3. Substitutions and symbolic dynamical systems

The aim of this chapter is to introduce the fundamental notions of ergodic theory, through the study of a few examples of symbolic dynamical systems.

In the course of this chapter, we shall use at some points notions of measure theory and spectral theory; when there is no explanation in the text, the reader is referred to any standard book, such as [103] and [285]; however, we have tried to isolate these places, and the reader can skip them without damage to his general understanding.

3.1 The Morse sequence

3.1.1 A geometrical problem

In 1920, M. Morse was studying *geodesics*, that is, the curves realizing the minimum distance between two points, on simply connected surfaces with constant negative curvature; he was looking at infinite geodesics which stay in a small part of the space; more precisely:

Definition 3.1.1. *A geodesic G is* recurrent *if for every $\varepsilon > 0$ there exists L such that for every segment S in G of length bigger than L, every point in G is at a distance smaller than ε from some point in S.*

Of course, a *closed* geodesic is a recurrent geodesic; but are there non-closed recurrent geodesics?

To answer this question, in [254], Morse did a *coding* of geodesics, by infinite sequences of 0 and 1, according to which boundary of the surface they meet: thus, we arrive in the space $\{0,1\}^{\mathbb{N}}$ of symbolic sequences, and to advance along a geodesic translates into looking at the next element of the sequence. The coding sends the topology of the surface onto the product topology in $\{0,1\}^{\mathbb{N}}$.

In these conditions, a closed geodesic corresponds to a nonperiodic sequence. In the same way, by replacing points by elementary segments, the reader shall be able to check that a recurrent geodesic corresponds to what is now called a *minimal* sequence (also called *uniformly recurrent*):

Definition 3.1.2. *A sequence* $u = (u_n)_{n\in\mathbb{N}}$ *is minimal (or uniformly recurrent) if for every* $i < j$*, there exists* s *such that for every* n*, there exists* $n < m < n + s$ *such that* $u_m = u_i$*, ...,* $u_{m-i+j} = u_j$*.*

Or else, this definition can be read by using language terminology:

Definition 3.1.3. *The sequence* u *is minimal if every word occurring in* u *occurs in an infinite number of places with bounded gaps.*

Hence what we need is a minimal non-periodic sequence, and the example Morse gave is the sequence that bears his name, though it had already been discovered by Prouhet in 1851 [284] and Thue [353] in 1907: we consider the map σ from $\{0,1\}$ to $\{0,1\}^\star$ defined by $\sigma(0) = 01$ and $\sigma(1) = 10$. We extend σ into a *morphism* of $\{0,1\}^\star$ for the concatenation by $\sigma(abcd...) = \sigma(a)\sigma(b)\sigma(c)\sigma(d)\ldots$ We can then iterate σ: $\sigma^2(0) = \sigma(0)\sigma(1) = 0110$, $\sigma^n(0) = \sigma^{n-1}(0)\sigma^{n-1}(1)$, and the nested words $\sigma^n(0)$ converge (in $\{0,1\}^{\mathbb{N}} \cup \{0,1\}^\star$) to the only infinite sequence which begins by $\sigma^n(0)$ for every n, the *Morse sequence* (also called the *Prouhet-Thue-Morse sequence*): $u = u_0u_1... = 0110100110010110...$ This sequence satisfies the fundamental equality $u = \sigma(u)$.

Lemma 3.1.1. *Let* u *be the Morse sequence; then* $u_{2n} = 0$ *and* $u_{2n+1} = 1$ *if* $u_n = 0$*,* $u_{2n} = 1$ *and* $u_{2n+1} = 0$ *if* $u_n = 1$*; and at place* $k2^n$ *of the sequence occurs always an* n*-word,* $\sigma^n(0)$ *if* $u_k = 0$*, and* $\sigma^n(1)$ *if* $u_k = 1$*.*

The proof is immediate and left as an exercise.

Proposition 3.1.1. *The Morse sequence* u *is minimal and neither periodic nor ultimately periodic.*

Proof. Any word W occurring in u must occur at a place between 0 and some 2^m, hence must occur in some $\sigma^m(0)$; but $\sigma^m(0)$ occurs in every $m+1$-word, hence infinitely often in u with gaps bounded by 2^{m+1}, and so does W. Hence u is minimal.

As always $u_{2n+1} \neq u_n$, $n + 1$ cannot be a period; and for any p and n_0, there exists $n \geq n_0$ such that $n+1$ is a multiple of p, hence u is not ultimately periodic. ∎

3.1.2 Combinatorial properties

Let us consider some combinatorial properties of the Morse sequence.

Definition 3.1.4. *A sequence is called* square-free *if in it any word of the form* UU*, for any word* U*, does not occur.*

There cannot be any infinite square-free sequence on two letters; but, as was shown in [352], the Morse sequence is free of powers $2 + \varepsilon$ for any $\varepsilon > 0$:

Proposition 3.1.2. *In the Morse sequence no word of the form* UUv *occurs, where* U *is any word and* v *is the first letter of* U*.*

Proof. Let l be the length of U; suppose for example that l is odd, and that UUv occurs at an odd place. Then $U = v|a_1 b_1| \dots |a_p b_p|$ where the bars indicate the separations between the 2-words occurring at places $2k$ in u. And a word of the form $|a_0 v|a_1 b_1| \dots |a_p b_p|v a_1| \dots |b_p v|$ must occur in the Morse sequence, with the indicated position of the bars. Hence $a_0 + v = a_1 + b_1 = \dots = a_p + b_p = v + a_1 = b_1 + a_2 = \dots = b_p + v = 1$, hence $1 - v = a_0 = \dots = a_p = v$, which is impossible as $v = 0$ or 1. The same reasoning applies if l is odd and UUv appears in an even place.

If l is even and UUv occurs at an even place, then $U = |vb_1|a_2 b_2| \dots |a_p b_p|$, and if $U' = va_2..a_p$, $\sigma(U'U'v)$ occurs in the Morse sequence u. As $\sigma(u) = u$ and σ is injective on the letters, then $U'U'v$ must occur, v is the first letter of U', and the length of U' is strictly smaller than l. The same reasoning applies to the case where l is even and UUv occurs in an odd place. And finally, if U is of length one, we check immediately that 000 and 111 do not occur. ■

An immediate consequence of this result is that two occurrences of the same word in the Morse sequence have to be disjoint; another one is that the Morse sequence does indeed provide a square-free sequence on a three-letter alphabet:

Proposition 3.1.3. *If u is the Morse sequence, and v is the sequence over the alphabet $\{-1, 0, 1\}$ defined by: $\forall n$, then $\forall n \in \mathbb{N}$, $v_n = u_{n+1} - u_n$, v is an infinite square-free sequence.*

Proof. If the word $a_1 \dots a_k a_1 \dots a_k$ occurs at place i in v, then $a_1 + \dots + a_k = 0$, as otherwise $u_{2k+i} - u_i = 2$ or -2; but this implies that for some $e = 0$ or 1, the words $e(e+a_1)\dots(e+a_1+..+a_{k-1})e(e+a_1)\dots(e+a_1+..+a_{k-1})e$ occurs in u, which is impossible. ■

The same techniques allow us to prove the following lemma, due to Del Junco ([118]), which is a simple example of the notion of *recognizability*: for the Morse sequence, we know where a given word should occur, or even occur approximately.

Definition 3.1.5. *The* Hamming distance *between two words of equal length is*

$$\overline{d}((v_1 \dots v_n), (w_1 \dots w_n)) = \frac{1}{n} Card\ \{1 \le i \le n;\ v_i \ne w_i\}.$$

Proposition 3.1.4. *If U and V are two n-words, if W is a word occurring in u at place i, and such that*

$$\overline{d}(UV, W) < \frac{1}{8},$$

then $W = UV$ and i is a multiple of 2^n.

Proof. The result is empty for $n = 0$; we suppose it is true at stage $n-1$; if U, V, W, i are as above, $W = W_1 W_2$, with either $\overline{d}(U, W_1) < \frac{1}{8}$ or $\overline{d}(V, W_2) < \frac{1}{8}$; as U and V are each made with two $n-1$-words, the recursion hypothesis implies that there exists an integer k with $i = k2^{n-1}$; hence W_1 and W_2 must be $n-1$-words, and are at a distance $\overline{d} < \frac{1}{2}$ of the corresponding $n-1$-words in U and V; as the distance between two $n-1$-words can only be 0 and 1, we must have $W_1 = U$ and $W_2 = V$.

It remains to be proved that k is even; if it is odd, let $W = abb'c$, $U = xx'$, $V = yy'$ their expressions as concatenation of $n-1$-words, where, if $m = m_1 \dots m_q$ is a word, the word m' is made with the letters $(1-m_1)\dots(1-m_q)$; then we must have $a' = b = c$, and $a'abb'cc' = bb'bb'bb'$ must occur, which is impossible as 000 and 111 do not occur. ■

Recognizability will be used several times in the sequel (see Lemma 3.1.7, 3.1.9, 3.1.10 below), and it is an important notion for the general study of substitutions ([258]). For more details, on recognizability, see Chap. 7.

3.1.3 Complexity

Let us recall the definition of the complexity function. For more details, see Chap. 1.

Definition 3.1.6. *Let u be a sequence; we call* complexity function *of u, and denote by $p_u(n)$, the function which to each integer $n \geq 1$ associates Card $\mathcal{L}_n(u)$, that is the number of different words of length n occurring in u.*

Furthermore, let us recall that a n-word is either $\sigma^n(a)$ or $\sigma^n(b)$.

Lemma 3.1.2. *In the Morse sequence, every word of a length of at least four has a unique decomposition into 2-words, possibly beginning by the last letter of a 2-word and possibly ending by the first letter of a 2-word.*

Proof. In other words, there is only one way to put the bars; it is true if our word contains 00 or 11, as they cannot occur between bars, so there must be a bar in the middle; if not, it must contain 0101 or 1010, but this implies also a unique way of putting the bars, as 010101 and 101010 do not occur. ■

Proposition 3.1.5. *[78, 111] For the Morse sequence, $p_u(1) = 2$, $p_u(2) = 4$, and, for $n \geq 3$, if $n = 2^r + p + 1$, $0 < p \leq 2^r$, then $p(n) = 6.2^{r-1} + 4p$ if $0 < p \leq 2^{r-1}$ and $p(n) = 8.2^{r-1} + 2p$ if $2^{r-1} < p \leq 2^r$.*

Proof. If $n \geq 4$, our lemma implies that $p_u(n) = p_0(n) + p_1(n)$, where $p_0(n)$ is the number of words of length n beginning just after a bar, and $p_1(n)$ just before. Let $n = 2k+1$ and W a word of length n; if W is in the first category, $W = |a_1 b_1| \dots |a_k b_k| a_{k+1}$, $a_i + b_i = 1$, and there are as many such words as words $a_1..a_{k+1}$, hence $p_0(2k+1) = p_u(k+1)$. In the same way

$p_1(2k+1) = p_u(k+1)$, $p_0(2k) = p_u(k)$, $p_1(2k) = p_u(k+1)$. The first values can be computed by hand. ∎

Note that $p_u(n)$ is smaller than $4n$ and that the differences $p_u(n+1) - p_u(n)$ take two values, 2 and 4.

3.1.4 The associated topological dynamical system

Recall that the topological dynamical system associated with a sequence u with values in the finite alphabet $\mathcal{A}$ is the system (X_u, T), where

- Ω is $A^{\mathbb{N}}$ equipped with the product topology of the discrete topology on each copy of A, or equivalently with the distance
$$d(x, y) = 2^{-\min\{n \in \mathbb{N}; x_n \neq y_n\}}.$$
- $T(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots)$,
- $X_u \subset \Omega$ is the closure of the set $(T^n u, n \in \mathbb{N})$.

We have seen in Chap. 1 (Lemma 1.3.2) that $x \in X_u$ if and only if there exists a sequence k_n such that $x_0 \ldots x_n = u_{k_n} \ldots u_{k_n+n}$ for every $n \geq 0$, which is also equivalent to $\mathcal{L}_n(x) \subset \mathcal{L}_n(u)$ for all n.

Proposition 3.1.6. *If u is minimal, $x \in X_u$ implies $u \in X_x$ and X_u is the set of all sequences having the same language as u.*

Proof. If u is minimal and $x \in X_u$, $u_0 \ldots u_p$ occurs in every long enough word occurring in u, hence in every possible $u_{k_n} \ldots u_{k_n+n}$ for n large enough, hence occurs in $x_0 \ldots x_n$; hence $u_0 \ldots u_p = x_{l_p} \ldots x_{l_p+p}$ and $u \in X_x$; and consequently $\mathcal{L}_n(u) = \mathcal{L}_n(x)$ for all n; the converse comes from Lemma 1.3.2. ∎

Lemma 3.1.3. *If u is minimal, the shift T is surjective onto X_u.*

Proof. If $x = x_0 x_1 \ldots$ is in X_u, then $x_0 \ldots x_n$ occurs in u, and not only in place 0; hence there exists $a \in \mathcal{A}$ such that $a x_0 \ldots x_n$ occurs in u; and there exists at least one a such that $a x_0 \ldots x_n$ occurs in u for infinitely many values of n; but that implies that $a x_0 x_1 \ldots$ is in X_u, and it is an antecedent of x. ∎

Remark. It is enough to suppose u recurrent in the above lemma.

Most sequences we shall consider have sub-linear complexity function $(\forall n \in \mathbb{N}, p(n) \leq Cn)$. In this case, even if the shift is not surjective, the following proposition shows that it can be made invertible up to a set which is at most countable.

Proposition 3.1.7. *Let u be a minimal sequence the complexity of which satisfies $p_u(n) \leq Cn$ for all n and some constant C; then there exists a finite set F such that, if D is the (at most countable, and T-invariant) set $\bigcup_{n \in \mathbb{Z}} T^n F$, T and all its iterates T^n, $n \geq 0$, are one-to-one from $X_u \setminus D$ to $X_u \setminus D$.*

Proof. We suppose first $p_u(n+1) - p_u(n) \leq C$ for all n; as every word w of length n has at least one *left extension* (that is a word aw occurring in u for some $a \in \mathcal{A}$), there can be no more than C words of length n which have two or more left extensions. Let F be the set of points in X_u such that $T^{-1}x$ has at least two elements; if the sequence $x = (x_n)_{n\in\mathbb{N}} \in F$, then $\exists a \neq b$ such that the sequences $ax_0x_1 \cdots \in X_u$ and $bx_0x_1 \cdots \in X_u$, and hence the word $x_0 \ldots x_n$ has at least two left extensions for every n; so F has at most C elements.
This is still true if $p_u(n) \leq Cn$ as then $p_u(n+1) - p_u(n) \leq C$ on a subsequence tending to infinity. ■

In the sequel, we may use the map T^{-1}, tacitly assuming that we use it only on $X_u \setminus D$.

Let us come back now to the notion of minimality. The minimality of a sequence is equivalent to that of the associated dynamical system:

Definition 3.1.7. *A topological dynamical system (X,T) is* minimal *if the only closed sets $E \subset X$ such that $T(E) \subset E$ are $\emptyset$ and E, or, equivalently, if for every point $x \in X$, the orbit of x, that is $\{T^n x, x \in \mathbb{N}\}$, is dense in X.*

Proposition 3.1.8. *The system (X_u,T) is minimal if and only if the sequence u is minimal.*

Proof. If u is minimal, we saw that for every $x \in X_u$, u, and hence its closed orbit, must be in X_x, the closed orbit of x, hence $X_x = X_u$.
If (X_u,T) is minimal; then, for any $x \in X_u$, the closed orbit X_x is equal to X_u, hence u is in X_x and $\mathcal{L}_n(u) \subset \mathcal{L}_n(x)$ for all n. Now, if W occurs in u, then W must occur in every $x \in X_u$, which implies $X_u = \cup_{n=0}^{+\infty} T^{-n}[W]$; hence, by compacity, $X_u = \cup_{i=1}^{p} T^{-n_i}[W]$ and hence for every $k \geq 0$, $T^k u \in \cup_{i=1}^{p} T^{-n_i}[W]$; this means that W occurs in u infinitely often with gaps bounded by $\max_i n_i$. ■

Definition 3.1.8. *Two topological dynamical systems (X,T) and (Y,S) are* topologically conjugate *if there exists a bicontinuous bijection ϕ from X to Y such that $\phi T = S\phi$.*

Lemma 3.1.4. *If u and v are sequences on finite alphabets and if (X_u,T) is topologically conjugate to (X_v,T), then the isomorphism ϕ must be a* finite code*: there exists a finite integer q such that $(\phi(x))_i$ depends only on $(x_i, \ldots x_{i+q})$.*

Proof. The idea of the proof is the following : ϕ associates with a sequence (x_n) a sequence (y_n); for $p \in \mathbb{Z}$, the map $\phi_p : X_u \to \mathcal{A}$, $(x_n)_{n\in\mathbb{N}} \mapsto y_p$ is continuous because of the product topology; by compacity, we deduce that $\phi_p^{-1}[a]$, which is open and closed, must be a finite union of cylinders; hence there exists a finite integer q such that $\phi_0(x)$ depends only on $x_0 \ldots x_q$, and hence also the other ϕ_p as ϕ commutes with the shift T. ■

Corollary 3.1.1. *Under the same hypothesis, if $p_u(n+1) - p_u(n)$ is bounded, so is $p_v(n+1) - p_v(n)$; and if $p_u(n) \leq Cn + C'$, then $p_v(n) \leq Cn + C''$.*

Proof. Then $p_v(n+1)-p_v(n) \leq p_u(n+q+1)-p_u(n-q'-1)$ where q' is the constant associated with ϕ_{-1}, and in the last case $p_v(n) \leq Cn+C'+C(q+1)$. ∎

3.1.5 Unique ergodicity

Let us recall that a *topological dynamical system* (X,T) is usually defined as a compact set X together with a continuous map T acting on the set X. A topological dynamical system (X,T) always has an *invariant probability measures* that is, a measure μ such that

$$\mu(X) = 1 \text{ and } \int f(Tx)d\mu(x) = \int f(x)d\mu(x)$$

for any borelian function f; we can take for μ any accumulation point for the vague convergence of the sequence of probability measures $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1}\delta_{T^{-k}x}$ for any point $x \in X$.

Definition 3.1.9. *A system is* uniquely ergodic *if it has only one invariant probability measure.*

Lemma 3.1.5. *Let σ be the Morse substitution, and $N(W,v)$ be the number of occurrences of the word W in the word v; then, for any factor W, when $n \to +\infty$, $\frac{N(W,\sigma^n(e))}{2^n}$ tends to a limit f_W, independent of $e = 0$ or 1.*

Proof. As $\sigma^{n+1}(0) = \sigma^n(0)\sigma^n(1)$ and $\sigma^{n+1}(1) = \sigma^n(1)\sigma^n(0)$, we have

$$N(W,\sigma^{n+1}(0)) + N(W,\sigma^{n+1}(1)) \geq 2N(W,\sigma^n(0)) + 2N(w,\sigma^n(1))$$

as we may have forgotten occurrences at the junction of two n-words. So the quantity

$$\frac{N(W,\sigma^n(0)) + N(W,\sigma^n(1))}{2^{n+1}},$$

which is smaller than 1, has a limit, denoted by f_W. But the quantity $2^{-n-1}N(W,\sigma^{n+1}(0))$ differs from the preceding one only because of the occurrences at the junction between n-words, and their contribution is smaller than $2^{-(n+1)}|W|$, so it has the same limit; and the same is true if we replace 0 by 1. ∎

Lemma 3.1.6. *Under the same hypotheses $\frac{N(W,u_k\ldots u_{k+n})}{n+1} \to f_W$, uniformly in k.*

Proof. Let $V = u_k \ldots u_{k+n}$. We write $V = A\sigma^p(u_j \ldots u_{j+l-1})B$, for some p much smaller than n, with $|A| < 2^p$ and $|B| < 2^p$. We have $n+1 = |A| + |B| + l2^p$, and, because of the previous result, $|N(W,\sigma^p(u_i)) - 2^p f_W| < \varepsilon 2^p$ for n and p large enough. And

$$N(W,V) \le N(W,A) + N(W,B) + \sum_{i=j}^{j+l-1} N(W, \sigma^p(u_i)) + l|W|,$$

the last term, which is necessary because of occurrences at the junctions between p-words, satisfying $l|W| \le \frac{n+1}{2^p}|W|$, while

$$N(W,V) \ge \sum_{i=j}^{j+l-1} N(W, \sigma^p(u_i)),$$

which yields the conclusion by dividing the inequalities by $n+1$ and taking first p then n large enough; note that p and l can be chosen independently of k, hence the uniform convergence. ■

Proposition 3.1.9. *For every sequence u satisfying the conclusion of last proposition, the system (X_u, T) is uniquely ergodic.*

Proof. For every word W, we define the measure of the associated cylinder by $\mu([W]) = \mu(T^{-n}[W]) = f_w$, for every $n \in \mathbb{N}$. This measure extends to all borelian sets and is a T-invariant probability measure. The conclusion of last proposition says that when $N \to +\infty$

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_{[W]}(T^{n+j}u) \to \mu(W)$$

uniformly in j for every cylinder $[W]$; hence

$$\frac{1}{N} \sum_{n=0}^{N-1} g(T^{n+j}u) \to \mu(g)$$

uniformly in j for every continuous function g; hence

$$\frac{1}{N} \sum_{n=0}^{N-1} g(T^{n+n_k}u) \to \mu(g)$$

uniformly in k for every continuous function g; hence

$$\frac{1}{N} \sum_{n=0}^{N-1} g(T^{n}x) \to \mu(g)$$

for every $x \in X_u$ and for every continuous function g.

But if there exists another T-invariant probability measure ν, then the last result and the dominated convergence theorem imply that $\mu(f) = \nu(f)$ for every continuous function, hence $\mu = \nu$. ■

Remark. The converse of this last proposition is true: if (X,T) is uniquely ergodic, μ its invariant probability measure, g a continuous function on X, then

$$\frac{1}{N}\sum_{n=0}^{N-1} g\circ T^n \to \mu(g) \quad \text{uniformly.}$$

An idea of the proof is the following: otherwise, there exist $\delta>0$, f continuous, a sequence $n_k\to+\infty$ and a sequence of points x_{n_k} such that

$$\left|\frac{1}{n_k}\sum_{n=0}^{n_k-1} g(T^n x)-\mu(g)\right| > \delta.$$

By compacity, there exists a subsequence m_k of n_k such that for every continuous g, $\lim_{k\to+\infty}\frac{1}{m_k}\sum_{n=0}^{m_k-1} g\circ T^n$ exists and defines a measure ν. The measure ν is then a T-invariant probability measure, hence $\nu=\mu$, which contradicts the assumption.

This result is not unrelated to the primordial theorem, known as Birkhoff's ergodic theorem, which will be stated in Sec. 3.4.3, but is in fact completely independent of it.

Corollary 3.1.2. *The measure-theoretic system* (X_u,T,μ) *is* ergodic*: every set* E *such that* $\mu(E\Delta TE)=0$ *has measure 0 or 1.*

Proof. Otherwise, the measure ν defined over the borelian sets by: $\nu(H)=\frac{\mu(H\cap E)}{\mu(E)}$ is another T-invariant probability. ■

Lemma 3.1.7. *If* W *is a word of length* l *occurring in the Morse sequence,* $\mu([w])\le\frac{4}{l}$.

Proof. Because of Proposition 3.1.4, $N(\sigma^n(e)\sigma^n(f),\sigma^p(0))\le 2^{-n}|\sigma^p(0)|$ if $p>n$ and hence $\mu([\sigma^n(e)\sigma^n(f)])\le 2^{-n}$ for any n, $e\in A$, $f\in A$. If $2^{n+1}<l\le 2^{n+2}$, then W must contain some $\sigma^n(e)\sigma^n(f)$, hence $\mu([W])\le\mu([\sigma^n(e)\sigma^n(f)])\le 2^{-n}\le\frac{4}{l}$. ■

Corollary 3.1.3. *The measure* μ *is a* non-atomic measure*: for any point* x, $\mu(\{x\})=0$.

Proof. For every point x, $\mu(\{x\})\le\mu([x_0\dots x_n])$. ■

Hence, because of Proposition 3.1.5, Proposition 3.1.7 and Corollary 3.1.3, the shift T is bijective on a subset of X_u of measure one.

3.1.6 Representation by Rokhlin stacks

Lemma 3.1.8. *The substitution* σ *defines a continuous map from* X_u *to* X_u.

The proof is left as an exercise.

Lemma 3.1.9. *Let $n \in \mathbb{N}$. We have*

$$X = \cup_{e=0,1} \cup_{k=0}^{2^n-1} T^k \sigma^n[e],$$

and this union is disjoint.

Proof. The set on the right is closed, so we have only to prove that $T^m u$ is in it, for every $m \in \mathbb{N}$. But if $m = 2^n a + b$, with $0 \le b < 2^n$, $T^m u = T^b T^{a2^n} u$ and $T^{a2^n} u = u_{a2^n} u_{a2^n+1} \cdots = \sigma^n(u_a u_{a+1} \dots) \in \sigma^n([u_a])$.

Suppose that $T^p \sigma^n[a] \cap T^q \sigma^n[b] \neq \emptyset$, with $q > p$; then $\sigma^n(x) = T^{q-p}\sigma^n(y)$ for a point x in $[a]$ and a point y in $[b]$; we have $y = \lim T^{m_i} u$ and $x = \lim T^{p_i} u$; hence for a fixed i large enough, we have $\sigma^n(u_{p_i} \dots u_{p_i+i}) = T^{q-p}(\sigma^n(u_{m_i} \dots u_{m_i+i})$. Hence, for i large enough, a word of the form $\sigma^n(u_q)\sigma^n(u_r)$ occurs in u at a position $j2^n + q - p$ and, because of the recognizability (Proposition 3.1.4), $q - p$ is a multiple of 2^n, and so must be 0; from which we deduce $a = b$. ■

Lemma 3.1.10. *We have*

$$\sigma^n[e] = [\sigma^n e \sigma^n 0] \cup [\sigma^n e \sigma^n 1]$$

for $e = 0$ or 1. For every $n \ge 0$,

$$\sigma^n[e] = \sigma^{n+1}[e] \cup T^{2^n}\sigma^{n+1}[e'],$$

where $e + e' = 1$. The measure of the cylinder $[0]$ is $\mu([0]) = 1/2$, and $\mu(\sigma^n[e']) = \mu(\sigma^n[e])$, if $e + e' = 1$.

Proof. If $x \in \sigma^n[e]$, we have $x = \sigma^n(x')$, with $x'_0 = e$, and so $x \in [\sigma^n(e)\sigma^n(x'_1)]$. Conversely, if $x \in [\sigma^n(e)\sigma^n(f)]$ and $x = \lim_{p\to+\infty} T_{k_p} u$, then for every large enough element M of the sequence k_p, $x_0 \dots x_{2^n-1} = \sigma^n(e)\sigma^n(f) = u_M \dots u_{M+2^n-1}$, and, because of Proposition 3.1.4, M is a multiple of 2^n. Hence $k_p = l_p 2^n$, and $x = \sigma^n(x')$, where $x' = \lim T^{l_p}(u)$, the sequence converging because $T^{k_p} u$ converges, and $x' \in [e]$.

The second assertion comes easily from the first, and the third one is immediate. ■

We may then build the sets $F_{n,e} = \sigma^n[e]$ and their images $T^k F_{n,e}$, $0 \le k \le 2^n - 1$ in the following way: conventionally, for fixed e and n, we denote the set $T^k F_{n,e}$ by dots drawn one above the other as k increases. At stage n, we have two *Rokhlin stacks* (each n-stack being $\cup_{k=0}^{2^n-1} T^k F_{n,e}$, $e = 0$ or 1), with *bases* $F_{n,0}$ and $F_{n,1}$, with *height* 2^n, whose *levels* $T^k F_{n,e}$ are disjoint sets of measure 2^{-n-1}, and T sends each level of each stack, except the top ones, onto the level immediately above; T is not explicit on the top levels. In the beginning, $F_{0,0}$ and $F_{0,1}$ are two disjoint sets of measure $1/2$; at stage n, we cut $F_{n,e}$ into two subsets of equal measure $F_{n+1,e}$ and $H_{n+1,e}$, and T becomes explicit on part of the levels where it was not yet so, as it sends $T^{2^n-1}F_{n+1,0}$ onto $H_{n+1,1}$ and $T^{2^n-1}F_{n+1,1}$ onto $H_{n+1,0}$. This defines the $n + 1$-stacks, which will have height 2^{n+1}.

Proposition 3.1.10. *There exists an at most countable set E such that if x and x' are elements of X/E and for every $n \geq 0$, x and x' are in the same stack and the same level $T^k F_{n,e}$, for $0 \leq k \leq 2^n - 1$, $e = 0, 1$, then $x = x'$.*

Proof. If x and x' are always in the same level of the same stack, they are in the same $T^{k_n}\sigma^n[e_n]$ for all n; and hence $x = x'$ between the indices 0 and $2^n - 1 - k_n$; this implies $x = x'$ if $2^n - 1 - k_n \to +\infty$. If not, there exists p such that $(T^{-m}x)_0 = (T^{-m}x')_0$ for $0 \leq m \leq 2^n - p$ for every n, hence for every positive m. Now, the Morse sequence has a sublinear complexity (Proposition 3.1.5), and the reasoning of Proposition 3.1.6, applied to right instead of left extensions, implies that there exists an at most countable set E on the complement of which $(T^m x)_0 = (T^m x')_0$ for every negative m implies $x = x'$. ■

Remark. As an exercise, the reader can check that E is made with two (non-closed, of course), orbits under T, the orbit of u and the orbit of the sequence $u' = (1 - u_n)_{n \in \mathbb{N}}$.

We say that the stacks of bases $F_{n,e}$ *generate* the system (X, T, μ) and that we have defined T by *cutting and stacking*, or else by *Rokhlin stacks*. We need two stacks at each stage, so T is said to be of *rank at most two*. In fact, it is shown in [118] that we cannot use only one stack at each stage: we say that T is of *rank two* . A survey of many examples and properties of finite rank systems can be found in [141].

Definition 3.1.10. *If Q is a partition of X in two sets $Q_i, i = 0, 1$; for every point $x \in X$, its (positive) Q-*name *is the sequence $Q(x)$ such that $Q(x)_n = i$ whenever $T^n x \in Q_i$, $n \geq 0$.*

For example, if $Q_i = F_{0,i}$, $i = 0$, 1, by definition $Q(x)_n = i$ whenever $x_n = i$ and $Q(x) = x$ for all x.

3.1.7 The dyadic rotation

The aim of the remainder of our study of the Morse system is to compare it with other "classical" systems, and particularly to geometric systems, living on manifolds, or, in the most elementary cases, on intervals.

First, we need to define new notions of isomorphism, as the topological conjugacy of Sec. 3.1.4 is too strong to be really useful.

Definition 3.1.11. *Two systems (X, T, μ) and (Z, R, ϱ) are* measure-theoretically isomorphic *if there exist $X_1 \subset X$, $Z_1 \subset Z$, and ϕ a (bimeasurable) bijection from X_1 to Z_1 such that $\mu(X_1) = \varrho(Z_1) = 1$, $\phi\mu = \varrho$ and $R\phi = \phi T$.*

This notion is in fact much weaker than the topological one, as the sets X_1 and Z_1 may be very wild. An intermediate notion, which requires both a topological and measure-theoretic structure, is the following:

Definition 3.1.12. *Two systems* (X, T, μ) *and* (Z, R, ϱ) *are* semi-topologically conjugate *if they are measure-theoretically conjugate, and, in the above definition, the complement of* X_1 *in* X *and the complement of* Z_1 *in* Z *are at most countable, and* ϕ *and* ϕ^{-1} *are continuous.*

In that case, if (X, T, μ) *is a symbolic system, we say that it is a* coding *of* (Z, R, ϱ).

The first classical systems we know are the rotations.

Definition 3.1.13. *We call* rotation *the dynamical system made with a compact group* G*, a translation* S *of* G *and the Haar measure* λ *on* G.

The most famous case is when G is the torus $\mathbb{T}^1$, and $Rx = x + \alpha$ modulo 1 , for an irrational α. We now introduce another important kind of rotation.

Let S_0 be the map from $Y_0 = [0, 1]$ to itself defined by

$$S_0(1 - \frac{1}{2^n} + x) = \frac{1}{2^{n+1}} + x, \quad \text{for} \quad 0 \le x < \frac{1}{2^{n+1}}, \quad n \in \mathbb{N};$$

under this form it is called the *Van der Corput map.* It preserves the Lebesgue measure.

We now send $[0, 1]$ onto $Y = \{0, 1\}^{\mathbb{N}}$ by the mapping $\chi(\omega) = (\omega_0\omega_1 \dots)$ whenever $\omega = \sum \omega_j 2^{-j-1}$ (there may be ambiguities at dyadic points, but still the coding is well-defined outside a countable, hence of measure zero, set). We see how the intervals of continuity of S_0 are coded, and notice that

$$S := \chi S_0 \chi^{-1}$$

sends points of the form $0abc\dots$ to $1abc\dots$, $10abc\dots$ to $01abc\dots$, $110abc\dots$ to $001abc\dots$, and hence S is the *addition in base two* of $100\dots0\dots$ with infinite carry. The Lebesgue measure is sent to the infinite product of measures $1/2(\delta_0 + \delta_1)$, denoted by λ (δ_0 and δ_1 denote here the Dirac measures).

The set Y, equipped with the addition in base two, is a compact group. It contains a copy of $\mathbb{N}$ by $\phi(1) = 10\dots0\dots, \phi(2) = \phi(1) + \phi(1), \dots$ and a copy of $\mathbb{Z}$ as we check that $10\dots0\dots + 11111\dots = 0000\dots$, hence we can define $\phi(-1) = 111\dots$. And we see that the integers are dense in Y equipped with the usual topology; we call $(Y, +)$ the *group of* 2*-adic integers.* Then, S is the rotation $x \mapsto x + 1$. It is immediate that (Y_0, S_0, λ) and (Y, S, λ) are semi-topologically conjugate, and we treat them as the same system: we call it the *dyadic rotation* or the *dyadic odometer.*

Proposition 3.1.11. *The system* (Y, S) *is uniquely ergodic, and hence ergodic.*

Proof. The *characters*, that is the homeomorphisms of Y to the torus $\mathbb{T}^1$, define for each measure ν on Y, a Fourier transform $\hat{\nu}(\gamma) = \int_{g\in Y} \gamma(g)d\mu$. If μ is invariant by S, $\hat{\mu}(\gamma) = \int_{g\in Y} \gamma(Sg)d\mu = \gamma(1)\hat{\mu}(\gamma)$, hence if $\hat{\mu}(\gamma) \neq 0$, $\gamma(1) = 1$, hence $\gamma(n) = 1$ for every integer, hence γ is identically 1. So $\hat{\mu}$ is zero on the characters which are not the identity, and one on the identity, hence μ is the Haar measure λ. ∎

For the measure-theoretic dynamical system (Y, S, λ), there is an underlying *spectral* structure: we consider the action on the space $\mathcal{L}^2(Y, \lambda)$ of the operator $U : \mathcal{L}^2(Y, \lambda) \to \mathcal{L}^2(Y, \lambda)$ defined by $Uf(x) = f(Sx)$.

Proposition 3.1.12. *The system* (Y, S) *has a* discrete spectrum, *which means that the vector space generated by its eigenfunction is dense in* $\mathcal{L}^2(Y, \lambda)$, *and its eigenvalues are all the* $e^{2i\pi\alpha}$ *for the dyadic rationals, i.e., the rationals of the form* $\alpha = p2^{-k}$, $p, k \in \mathbb{Z}$.

Proof. As Y is a compact group and λ is the Haar measure, we know that the characters generate a dense subspace of $\mathcal{L}^2(Y, \lambda)$. But if γ is a character, $\gamma(Sg) = \gamma(1)\gamma(g)$. Hence all the characters are eigenfunctions, we get thus all the eigenfunctions, and the eigenvalues are all the $\gamma(1)$.

Hence we have to find all the characters of Y. Such a character γ must be also a character of $\mathbb{Z}$, hence $\gamma(n) = e^{2\pi in\alpha}$ for an α in $[0, 1]$; if γ can be extended in a continuous way to Y, it will remain a character. But $\omega = \omega_0\omega_1 \cdots = \lim_{k\to+\infty} \omega_0 \ldots \omega_k 00 \ldots 0 \cdots = \lim_{k\to+\infty} n_k(\omega)$, where $n_k(\omega) = \sum_{j=0}^{k} \omega_j 2^j$. Hence we have to find the α such that $\forall\omega, e^{2\pi i\alpha n_k(\omega)}$ converges when $k \to +\infty$.

We write $\alpha = \sum_{k=0}^{+\infty} \alpha_k 2^{-k-1}$; if the α_i are ultimately equal to 1, the expansion is improper; otherwise, either they are ultimately equal to 0, or for infinitely many k, $\alpha_k = 1, \alpha_{k+1} = 0$. We choose an ω such that $\omega_k = 1$ for all these values of k. Then we have, for this ω, $\gamma(n_k(\omega)) = \gamma(n_{k-1}(\omega))e^{2\pi i\omega_k 2^k\alpha} = \gamma(n_{k-1}(\omega))e^{2\pi i 2^k\alpha}$, and $2^k\alpha = m + 1/2 + 1/8 + \ldots$, where m is an integer, and hence $2^k\alpha$ falls between $1/2$ et $3/4$, modulo 1 and the $\gamma(n_k(\omega))$ cannot converge.

Hence α must have a dyadic expansion consisting ultimately of zeros, hence it must be a dyadic rational number; conversely, it is clear that every dyadic rational number is an eigenvalue. ∎

Proposition 3.1.13. *The system* (Y, S) *is a rank one system.*

Proof. Let G_0 the interval $[0, 1[$, G_n the interval $[0, 2^{-n}[$. We see that S sends in an affine way S^iG_n onto $S^{i+1}G_n$ for $i = 0, 1, \ldots 2^n - 1$. We can write these intervals one above the other, and this makes a stack filling all the space; the action of S is to climb up one level in the n-stack and is not defined on the top level. The $n + 1$-stack is built by cutting the n-stack vertically into two equal halves and stacking the right half on the left half; this operation defines S on a greater part of the space, and allows us to know eventually S on the whole space. As the $T^k G_n$ are arbitrarily small intervals, outside a countable set (the set of all the endpoints of the $S^k G_n$), a point x is determined by the sequence $0 \leq k_n(x) \leq 2^n - 1$ such that $x \in S^{k_n(x)} G_n$, $n \in \mathbb{N}$. Hence S is generated by the stacks of bases G_n, one stack at each stage, and is of rank one. ■

Under that form, S is known as the *Von Neumann–Kakutani map.*

3.1.8 Geometric representation of the Morse system

In the Morse system (X, T) we introduce the following equivalence relation: $x \sim x'$ if either $x = x'$ or $x_n + x'_n = 1$ for every $n \in \mathbb{N}$. Each equivalence class has two elements, and we check that on the quotient space X', the shift and the measure define naturally a measure μ' and a shift T'; we say that (X', T', μ') is a *factor with fiber two* of (X, T, μ).

To come back from X' to X, we associate with $x' \in X'$ the two elements of the class x': we denote by $(x', 0)$ the one whose first coordinate is 0, by $(x', 1)$ the other one. So we have $X = X' \times \{0, 1\}$, as topological spaces too, and μ is the tensor product of μ' by $1/2(\delta_0 + \delta_1)$. And $T(x, e) = (T'x, z(e, x))$, where $z(0, x) + z(1, x) = 1$; hence we can write, in $\mathbb{Z}/2\mathbb{Z}$, that $Z(e, x) = e + \psi_0(x)$, where ψ_0 is a measurable map from X' to $\mathbb{Z}/2\mathbb{Z}$. We say that (X, T, μ) is a *skew product* of (X', T', μ') by $\mathbb{Z}/2\mathbb{Z}$ built with the map ϕ, or a *two-point extension* of (X', T', μ').

For fixed k and n, both levels $T^k F_{n,e}$ project into X' on the same level ${T'}^k F_n$; because of Proposition 3.1.11, the stacks F_n generate the system (X', T'), which is of rank one. And, from the construction of the stacks with bases $F_{n,e}$, we deduce that the stacks with base F_n are built in the same way as the stacks for (Y, S): the $n+1$-stack is built by cutting the n-stack vertically into two equal halves and stacking the right half on the left half, starting from $F_0 = X'$.

Proposition 3.1.14. *The Morse system is a coding of a two-point extension of the dyadic rotation.*

Proof. Let $(Y", S", \lambda)$ be the following system: $Y"$ is the interval $[0, 1[$, λ the Lebesgue measure; for $n \geq 0$, and $0 \leq k \leq 2^n - 1$ let $G_{n,0,k}$ be the left half (closed on the left, open on the right) of the interval $S^k G_n$ and $G_{n,1,k}$ the right half of $S^k G_n$. The $G_{n,e,k}$, $e = 0, 1$, $0 \leq k \leq 2^n - 1$, constitute for fixed n a partition of $Y"$ into sets of measure 2^{-n-1}. We define $S"$ by sending in an affine way each $G_{n,e,k}$ onto $G_{n,e,k+1}$ if $k \leq 2^n - 1$; $S"$ is not explicitly defined on the $G_{n,e,2^n-1}$. We check that these definitions are compatible, and that at stage n, we cut $G_{n,e,0}$ into two halves which are $G_{n+1,e,0}$ and $G_{n+1,1-e,2^n}$, and $S"$ becomes explicit on part of the levels where it was not yet so in the same way as for the stacks of the Morse system (Sec. 7), $G_{n,e,k}$ playing the role of $T^k F_{n,e}$ and $G_{n+1,1-e,2^n}$ the role of the auxiliary $H_{n+1,e}$.

$(Y", S")$ has rank at most two, and is generated by stacks whose levels are $G_{n,e,k} = S"^k G_{n,e,0}$: except on a countable set $E"$ (which is just the set of all endpoints of the $G_{n,e,k}$), if y and y' are in the same $G_{n,e(n),k(n)}$ for all n, then $y = y'$.

We define now a bijection ϕ between $X \setminus E$ (of Proposition 3.1.11) and $Y" \setminus E"$ by associating with a point x such that $x \in T^{k(n)} F_{e(n)}$ for all n the unique point y such that $y \in S"^{(k(n)} G_{n,e(n),0}$ for all n, and conversely from $Y"$ to X; that is possible as, the stacks being defined by the same recursion formulas, the possible sequences $k(n), e(n)$ are the same, and we check by a closer scrutiny of E and $E"$ that by this process all points in $X \setminus E$ are indeed sent to points in $Y" \setminus E"$, and conversely. The construction of T and definition of $S"$ ensure also that $\phi(Tx) = S"(\phi x)$; ϕ sends μ to λ as all the levels at each stage have the same measure; and ϕ is bicontinuous because all the $T^k F_{n,e} \cap X \setminus E$ and the $S"^k G_{n,e,0} \cap X" \setminus E"$ are open sets. Hence ϕ is a coding.

And, by construction, ϕ associates the factor of T generated by the stacks with base F_n with the factor of $S"$ generated by the stacks with base G_n, and provides a semi-topological conjugacy between the dyadic rotation (Y, S) and the factor (X', T') in (X, T). Hence we can carry the map ψ_0 to Y: $(Y", S", \lambda)$ is a skew product of (Y, S, λ) by $\mathbb{Z}/2\mathbb{Z}$ built with the map $\psi = \psi_0 \circ \phi^{-1}$, or a two-point extension of the dyadic rotation. ∎

Note that for a point $y \in Y" \setminus E"$, $\phi^{-1}(y)$ is simply its Q'-name, where $Q'_e = G_{0,e,0}$, $e = 0, 1$.

Thus we have a geometric model for the Morse system. Note that (X, T) and $(Y", S")$ cannot be topologically conjugate, as $S"$ is not continuous while T is.

Note also that for the dyadic rotation S, we do not know any simple coding, for example using one sequence of words corresponding to the sequence of stacks (partitions into levels give periodic sequences, while S is ergodic and hence aperiodic); it can be shown that a coding of S is given by the system associated with the *Toeplitz sequence*, the fixed point of the substitution $0 \to 11$, $1 \to 10$.

3.1.9 Spectral properties

We begin by recalling some facts about the *spectral study* of a dynamical system (X, T, μ), that is the study of the unitary operator $Uf = f \circ T$ over the space $\mathcal{L}^2(X)$.

Lemma 3.1.11. *The eigenvalues of U are of modulus one; T is ergodic if and only if the eigenvalue one is a simple eigenvalue for U; in that case, every eigenvalue is simple and every eigenfunction is of constant modulus.*

Proof. The eigenvalues have modulus one because U is unitary.

If E is an invariant set of nontrivial measure, 1_E is a nonconstant invariant function; if f is a nonconstant invariant function, we can cut its image into two disjoint sets, whose inverse images have a nontrivial measure and are invariant sets.

If f is an eigenfunction for the eigenvalue β, $|f|$ is an eigenfunction for the eigenvalue $|\beta| = 1$ and hence is a constant. If f_1 and f_2 are eigenfunctions for β, $|f_2|$ is a nonzero constant, and f_1/f_2 is an eigenfunction for 1 and hence a constant. ■

Definition 3.1.14. *Given a unitary operator U on a Hilbert space H, a* cyclic space *is the closure of the set $\{U^n(f), n \in \mathbb{Z}\}$ for an element f in H. The* spectral type *of f, or of the generated cyclic space, is the probability measure ϱ on the torus $\mathbb{T}^1$ defined by $\hat{\varrho}(n) = (U^n f, f) = \int_X U^n f.fd\mu$. We say that U has a spectrum of* multiplicity *at most k if H is the direct sum of k cyclic spaces; if $k = 1$, we say the spectrum is* simple.

A useful criterion to show that a space is cyclic is the following:

Lemma 3.1.12. *If L is a separable Hilbert space approximated by an increasing sequence of cyclic spaces, L itself is a cyclic space.*

Proof. We denote by $H(f)$ the cyclic space generated by the function f. Then we have $L = \lim(H(f_n))$, and we want to find g such that $L = H(g)$. It suffices to show that for a dense sequence of functions g_n in L, $\cap_{n,p}\{g; d(g_n, H(g)) < 1/p\}$ is nonempty. Because of Baire's theorem, it is enough to show that for given f and ε, $\{g; d(f, H(g)) < \varepsilon\}$ is dense. So we take a function k, which we may choose of norm one, as well as f. For some m, $d(f, H(f_m)) < \varepsilon$ and $d(k, H(f_m)) < \varepsilon$. Hence $d(k, P(U)f_m) < \varepsilon$ for some polynomial P; hence we can find a polynomial Q, nonzero on the unit circle, such that $d(k, Q(U)f_m) < 2\varepsilon$. Then $Q(U)$ is invertible and its inverse is approximated by polynomials in U, hence $g = Q(U)f_m$ satisfies $d(k, g) < 2\varepsilon$ and $d(f, H(g)) < 2\varepsilon$. ■

We have to admit the following result of spectral theory, as its proof would need too many prerequisites:

Lemma 3.1.13. *If H and H' are cyclic spaces whose spectral types ϱ and ϱ' are* mutually singular, *that is $\mathbb{T}^1 = E \cup E', \varrho(E) = \varrho'(E') = 1, \varrho'(E) = \varrho(E') = 0$, then $H \perp H'$ et $H + H'$ is a cyclic space, of spectral type $\varrho + \varrho'$. This lemma is still valid for a countable sum of subspaces.*

Corollary 3.1.4. *If U has a discrete spectrum and all its eigenvalues are simple, then U has simple spectrum and its spectral type is the sum over all the eigenvalues β of the Dirac masses δ_β.*

Proof. We write $\mathcal{L}^2(X)$ as the sum of the cyclic spaces generated by the eigenfunctions f_β, whose spectral types δ_β are mutually singular. ∎

Proposition 3.1.15. *[118] The Morse system has a nondiscrete simple spectrum; its eigenvalues are all the dyadic rational numbers.*

Proof. Let R be the map on X which associates with $(x_n)_{n\in\mathbb{N}}$ the point $(1-x_n)_{n\in\mathbb{N}}$. The map R sends each $T^jF_{i,n}$ onto $T^jF_{i',n}$ where $i+i'=1$, and commutes with T (R is called the *flip*). Let V be the operator on $\mathcal{L}^2(X')$ defined by $Vf = f\circ R$, H the space of functions f such that $Vf=f$, K the space of functions such that $Vf=-f$. The sets H and K are U-invariant. The action of U on H is also the action of the spectral operator associated with T' on $\mathcal{L}^2(X')$; as measure-theoretic isomorphism implies spectral isomorphism, it is isomorphic to the action of the spectral operator associated with S on $\mathcal{L}^2(Y)$, hence the restriction U_H has a discrete spectrum, and simple spectrum as S is ergodic; the eigenvalues are the dyadic rationals.

Let $f_n = 1_{F_{1,n}} - 1_{F_{0,n}}$, $K(f_n)$ the cyclic space it generates under U; these increase to K, hence K is cyclic.

Let g be an eigenfunction for U_K; then $g\in K, U_Kg=\beta g$, hence $g^2\in H, \beta^2g^2 = U_Hg^2$. But then β^2 is a dyadic rational, and so is β, but then U has two orthogonal eigenfunctions for the eigenvalue β, which contradicts ergodicity. Hence U_K has a continuous spectrum, and the spectral type of K is singular with every discrete measure.

But $\mathcal{L}^2(X) = H+K$, by writing $2f = (f+Vf)+(f-Vf)$, hence U has a simple spectrum; its eigenvalues and eigenvectors are those of U_K, and the eigenvectors are not dense. ∎

3.2 The Rudin-Shapiro sequence

3.2.1 Automata

For all the definitions concerning automata and automatic sequences, see Chap. 1.

Recall that if σ is a substitution of length q over the alphabet $\mathcal{A}$ and u a fixed point of σ, then u is recognized by a q-automaton (Proposition 1.5.1).

3.2.2 The Rudin-Shapiro substitution

The Rudin-Shapiro sequence was introduced independently in [304] and [326] for some estimations in harmonic analysis (see Chap. 2 and Sec. 3.2.3 below); it is generated by taking the fixed point u beginning by a of the substitution

$$a\to ab \qquad b\to ac \qquad c\to db \qquad d\to dc.$$

The associated automaton is:

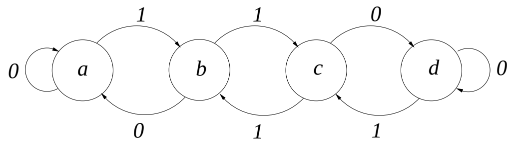

Fig. 3.1. Automaton associated with the Rudin-Shapiro substitution

Lemma 3.2.1. *We have $u_n = a$ or b if and only if there is an even number of words* 11 *in the expansion of n in base* 2.

The proof is left as an exercise (see also Chap. 2).

Let v be the sequence deduced from u by the letter-to-letter projection ϕ, which replaces a and b by 1, c and d by -1. The sequence v is then the limit in the usual sense of the words $\phi(\sigma^n(a))$. We check by induction that $\phi(\sigma^n(c)) = -\phi(\sigma^n b)$; hence if $A_n = \phi(\sigma^n(a))$ and $B_n = \phi(\sigma^n(b))$, the sequence v is the limit A_n given by the recursion formula $A_{n+1} = A_n B_n$, $B_{n+1} = A_n(-B_n)$. We remark that v is a q-automatic sequence (i.e., recognized by a q-automaton) though it is not a priori a fixed point of a substitution; such a sequence, deduced from a fixed point by identifying letters, is called a *substitutive sequence*.

Like the Morse sequence, u is minimal. Its complexity is $8n-8$ for every n large enough, and is easy to compute by the same method as for Morse. The complexity of v is also $8n-8$ ultimately, and X_u and X_v are topologically conjugate ([23]). The same methods as for Morse (but see Sec. 3.3.1 below for the general setting) allow us also to prove that the system is uniquely ergodic, with a non-atomic invariant probability measure μ.

A similar construction also shows that the associated shift transformation T has rank at most four, and is an (at most) four-point extension of the dyadic rotation: after the same kind of isomorphism as in Sec. 3.1.8, X is the product of Y by the set (a, b, c, d), and $T(x, e) = (Sx, \phi(x)e)$ for a map ϕ of Y into the group of permutations on four points. Note that we do not prove that this extension is non-trivial; (X, T) might be isomorphic to an extension of (Y, S) with less than four points, and actually is a two-point extension of the dyadic rotation (this is a particular case of a result in [220]).

3.2.3 Spectral properties

Though we do not want to give a complete description of the spectral structure, it has a remarkable property: the question whether the Lebesgue spectrum and finite multiplicity could co-exist stayed open for a long time, and this system solved it.

Lemma 3.2.2. *Let f be the function defined over X taking the value* 1 *whenever $x_0 = a$ or b,* 0 *otherwise; its spectral type is the vague limit of the* *measures with density* $1/(N+1)|f_N|^2$ *with respect to the Lebesgue measure* λ, *where* $f_N(x) = \sum_{n=0}^{N-1} z_n e^{2\pi inx}$, *with* $z_n = 1$ *whenever* $u_n = a$ *or* b, 0 *otherwise.*

Proof. We use a standard method to compute this spectral type ϱ_f. The measure ϱ_f is defined by $\hat{\varrho_f}(p) = (f, U^p f) = \mu(x; x_0 = a \text{ or } b, x_p = a \text{ or } b)$. By unique ergodicity, we read this measure on the sequence u as the sum of the frequencies of the words $a_0 \ldots a_p$, where $a_0 = a$ or b, $a_p = a$ or b; we deduce, remembering what is v, that it is exactly $\lim_{N\to+\infty} 1/(N+1)\sum_{n=0}^{N} 1_{v_n=v_{n+p}=1}$.

Let then f_N be the function defined on the torus $\mathbb{T}^1$ by $f_N(x) = \sum_{n=0}^{N-1} z_n\, e^{2\pi inx}$, where $z_n = (v_n+1)/2$; and let $d\varrho_N = 1/(N+1)|f_N|^2 d\lambda$. We check immediately that $\hat{\varrho}(p) = 1/(N+1)\sum_{n=0}^{N-p} z_n z_{n+p}$; but $z_n z_{n+p} = 1$ if $v_n = v_{n+p} = 1$, 0 otherwise. The convergence of Fourier coefficients is by definition equivalent to the vague convergence of measures, hence we have identified ϱ_f as the vague limit of the ϱ_N. ∎

Proposition 3.2.1. *The associated operator* U *has a spectrum with multiplicity at most four; all the dyadic rationals are eigenvalues; there exist functions the spectral type of which is equivalent to the Lebesgue measure.*

Proof. Because (X, T) is a four-point extension of a system with a simple spectrum, the Hilbert space $\mathcal{L}^2(X)$ is a limit of $H_n^1 + H_n^2 + H_n^3 + H_n^4$, where the H_n^i are (not necessarily nonempty or orthogonal) cyclic space, and a generalization of Lemma 3.1.12 shows that $\mathcal{L}^2(X)$ is generated by at most four cyclic spaces. One of them, as for the Morse system, contains the eigenvectors of the dyadic rotation, hence the dyadic rationals are eigenvalues.

We normalize f to get a function which is orthogonal to the constants: it is immediate that $\int f d\mu = 1/2$, hence we take $g = 2f - 1$. We deduce from the previous lemma that ϱ_g is the vague limit of $\varrho_n' = 2^{-n}|g_n|^2 d\lambda$, where $g_n(x) = \sum_{k=0}^{2^n-1} v_k e^{2\pi ikx}$. If $B_n = w_{n,0} \ldots w_{n,2^n-1}$, we take $h_n(x) = \sum_{k=0}^{2^n-1} w_(n,k) e^{2\pi ikx}$. From the structure of v, we deduce that $g_{n+1}(x) = g_n(x) + e^{2i\pi 2^n x} h_n(x)$ and $h_{n+1}(x) = g_n(x) - e^{2i\pi 2^n x} h_n(x)$; hence $|g_{n+1}^2(x)| + |h_{n+1}^2(x)| = 2(|g_n^2(x)| + |h_n^2(x)|)$, and, going back to $n = 0$, that this quantity is equal to 2^n for all x. So $2^{-n}|g_n|^2 \le 1$, for all n; by dominated convergence, the densities of the ϱ_n' must converge in $\mathcal{L}^2(\mathbb{T}^1, \lambda)$; hence ϱ_g has a density with respect to Lebesgue measure. Furthermore, if $2^{-n}|g_n|^2(x) \to 0$, then $2^{-n}|h_n|^2(x) \to 1$, which is incompatible with the recursion formula, so the density of ϱ_g is strictly positive, and ϱ_g is equivalent to Lebesgue measure. ∎

In particular, this implies that there exists a function f', of norm one, such that $(U^n f', f') = 0$ if $n \neq 0$; T is strongly mixing (see Definition 3.4.1 below), and even more, on some functions. The complete analysis ([286]) shows that the system is generated by two cyclic spaces with spectral types λ and $\lambda + \delta$, λ being the Lebesgue measure and δ the sum of the atomic measures on the dyadic rationals. By the way, let us recall (see Chap. 2) that the original Rudin-Shapiro sequence, v, was precisely introduced to minimize $\sup_{x\in[0,1]} |\sum_{n=0}^{N-1} \varepsilon_n e^{2\pi inx}|$, when ε_n is a sequence taking values $+1$ and -1. For the sequence v_n, there is a bound $(2+\sqrt{2})\sqrt{N}$, and v is the sequence which gives the lowest bound.

3.3 The Fibonacci sequence

3.3.1 Unique ergodicity of primitive substitutions

Let us recall some properties of primitive substitutions; see also Chap. 1.

Definition 3.3.1. *The substitution σ over the alphabet $\mathcal{A}$ is* primitive *if there exists k such that, for every a and b in $\mathcal{A}$, the letter a occurs in $\sigma^k(b)$.*

Let us recall the statement of Proposition 1.4.1: *if σ is primitive, any of its fixed points is a minimal sequence.*

Lemma 3.3.1. *With the notations of the last section, if σ is primitive, the $\frac{N(e,\sigma^n a)}{|\sigma^n(a)|}$ tends to a limit f_e when $n \to +\infty$, for every $a \in \mathcal{A}$ and every word e of one letter.*

Proof. Let $N(V)$ be the vector $(N(e,V), e \in \mathcal{A})$ for any word V, and let $\mathbf{M}_\sigma$ be the *matrix of the substitution* (also called *incidence matrix*) σ defined by $M = ((m_{i,j}))$, $i \in \mathcal{A}$, $j \in \mathcal{A}$, where $m_{i,j} = N(j, \sigma(i))$. For any V we have $N(\sigma(V)) = \mathbf{M}_\sigma.N(V)$.

Primitivity implies that all coefficients of the matrix $\mathbf{M}_\sigma^k$ are positive, and by *Perron-Frobenius theorem* (see Chap. 1 a nd [285] for example) the matrix $\mathbf{M}_\sigma$ has a real positive eigenvalue θ, which is bigger than 1 as $\mathbf{M}_\sigma$ has integer coefficients, simple, corresponding to a positive eigenvector g, and such that $\theta > |\lambda|$ for any other eigenvalue λ; we thus deduce that $\frac{\mathbf{M}_\sigma^n N(a)}{\theta^n}$ converges to a positive multiple g_a of g for every $a \in \mathcal{A}$.

Thus, $|\sigma^n(a)|$ is the scalar product of $N(\sigma^n(a))$ with the vector $h = (1,\ldots,1)$, hence $\frac{N(\sigma^n(a))}{|N(\sigma^n(a))|} \to \frac{g_a}{(g_a,h)}$, which is the multiple of g whose sum of coordinates equals 1, and hence is independent of a. ∎

Lemma 3.3.2. *If σ is primitive, $\frac{N(W,\sigma^n(a))}{|\sigma^n(a)|}$ tends to a limit f_w when $n \to +\infty$, for every $a \in \mathcal{A}$ and every word W.*

Proof. We write each $V = v_1 \ldots v_l$ as a one-letter word for a new primitive substitution ζ_l over the alphabet $\mathcal{A}_l$ whose letters are the words of length l occurring in u: if $W = w_1 \ldots w_l$, $\sigma(w) = y_1 \ldots y_m$ and $q = |\sigma(w_1)|$, we define $\zeta_l(w) = (y_1 \ldots y_l)(y_2 \ldots y_{l+1}) \ldots (y_q \ldots y_{q+l-1})$. We check that $U_l = (u_0 \ldots u_{l-1})(u_1 \ldots u_l) \ldots (u_n \ldots u_{n+l-1})$ is a fixed point for ζ_l.

The substitution ζ_l is primitive: if V and W are in $\mathcal{A}_l$, $\exists p; W$ is a subfactor of $\sigma^p(a)$ by definition of u, and $\exists m; a$ is a subfactor of $\sigma^m(v_1)$ by primitivity of σ, hence W is a letter of $\zeta_l^{m+p}(V)$. Then we check that $|\sigma^n(v_1)| = |\zeta_l^n(V)|$ and that $N(W, \zeta_l^n(V))$ is close to $N(W, \sigma^n(v_1))$ when n is large. Hence the previous lemma applies. ■

Proposition 3.3.1. *If u is a fixed point of a primitive substitution, the system (X_u, T) is uniquely ergodic.*

Proof. We can prove again Lemma 3.1.6 in the same way, except that $\sigma^p(u_i)$ is no longer equal to 2^p; but, by Perron–Frobenius, if p is large enough,

$$c\theta^p < \inf_{a \in \mathcal{A}} |\sigma^p(a)| < \sup_{a \in \mathcal{A}} |\sigma^p(a)| < d\theta^p,$$

which is enough to yield the same conclusion. Then we apply Proposition 3.1.9 ■

Proposition 3.3.2. *If u is a non-periodic fixed point of a primitive substitution, the unique invariant probability measure of the system (X_u, T) is non-atomic.*

Proof. If $\mu(\{x\}) = \alpha > 0$, then $\mu([W_n]) \geq \alpha$ for all n, where $W_n = x_0 \dots x_n$. Hence, because of the explicit expression of $\mu([W_n])$, for every n there must exist $j_n > i_n$ such that $j_n - i_n < \frac{1}{\alpha}$ and W_n occurs in u at places i_n and j_n; if $n > \frac{2}{\alpha}$, this implies that $x_0 \dots x_{[\frac{n}{2}]} = x_{j_n - i_n} \dots x_{[\frac{n}{2}] + j_n - i_n}$; as $j_n - i_n$ takes only a finite number of values, there exists k such that $x_0 \dots x_{[\frac{n}{2}]} = x_k \dots x_{[\frac{n}{2}]+k}$ for infinitely many n, hence for every n. Hence x is periodic, and so is u which has the same language. ■

The following result is proved in [269], using the results in [133], or [101] in the particular case of constant length:

Proposition 3.3.3. *If σ is primitive, its fixed points have a sub-linear complexity.*

Proof. For every n, we find p such that

$$\inf_{a \in \mathcal{A}} |\sigma^{p-1}(a)| \leq n \leq \inf_{a \in \mathcal{A}} |\sigma^p(a)|.$$

Every word of length n has to be included either in a $\sigma^p(a)$ or in a $\sigma^p(ab)$; for fixed a and b, there are at most $|\sigma^p(ab)|$ such words, according to the position of the first letter; and there are at most $K = (\text{ Card } \mathcal{A})^2$ possible choices for a and b. Hence for n large enough

$$p_u(n) < 2Kd\theta^p < 2K\frac{d}{c}\theta n$$

with c and d as in the previous proposition. ■

3.3.2 A trajectory of a rotation

We are interested here in a problem which has, at first sight, no connection with substitutions: given an irrational rotation of the torus $\mathbb{T}^1$, of angle α, and a particular partition (corresponding to the point of discontinuity of the rotation when it is considered as acting on the fundamental domain $[0,1[$), we want to find explicitly the name of a point x under that partition. This problem is solved for any α and x in [211], using formal power series, and in [35], using the dynamical notion of the *Rauzy induction* ([291] and see also Chap. 6). We shall make here this computation for a particular value of α and x. About the general case, and its link with the continued fraction approximation of α, more information can be found in Chap. 6.

We will define first what is the induction:

Definition 3.3.2. *For a transformation T on X, a set $A \subset X$, and a point $x \in A$, we call* first return time *of x in A and denote by $n_A(x)$ the (possibly infinite) smallest integer $m > 0$ such that $T^m x \in A$. The* induced map *of T on A is the map $T^{n_A(x)}x$ defined on $A \cap \{n_A(x) < +\infty\}$.*

Each time we use these notions in this course, $n_A(x)$ will be finite and the induced map defined everywhere.

We look at the irrational rotation of angle $\theta = 1/2(\sqrt{5}-1)$, defined on the torus $\mathbb{T}^1$ by $\overline{R}x = x + \theta \bmod 1$, and, after an immediate semi-topological conjugacy, on the interval $[0,1[$ by

$$\begin{aligned} Rx &= x + \theta \quad \text{if} \quad x \in [0, 1-\theta[, \\ Rx &= x + \theta - 1 \quad \text{if} \quad x \in [1-\theta, 1[. \end{aligned}$$

Let P_0 be the set $[0, 1-\theta[$ and P_1 the set $[1-\theta, 1[$; let v be the P-name of the point θ under R (see Definition 3.1.10 above). We shall now prove the following result:

Proposition 3.3.4. *The sequence v is the image by $T\tau_0$ of the fixed point of the substitution τ, where T is the shift, $\tau_0(0) = 10, \tau_0(1) = 0$, $\tau(0) = 001, \tau(1) = 01$.*

Proof. We first write $v = Tv"$, where $v"_n = 0$ if $R^n 0 \in [1-\theta, 1[$, $v"_n = 1$ if $R^n 0 \in [0, 1-\theta[$, T being the shift. We make now an induction on a particular interval, which is the so-called Rauzy induction; let I be the interval $[0,\theta[$ and R' the induced map of R on I, $n(x)$ being the first return time of x in I; we compute easily that:
if $x \in [0, 1-\theta[$, $n(x) = 2$, $R'x = x + 2\theta - 1$,
if $x \in [1-\theta, \theta[$, $n(x) = 1$, $R'x = x + \theta - 1$.

Let v' be the sequence defined by $v'_n = 0$ if $R'^n 0 \in [0, 1-\theta[$, $v'_n = 1$ if $R'^n 0 \in [1-\theta, \theta[$.

We compare v' with v". If $n(x)$ is the first return time of x in I, $n_2(x) = n(x) + n(R'x)$ is the second return time, $n_i(x) = n_{i-1}(x) + n({R'}^{i-1}x)$ is the i-th return time of x in I. Suppose $v'_i = 0$; then ${R'}^i 0$ is in $[0, 1-\theta[$; hence $R^{n_i(0)}0$, which is the same point, is in $[0, 1-\theta[$, $R^{n_i(0)+1}0$ is in $[\theta, 1[\subset [1-\theta, 1[$, and $R^{n_i(0)+2}$ is again in $[0, \theta[$, hence must be ${R'}^{i+1}0$; hence we can write that $n_{i+1}(0) - n_i(0) = 2$, $v"_{n_i(0)} = 1$, and $v"_{n_i(0)+1} = 0$. Suppose on the contrary that $v'_i = 1$; then ${R'}^i 0$ is in $[1-\theta, \theta[$, hence $R^{n_i(0)}0$ is in $[1-\theta, \theta[\subset [1-\theta, 1[$ and $R^{n_i(0)+1}$ is still in $[0, \theta[$. Hence we can write that $n_{i+1}(0) - n_i(0) = 1$ and $v"_{n_i(0)} = 0$. The previous analysis allows us, from the knowledge of v', to determine v" between two consecutive $n_i(0)$: we have proved that $v" = \tau_0(v')$, where τ_0 is the substitution $\tau_0(0) = 10, \tau_0(1) = 0$. Hence the initial v is $T\tau_0(v')$.

We do not change v' if we make a homothety of ratio $1/\theta$; R' becomes a rotation of $2 - 1/\theta = 1 - \theta$ on the interval $[0, 1[$. And $v'_n = 1$ if ${R'}^n 0 \in [\theta, 1[$, $v'_n = 0$ if ${R'}^n 0 \in [0, \theta[$.

Let Q be the induced transformation of R' on $[0, \theta[$, $m(x)$ being the first return time; we have:
if $x \in [0, 2\theta - 1[$, $m(x) = 1$, $Qx = x + 1 - \theta$,
if $x \in [2\theta - 1, \theta[$, $n(x) = 2$, $Qx = x + 1 - 2\theta$.

Let w be the sequence defined by $w_n = 0$ if $Q^n 0 \in [0, 2\theta - 1[$, $w_n = 1$ if $Q^n 0 \in [2\theta - 1, \theta[$. By the same method as above, we check that $v' = \tau_1(w)$, where $\tau_1(0) = 0, \tau_1(1) = 01$.

We normalize again, dividing by θ. We see that Q becomes the rotation R itself, and that $w_n = 1$ if $R^n 0 \in [1-\theta, 1[$, $w_n = 0$ if $R^n 0 \in [0, 1-\theta[$.

By a new Rauzy induction, we come back to R' and v', and show that $w = \tau_2(v')$, where $\tau_2(0(= 01, \tau_2(1) = 1$.

Hence $v' = \tau(v')$, where $\tau(0) = 001, \tau(1) = 01$; and $v = T\tau_0(v')$, which proves our proposition. ■

3.3.3 The Fibonacci substitution: geometric representation

In this section, σ is the *Fibonacci substitution* $\sigma(0) = 01$, $\sigma(1) = 0$. Its unique fixed point u is the *Fibonacci sequence* $0100101\ldots$ Note that the length of $\sigma^n(0)$ is the n-th *Fibonacci number* f_n, given by the recursion formulas $f_0 = 1$, $f_1 = 2$, $f_{n+1} = f_n + f_{n-1}$.

The substitution σ is primitive, hence the system (X_u, t) is uniquely ergodic. The dominant eigenvalue of the matrix is the *golden ratio number* $\theta_0 = \frac{1+\sqrt{5}}{2}$. Let $\theta = \frac{\sqrt{5}-1}{2}$.

Let R and $\overline{R}$ be the irrational rotation of angle θ defined respectively on the interval $[0, 1[$ and on the torus $\mathbb{T}^1$ as in Sec. 3.3.2. The irrational rotation $\overline{R}$ is a translation of a compact group; hence, by the same proof as in Sec. 3.1.7, it is uniquely ergodic; as the invariant measure, which is the Lebesgue measure, gives a strictly positive measure to every open set, this implies by a standard argument ([285]) that $\overline{R}$ is minimal (which is also a consequence of Kronecker's theorem). We check that these properties are shared by R.

Proposition 3.3.5. *We have $u_n = 0$ whenever $R^n\theta \in [1-\theta, 1[$, $u_n = 1$ whenever $R^n\theta \in [0, 1-\theta[$.*

Proof. We just have to identify the sequence u with the sequence v in the previous proposition. But $\tau_0\tau(0) = 10100$, $\tau_0\tau(1) = 100$; we check that $\tau_0\tau^n(0)$ is made with a 1 followed by $\sigma^{2n+1}(0)$ (minus its last letter), and that $\tau_0\tau^n(1)$ is made by the f_{2n+1} last letters of $\tau_0\tau^n(1)$; this is achieved by using the reverse decomposition of u: $\sigma^n(0) = \sigma^{n-1}(0)\sigma^{n-2}(0) = \sigma^{n-2}(0)(\sigma^{n-3}(0)\sigma^{n-2}(0))$; hence v begins by $\sigma^{2n+1}(0)$, and $v = u$. ■

Corollary 3.3.1. *The complexity function of the Fibonacci sequence is*

$$p(n) = n + 1 \textit{ for every } n.$$

Proof. A word $w_0 \dots w_{n-1}$ occurs in u if and only if $\cap_{i=0}^{n-1} R^{-i}P_{w_i} \neq \emptyset$ (by minimality). The partition of the interval $[0, 1[$ by the sets $\cap_{i=0}^{n-1} R^{-i}P_{w_i}$ when w ranges over $\mathcal{L}_n(u)$, is the partition of the interval by the points $R^{-i}0, 1 \leq i \leq n$; hence there are $n+1$ nonempty intervals. ■

As an exercise, we propose a direct and combinatorial proof of this result.

Exercise 3.3.1. • Prove that every factor W of the Fibonacci sequence can be uniquely written as follows:

$$W = R\sigma(V)S,$$

where V is a factor of the Fibonacci sequence, $R \in \{\varepsilon, 1\}$, and $S = 0$, if the last letter of W is 0, and $R_2 = \varepsilon$, otherwise.

- Prove that if W is a left special factor distinct from the empty word, then there exists a unique left special factor V such that $W = \sigma(V)S$, where $S = 0$, if the last letter of W is 0, and $R_2 = \varepsilon$, otherwise. Deduce the general form of the left special factors.
- Prove that the Fibonacci word is not ultimately periodic.
- Prove that the complexity function of the Fibonacci sequence is $p(n) = n+1$ for every n.

We say that the Fibonacci sequence is a *Sturmian sequence*; it has the lowest possible complexity for a nonperiodic sequence.

Proposition 3.3.6. *The system (X_u, T, μ) associated with the Fibonacci sequence is a coding of the rotation R on the interval, or $\overline{R}$ on the torus, preserving the Lebesgue measure λ.*

Proof. Let P be the partition ($P_1 = [0, 1-\theta[, P_0 = [1-\theta, 1[$) and $P(x)$ be the P-name of x. We have $P(\theta) = u, P(R^n\theta) = T^n u$; as R is continuous, when $R^{n_k}\theta \to x$, we have $P(R^{n_k}\theta) \to P(x)$, for the product topology on $\{0,1\}^{\mathbb{N}}$, except if x is in the orbit of θ (because the atoms of P are semi-closed); and $P(x) = \lim T^{n_k}u$ is then a point of X_u; as the set $\{n\theta, n \in \mathbb{N}\}$ is dense in $[0,1[$, we see that $P([0,1[/D) \subset X$, where D is a countable set.

Conversely, two points with the same P-name are not separated by arbitrarily small intervals, hence are identical; and every point in X_u, written under the form $\lim T^{n_k}u$, may be written, after taking a subsequence such that $R^{n'_k}\theta$ converges, and deleting a countable number of points, as $P(\lim R^{n'_k}\theta)$; hence $P(x)$ is a bicontinuous bijection, except on a countable set, and $PR = TP$, while P sends the only invariant measure λ for R to the only invariant measure for T.

As for $\overline{R}$, it is semi-topologically conjugate to R. ∎

The systems (X_u, T), $([0,1[, R)$ and $(\mathbb{T}^1, \overline{R})$ are not mutually topologically conjugate: $\overline{R}$ and T are continuous while R is not; and between T and $\overline{R}$, the topology of X_u is generated by clopen sets while the topology of $\mathbb{T}^1$ is not.

It is easy to show that $\overline{R}$ has a discrete spectrum and that the eigenvalues are all the $e^{2i\pi n\alpha}, n \in \mathbb{Z}$. Also, $\overline{R}$ has rank one (see for example [141]), though the approximation by stacks is less simple than in the cases we have seen (the stacks do not fill the whole space at every stage). All these properties are of course shared by R and the Fibonacci system.

3.4 The Chacon sequence

3.4.1 Elementary properties

The *Chacon substitution*

$$\sigma(0) = 0012 \qquad \sigma(1) = 12 \qquad \sigma(2) = 012$$

defines a primitive substitution, hence the system (X_0, T) associated with the fixed point v beginning by 0 is uniquely ergodic.

This system is a new form ([139]) of an historical system, which we shall study in its traditional form: we remark that v begins by the word w_n, where $w_0 = 0$ and $w_{n+1} = w_n w_n 1 w'_n$, where w'_n is deduced from w_n by changing the first letter from 0 to 2.

Let now u be the sequence beginning with b_n, where $b_0 = 0, b_{n+1} = b_n b_n 1 b_n$, for all $n \in N$, and (X, T) the dynamical system associated with u; it is called *Chacon's map* ([95]). Note that u is also a fixed point of a substitution ($0 \to 0010$, $1 \to 1$) but a non-primitive one. We send X_0 onto X by replacing every 2 by 0 in every point; to come back from a point in X to its pre-image in X_0, we replace 0 by 2 whenever there is a 1 just before. So the two systems are finitarily hence topologically conjugate, and in particular (X,T) also is uniquely ergodic. Let μ be the invariant measure given by the frequencies of words in u.

The system (X,T) is a rank one system, with the following construction: we cut the interval $[0,1[$ into two intervals P_0 and P_1, of length d_0 and d_1; we take $F_0 = P_0$; then we cut F_0 into three intervals of equal length, and send by T in an affine way the first one onto the second one, the second one onto a sub-interval of P_1, beginning at the left end of P_1 and of corresponding length, namely $d_0/3$, and this sub-interval of P_1 is sent onto the third piece of P_0. At stage n, we cut the n-stack vertically into three equal *columns*, and send (affinely) the top of the first one onto the bottom of the second one, the top of the second one onto a sub-interval of P_1 of corresponding length, beginning at the left end of the yet unused part of P_1, and this sub-interval onto the bottom of the first column. We check that by associating with a point its P-name, we send F_n, the basis of the n-stack, onto the cylinder $x_0 = u_0, \ldots, x_{h_n-1} = u_{h_n-1}$, where $h_0 = 0$, $h_{n+1} = 3h_n + 1$, which gives $h_n = \frac{3^{n+1}-1}{2}$.

The word $b_n u_0 \ldots u_{h_n-1}$ is called the n-block; every point in X is a concatenation of n-blocks (that is: copies of b_n) and letters 1; the 1 are called *spacers*, or n-spacers if they are between two n-blocks.

Lemma 3.4.1. *Chacon's system is* weakly mixing *: there are no eigenfunctions in $\mathcal{L}^2(X)$ except the constants.*

Proof. Let f be an eigenfunction for the eigenvalue λ, of norm 1; we approximate f in $\mathcal{L}^2$ by a sequence of functions f_n constant on the levels of the n-stack. Let C_1 and C_2 be the first and second column of the n-stack, their measure is close to $1/3$ and bigger than $1/4$, and
$\int_{C_1} |T^{h_n} f_n - \lambda^{h_n} f_n|^2 \leq \int_{C_1} |T^{h_n} f_n - T^{h_n} f|^2 + \int_{C_1} |\lambda^{h_n} f - \lambda^{h_n} f_n|^2 < 2\varepsilon$ if n is large enough. But, as f_n is constant on the levels of the n-stack, $T^{h_n} f_n = f_n$ on C_1 and so $|\lambda^{h_n} - 1|^2 < 8\varepsilon$ if n is large enough.
In the same way $T^{h_n+1} f_n = f_n$ on C_2 implies that $|\lambda^{h_n+1} - 1|^2 < 8\varepsilon$ if n is large enough. We conclude that $\lambda = 1$, and it is a simple eigenvalue because unique ergodicity implies ergodicity. ■

Weak mixing and ergodicity can be expressed also as convergences of Cesaro averages; the following lemma is proved for example in [103] or [285]:

Lemma 3.4.2. *A transformation T is ergodic if and only if, for every pair of measurable sets A and B, $(1/n)\sum_{k=0}^{n-1} \mu(A \cap T^k B) \to \mu(A)\mu(B)$ when $n \to \infty$.*

A transformation T is weakly mixing *if and only if, for every pair of measurable sets A and B, $(1/n)\sum_{k=0}^{n-1} |\mu(A \cap T^k B) - \mu(A)\mu(B)| \to 0$ when $n \to \infty$.*

Definition 3.4.1. *A transformation T is* strongly mixing *if and only if, for every pair of measurable sets A and B, $\mu(A \cap T^n B) \to \mu(A)\mu(B)$ when $n \to \infty$.*

Lemma 3.4.3. *Chacon's map is not strongly mixing.*

Proof. We have $\mu(T^{h_p} F_n \cap F_n) > (1/4)\mu(F_n)$ if $p \geq n$, while $\mu(F_n) < 1/4$ if $n \geq 2$. ∎

3.4.2 Complexity and geometric representation

Proposition 3.4.1. *The complexity of the sequence u is $p(n) = 2n - 1$ if $n \geq 2$.*

Proof. Let $n \geq h_k + 1$; every word of length n occurring in u begins somewhere in the k-block, and then meets k-blocks and k-spacers, or else begins with a k-spacer; with such an occurrence of w we associate its initial position $0 \leq a \leq h_k - 1$, the number d of holes (between two k-blocks) which it crosses, and for each hole the number $s_d = 0$ or 1 of k-spacers inside it, or else, if the occurrence of w begins by a k-spacer, we write that $a = -1$ and there is a hole at the beginning, with $s_1 = 1$ (in the same way, if the occurrence ends with a spacer, we write that there is a hole at the end). We can rebuild w from n, d, a, s_1, ... , s_d; these parameters, which we call the *k-configuration*, depend a priori on the particular occurrence and not only on the word w; we define the hypothesis $H(k, n)$ by: *each word of length n occurring in u has only one k-configuration.*

For $n = h_k + 1$, for each initial position between 0 and $h_k - 1$ there are two possible k-configurations, according to whether after the k-block we see another k-block or a spacer; there is also a configuration beginning by a k-spacer. This makes $2h_k + 1$ possible configurations. Hence the hypothesis $H(k, h_k + 1)$ is equivalent to $p(h_k + 1) = 2h_k + 1$.

Let $h_k + 1 \leq n \leq h_{k+1} = 3h_k + 1$, and suppose $H(k, n)$ is satisfied; a word of length n corresponds to one configuration, with $1 \leq d \leq 3$, we call $(s_1, \ldots, s_d)$ its *spacer configuration*; we check that every spacer configuration is allowed, with zero or one spacer in each hole, except that there cannot be three consecutive holes without a spacer nor three consecutive holes with one spacer.

We say a spacer configuration with c holes is *bi-extensible* if, by adding at its end another hole with either zero or one spacer, we get two allowed spacer configurations with $c+1$ holes. For $c = 1, 2$, there are two bi-extensible spacer configurations with c holes.

Because of $H(k, n)$, a word W is bi-extensible (that is: has two right extensions $W0$ and $W1$ occurring in u) if and only if it ends with a k-block, and its spacer configuration is bi-extensible. So, if n is such that all the words of length n ending with a k-block have the same number of holes, there are two bi-extensible words. Otherwise, there remains the case $n = 2h_k + 1$, where the only words ending with a k-block have two holes if they don't contain any spacer and one hole if they contain one spacer, and there are still two bi-extensible words. The two right extensions of a bi-extensible word are two different words by definition; hence $H(k, n)$ implies simultaneously $H(k, n+1)$ and $p(n+1) - p(n) = 2$ as long as $n \leq h_{k+1}$. And so, starting from $H(k, h_k + 1)$, we deduce that $p(h_{k+1} + 1) = 2h_{k+1} + 1$, which implies $H(k+1, h_{k+1} + 1)$ and we may continue the recursion; we check that from $n = 2, 3, 4$ and $n = h_1 + 1 = 5$, $p(n) = 2n - 1$. ∎

We check that the complexity of the sequence v is $2n+1$ (for each n, two words of length n of u correspond each to two words of v, and every other one corresponds to one word); this provides an example of a complexity function changed by a topological conjugacy.

Proposition 3.4.2. *The induced map of Chacon's transformation on the cylinder $x_0 = 0$ is semi-topologically conjugate to the triadic rotation.*

Proof. We consider the substitution τ defined by $a \to aab$, $b \to bab$, and w, its fixed point beginning by a; $w = aabaabbabaabaabbabbabaabbab\ldots$, is deduced from u by $0 \to a$, $10 \to b$. Let S be the induced map (see Definition 3.3.2 above) of T on $x_0 = 0$; we see that $Sx = Tx$ if $x_0 = x_1 = 0$, $Sx = T^2x$ if $x_0 = 0, x_1 = 1$. We call A and B these cylinders, and we code the trajectory of a point y under S by $y_n = a$ if $S^n y \in A$, and $y_n = b$ if $S^n y \in B$. We check, as in the proof of Proposition 3.3.4, that if x is a point in x and the sequence $v(x)$ its Q-name (under S) for the partition $\{A, B\}$, we deduce Tx from $v(x)$ by $a \to 0$, $b \to 10$. By the same reasoning as in Proposition 3.3.6, going from the fixed point to every point, we check that the dynamical system associated with τ is topologically conjugate with the system (Y, S).

The structure of the system associated with τ can be analyzed in the same way as for the Morse system: we get two stacks, with bases $F_{n,0}$ and $F_{n,1}$, and, if we identify corresponding levels of the two stacks, we get the sequence of stacks with bases F_n associated with the triadic rotation (which is a straightforward generalization of the dyadic one). But here $\tau^n(0)$ and $\tau^n(1)$ differ only by the first letter; the reasoning of Proposition 3.1.10 shows then that, except on a countable set, a point is determined by the sequence of levels of F_n to which it belongs (the knowledge that x is in the 0 or 1 stack does not provide extra information). Hence the system associated with τ is a coding of the triadic rotation. ∎

Thus we have a geometric representation of Chacon's map: it is a *Rokhlin-Kakutani exduction* of the triadic rotation (Y, S): there exists a partition (Y_0, Y_1) of Y such that $X = Y_0 \cup (Y_1 \times (0,1))$ and $Tx = Sx$ if $x \in Y_0$, $T(x, 0) = (x, 1)$ and $T(x, 1) = (Sx, 0)$ if $x \in Y_1$. The set Y_0 is the cylinder $x_0 = a$ in the system associated with τ.

However, being an exduction is a much weaker property than being a two (or four)-point extension, as many properties are lost; for example, the triadic rotation has a discrete spectrum, while Chacon's transformation has no eigenvalue. If the exduction is only on a measurable set, the theory of *Kakutani equivalence* ([266]) says we can get, from the triadic rotation, many different systems, such as any ergodic substitution or even the *horocycle flow*; the interest of the above construction is that the exduction is on an explicit and, after a coding, clopen set.

The study of Morse, Rudin-Shapiro and τ show us a particular case of the following theorem ([113]), which we state without proof. For more details, see Chap. 7.

Proposition 3.4.3. *If ζ is a primitive substitution of constant length q, that is $|\zeta a| = q$, $\forall a \in \mathcal{A}$, then the associated system has a discrete spectrum if and only if there exist k and i such that the i-th letter of $\zeta^k(a)$ is the same letter for every a in $\mathcal{A}$.*

The Fibonacci substitution, on the other side, is a *Pisot substitution*, as the dominant eigenvalue is a Pisot number (it is strictly larger than 1 and all the other eigenvalues have modulus strictly smaller than 1); the existence of an analogous theorem for this family of substitutions is an open question. The general study of eigenvalues for any substitution can be found in [147], see also Chap. 7.

3.4.3 Joinings

Our aim is now to find every *joining* of T with itself, that is every ergodic measure on $X \times X$, invariant by $T \times T$, and whose marginals on X are μ; T has one invariant measure, μ, while $T \times T$ preserves already the measure $\mu \times \mu(A \times B) = \mu(A)\mu(B)$ (it is ergodic because T is weakly mixing - this is a nontrivial result), but also $\nu(A \times B) = \mu(A \cap B)$ and others. The joinings have been studied in [120]; we need first to recall a classical ergodic result:

Definition 3.4.2. *For a transformation S, defined as the shift over a space of sequences, and an invariant measure ν, a point x is* generic *if for every cylinder A, $(1/N)\sum_{n=1}^{N} 1_A(S^n x) \to \nu(A)$ if $N \to +\infty$, and the analogous for $N \to -\infty$.*

Lemma 3.4.4. *If (S, ν) is ergodic, ν-almost every point is generic.*

We shall not prove this result, which is a form of *Birkhoff's ergodic theorem*; the usual form is that for every measurable f, $(1/n)\sum_{i=0}^{n-1} f(S^n x) \to \int f d\nu$, ν-almost everywhere; note that we have proved it if S is a primitive substitution, and in that case every point is generic. We shall use Lemma 3.4.4 for the Cartesian powers of Chacon's map; in fact, there is a deep result ([119]) showing that for $T \times T$, all but a countable set of points are generic.

In Chacon's system (X, T), let $R(k,n)$ be the set of x such that x_0 is in the k-th n-block inside its $n+1$-block, $k = 1, 2, 3$.

Lemma 3.4.5. *The sets $R(k,n)$ and $R(k', n+1)$ are, for fixed k and k', independent sets for all n. Almost every point x of X is in $R(k,n) \cap R(k', n+1)$ for a set of integers of density at least 1/10. Every point x with these properties is called* admissible.

Proof. The first assertion is immediate. The second comes from the strong law of large numbers applied to the variables

$$X_n(k,k') = 1_{R(k,2n)} 1_{R(k',2n+1)} / \mu(\cup_{i=0}^{2n-1} T^i F_{2n}) \mu(\cup_{i=0}^{2n} T^i F_{2n+1}),$$

which are independent and equidistributed, and the same after exchanging $2n$ and $2n+1$. ■

Lemma 3.4.6. *If x and y are admissible points belonging to different orbits, then, for infinitely many values of n, x_0 and y_0 are in different n-blocks inside their $n+1$-block, and these n-blocks overlap on a length bigger than $h_n/10$.*

Proof. We take m such that x_0 is in the second m-block. If, for every $n > m$, x_0 and y_0 are in the same n-block, they are always in the same column of the n-stack, and their difference of level $i(x,y)$ is constant: then x and $T^i y$ are not separated by the partition into levels of the stacks and x and y are on the same orbit. We take the first $n > m$ such that x_0 and y_0 are in different n-blocks. Then , either x_0 is in the second $n-1$-block (if $n = m+1$) or x_0 and y_0 are in the same $n-1$-block, which guarantees an overlap of at least h_{n-1} to the left or to the right of (x_0, y_0). ■

Lemma 3.4.7. *If (Y, S, ϱ) is an ergodic transformation, ν an ergodic measure on $Y \times Y$, with marginals ϱ on each copy of Y, invariant under $S \times I$, where I is the identity, then $\nu = \varrho \times \varrho$.*

Proof. Let $B \subset Y$ measurable, and $P_B(x)$ a *conditional probability* of $Y \times B$ relatively to the product of the borelian σ-algebra Y by the σ-algebra $(Y, \emptyset)$. We have $\nu(A \times B) = \int_A P_B(x) d\varrho(x)$. But $\nu(SA \times B) = \nu(A \times B)$, hence $\int_A P_B(x) d\varrho(x) = \int_{SA} P_B(x) d\varrho(x) = \int_A P_B(S^{-1}x) d\varrho(x)$,and, by unicity of the conditional expectation, P_B is S-invariant, and hence constant by ergodicity. Its value can only be $\nu(B)$ and hence $\nu(A \times B) = \nu(A)\nu(B)$. ■

Proposition 3.4.4. *T has* minimal self-joinings of order two*: its joinings with itself are the product measure $\mu \times \mu$ and the* diagonal *measures* $\nu(A \times B) = \mu(A \cap T^i B)$ *for an integer i.*

Proof. Let ν be a joining; we choose a point (x, y) generic for ν, such that x and y are admissible (it is possible because the marginals are μ). If x and y are on the same orbit under T, we check that ν is diagonal. Henceforth we suppose x and y are not on the same orbit.

Let $P(k) = (P_1(k), \ldots, P_{h_k}(k), P_s(k))$ be the partition into levels of the k-stack and the complement of the whole k-stack. We shall show that for all k , and all $1 < i \leq h_k, 1 \leq j \leq h_k$, $\nu(P_i \times P_j) = \nu(P_{i-1} \times P_j)$. As unions of these atoms approximate every measurable set, we shall deduce that ν is $T \times I$-invariant, and so is the product measure.

We fix a k, $P = P(k)$, and an ε.

We choose n such that x_0 and y_0 are in different n-blocks, with overlap on at least $h_n/10$ indices. If n is large enough, genericity ensures that on every segment $(x, y)_0, \ldots, (x, y)_l$ or $(x, y)_{-l}, \ldots, (x, y)_0$, where $l > \varepsilon h_n/1000$, the pairs (i, j) appear with a frequency $\varepsilon/100$-close to $\nu(P_i \times P_j)$. We deduce that these pairs appear with a frequency $\varepsilon/50$-close to $\nu(P_i \times P_j)$ in every segment of (x, y) of length at least $h_n/10$ containing the origin.

Suppose for example that x is in the second and y in the first n-block, and that the overlap between the n-blocks of x and y goes from $-q$ to r, with $q + r + 1 \geq h_n/10$. Let $Q_{i,j}$ be the density of l in $(-q, r)$ for which $T^l x \in P_i, T^l y \in P_j$; we have $|Q_{i,j} - \nu(P_i \times P_j)| < \varepsilon/50$. Let $R_{i,j}$ be the density of l in $(-q + h_n, r + h_n)$ for which $T^l x \in P_i, T^l y \in P_j$: we have $|R_{i,j} - \nu(P_i \times P_j)| < \varepsilon/2$, for otherwise there would be a proportion bigger than $\varepsilon/(2\times 11) - \varepsilon/50$ of errors on the segment $(-q, r+h_n)$. But for $l \in (-q, r)$, because of the position of x and y in their $n + 1$-block, if $T^l x \in P_i$ and $T^l y \in P_j$, then $T^{l+h_n} x \in P_{i-1}$ and $T^{l+h_n} y \in P_j$. Hence, if $i \neq 1$, $R_{i-1,j} = Q_{i,j}$ and $|\nu(P_i \times P_j) - \nu(P_{i-1} \times P_j)| < \varepsilon$. We conclude in the same way for other positions of x and y, possibly replacing the translation of h_n by $-h_n$ ou $2h_n$. ■

It is worth mentioning that, while the above proof uses the presence of isolated spacers between n-blocks, a similar property, called the *R-property*, was used in [289] to compute the joinings of the *horocycle flows*, and was the basis of the famous papers of Ratner which culminate in the proof of Ragunathan's conjecture ([290]).

The interest of minimal self-joinings, as well as the notion itself, appear in [305]:

Proposition 3.4.5. *Let (X, T, μ) be a system with minimal self-joinings of order two and such that T^n is ergodic for all n. Then every transformation S preserving μ and commuting with T is a power T^n, and every borelian sub-σ-algebra invariant by T is trivial.*

Proof. Suppose that S is measure-preserving and commutes with T; then $\mu(S^{-1}A) = \mu(A)$ for every borelian set. If S is invertible, we have also $\mu(SA) = \mu(A)$ and $\nu(A \times B) = \mu(A \cap SB)$ is a joining, and so $S = T^n$;

if S is not invertible, and if ξ is the full borelian σ-algebra, $S^{-1}\xi$ is a non-trivial invariant sub-σ-algebra and the second assertion will imply the first.

Let H be an invariant sub-σ-algebra for T; we write $x \equiv y$ if x and y are not separated by H; the set Z of equivlence classes is equipped with the trace measure of μ, and the dynamical system (Z, T, μ) is called a *factor* of (X, T, μ) (as in Sec. 3.1.8, but here not necessarily with a finite fiber); by abuse of language, we speak of the factor H. We define $\nu(A \times B) = \int_X E_H(1_A)(x)E_H(1_B)(x)d\mu(x)$. We check that a set A is measurable for H if and only if $A \times X = X \times A$, ν-almost everywhere: $\nu(A \times X \Delta X \times A) = \nu(A \times A^c) + \nu(A^c \times A) = 0$ if and only if $E_H(1_A)$ takes only values 0 or 1, that is $A \in H$.

The measure ν is an invariant measure for $T \times T$, not necessarily ergodic, but it admits a decomposition into *ergodic components*, and these are joinings. Hence $\nu = a\mu \times \mu + (1-a)\sum_{j \in Z} b_j \Delta_j$, where $\Delta_j(A \times B) = \mu(A \cap T^j B)$.

Suppose there exists a set $A_0 \subset Z$ of non-trivial measure, and let $A \subset X$ be the union of all elements in A_0, which has the same measure. We have $\mu(A) = \nu(A \times A) = a(\mu(A)^2) + (1-a)\sum b_j \mu(A \cap T^j A)$. Because of the ergodicity of every T^j, we have a convex combination of terms $< \mu(A)$, except the term for $j = 0$; hence only this last one can have a nonzero coefficient, and $\nu = \Delta_0$. Hence $\nu(A \times B) = \mu(A \cap B)$, and by the above criterion, every set is H-measurable. ■

Corollary 3.4.1. *Such a transformation has no measure-preserving roots.*

Proof. Suppose $S^n = T$; then $ST = TS = S^{n+1}$ and we conclude. Even if we know only that S^n is measure-theoretically isomorphic with T, $S^n = T'$ still has minimal self-joinings and we can conclude. ■

Furthermore, it can be shown also [120]) that Chacon's map T has *minimal self-joinings of all orders*: for every p-uple of nonzero integers $(k_1, \ldots, k_p)$,the ergodic measures invariant by $T^{k_1} \times \cdots \times T^{k_p}$,and with marginals μ are all the products of diagonal measures, defined here by $\nu(A_1 \times \ldots A_n) = \mu(T^{l_1}A_1 \cap \ldots . T^{l_n}A_n)$; these include the product measure $\mu \times \ldots \mu$.

A transformation having minimal self-joinings of all orders can be used in the so-called *counter-example machine* ([305]): by using Cartesian products of the T^{k_n}, we can build a large family of transformations with surprising properties, for example a transformation with a continuum of non-isomorphic square roots, or two non-isomorphic systems where each one is a factor of the other.