{H}}(\delta(W)) = \pi(\mathbf{l}(W)).$$

5. Let $\mathcal{R}$ be the generalized Rauzy fractal of σ given in Definition 7.5.1 and $\mathcal{R}_1$ be the Rauzy fractal as defined in Sec. 7.4.3. Prove that $q_{\mathcal{H}}(\mathcal{R}) = \mathcal{R}_1$.

The result proved in this exercise is general:

Lemma 7.5.3. *Let σ be a substitution of Pisot type over a d-letter alphabet. Let π be the linear projection parallel to the expanding direction of $\mathbf{M}_\sigma$, defined from $\mathbb{R}^3$ onto the contracting hyperplane $\mathbf{M}_\sigma$. There exists a linear embedding $q_{\mathcal{H}} : \mathbb{R}^{r-1} \times \mathbb{D}^s \to \mathcal{H}$ such that: $\forall W \in \mathcal{A}$, $q_{\mathcal{H}}(\delta(W)) = \pi(\mathbf{l}(W))$.*

This implies that Definition 7.5.1 generalizes the Rauzy fractal defined by G. Rauzy for the Tribonacci substitution:

Corollary 7.5.4. *Let $\mathcal{R}$ be the generalized Rauzy fractal of σ. Let $\mathcal{R}_1 = \{\pi\left(\sum_{i=0}^{n} \mathbf{e}_{u_i}\right);\ n \in \mathbb{Z}\}$, where u is a periodic point for σ. Then $\mathcal{R}$ and $\mathcal{R}_1$ are linearly isomorphic: $q_{\mathcal{H}}(\mathcal{R}) = \mathcal{R}_1$.*

The reason that led us to define generalized Rauzy fractals by a numerical method instead of a geometric method is that the map δ allows a better understanding of the action of the shift map and of the substitution σ on the Rauzy fractal. Indeed, the following relationships are easy to get:

$$\forall\, W_1, W_2 \in \mathcal{A}^*, \quad \delta(W_1\, W_2) = \delta(W_1) + \delta(W_2);$$
$$\forall\, W \in \mathcal{A}^*, \ \ \delta(\sigma(W)) \ = \begin{pmatrix} \alpha_2 & & (0) \\ & \ddots & \\ (0) & & \alpha_{r+s} \end{pmatrix} \delta(W). \tag{7.7}$$

For instance, such equalities are the main argument to get the formulas of Sec. 7.4.5 that describe the self-similar structure of the Rauzy fractal for the Tribonacci substitution.

Cylinders. As for the Tribonacci substitution, let us define the *cylinders*:

$$\forall\, i = 1 \dots d, \quad \mathcal{R}_i = \overline{\{\delta\,(u_0 \dots u_j)\,;\ j \in \mathbb{N};\ u_j = i\}}.$$

Exercise 7.5.5. Prove that the Rauzy fractal is the union of cylinders:

$$\mathcal{R} = \bigcup_{i=1}^{d} \mathcal{R}_i.$$

Remark. Note that we do not know at the moment whether the union is in general disjoint in measure. We will see in the next sections that this is one of the main problems raised by the study of generalized Rauzy fractals.

7.5.2 Topological properties of generalized Rauzy fractals

Many authors investigated topological properties of Rauzy fractals: V. F. Sirvent and Y. Wang prove in [409] that generalized Rauzy fractals have nonempty interior and are the closure of their interior.

Conditions for simple connectedness are given in [224] and in [187]. Conditions for connectedness are given by V. Canterini in [100, 101], and restated in a more algorithmic way in [398]. An estimation for the Haussdorf dimension of the boundary of Rauzy fractals is computed in [163].

C. Holton and L. Q. Zamboni define in [208, 211, 210] real and complex representations associated with every substitution, the incidence matrix of which has a nonzero eigenvalue of modulus less than one. Their work gives explicit bounds for the Haussdorf dimension of the projections on axes of the Rauzy fractal associated with a substitution. More precisely, for any $k = 2, \ldots, r+s$, they prove that the projection of the Rauzy fractal on the axis associated with the eigenvalue α_k, that is, the k-coordinate of the Rauzy fractal in $\mathbb{R}^{r-1} \times \mathbb{D}^s$ (this axis can be $\mathbb{R}$ or $\mathbb{C}$), admits a Haussdorf dimension which is less than $-\log(\alpha_1)/\log(\alpha_k)$.

Thus, for any $k = 2, \ldots, r+s$, the projection is a Cantor set. If $k = s+1, \ldots, s+d$ and $|\alpha_k|^2\alpha_1 < 1$, then the projection has nonempty interior.

Let us also mention that Rauzy fractals appear in [96], as continuous factors of odometers, derived from sequence of cutting time of an interval transformation restricted to a wild attractor.

Finally, some precise results exist about the n-dimensional generalization of the Tribonacci substitution: the n-bonacci substitution is the n-letter alphabet substitution defined by

$$1 \mapsto 12, \quad 2 \mapsto 13, \quad \ldots \quad n-1 \mapsto 1n, \quad n \mapsto 1.$$

The spectral study of this substitution was made by B. Solomyak (see Sec. 7.3.3). In [400, 402], V. F. Sirvent studies dynamical and geometrical properties of the Rauzy fractals associated with these substitutions. Other properties, mainly arithmetical, of the substitutions are given in [405, 403, 408]. In [406] are studied the dynamical and geometrical properties which are common to all the k-bonacci substitutions.

The construction of the Rauzy fractal being generalized, the question is to obtain conditions for a substitutive system to be semi-topologically conjugate to a domain exchange on the Rauzy fractal associated with the substitution, that is, to generalize the results of [350].

The problem of domain exchange. As for the Tribonacci substitution, one can move along the broken line associated with a periodic point of a substitution thanks to a translation by one of the canonical vectors. This can be expressed in the following way:

Exercise 7.5.6. Let $\mathcal{R}$ be the Rauzy fractal associated with a unimodular substitution of Pisot type, as defined in Definition 7.5.1. Let $\mathcal{R}_i$ be the cylinders, the union of which is equal to $\mathcal{R}$.

Prove that for every $i = 1 \ldots d$, the following inclusion is satisfied:

$$\mathcal{R}_i - \delta(i) \subset \mathcal{R}. \tag{7.8}$$

The inclusion (7.8) means that one may define dynamics over the Rauzy fractal, in terms of a domain exchange, as soon as the cylinders $\mathcal{R}_i$ are disjoint up to a set of zero measure. Unfortunately, whereas we know that the union of the d cylinders $\mathcal{R}_i$ is equal to $\mathcal{R}$ (Exercise 7.5.5), nothing allows us to say that these cylinders partition $\mathcal{R}$ up to a set of zero measure.

A necessary condition for this is the substitution to be unimodular: if the substitution is not unimodular, the substitutive system may have a p-adic component and the cylinders should intersect on a nonzero measure set.

Unfortunately again, the restriction to the class of unimodular substitution is still not sufficient to define dynamics on the Rauzy fractal of a substitution: for substitutions over a two-letter alphabet, B. Host proved explicitly that the coincidence condition (see Sec. 7.3.3) is necessary and sufficient for the cylinders to be disjoint in measure. This is developed in Sec. 7.5.4.

For substitutions over a finite alphabet, P. Arnoux and S. Ito give an alternative construction of the Rauzy fractal thanks to induction and successive approximations (see also Chap. 8). This allowed them to generalize the coincidence condition and to prove that this condition is sufficient for the cylinders to intersect on a set of zero measure. Their main results are presented in Sec. 7.5.4.

The problem of tilings. The main problem of the geometric realization is to know whether a substitutive dynamical system is isomorphic to a compact group translation. Again, for two-letter alphabet substitutions, B. Host answered that question in terms of coincidences (see Sec. 7.5.4).

Another approach to this problem was made in [103] by formalizing and generalizing the ideas of G. Rauzy in [353]: thanks to recognizability and formal power series, it is possible to define an explicit conjugacy between a substitutive system and the domain exchange over the Rauzy fractal. This conjugacy is the reverse from the map which maps the Rauzy fractal onto the symbolic space $\{1, \ldots, d\}$, by coding the action of the domain exchange over the Rauzy fractal. This brings a characterization of substitutive systems with a purely discrete spectrum. It will be presented in Sec. 7.5.5.

7.5.3 The coincidence condition

As explained in Sec. 7.3, the spectrum of a substitutive dynamical system is related to the notion of coboundary introduced by B. Host. In two specific cases, that is, constant length substitutions and two-letter substitutions, the spectrum is purely discrete if and only if the substitution satisfies the combinatorial condition of coincidence (Theorem 7.3.4 and 7.3.35).

Let us point out the general definition of coincidences as introduced in [43], and its relationship with coboundaries. We will see later that the general condition of coincidence plays an important role in the problem of the domain exchange.

As before, $\mathbf{l} : \{1, \dots, d\} \to \mathbb{N}^d$ is the abelianization map over $\{1, \dots, d\}$.

Definition 7.5.7 (Arnoux, Ito [43]). *A substitution σ satisfies the* coincidence condition *on prefixes (respectively suffixes) if for every couple of letters (j, k), there exists a constant n such that $\sigma^n(j)$ and $\sigma^n(k)$ can be decomposed in the following way:*

$$\sigma^n(b_1) = p_1 a s_1 \quad \textit{and} \quad \sigma^n(b_2) = p_2 a s_2, \quad \textit{with} \quad \mathbf{l}(p_1) = \mathbf{l}(p_2)$$
$$(\textit{respectively} \quad \mathbf{l}(s_1) = \mathbf{l}(s_2))\,.$$

Example 7.5.8. The Tribonacci substitution satisfies the coincidence condition, whereas the Morse sequence does not satisfy this condition.

The coincidence condition is connected with the notion of coboundary introduced by B. Host and detailed in Sec. 7.3.2:

Lemma 7.5.9 (Host). *Let σ be a substitution with a nontrivial coboundary $g : \mathcal{A} \to \mathbb{U}$. Let f be the function of modulus 1 which satisfies $f(b) = g(a)f(a)$ as soon as the word ab belongs to the language of the fixed point of the substitution.*

If there exist two letters a and b and a rank k such that

- $f(a) \neq f(b)$,
- $\sigma^k(a)$ *begins with* a *and* $\sigma^k(b)$ *begins with* b,

then σ does not satisfy the coincidence condition on prefixes.

Proof. If σ satisfies the coincidence condition, there exist a multiple k_1 of k and a letter c such that

$$\sigma^{k_1}(a) = p_1 c s_1 \quad \text{and} \quad \sigma^n(b) = p_2 c s_2, \quad \text{with} \quad \mathbf{l}(p_1) = \mathbf{l}(p_2).$$

Let $p_1 = a u_1 \dots u_{n-1}$ and $p_2 = b v_1 \dots v_{n-1}$. Since g is a coboundary we have

$$\frac{f(c)}{f(b)} = g(v_{n-1}) \dots g(v_1) = g(u_{n-1}) \dots g(u_1) = \frac{f(c)}{f(a)}.$$

Thus $f(a) = f(b)$ which is impossible. ∎

By using the following method of B. Host to construct substitutions with nontrivial coboundaries, one gets a family of substitutions without coincidence.

Proposition 7.5.10 (Host). *Let σ be a primitive substitution over an alphabet $\mathcal{A}$. Let $\mathcal{L}_2(X_\sigma)$ be the set of words of length 2 in the language of the substitution.*

The following substitution τ on the alphabet $\mathcal{L}_2(X_\sigma)$ is primitive and has a nontrivial coboundary:

$$\forall\, ab \in \mathcal{L}_2(X_\sigma), \quad \text{if } \begin{cases} \sigma(a) = a_1 \ldots a_k \\ \sigma(b) = b_1 \ldots b_l \end{cases},$$
$$\text{then } \tau(ab) = (a_1 a_2)(a_2 a_3) \cdots (a_{k-1} a_k)(a_k b_1) \in \mathcal{L}_2(X_\sigma)^*.$$

Proof. For any letter $a \in \mathcal{A}$, choose a number $\psi(a) \neq 1$ of modulus 1. For any $(ab) \in \mathcal{L}_2(X_\sigma)$, let $f(ab) = \psi(a)$ and $g(ab) = \psi(b)/\psi(a)$.

Let (ab) and $(a'b')$ be two elements of $\mathcal{L}_2(X_\sigma)$ such that the word $(ab)(a'b')$ (of length 2 in $\mathcal{L}_2(X_\sigma)^*$) belongs to the language of τ. By construction, we have $b = a'$, so that

$$f(a'b') = \psi(a') = \psi(b) = g(ab)\psi(a) = g(ab) f(ab).$$

Thus, g is a nontrivial coboundary for τ. ∎

Remark. This last method to get substitutions with nontrivial coboundaries shows that the notion of coboundary is the analogous of the notion of height in the constant length case (see Sec. 7.3.1). ∎

Substitutions of Pisot type. This last method does not provide Pisot type substitutions with nontrivial coboundary: the characteristic polynomial of the substitution that is obtained is never irreducible. We still do not know whether there exists a Pisot type substitution which does not satisfy the coincidence condition.

A negative answer to that question was recently given by M. Barge and B. Diamond [54], for substitutions over a two-letter alphabet, as stated in Theorem 7.3.36. In the general context of substitutions of Pisot type over more than 2 letters, they prove that the coincidence condition is satisfied by at least two letters:

Theorem 7.5.11 (Barge, Diamond [54]). *Let σ be a substitution of Pisot type on the alphabet $\mathcal{A} = \{1 \ldots d\}$. There exist two distinct letters i, j for which there exist integers k, n such that $\sigma^n(i)$ and $\sigma^n(j)$ have the same k-th letter, and the prefixes of length $k-1$ of $\sigma^n(i)$ and $\sigma^n(j)$ have the same image under the abelianization map.*

This is the most complete result on coincidence that is know at the moment.

7.5.4 Condition for non-overlapping in generalized Rauzy fractals

Construction by successive approximations. P. Arnoux and S. Ito generalize in [43] the alternative construction given by S. Ito and M. Kimura [222] of the Rauzy fractal associated with the Tribonacci substitution. Indeed, they construct by successive approximations the Rauzy fractal for any unimodular Pisot type substitution. The construction is described precisely in Chap. 8.

Thanks to this construction and by generalizing some of the techniques of B. Host in [216], P. Arnoux and S. Ito prove that, under the condition of coincidence, the d cylinders of the generalized Rauzy fractal associated with a substitution over d letters are disjoint in measure. The first consequence of this is that the Rauzy fractal of such a substitution has a self-similar structure: a generalized Rauzy fractal satisfies formulas such as those stated in Sec. 7.4.5 for the Tribonacci substitution.

The second consequence is that, now, nothing prevents us to define a domain exchange on the Rauzy fractal, as it was the case for the Tribonacci substitution.

$$\forall x \in \mathcal{R}, \quad \varphi_\sigma(x) = x - \delta(i), \quad \text{ if } x \in \mathcal{R}_i.$$

P. Arnoux and S. Ito prove that coding this action on the Rauzy fractal leads to the symbolic dynamical system generated by the substitution:

Theorem 7.5.12 (Arnoux, Ito [43], see Chap. 8). *Let σ be a unimodular Pisot type substitution over a d-letter alphabet which satisfies the condition of coincidence. Then the substitutive dynamical system associated with σ is measure-theoretically isomorphic to the exchange of d domains defined almost everywhere on the self-similar Rauzy fractal of σ.*

Remark. This result was previously proved by B. Host in an unpublished work [216] for substitutions over a two-letter alphabet.

Appropriate quotient: toral translation factor. The Euclidean space $\mathbb{R}^{d-1}$ projects onto a $d-1$-dimensional torus modulo the lattice $\mathcal{L} = \sum_{i=2}^{d} \mathbb{Z}(\delta(i) - \delta(d))$. Through the quotient map, all the vectors $\delta(i)$ map onto the same vector, so that the domain exchange φ_σ maps to a toral translation. In other words, the diagram shown in Fig. 7.5.4 commutes, where the symbol $\simeq$ on an arrow denotes a measure-theoretic isomorphism.

Thus the domain exchange can be factorized by an irrational translation. This implies a more geometrical version of Corollary 7.3.33:

Proposition 7.5.13 (Arnoux, Ito [43]). *Any unimodular Pisot type substitutive dynamical system which satisfies the coincidence condition admits as a continuous factor an irrational translation on the torus $\mathbb{T}^{d-1}$, the fibers being finite almost everywhere.*

Remark. The fibers are finite almost everywhere as the Rauzy fractal is bounded.

$$\begin{tikzcd}
\Omega \arrow[rr, "\text{shift map}"] & & \Omega \\
\mathcal{R} \arrow[u, "\simeq"', "\substack{\text{coding map}\\ \text{up to the}\\ \text{cylinder partition}}"] \arrow[rr, "\tau\ \substack{\text{domain}\\ \text{exchange}}"'] \arrow[d, "\text{quotient map}"'] & & \mathcal{R} \arrow[u, "\simeq", "\substack{\text{coding map}\\ \text{up to the}\\ \text{cylinder partition}}"'] \arrow[d, "\text{quotient map}"] \\
\left(\mathbb{R}^{r-1} \times \mathbb{C}^s\right)/\mathcal{L} \arrow[rr, "\text{addition of } \delta(1)"'] \arrow[d, "\simeq"] & & \left(\mathbb{R}^{r-1} \times \mathbb{C}^s\right)/\mathcal{L} \arrow[d, "\simeq"'] \\
\mathbb{T}^{d-1} \arrow[rr, "\substack{\text{minimal}\\ \text{translation}}"'] & & \mathbb{T}^{d-1}
\end{tikzcd}$$

Fig. 7.9. Relations between the different maps related to Rauzy fractals.

Cardinality of the fibers of the quotient map. According to a proof from B. Host, one can easily prove that the fibers are not only finite almost everywhere: they are constant almost everywhere.

Exercise 7.5.14. Let σ be a unimodular substitution of Pisot type. Let $\varphi : \Omega \to \mathbb{T}^{d-1}$ be a continuous toral representation of the substitutive dynamical system associated with σ, that is, there exists $\mathbf{t} \in \mathbb{T}^{d-1}$ such that $\varphi S(w) = \varphi(w) + \mathbf{t}$, for all $w \in \Omega$.

Let $\psi : \mathbb{T}^{d-1} \mapsto \mathbb{N} \cup \infty$ be the cardinality of the fibers:

$$\forall \mathbf{x} \in \mathbb{T}^{d-1},\ \psi(\mathbf{x}) = \mathrm{Card}\,\{w \in \Omega,\ \varphi(w) = \mathbf{x}\}.$$

1. Let $w_1, \ldots, w_n$ be n distinct elements in Ω and $\varepsilon \in \mathbb{R}$. Let $\mathcal{A}_{n,\varepsilon}$ be the set:
$$\mathcal{A}_{n,\varepsilon} = \{\mathbf{x} \in \mathbb{T}^{d-1};\ \forall i \le n,\ d(\varphi(w_i), \mathbf{x}) < \varepsilon\}.$$
Prove that $\mathcal{A}_{n,\varepsilon}$ is an open set for the Borel topology.
2. Deduce that ψ is a measurable map for the Lebesgue measure over $\mathbb{T}^{d-1}$ and the discrete measure over $\mathbb{N}$.
3. Prove that ψ is constant almost everywhere.

A more explicit version of the tiling problem. Note that the toral translation obtained in the diagram below is precisely the maximal equicontinuous factor of the substitutive dynamical system, up to the hypothesis of no nontrivial coboundary:

Exercise 7.5.15. Let σ be a unimodular substitution of Pisot type with no nontrivial coboundary.

1. Extend δ as a linear map over $\mathbb{R}^d$.
2. Prove that the kernel of δ is generated by any expanding eigenvector $\mathbf{u}$ of the incidence matrix of σ.
3. Consider the decomposition $\delta(1) = \sum_{i\geq 2} \lambda_i(\delta(i) - \delta(1))$. Prove that
$$\forall\, i \geq 2, \quad \lambda_i = \frac{(\mathbf{u})_i}{\sum_j (\mathbf{u})_j}.$$
4. Deduce from Prop. 7.3.29 that the addition of $\delta(1)$ over $\left(\mathbb{R}^{r-1} \times \mathbb{C}^s\right)/\mathcal{L}$ is the maximal equicontinuous factor of the substitutive dynamical system associated with σ.

Corollary 7.5.16. *A unimodular substitutive system of Pisot type, with coincidence and no nontrivial coboundary, has a purely discrete spectrum if and only if the quotient map from the Rauzy fractal onto the torus is one-to-one. In particular, in this case, the Rauzy fractal generates a periodic tiling.*

An answer to the tiling problem? For two-letter substitutions, B. Host proves that the problem of domain exchange is equivalent to the tiling problem. Indeed, in an unfortunately unpublished work [216], B. Host explicitly realizes the inverse map of the coding map up to the partition by cylinder (that is, the arrow $\mathcal{R} \to X_\sigma$ on the lower commutative diagram). He states that there exists an explicit continuous mapping $f : X_\sigma \to \mathcal{R}$ such that

$$\exists\, a_0, a_1 \in \mathbb{R}; \quad \forall w \in X_\sigma,\ f(Sw) = f(w) + a_{w_0}. \tag{7.9}$$

An important fact is the following:

Lemma 7.5.17 (Host [216]). *Let σ be a substitution over a two-letter alphabet. Let $f : X_\sigma \to \mathcal{R}$ be a continuous function which satisfies (7.9). Then f is almost everywhere one-to-one if and only if its projection onto the torus modulo $(a_1 - a_0)$ is also almost everywhere one-to-one.*

As we know that f is almost everywhere one-to-one as soon as the substitution satisfies the coincidence condition (Theorem 7.5.12), we deduce that the dynamical system associated with a unimodular substitution of Pisot type with coincidence over a two-letter alphabet is explicitly semi-topologically conjugate to a toral translation. In other words:

Corollary 7.5.18 (Host [216]). *The generalized Rauzy fractal of a two-letter unimodular substitution of Pisot type with coincidence generates a periodic tiling of $\mathbb{R}$.*

Remark. The "toral" version of this corollary is included in Hollander's result in [206] (see Sec. 7.3.3). However, it was obtained much before Hollander's work. It is less general, since it depends on the unimodular condition, but it has the advantage of being constructive.

Unfortunately, in the general case, these techniques do not allow one to prove that the projection on the torus preserves the injectivity: Lemma 7.5.17 is not true when the alphabet contains strictly more than two letters.

P. Arnoux and S. Ito give in [43] a sufficient condition for the rotation factor to be isomorphic to the substitutive system: if there exists a set with nonempty interior which is included in all the domains (after renormalization) which appear in the construction by successive approximation of the Rauzy fractal of the substitution, then we get an isomorphism (see Chap. 8).

7.5.5 Rauzy fractals and formal power series: condition for tiling

Thus, the work of P. Arnoux and S. Ito generalizes [350] to a large class of substitutions, but two fundamental questions remain unsolved: first, the mapping from the substitutive dynamical system onto the Rauzy fractal which conjugates the shift map and the domain exchange is not explicit, and the same for the mapping onto the torus. Secondly, there is no result about the injectivity of the mapping onto the torus. Both questions were solved by G. Rauzy in [350] for the Tribonacci substitution. He gave in [353] ideas to generalize his work for any Pisot type substitution. This approach has been partially realized in [101] and [397], to obtain explicit representation mappings and conditions for injectivity of the toral mapping.

Prefix-suffix expansion. Indeed, in [102], the prefix-suffix expansion defined in Sec. 7.2.2 is used to produce a measure-theoretic isomorphism between the shift map on a substitutive system and an adding machine over the shift of finite type, considered as a Markov compactum in which the order is naturally defined by the substitution. This map is expressed under the form $\Gamma(w) = (p_i, a_i, s_i)_{i\geq 0}$ for all w in X_σ, so that we have:

$$w = \dots \sigma^i(p_i)\sigma^{i-1}(p_{i-1}) \dots \sigma(p_1)p_0 \cdot a_0 s_0 \sigma(s_1) \dots \sigma^i(s_i) \dots .$$

We already explained that this provides a new proof of A. V. Vershik's results, originally proved thanks to a stationary Markov compactum with a partial ordering on the labels of the edges. This ordering was arbitrary, so that the representation by the Markov compactum was not explicit. On the contrary, the definition of labels through prefixes of images of letters under the substitution induces a natural ordering between the edges and provides a well determined combinatorial representation.

Geometric realization of the substitutive system. If the substitution is of Pisot type over a d-letter alphabet, then this prefix-suffix expansion allows one to realize explicitly any element of X_σ as a point in $\mathbb{R}^{d-1}$ expressed in terms of power series. More precisely, a consequence of (7.5.1) is the following, where $\delta_2, \dots, \delta_{r+s}$ denote the coordinates of the map δ:

$$\psi_\sigma(w) = \lim_{n\to+\infty} \delta(\,\sigma^n(p_n)\ldots\sigma^0(p_0)\,) = \begin{pmatrix} \sum_{i\geq 0}\delta_2(p_i)\,{\alpha_2}^i \\ \vdots \\ \sum_{i\geq 0}\delta_{r+s}(p_i)\,{\alpha_{r+s}}^i \end{pmatrix} \in \mathcal{R}.$$

On each coordinate, this mapping coincides with the real or complex representations of C. Holton and L. Q. Zamboni in [208, 210]. Gathering all the one-dimensional realizations together is natural: indeed each of them is a different expression of the same formal numeration system (the one generated by the characteristic polynomial of the incidence matrix of the substitution).

It is proved in [103] that this representation map is the reverse map of the coding map, up to the cylinders partition, from the Rauzy fractal onto the substitutive dynamical system (see Theorem 7.5.12). In particular, this map projects onto the $(d-1)$-dimensional torus and satisfies a commutation relation with the translation described in the preceding section. This gives a new proof of Corollary 7.3.33 and 7.5.13 which makes the toral representation totally explicit.

This also gives a new proof of Theorem 7.5.13. The interest of this method lies in the fact the Rauzy fractal is obtained as the image of an explicit map, so that it makes it possible to derive various topological properties of the fractal, such as a criterion for the connectedness of the fractal [101], or a description of the boundary of the Rauzy fractal [398].

Application to the tiling question. The main interest of such a construction is that the explicit expression of the mapping from the substitutive dynamical system onto the Rauzy fractal on one side, and onto the torus one the other side, gives an efficient approach for a detailed study of the injectivity of the realizations. Indeed, in [398], the question of injectivity is not treated as usual from a measurable point of view [216, 43], but from a formal and algebraic point of view, based on two points. The first one is the combinatorial structure of the substitutive system, deduced from the prefix-suffix automaton. The second one is the algebraic structure of the system, produced by the incidence matrix of the substitution.

More precisely, one can note that the representation map ψ_σ onto the Rauzy fractal of a Pisot type substitution shall be expressed with a formal power series with coefficients into the splitting field of the characteristic polynomial of the incidence matrix of the substitution: the representation map is just obtained by gathering the set of finite values which can be taken for any Archimedean topology by the formal power series.

In that way, the study of the injectivity of the representation map is reduced to the study of sequences of digits such that the associated formal power series tends to zero for all the Archimedean metrics for which the power series has a limit. This implies that the sequences of digits are labels of paths in a finite automaton. The understanding of the structure of the automaton is fundamentally connected with the injectivity of the geometric (onto the Rauzy fractal) and toral realizations.

For instance, the injectivity in measure of the geometric representation of a unimodular Pisot type substitutive system can be characterized in terms of explicit combinatorial conditions on graphs. Contrary to the coincidence condition, it is a necessary and sufficient condition, but it is quite difficult to verify this condition quickly. Furthermore, this condition is satisfied by substitutions with coincidence, but we do not know whether equivalence holds.

The main interest of this approach is that, contrary to what happens with the coincidence condition, these methods also give results about the injectivity of the toral realization.

Theorem 7.5.19 (Siegel [398]). *The injectivity in measure of the toral representation of a unimodular Pisot type substitutive dynamical system deduced from the Rauzy fractal construction is equivalent to an explicit combinatorial condition which depends only on the definition of the substitution which generates the system.*

Thus one can test whether the spectrum of a given unimodular Pisot type substitutive dynamical system is purely discrete.

Corollary 7.5.20 (Siegel [398]). *There exists an explicit and effective sufficient combinatorial condition for a unimodular Pisot type substitutive dynamical system to be measure-theoretically isomorphic to a toral translation, that is, to have a purely discrete spectrum. This condition is necessary when the substitution has no non-trivial coboundary.*

For instance, the systems generated by the substitutions $1 \mapsto 12$, $2 \mapsto 31$, $3 \mapsto 1$, and $1 \mapsto 12$, $2 \mapsto 13$, $3 \mapsto 132$ have a purely discrete spectrum: both are measure-theoretically isomorphic to a translation on the two-dimensional torus, both conjugacy maps being explicit semi-topological conjugacies.

In more geometrical terms, the injectivity of the toral realization has the following interpretation in terms of tilings:

Corollary 7.5.21 (Siegel [398]). *There exists an explicit and effective combinatorial condition for the Rauzy fractal associated with a unimodular Pisot type substitution over d letters to generate a periodic tiling of $\mathbb{R}^{d-1}$.*

The condition is too long to be given here. The difficulty in checking whether this condition holds prevents one from obtaining general results about the injectivity of the toral realization. However, even if the condition is too long to be checked by hand, it can be checked in a finite time for any explicit example, thanks to the effective algorithm given in [398]. For instance, this algorithm allows one to state that the substitution $1 \mapsto 21$, $2 \mapsto 13$, $3 \mapsto 1$ generates a periodic tiling shown in Fig. 7.10. Notice that all the substitutions that have been tested give a positive answer to the tiling question.

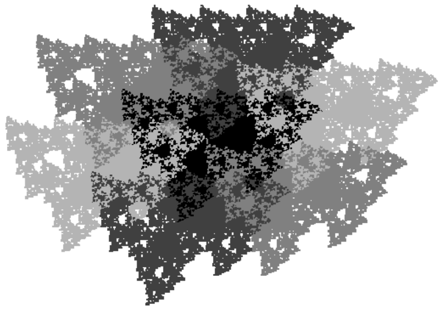


Fig. 7.10. Periodic tiling for the Rauzy fractal of $1 \mapsto 21$, $2 \mapsto 13$, $3 \mapsto 1$.

About unimodular substitutions of Pisot type, an important remaining question at the moment is to obtain a general sufficient condition, for instance in terms of coincidences, for a unimodular Pisot type substitutive system to have purely discrete spectrum, or equivalently, to be measure-theoretically isomorphic to a toral rotation.

7.6 Extensions and applications

7.6.1 Non-unimodular substitutions

The formal point of view of substitutive dynamical systems can be extended to non-unimodular substitutions. Indeed, in the unimodular case, the Rauzy fractal of a substitution of Pisot type over d letters is made of the set of finite values which can be taken for any Archimedean topology by the formal power series. But, in the non-unimodular case, there exists a new type of topology for which the formal power series converges: the p-adic topology. One can define a new Rauzy fractal, with p-adic components, which consists of the set of finite values which can be taken for any topology by the formal power series (Archimedean and p-adic). Then, one obtains a set in the product of $\mathbb{R}^{d-1}$ and finite extensions of p-adic fields $\mathbb{Q}_p$, where p are prime divisors of the determinant of the incidence matrix of the substitution [396].

The techniques developed in the unimodular case can be generalized:

- there exists a combinatorial condition on the substitution (in terms of graphs) which is sufficient for the substitutive system to code the action of a domain exchange in the new Rauzy fractal;
- there exists a lattice such that the quotient map modulo this lattice transposes the domain exchange into a translation over a compact group;
- there exists an automaton which decides whether the quotient map is almost everywhere one-to-one.

Thus, we get the following result:

Theorem 7.6.1 (Siegel [398]). *There exists an explicit and effective combinatorial condition for a Pisot type substitutive dynamical system to be measure-theoretically isomorphic to a toral translation, that is, to have a purely discrete spectrum.*

Example 7.6.2. The system generated by the substitution $1 \mapsto 1112$ et $2 \mapsto 12$, which is known to have a purely discrete spectrum (Theorem 7.3.35), is semi-topologically conjugate to a translation over the direct product of $\mathbf{Z}_2$ with two-adic solenoid $\mathcal{S}_2$(see [396, 398]).

The system generated by the substitution $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 11233$ has a purely discrete spectrum.

Many questions remain for non-unimodular Pisot substitutions: what is in general the structure of the maximal equicontinuous factor? Even in the two-dimensional case, where a condition for a purely discrete spectrum is known, the structure is difficult to understand. As a final question, we know that the substitution defined by $1 \mapsto 11222$ and $2 \mapsto 1222$ is isomorphic to a compact group translation. We know that this group contains a one-dimensional torus, and no p-adic factors. But what is precisely this group?

7.6.2 Substitutions and Markov partitions

The prefix-suffix expansion and the Rauzy fractal associated with a unimodular substitution of Pisot type over a d-letter alphabet, naturally lead to define a domain in $\mathbb{R}^d$.

This domain consists of finite unions of cylinders. The base of each cylinder is a sub-domain of the Rauzy fractal of the substitution embedded in the contracting hyperplane of the incidence matrix. Each cylinder is directed by the expanding direction of the matrix. Examples are given in Fig. 7.11 for the Fibonacci substitution and the Tribonacci substitution.

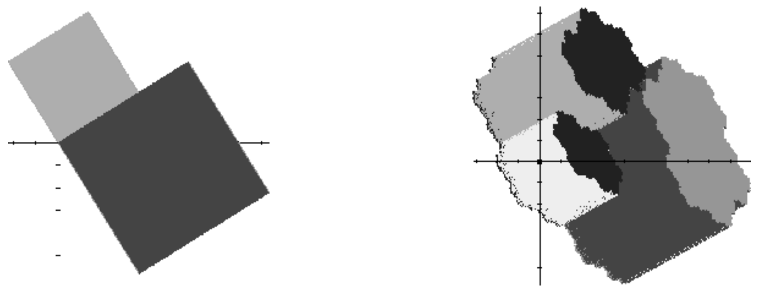


Fig. 7.11. Domains in $\mathbb{R}^d$ associated with the Fibonacci and Tribonacci substitutions.

If the Rauzy fractal of the substitution generates a tiling of $\mathbb{R}^{d-1}$ (see Corollary 7.5.21), then the domain is a fundamental domain identified with the d-dimensional torus $\mathbb{T}^d$ [397]. Illustrations are shown in Fig. 7.12 for the Fibonacci and Tribonacci substitutions.

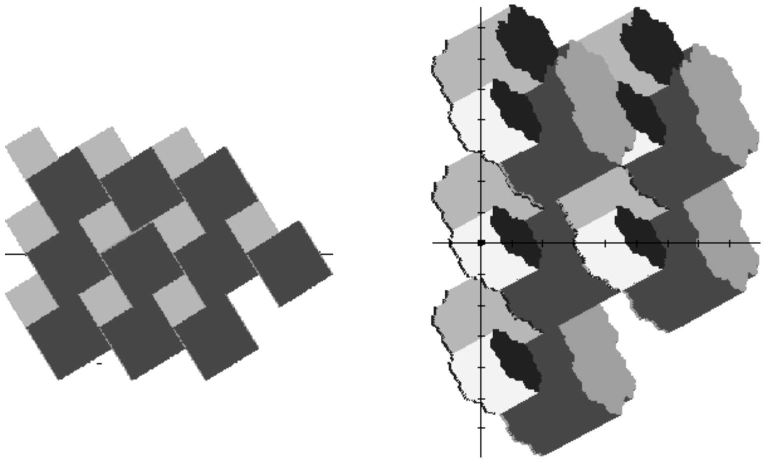


Fig. 7.12. Tilings of $\mathbb{R}^2$ and $\mathbb{R}^3$ associated with the Fibonacci and Tribonacci substitutions.

Moreover, the prefix-suffix automaton induces a natural partition of this fundamental domain such that the action of the toral automorphism associated with the incidence matrix of the substitution is coded, with respect to this partition, by the set of biinfinite walks of the prefix-suffix automaton.

Thus, the Rauzy fractal of a unimodular substitution of Pisot type together with the prefix-suffix automaton produce a completely explicit Markov partition for the toral automorphism associated with the incidence matrix of the substitution, as soon as the Rauzy fractal generates a periodic tiling.

This produces a new method to build Markov partition, with the use of substitutions. The advantage of this constructive method, compared to those of R. Kenyon and A. Vershik [245] or S. Leborgne [261], is to permit to characterize some topological properties of the Markov partition (boundary parametrization, connectedness, simple connectedness).

Given a toral automorphism of Pisot type with nonnegative coefficients, it is natural to consider the substitution which has the same incidence matrix, and such that the image of each letter through the substitution consists of letters in an increasing order. Such a substitution satisfies the coincidence condition, so that there exists a Rauzy fractal associated with the substitution. It remains to check that the Rauzy fractal generates a periodic tiling to finally obtain an explicit Markov partition of the toral automorphism considered at the beginning.