7

Infinite words with uniform frequencies, and invariant measures

S. Ferenczi

IML - Université Aix-Marseille II - CNRS UMR 6206
163, avenue de Luminy, Case 907, F-13288 Marseille cedex 09, France

T. Monteil

LIRMM - Université Montpelier II - CNRS UMR 5506
161 rue Ada, F-34392 Montpellier cedex 5, France

A fruitful use of word combinatorics is its contribution to the study of dynamical systems, through the symbolic dynamical systems defined in Section 1.6. Indeed, the study of most dynamical systems in the *topological* and the *measure-theoretic* categories can be reduced, by appropriate coding techniques, to the study of a suitable symbolic system X_x; and the topological properties of a symbolic system X_x (equipped with the product topology on $A^{\mathbb{N}}$) can be translated into combinatorial properties of the infinite word x.

In this chapter, we shall study the first two combinatorial properties of infinite words which are significant (and indeed, primordial) for symbolic dynamical systems. The first one is the well-known *uniform recurrence* which translates the dynamical property of *minimality*, that is the fact that the topological system cannot be split into smaller systems. The second one is the fact that the topological system has one invariant probability measure; this is called *unique ergodicity*, a somewhat unhappy expression as it suggests a close association with the classical (*i.e.*, measure-theoretic) ergodic theory, though in fact it is a purely topological notion. Thus, for symbolic systems, unique ergodicity translates into the existence of *frequencies* for every finite factor of the infinite word x, and the limit defining these frequencies is a uniform one; thus we propose to say that the infinite word x has *uniform frequencies*. Similarly, the set of *invariant measures* depends only on the topological structure, and combinatorial properties of the word x will give informations on its structure.

Thus we want to explain how the notions of minimality/uniform recurrence and unique ergodicity/uniform frequencies provide an interaction between dynamical systems and word combinatorics, in most cases the dynamics being the main source of questions and the combinatorics the main tool for answers. We shall focus more on the second couple of notions, as it has been less studied, and also because it constitutes a very strong property, for the system and for the infinite word: in particular it means that dynamical results, which ergodicians are happy to know *almost everywhere* for a given system, will be valid *everywhere* for a uniquely ergodic system.

In Section 7.1 we study the relationship between symbolic systems and languages, and between minimality and uniform recurrence. In Section 7.2 we describe what is known of the set of invariant measures for a general symbolic system, and detail the notion of unique ergodicity and its consequences. Section 7.3 presents some achievements of word combinatorics, initiated by M. Boshernitzan, which allow us to deduce uniform frequencies (or, more generally, to bound the number of ergodic invariant measures of the system) from simple combinatorial properties of the words. In Section 7.4 we review the known examples of words with uniform frequencies, and in Section 7.5 we give important examples which do not have uniform frequencies. We finish in Section 7.6 by hinting how these basic notions have given birth to very deep problems and high achievements in dynamical systems.

7.1 Basic notions

7.1.1 Languages and subshifts

Let us recall some definitions introduced in Section 1.6.2. Proofs are sketched. Let A be a finite alphabet. The set of infinite words $A^{\mathbb{N}}$ is endowed with the product topology defined in Section 1.2.10. It is a compact metrisable space. Let S denote the *shift map* defined on $A^{\mathbb{N}}$ by

$$S(x_0x_1x_2\cdots) = x_1x_2x_3\cdots \ .$$

This map is continuous for the product topology.

Definition 7.1.1 A *subshift* (also called *symbolic dynamical system*) is a couple (X, S), where X is a non-empty closed subset of $A^{\mathbb{N}}$, which is stable under S. We still denote by S the restriction $S|_X$. The *orbit* of a point x in X under the shift map S is the set $\mathcal{O}(x) = \{S^n(x) \mid n \in \mathbb{N}\}$.

Given a subshift (X, S), let $L(X)$ denote the *language* of the finite words which occur in some element of X, and let $L_n(X) = L(X) \cap A^n$ (for $n \in \mathbb{N}$). This language is a non-empty subset of A^* with the additional properties of being

(i) *factorial*: any factor of any element of $L(X)$ is also in $L(X)$,
(ii) *extendable*: for any element w of $L(X)$, there exists a letter a in A such that wa is also in $L(X)$.

A subshift is determined by its language, more precisely, we have the following correspondence:

Proposition 7.1.2 *The map $(X, S) \mapsto L(X)$ is a bijection from the set of subshifts to the set of non-empty factorial and extendable languages.*

Proof We can define an inverse map by sending a non-empty factorial and extendable language L to the subshift $(\{y \in A^{\mathbb{N}} \mid L(y) \subseteq L\}, S)$. □

When x is an infinite word, we denote by (X_x, S) the smallest subshift containing x. The set X_x is equal to the closure (in $A^{\mathbb{N}}$) of the orbit of x under S:

$$X_x = \overline{\mathcal{O}(x)} = \overline{\{S^n(x) \mid n \in \mathbb{N}\}}.$$

In terms of languages, if $L(x)$ denotes the set of all finite factors of the word x (see Definition 1.2.8), then (X_x, S) is the subshift whose associated language is $L(x)$: $L(x) = L(X_x)$. Hence, we have,

$$X_x = \{y \in A^{\mathbb{N}} \mid L(y) \subseteq L(x)\}.$$

If $u = u_0 \cdots u_{n-1} \in A^n$ is a finite word, we can define the *cylinder* $[u]_X$ (or simply $[u]$ when the context is clear) as the set of elements of X which admit u as a prefix:

$$[u]_X = \{x \in X \mid \forall i \leq n-1,\ x_i = u_i\}.$$

The cylinders are clopen sets and form a basis of the topology of X. In particular, the characteristic function $\chi_{[u]}$ of a cylinder $[u]$ is continuous. We have,

$$L(X) = \{u \in A^* \mid [u]_X \neq \emptyset\}.$$

7.1.2 Uniform recurrence and minimality

Definition 7.1.3 A symbolic dynamical system (X, S) is *minimal* if it does not contain a smaller subshift.

Definition 7.1.4 The infinite word x is *uniformly recurrent* if every factor of x occurs in an infinite number of places with bounded gaps.

Those two notions are in correspondence, we recall this with some more details:

Proposition 7.1.5 *Let x be an infinite word on the alphabet A. The following assertions are equivalent:*

(i) *The word x is uniformly recurrent.*
(ii) *For any factor $u \in L(x)$, there exists an integer $n \in \mathbb{N}$ such that any element of $L_n(x)$ contains an occurrence of u.*
(iii) *Any word y of X_x satisfies $L(y) = L(x)$.*
(iv) *Any word y of X_x satisfies $X_y = X_x$.*
(v) *The orbit of any element y of X_x under the shift map S is dense in X_x.*
(vi) *The subshift (X_x, S) is minimal.*

Proof (i) $\Leftrightarrow$ (ii) and (iii) $\Leftrightarrow$ (iv) are direct reformulations. For (iv) $\Leftrightarrow$ (v) $\Leftrightarrow$ (vi), it suffices to notice that, for any y in X_x, $X_y = \overline{\mathcal{O}(y)}$ is the smallest subshift containing y and is included in X_x. For (ii) $\Rightarrow$ (iii), let y be an element of X_x, and let u be an element of $L(x)$. Because of the hypothesis, there exists an integer sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that $(S^{\alpha_n}(x))$ converges to y, and there exists an integer $n_0 \in \mathbb{N}$ such that any element of $L_{n_0}(x)$ contains an occurrence of u. Since the sequence $(S^{\alpha_n}(x))$ converges to y, for N big enough, the prefix of length n_0 of $S^{\alpha_N}(x)$ coincides with the prefix of length n_0 of y. Hence y contains an occurrence of u. For $\neg$(ii) $\Rightarrow$ $\neg$(iii), because of the hypothesis, there exists a factor $u \in L(x)$ such that for any integer $n \in \mathbb{N}$, there exists an integer α_n such that u does not occur in $x_{\alpha_n} \cdots x_{\alpha_n+n}$. Let y be an adherent point of the sequence $(S^{\alpha_n}(x))_{n\in\mathbb{N}}$ in X_x (which exists by compactness). The finite word u is not a factor of y, hence $L(y) \subset L(x)$. $\square$

7.1.3 Uniform frequencies

Definition 7.1.6 The infinite word x has *uniform frequencies* if, for every factor w of x, the ratio $\frac{|x_k \cdots x_{k+n}|_w}{n+1}$ (see 1.2.3 for the notation) has a limit $f_w(x)$ when $n \to +\infty$, uniformly in k.

The aim of the next section is to provide a dynamical equivalent of uniform frequencies (see in particular Proposition 7.2.10 below).

7.2 Invariant measures and unique ergodicity

7.2.1 Two frameworks in dynamical systems

We first insist on the difference between topological and measure-theoretic dynamics, unique ergodicity being a topological notion. A good reference for this section is (Denker, Grillenberger, and Sigmund 1976).

Definition 7.2.1 A *measure-theoretic dynamical system* is a quadruple

$(X, \mathcal{B}, \mu, T)$ where $(X, \mathcal{B}, \mu)$ is a probability Lebesgue space and $T : X \to X$ is a measurable function that preserves the measure μ:

$$\forall B \in \mathcal{B},\ \mu(T^{-1}(B)) = \mu(B).$$

Such a system is *ergodic* if the only T-invariant measurable subsets of X have measure 0 or 1 (a subset B of X is T*-invariant* if $T^{-1}(B) = B$).

A measure-theoretic ergodic dynamical system satisfies the Birkhoff ergodic theorem (see Theorem 1.6.7). We refer the reader to Chapter 5 of (Pytheas Fogg 2002) for a presentation of this fundamental result in the present framework:

$$\forall f \in L^1(X, \mathbb{R}) \ , \quad \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \xrightarrow[n\to\infty]{\mu-a.e.} \int_X f d\mu.$$

Definition 7.2.2 A *topological dynamical system* is a couple (X, T), where X is a compact metric space and $T : X \to X$ is a continuous function.

Hence, a subshift can be seen as a topological dynamical system.

7.2.2 The set of invariant measures of a subshift

A way to understand a topological dynamical system (X, T) is to study the set $\mathcal{M}(X, T)$ of Borel probability measures on X that are preserved by T: this corresponds to the measure-theoretic dynamical systems that are housed by (X, T). Let $\mathcal{E}(X, T)$ denote the subset of ergodic invariant measures.

We will need some basic material coming from measure theory and functional analysis (the Riesz representation theorem and the Banach–Alaoglu theorem) (Rudin 1987) (Rudin 1991), it is summarised in the following proposition:

Proposition 7.2.3 *The set of Borel probability measures on a compact metrisable space X can be identified with a convex subset of the topological dual of $C^0(X, \mathbb{R})$, endowed with the weak-star topology. This topology is metrisable and compact. A sequence $(\mu_n)_{n\in\mathbb{N}}$ of such measures converges to a measure μ if, and only if, for each continuous function $f \in C^0(X, \mathbb{R})$, $\int_X f d\mu_n$ converges to $\int_X f d\mu$. When X is a closed subset of $A^{\mathbb{N}}$, continuous functions can be replaced with characteristic functions of cylinders,* i.e., *μ_n converges to μ if, and only if, for each finite word u, $\mu_n([u])$ converges to $\mu([u])$.*

We can now describe some properties and the geometry of the set of invariant measures of a topological dynamical system and its subset of ergodic measures:

Proposition 7.2.4 *Let (X, T) be a topological dynamical system.*

(i) *The set $\mathcal{M}(X, T)$ is a non-empty convex compact subspace of the space of Borel probability measures on X.*
(ii) *Let $\mu \in \mathcal{E}(X, T)$ and $\nu \in \mathcal{M}(X, T)$. If ν is absolutely continuous with respect to μ, then $\mu = \nu$.*
(iii) *The extreme points of $\mathcal{M}(X, T)$ are the ergodic measures.*
(iv) *Two distinct elements of $\mathcal{E}(X, T)$ are mutually singular.*
(v) *A set of n distinct ergodic measures generates an affine space of dimension $n - 1$,* i.e.*, such measures are affinely independent.*

Proof

(i) The set $\mathcal{M}(X, T)$ is clearly convex and closed in the set of Borel probability measures on X. Let us prove that the system (X, T) admits at least one invariant probability measure. Pick a point x in X. Since the set of Borel probability measures on X is metrisable and compact, we can take for μ any accumulation point of the sequence of probability measures defined by $\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k(x)}$, where δ stands for the one-point *Dirac measure*. The average ensures that μ is preserved by T.
(ii) Since ν is absolutely continuous with respect to μ, the Radon–Nikodym theorem ensures that there exists a map $f \in L^1(X, \mathbb{R}_+)$ such that for any Borel subset B of X, $\nu(B) = \int_B f d\mu$. Let us show that f is constant μ-almost everywhere. Assume, by contradiction, that the measure of the Borel set $B = \{x \in X \mid f(x) > \int_X f d\mu\}$ belongs to $(0, 1)$. Since μ is ergodic, B is not T-invariant. Therefore, $\mu(T^{-1}B \setminus B) = \mu(B \setminus T^{-1}B) > 0$. We have, $\mu(B \setminus T^{-1}B) \int_X f d\mu < \int_{B \setminus T^{-1}B} f d\mu = \nu(B \setminus T^{-1}B) = \nu(T^{-1}B \setminus B) = \int_{T^{-1}B \setminus B} f d\mu \leq \mu(T^{-1}B \setminus B) \int_X f d\mu$, which is absurd. Hence f is constant with value $\nu(X) = 1$, so $\mu = \nu$.
(iii) Let μ be an element of $\mathcal{E}(X, T)$ that can be written as a convex combination $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$, where μ_1 and μ_2 are two elements of $\mathcal{M}(X, T)$ and $\lambda \in (0, 1)$. Since $\lambda\mu_1 \leq \mu$, the measure μ_1 is absolutely continuous with respect to μ, hence (ii) ensures that $\mu_1 = \mu$, so μ is an extreme point of $\mathcal{M}(X, T)$. Conversely, let μ be an extreme point of $\mathcal{M}(X, T)$. Assume by contradiction that μ is not ergodic: there exists a Borel set B of X which is T-invariant and satisfies $\mu(B) \in$

$(0,1)$. Let us define, for any two Borel sets C and D of X with $\mu(C)>0$, $\mu_C(D)=\frac{\mu(C\cap D)}{\mu(C)}$. We have $\mu=\mu(B)\mu_B+(1-\mu(B))\mu_{X\setminus B}$, hence μ can be written as a non-trivial convex combination of two elements of $\mathcal{M}(X,T)$, a contradiction.

(iv) Let μ and ν be two distinct ergodic measures for (X,T). Since μ and ν are distinct, there exists a measurable function f such that $\int_X fd\mu \neq \int_X fd\nu$. Let $G_{\mu,f}$ be the set of points of X for which the Birkhoff ergodic sums for f converge to $\int_X fd\mu$ and $G_{\nu,f}$ be the one corresponding to $\int_X fd\nu$. Those two sets are disjoint and satisfy $\mu(G_{\mu,f})=\nu(G_{\nu,,f})=1$ and $\mu(G_{\nu,f})=\nu(G_{\mu,f})=0$.

(v) Let $\mu_1,\dots,\mu_n$ be n distinct elements of $\mathcal{E}(X,T)$. Let $(\beta_1,\dots,\beta_n)$ be a tuple in $\mathbb{R}^n$ such that $\sum_{k=1}^n\beta_k\mu_k=0$. Let $1\le i\le n$ be an integer. By (iv), for any $j\neq i$, there exists a Borel set G_j such that $\mu_i(G_j)=1$ and $\mu_j(G_j)=0$. We have $\beta_i=\sum_{k=1}^n\beta_k\mu_k(\cap_{j\neq i}G_j)=0$. Hence, the μ_k are linearly independent, therefore affinely independent.

□

Those results remain valid for the particular case of symbolic systems, and can again be interpreted in terms of word combinatorics.

Definition 7.2.5 A *weight function* on a subshift (X,S) is a map $\varphi: L(X)\to\mathbb{R}^+$ such that:

(i) $\varphi(\varepsilon)=1$,
(ii) $\forall w\in L(X),\ \varphi(w)=\sum_{a\in A,\ wa\in L(X)}\varphi(wa)$,
(iii) $\forall w\in L(X),\ \varphi(w)=\sum_{a\in A,\ aw\in L(X)}\varphi(aw)$.

Let us denote by $\mathcal{W}(X,S)$ the set of *weight functions* on (X,S).

Again, we have a correspondence:

Proposition 7.2.6 *The following map is a bijection.*

$$\left(\begin{array}{ccc}\mathcal{M}(X,S) & \longrightarrow & \mathcal{W}(X,S)\\ \mu & \longmapsto & \left(\begin{array}{ccc} L(X) & \longrightarrow & \mathbb{R}^+\\ w & \longmapsto & \mu([w])\end{array}\right)\end{array}\right)$$

Proof Despite the symmetry between the second and the third item, they do not play the same role: (ii) follows from the additivity of the measure, whereas (iii) follows from its shift-invariance. The first item tells that the measure is a probability measure, *i.e.*, a measure of total mass 1. Therefore, the map is well defined. Since the set of cylinders generates the Borel σ-algebra on X (it is a basis of the topology of X) and is closed under finite intersection, the map is injective (acccording to Dynkin theorem, see *e.g.*, (Rudin 1987)). The surjectivity is guaranteed by the Carathéodory extension theorem (construction of an outer measure), for example any open set is a disjoint (possibly infinite) union of cylinders, its measure is given by summing the weights of those cylinders. □

The Birkhoff ergodic theorem applied to characteristic functions of cylinders can be restated in terms of frequencies:

Proposition 7.2.7 *Let (X, S) be a subshift, and let μ be an element of $\mathcal{E}(X, S)$. Then for μ-almost any x in X, and for any finite word w in $L(X)$, the frequency $f_w(x)$ exists and is equal to $\mu([w])$.*

Proof Apply the Birkhoff ergodic theorem with the countable family of maps $f = \chi_{[w]}$ and notice that

$$|x_0 \cdots x_{n+|w|-2}|_w = \sum_{k=0}^{n-1} \chi_{[w]} \circ S^k(x).$$

□

The points which satisfy this convergence are said to be *generic* for μ. With the notations of the proof of Proposition 7.2.4(iv), the set of generic points for μ is equal to $\cap_{w \in L(X)} G_{\mu,\chi_{[w]}}$. This proposition lets us imagine a proto-strategy to prove that an infinite word x has frequencies, by considering the subshift X_x and looking for the ergodic invariant measures on it. Unfortunately, it is possible that the word x is generic for no element of $\mathcal{E}(X_x, S)$ (see Proposition 7.2.11).

7.2.3 Unique ergodicity

Definition 7.2.8 A topological dynamical system is *uniquely ergodic* if it has only one invariant probability measure.

The unique invariant measure μ is ergodic (it is an extreme point of a singleton by Proposition 7.2.4(iii)). In this extreme case of a uniquely ergodic dynamical system, *all* the orbits are equidistributed:

Proposition 7.2.9 *Let (X, T) be a uniquely ergodic topological dynamical system, whose unique invariant measure is denoted by μ. Let $f : X \to \mathbb{R}$ be a continuous function. Then, the sequence of functions $(\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k)_{n\in\mathbb{N}}$ converges uniformly to the function with constant value $\int_X f d\mu$.*

Proof Assume by contradiction that there exist $\varepsilon > 0$, a sequence $(x_n)_{n\in\mathbb{N}}$ in X and an *extraction* (*i.e.*, a strictly increasing integer sequence) α such that for any integer n, $\frac{1}{\alpha(n)}\sum_{k=0}^{\alpha(n)-1} f(T^k(x_\alpha(n))) - \int_X f d\mu \geq \varepsilon$. As in the proof of Proposition 7.2.4(i), let ν be an adherent point of the sequence of probability measures $\nu_n = \frac{1}{\alpha(n)}\sum_{k=0}^{\alpha(n)-1} \delta_{T^k(x_\alpha(n))}$. The measure ν is T-invariant and satisfies $\int_X f d\nu - \int_X f d\mu \geq \varepsilon$, this contradicts the uniqueness of μ. □

In the symbolic case, when we apply this result to characteristic functions of cylinder sets as in Proposition 7.2.7, we get the following proposition:

Proposition 7.2.10 *A symbolic system (X_x, S) is uniquely ergodic if, and only if, x has uniform frequencies.*

Given a symbolic dynamical system (X, S), one can imagine the situation where there exists more than one ergodic invariant measure but any x in X has frequencies (that can depend on the point x). A theorem of J. Oxtoby (Oxtoby 1952) ensures that this is not possible in the minimal case, making unique ergodicity a necessary condition for global existence of frequencies:

Proposition 7.2.11 *If (X, S) is a minimal non-uniquely ergodic symbolic system, then there exist an infinite word $x \in X$ and a finite word w such that the frequency of w in x is not defined.*

Proof Assume by contradiction that any word in X has frequencies. Therefore, the function

$$f_w(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \chi_{[w]}(S^k(x))$$

is well defined on X, for any finite word w. Since X is not uniquely ergodic, there exist two ergodic measures $\mu \neq \nu$, hence there is a finite word w such that $\mu([w]) \neq \nu([w])$. Let x be a generic point for μ and y be a generic point for ν. We have $f_w(x) = \mu([w]) \neq \nu([w]) = f_w(y)$ and since f_w is constant along the orbits, f_w is nowhere continuous (the minimality implies that both orbits of x and y are dense). But f_w is a simple limit of continuous functions on the complete metric space X (which is even compact), hence the Baire category theorem implies that the points of continuity of f_w must be dense in X. A contradiction. □

It is important to keep in mind that unique ergodicity is *not* a measure-theoretic notion but a topological one. The Jewett, Krieger and Rosenthal theorem (Jewett 1969) (Krieger 1972) (Rosenthal 1988) asserts that unique ergodicity does not imply any restriction on the (unique) associated measure-theoretic ergodic dynamical system:

Proposition 7.2.12 *Every measure-theoretic ergodic dynamical system has a uniquely ergodic topological model,* i.e., *for any measure-theoretic ergodic dynamical system $(X, \mathcal{A}, \mu, S)$, there exists a uniquely ergodic topological dynamical system (Y, T) such that $(X, \mathcal{A}, \mu, S)$ is isomorphic to $(Y, \mathcal{B}, \mu, T)$, where $\mathcal{B}$ denotes the set of Borel subset of X and μ denotes the unique element of $\mathcal{M}(X, T)$.*

7.2.4 Finitely many ergodic invariant measures

When a subshift admits more than one ergodic invariant measure, we can still get some information from a bound on the number of its ergodic invariant measures. Let us see how the cardinality of $\mathcal{E}(X, S)$ measures the diversity of the behaviours of the orbits in the subshift (X, S).

First, Proposition 7.2.4(iii) tells that $\mathcal{E}(X, S)$ is the set of extreme points of $\mathcal{M}(X, S)$. The Krein–Milman theorem (see *e.g.*, (Rudin 1991)) ensures that, in such a situation, $\mathcal{M}(X, S)$ is the closure of the convex hull of $\mathcal{E}(X, S)$. Moreover, the Choquet theorem (see also (Rudin 1991)) ensures that any invariant measure can be written as an average of ergodic measures. Hence, any invariant measure can be recovered from the ergodic ones. Conversely, Propositions 7.2.4(iv) and 7.2.4(v) roughly tell us that no ergodic measure can be recovered from finitely many other ones (though some of them could be obtained as limit points of infinitely many other ones): all of them are necessary to describe $\mathcal{M}(X, S)$. Concerning the orbits, since any ergodic measure admits some generic points, ergodic measures are also all needed to describe the different behaviours of the typical orbits.

Let us now focus on the case when $\mathcal{E}(X, S)$ is known to be finite. Any invariant measure can be written, in a single manner, as a (finite) convex combination of ergodic invariant measures. However, Proposition 7.2.11 tells us that some orbits may be generic for no ergodic measure (see also Exercises 7.8 and 7.9). Their behaviours are nevertheless not out of control:

Proposition 7.2.13 *Let (X, S) be a subshift and let x be an element of X. Let $(w_j)_{j \in J}$ be a family of elements of $L(X)$. Let α be an extraction such that, for any $j \in J$, the sequence $(|x_0 \cdots x_{\alpha(n)-1}|_{w_j}/\alpha(n))$ converges to a number denoted by $f_{w_j}(x, \alpha)$. Then there exists an invariant measure $\mu \in \mathcal{M}(X, S)$ such that the vector $(f_{w_j}(x, \alpha))_{j \in J}$ is equal to $(\mu[w_j])_{j \in J}$.*

In particular, if (X, S) admits at most K ergodic invariant measures, and if $(x^{(i)})_{i \in I}$ is a family of elements of X, then the set of vectors

$\{(f_{w_j}(x^{(i)}, \alpha^{(i)}))_{j\in J} \mid i \in I\}$ *spans an affine space of dimension at most* $K-1$.

Another direct consequence is that, for any $x \in X$ *and any* $w \in L(X)$*,*

$$\min_{\mu\in\mathcal{E}(X,S)} \mu([w]) \leq \liminf_{n\to\infty} \frac{|x_0\cdots x_n|_w}{n+1} \leq \limsup_{n\to\infty} \frac{|x_0\cdots x_n|_w}{n+1} \leq \max_{\mu\in\mathcal{E}(X,S)} \mu([w]).$$

Proof As in the proof of Proposition 7.2.4(i), the sequence of probability measures $\mu_n = \frac{1}{\alpha(n)}\sum_{k=0}^{\alpha(n)-1}\delta_{T^k(x)}$ admits an adherent point $\mu \in \mathcal{M}(X,S)$. We have, $(f_{w_j}(x,\alpha))_{j\in J} = (\mu[w_j])_{j\in J}$. □

7.3 Combinatorial criteria

We describe combinatorial criteria implying unique ergodicity, or a bound on the number of ergodic invariant measures.

7.3.1 Complexity and Boshernitzan's criteria

Definition 7.3.1 The *complexity function* is the function that maps any integer n to the number $p_X(n) = \mathrm{Card}(L_n(X))$.

Theorem 7.3.2 (Boshernitzan 1984) *Let* $K \geq 1$ *be an integer. A minimal symbolic system* (X,S) *such that* $\left\lfloor \liminf_{n\to\infty} \frac{p_X(n)}{n} \right\rfloor \leq K$ *admits at most* K *ergodic invariant measures.*

Theorem 7.3.2 will follow from Theorem 7.3.7.

Theorem 7.3.3 (Boshernitzan 1984) *A minimal symbolic system* (X,S) *such that* $\limsup_{n\to\infty} \frac{p_X(n)}{n} < 3$ *is uniquely ergodic.*

The original proof of Theorem 7.3.3 is too long and technical to be given here. With a careful study of the evolution of the Rauzy graphs (see Section 7.3.2 below), it can be generalised (and the proof simplified) in the following way:

Theorem 7.3.4 (Monteil 2009) *Let* $K \geq 3$ *be an integer. A minimal symbolic system* (X,S) *such that* $\limsup_{n\to\infty} \frac{p_X(n)}{n} < K$ *admits at most* $K-2$ *ergodic invariant measures.*

7.3.2 Deconnectability of the Rauzy graphs

Theorem 7.3.2 (and Boshernitzan's proof) can be generalised as follows. Let (X, S) be a symbolic system. Any word in $L_{n+1}(X)$ is naturally linked to two words in $L_n(X)$: its prefix and its suffix of length n. A way to keep track of this factorial structure of the language associated with X is the use of *Rauzy graphs*.

Definition 7.3.5 For any integer n, the nth *Rauzy graph* $G_n(X)$ is the directed graph such that:

(i) the set of vertices of $G_n(X)$ is $L_n(X)$,
(ii) there is an (oriented) edge from u to v in $G_n(X)$ if there exists w in $L_{n+1}(X)$ such that w begins with u and ends with v.

When X is the subshift associated with an infinite word x, the Rauzy graphs can be denoted by $G_n(x)$.

These graphs were first defined by N. G. de Bruijn (de Bruijn 1946) in a particular case: the *de Brujn graphs* are the Rauzy graphs when $L_n(X)$ is made of all the possible words of length n on the alphabet, or equivalently when X is the *full shift* defined in Section 7.5.1 below. In their full generality, they were defined by G. Rauzy in (Rauzy 1983) and independently by M. Boshernitzan, in the first published reference in a refereed journal, (Boshernitzan 1985); they were then named by Rauzy's followers. They should not be confused with the *Rauzy diagrams*, see (Yoccoz 2005) for example, which are graphs describing classes of permutations for interval-exchange transformations.

Definition 7.3.6 If $K \geq 1$, a symbolic system (X, S) is said to be K-*deconnectable* if there exist an extraction α and a constant $K' \geq 1$ such that for all $n \geq 1$ there exists a subset $D_{\alpha(n)} \subseteq L_{\alpha(n)}(X)$ of at most K vertices such that every path in $G_{\alpha(n)}(X) \setminus D_{\alpha(n)}$ is of length less than $K'\alpha(n)$ (in particular it does not contain any cycle).

This means that we can disconnect (in a specific way) infinitely many Rauzy graphs by removing at most K vertices.

Theorem 7.3.7 (Monteil 2005) *A K-deconnectable symbolic system (X, S) has at most K ergodic invariant probability measures.*

Proof We will first build at most K possible candidates and then prove that they are the only ones.

Step 1 *We build the candidates to be the only ergodic invariant probability measures.*

For any integer n, let $d_{1,\alpha(n)}, d_{2,\alpha(n)}, \ldots, d_{K,\alpha(n)}$ be an enumeration of $D_{\alpha(n)}$ (there is no loss of generality to consider that all the $D_{\alpha(n)}$ have exactly K elements). For now, we work in the subshift $(A^{\mathbb{N}}, S)$. We approximate X from the outside by K sequences of periodic subshifts as follows: for $i \leq K$ and $n \in \mathbb{N}$, we define

$$\mu_{i,\alpha(n)} = \frac{1}{\alpha(n)} \sum_{k=0}^{\alpha(n)-1} \delta_{S^k(d^{\omega}_{i,\alpha(n)})},$$

where $d^{\omega}_{i,\alpha(n)}$ denotes the infinite word $d_{i,\alpha(n)} d_{i,\alpha(n)} d_{i,\alpha(n)} \cdots$ and δ stands for the one-point Dirac measure. The measure $\mu_{i,\alpha(n)}$ is the only element of $\mathcal{M}(A^{\mathbb{N}}, S)$ that gives full measure to the periodic subshift generated by the periodic word $d^{\omega}_{i,\alpha(n)}$. By compactness of $\mathcal{M}(A^{\mathbb{N}}, S)^K$, there exists an extraction β such that for each $i \leq K$,

$$\mu_{i,\alpha\circ\beta(n)} \xrightarrow[n\to\infty]{} \mu_i,$$

for some μ_i in $\mathcal{M}(A^{\mathbb{N}}, S)$. Note that if X is aperiodic (that is, if it contains no periodic orbit), the measures $\mu_{i,\alpha(n)}$ give measure 0 to X. Anyway,

Step 2 *We show that for $i \leq K$, $\mu_i(X) = 1$.*

Since X is closed in $A^{\mathbb{N}}$, we have the following approximation by open sets:

$$X = \overline{X} = \bigcap_{n\in\mathbb{N}} \bigcup_{u\in L_n(X)} [u],$$

where $[u]$ is considered as a cylinder in the dynamical system $(A^{\mathbb{N}}, S)$. Let $k \geq n \geq 1$ be two integers. For $i \in \{0, \ldots, \alpha \circ \beta(k) - n\}$, the finite word $(d^{\omega}_{i,\alpha\circ\beta(k)})_i \cdots (d^{\omega}_{i,\alpha\circ\beta(k)})_{i+n-1}$ is a factor of the finite word $d_{i,\alpha\circ\beta(k)}$, hence it belongs to $L_n(X)$. Hence, $\mu_{i,\alpha\circ\beta(k)}(\bigcup_{u\in L_n(X)}[u]) \geq (\alpha\circ\beta(k) - n + 1)/\alpha\circ\beta(k)$.

Letting k tend to infinity, since the characteristic function of $\bigcup_{u\in L_n(X)}[u]$ is continuous, we have $\mu_i(\bigcup_{u\in L_n(X)}[u]) = 1$. By countable intersection (n is arbitrary), we have $\mu_i(X) = 1$.

Hence, we can consider μ_i as an element of $\mathcal{M}(X, S)$.

Step 3 *Let μ be an ergodic measure on X. We show that μ is one of the μ_i.*

Since μ is ergodic, we can pick a point x in X that is generic for it, that is, for any u in $L(X)$,

$$\mu([u]) = \lim_{n\to\infty} \frac{|x_0 \cdots x_{n-1}|_u}{n}.$$

Let n be a fixed positive integer. We decompose x into blocks of length $\ell = (K'+1)\alpha \circ \beta(n)$, *i.e.*, $x = b_0.b_1.b_2.b_3.b_4 \cdots$ with $b_j = x_{\ell j} \cdots x_{\ell(j+1)-1}$. Because of the hypothesis, any b_j contains an occurrence of one of the $d_{i,\alpha\circ\beta(n)}$ (b_j can be viewed as a path of length $K'\alpha \circ \beta(n)$ in $G_{\alpha\circ\beta(n)}(X)$). So, there exists some $i_{\alpha\circ\beta(n)}$ such that the upper density of the set

$$\{j \in \mathbb{N} \mid |b_j|_{d_{i_{\alpha\circ\beta(n)},\alpha\circ\beta(n)}} \geq 1\}$$

is at least $1/K$. Let γ be an extraction such that $i_{\alpha\circ\beta\circ\gamma(.)}$ is constant with value denoted by i. We denote by $\tilde{\alpha}$ the extraction $\alpha \circ \beta \circ \gamma$.

Let us show that $\mu = \mu_i$. Let u be a finite word in $L(X)$. Let n be an integer greater than $|u|$. There is an extraction δ such that for any integer m,

$$\mathrm{Card}\{j < \delta(m) \mid |b_j|_{d_{i,\tilde{\alpha}(n)}} \geq 1\} \geq \delta(m)/2K.$$

Therefore,

$$|b_0.b_1 \cdots b_{\delta(m)-1}|_u \geq \frac{\delta(m)}{2K} |d_{i,\tilde{\alpha}(n)}|_u.$$

Hence,

$$\mu([u]) = \lim_{m\to\infty} \frac{|b_0.b_1 \cdots b_{\delta(m)-1}|_u}{(K'+1)\tilde{\alpha}(n)\delta(m)} \geq \frac{|d_{i,\tilde{\alpha}(n)}|_u}{2K(K'+1)\tilde{\alpha}(n)}.$$

Moreover, we can control the frequency of occurrences of u in $d^{\omega}_{i,\tilde{\alpha}(n)}$ by counting separately the occurrences of u that fall in some $d_{i,\tilde{\alpha}(n)}$ and the occurrences of u that appear between two consecutive occurrences of $d_{i,\tilde{\alpha}(n)}$:

$$\mu_{i,\tilde{\alpha}(n)}([u]) = \frac{1}{\tilde{\alpha}(n)} \sum_{k=0}^{\tilde{\alpha}(n)-1} \delta_{S^k(d^{\omega}_{i,\tilde{\alpha}(n)})}([u]) \leq \frac{1}{\tilde{\alpha}(n)}(|d_{i,\tilde{\alpha}(n)}|_u + |u|).$$

Therefore,

$$\mu_{i,\tilde{\alpha}(n)}([u]) \leq \frac{|d_{i,\tilde{\alpha}(n)}|_u}{\tilde{\alpha}(n)} + \frac{|u|}{\tilde{\alpha}(n)} \leq 2K(K'+1)\mu([u]) + \frac{|u|}{\tilde{\alpha}(n)}.$$

Letting n tend to infinity, we have $\mu_i([u]) \leq 2K(K'+1)\mu([u])$, so μ_i is absolutely continuous relatively to μ. Since μ_i is S-invariant and μ is ergodic, Proposition 7.2.4(ii) ensures that $\mu_i = \mu$, hence there are at most K ergodic invariant measures.

□

Proof of theorem 7.3.2 Because of the hypothesis, there is an extraction α such that for all integer n, $p_X(\alpha(n)) \leq (K+1)\alpha(n)$ and $p_X(\alpha(n)+1) - p_X(\alpha(n)) \leq K$. For n in $\mathbb{N}$, let $D_{\alpha(n)}$ be the set of left special factors (see Section 4.5) of length $\alpha(n)$ of X, whose cardinality is not greater than K.

Let $n \in \mathbb{N}$. Any loop O in $G_{\alpha(n)}(X)$ must contain a left special factor. Indeed, since we can assume that X is aperiodic, there exists a finite word u in $L_{\alpha(n)+1}(X) \setminus O$, and since X is minimal, there exists a path from the edge u to any vertex of O, so the first vertex of O that this path meets is a left special factor. Therefore, $G_{\alpha(n)}(X) \setminus D_{\alpha(n)}$ does not contain any loop.

So, a path in $G_{\alpha(n)}(X) \setminus D_{\alpha(n)}$ is necessarily injective and cannot be of length greater than $\mathrm{Card}(L_{\alpha(n)}(X)) \leq (K+1)\alpha(n)$. Therefore, (X,S) is K-deconnectable and Theorem 7.3.7 applies.

□

7.3.3 Boshernitzan's ne_n condition

Let (X,S) be a minimal symbolic system. If $\mu \in \mathcal{M}(X,S)$ is a S-invariant probability measure and n is an integer, we denote by $e_n(\mu)$ the minimal measure of the cylinder sets of length n of X.

Theorem 7.3.8 (Boshernitzan 1992) *Let (X,S) be a minimal symbolic system. If there exists μ in $\mathcal{M}(X,S)$ such that $\limsup_{n\to\infty} ne_n(\mu) > 0$, then (X,S) is uniquely ergodic.*

Proof Because of the hypothesis, there exists $c > 0$ such that, for infinitely many integers $n \in \mathbb{N}$, we have $ne_n(\mu) \geq c$. For such n and $w \in L_n(X)$, we have $\mu([w]) > c/n$ and since all those cylinders are disjoint, we have $p_X(n) = \mathrm{Card}\, L_n(X) \leq n/c$. Hence, $\liminf_{n\to\infty} p_X(n)/n \leq 1/c < \infty$. So, Theorem 7.3.2 tells us that (X,S) admits a finite number of ergodic invariant measures.

Assume by contradiction that (X,S) is not uniquely ergodic: it admits at least two distinct ergodic invariant measures. Hence, there exists a finite word w such that the set $E = \{\mu([w]) \mid \mu \in \mathcal{E}(X,S)\}$ has a finite cardinality which is greater than one. We choose two ergodic measures μ_1 and μ_2 which correspond to consecutive elements in E, *i.e.*, such that $\mu_1([w]) < \mu_2([w])$ and $(\mu_1([w]), \mu_2([w])) \cap E = \emptyset$. Let r and s be two real numbers such that $\mu_1([w]) < r < s < \mu_2([w])$.

For $n \geq 1$, let F_n denote the set $\{x \in X \mid \frac{|x_0 \cdots x_{n-1}|_w}{n} \in [r,s]\}$. If ν is an ergodic invariant measure of (X,S), then the Birkhoff ergodic theorem tells us that ν-almost every point $x \in X$ satisfies $\frac{|x_0 \cdots x_{n-1}|_w}{n} \xrightarrow[n\to\infty]{} \nu([w]) \notin [r-s]$, which implies $\nu(\bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} X \setminus F_n) = 1$. Hence,

$\nu(F_N) \leq \nu(\cup_{n \geq N} F_n) \xrightarrow[N\to\infty]{} \nu(\cap_{N\in\mathbb{N}} \cup_{n\geq N} F_n) = 0$. Since μ is a (finite) convex combination of such ergodic invariant measures, we also have $\mu(F_N) \xrightarrow[N\to\infty]{} 0$.

Let y be a generic point for μ_1 and z be a generic point for μ_2: there exists an integer N such that for any $n \geq N$, $\frac{|y_0\cdots y_{n-1}|_w}{n} < r$ and $\frac{|z_0\cdots z_{n-1}|_w}{n} > s$. Let n be an integer which is greater than N. Since (X, S) is minimal, the Rauzy graph $G_n(X)$ is strongly connected, so there is a path $y_0 \cdots y_{n-1} = v_1 \to v_2 \to \cdots \to v_{p-1} \to v_p = z_0 \cdots z_{n-1}$ from $y_0 \cdots y_{n-1}$ to $z_0 \cdots z_{n-1}$. We choose a shortest such path (in particular, it has no loop).

For $i \leq p-1$, we have $|v_{i+1}|_w \leq |v_i|_w + 1$, hence $\frac{|v_{i+1}|_w}{n} \leq \frac{|v_i|_w}{n} + \frac{1}{n}$. Since $\frac{|v_1|_w}{n} < r$ and $\frac{|v_p|_w}{n} > s$, the set $I = \{i \leq p \mid \frac{|v_i|_w}{n} \in [r,s]\}$ has cardinality greater than or equal to $n(s-r) - 2$. Hence, F_n contains $\cup_{i\in I}[v_i]$, so its measure is at least $\mu(F_n) \geq (n(s-r)-2)e_n(\mu) = (s-r)ne_n(\mu) - 2e_n(\mu)$. Hence, $ne_n(\mu) \xrightarrow[n\to\infty]{} 0$, a contradiction.

$\square$

7.4 Examples

7.4.1 Classical symbolic systems

Because of Theorem 7.3.3, every *Sturmian* infinite word (see Definition 1.2.13) has uniform frequencies. Exercises 7.12, 7.13 and 7.14 in Section 7.7 deal with some applications of the results of Section 7.3 to some known families of infinite words.

It follows from Theorem 1.6.9 above that a *fixed point of a primitive substitution* has uniform frequencies: this was shown in (Queffélec 1987), and, from the same proof, we can deduce that these words satisfy a strong version of the ne_n condition of Boshernitzan (see also Exercise 7.14). Thus the Thue–Morse word (Example 1.2.21), the Fibonacci word (Example 1.2.22), or the *Rudin–Shapiro word* (defined as the fixed point beginning with a of the substitution $a \mapsto ab$, $b \mapsto ac$, $c \mapsto db$, $d \mapsto dc$) have uniform frequencies.

For the Chacon word defined after Proposition 1.4.6, the substitution is not primitive, but this word has also uniform frequencies; this can be proved directly, see Exercises 1.8 and 7.5; we could also notice that the complexity is $2n-1$ for $n \geq 2$ (Ferenczi 1995) and apply Theorem 7.3.3, or notice that the dynamical system is topologically isomorphic to the symbolic system associated with the fixed point of a primitive substitution (Ferenczi 1995) and use Theorem 1.6.9. Note that Proposition 6.5.6 generalises this fact, implying that the fixed points of many non-primitive substitutions have uniform frequencies.

7.4.2 Non-uniformly recurrent words

It is a common mistake to believe that uniform frequencies imply uniform recurrence. A simple counter-example is the word $x = baaa\cdots$, which has uniform frequencies ($f_w(x) = 1$ if $w = a^k$ for any k, $f_w(x) = 0$ if w is any other word) but is not uniformly recurrent (b occurs only once).

A more elaborate counter-example was suggested by E. Lesigne: let y be an infinite word which is known to be uniformly recurrent and to have uniform frequencies, for example the Thue–Morse word. Let w be a finite word not occurring in y, for example here $w = aaa$. We build a new infinite word x by inserting the word w into y at places along a sequence of uniform density 0: for example here, for every $k \geq 0$, we put

(i) $x_{2^k+3k+i} = a$, $i = 1, 2, 3$,
(ii) $x_{2^k+3k+3+j} = y_{2^k+j}$, $1 \leq j \leq 2^{k+1} - 2^k$.

Then x is not uniformly recurrent as w does not occur with bounded gaps; but if we compute $\frac{|x_k\cdots x_{k+n}|_{w'}}{n+1}$ for a finite word w' occurring in x, we see that it converges uniformly to $\lim_{n\to+\infty}\frac{|y_k\cdots y_{k+n}|_{w'}}{n+1}$, which is 0 if w' does not occur in y, and is known to exist otherwise. Hence, x has uniform frequencies. See also Exercise 7.16.

This example is fairly typical, and assuming that uniform frequencies imply uniform recurrence is not a very serious mistake, as

Proposition 7.4.1 *Let x be an infinite word with uniform frequencies, with*

$$f_w(x) = \lim_{n\to+\infty}\frac{|x_k\cdots x_{k+n}|_w}{n+1}\ .$$

Then, there exist infinite words y such that $L(y) = \{w \in L(x) \mid f_w(x) > 0\}$. Any of these y is uniformly recurrent and has uniform frequencies, with

$$f_w(y) = \lim_{n\to+\infty}\frac{|y_k\cdots y_{k+n}|_w}{n+1} = f_w(x)$$

for every $w \in L(y)$.

Moreover, (X_y, S) is the only minimal subshift of (X_x, S), and X_y has full measure in X_x for the invariant measure. For any $z \in X_x$, there exists an infinite sequence n_k such that $\lim_{k\to+\infty} S^{n_k}(z) \in X_y$.

Proof The subshift (X_x, S) is uniquely ergodic, let μ denotes its unique invariant probability measure. The language $L = \{w \in L(x) \mid f_w(x) > 0\}$ is factorial and extendable, hence Proposition 7.1.2 ensures the existence of a subshift $Y \subseteq X_x$ such that $L(Y) = L$.

Let Z be a subshift which is included in X_x. Proposition 7.2.4(i) ensures that (Z, S) admits an invariant measure ν, which can be considered as an invariant measure on X_x, hence $\mu = \nu$. In particular, ν gives positive measure to the cylinders $[w]$ for any $w \in L$, hence $L(Y) \subseteq L(Z)$, hence $Y \subseteq Z$. Therefore, Y is the only minimal subshift which is included in X_x.

Let y be any element of Y: y is uniformly recurrent, and since $y \in X_x$, it has uniform frequencies with $f_w(y) = f_w(x)$ for any $w \in L(y) \subseteq L(x)$. The last assertion comes from the fact that for any $z \in X_x$, X_z contains $Y = X_y$. □

Corollary 7.4.2 *If x is an infinite word with uniform frequencies, x is uniformly recurrent if, and only if, $f_w(x) > 0$ for every factor w of x.*

Proof X_x is minimal if, and only if, $X_x = X_y$, or $L(x) = L(y)$. □

A minimal and uniquely ergodic dynamical system is sometimes called *strictly ergodic*; the symbolic translation of this notion could be *positive uniform frequencies.*

7.4.3 Positive entropy and Grillenberger words

The *topological entropy* of a symbolic dynamical system (X_x, S) can be defined as

$$h(x) = \lim_{n \to +\infty} \frac{\log p_x(n)}{n}.$$

For a word with uniform frequencies, the same limit is also the *metric* (or *measure-theoretic*) *entropy* of the system (X_x, S, μ) equipped with its unique invariant probability measure. All the examples in Sections 7.4.1 and 7.4.2 have entropy zero.

Though the Jewett–Krieger theorem (see Proposition 7.2.12) implies that there exist uniquely ergodic systems of every given topological entropy, it seems difficult to find an infinite word with uniform frequencies and positive entropy, which implies exponential complexity. The standard way to ensure exponential complexity is to concatenate words independently: starting with a family of words $B_1, \ldots, B_r$, we decide that our infinite word will have all the factors $B_{i_1} \cdots B_{i_s}$ for some s and every sequence $(1 \leq i_1 \leq r, \ldots, 1 \leq i_s \leq r)$, and then iterate this process. But the resulting infinite word will not have uniform recurrence or frequencies (if we start from all the 1-letter words, we get the counter-examples of Section 7.5.1).

The first explicit examples with uniform frequencies (and recurrence) and exponential complexity (indeed, with arbitrarily high entropy) were built by

F. Hahn and Y. Katznelson (Hahn and Katznelson 1967); a much simpler construction is due to C. Grillenberger (Grillenberger 1972), who proved

Proposition 7.4.3 (Grillenberger 1972) *For every integer k and every real number $0 \leq h < \log k$, there exists a uniformly recurrent infinite word x on an alphabet of k letters which has uniform frequencies and for which $h(x) = h$.*

The basic construction falls into the general framework of *adic* words studied in Section 7.5.3; it is a clever replacement of the independent concatenation mentioned above by permutations, and of exponentials by factorials; indeed, to build a one-sided infinite word x (as a slight variation from (Grillenberger 1972) which considers two-sided infinite words), we build inductively families of words $B_{n,i}$, $1 \leq i \leq k_n$, where $B_{0,i} = i$, $1 \leq i \leq k$, and, for any permutation π on $\{1, \ldots, k_n\}$

$$B_{n+1,\pi} = B_{n,\pi(1)} \cdots B_{n,\pi(k_n)},$$

the permutations are then ordered lexicographically to number the new words from $B_{n+1,1}$ to $B_{n+1,k_n!}$.

Then the infinite word x beginning with $B_{n,1}$ for every n has uniform frequencies and recurrence and, if $k \geq 3$, exponential complexity. The infinite words in Proposition 7.4.3 are then deduced from one of these x by applying a suitable substitution. Note that on two letters, the above construction yields the Thue–Morse word, which is of entropy zero: if we want exponential complexity on two letters, we can start with the above construction on three letters, and replace them by aa, ab, ba.

7.4.4 Interval exchange maps

Interval exchange maps are one-dimensional geometrical systems introduced by V. I. Oseledec (Oseledec 1966), for the study of which symbolic dynamics proved to be a very efficient tool.

Definition 7.4.4 Let $r \geq 3$. Let Λ_r be the set of vectors $(\lambda_1, \ldots, \lambda_r)$ in $\mathbb{R}^r$ such that $0 \leq \lambda_i \leq 1$ for all i and $\Sigma_{i=1}^r \lambda_i = 1$. An *$r$-interval exchange map* is given by a vector $\lambda \in \Lambda_r$ and a permutation π of $\{1, 2, \ldots, r\}$. The map $T_{\lambda,\pi}$ is the piecewise translation defined by partitioning the interval $[0, 1)$ into r sub-intervals of lengths $\lambda_1, \lambda_2, \ldots, \lambda_r$ and rearranging them according to the permutation π or, formally,

$$\Delta_i = \left[\sum_{j<i} \lambda_j, \sum_{j \leq i} \lambda_j \right),$$

$$T_{\lambda,\pi}\xi = \xi + \sum_{\pi^{-1}j<\pi^{-1}i} \lambda_j - \sum_{j<i} \lambda_j$$

if $\xi \in \Delta_i$.

The map $T_{\lambda,\pi}$ satisfies the *i.d.o.c. property* (Keane 1975) if the negative orbits of the discontinuity points $\sum_{j\leq i} \lambda_j$, $1 \leq i \leq r-1$, are infinite and disjoint.

Warning: roughly half the texts on interval exchange maps re-order the subintervals by π^{-1}; as it is not always clear to which half a given text belongs, we insist that the present definition corresponds to the following ordering of $T\Delta_i$: from left to right, $T\Delta_{\pi(1)}, \ldots, T\Delta_{\pi(r)}$. It makes sense to re-order also the Δ_i, thus defining T by two permutations π_0 and π_1 (though of course sometimes π_0^{-1} and π_1^{-1} are used...), see (Yoccoz 2006) for example. Though it might have come useful in the definition of Keane's examples in Section 7.5.2.2 below, we prefer to stick to one permutation.

Definition 7.4.5 A *natural coding* of an r-interval exchange map is any of the words $x(\xi)$ for a point $\xi \in [0,1)$, where $x_i(\xi) = j$ whenever $T^i\xi \in \Delta_j$.

If x is a natural coding of $T_{\lambda,\pi}$, we can consider the symbolic system (X_x, S). Though this system is not topologically conjugate to $([0,1), T_{\lambda,\pi})$, it shares all its properties of minimality and unique ergodicity, and any invariant measure for one of these systems can be carried to the other one. The i.d.o.c. condition ensures that (X_x, S) and $([0,1), T_{\lambda,\pi})$ are minimal, each natural coding x is uniformly recurrent and the language $L(x)$ is the same for all the natural codings. Then the complexity function of any natural coding is $(r-1)n+1$ and thus the number of ergodic invariant probability measures on (X_x, S) or $([0,1), T_{\lambda,\pi})$ is at most $r-2$ by Theorem 7.3.4 (geometrical methods can improve this bound to $\frac{r}{2}$(Katok 1973)).

We shall call m the normalised Lebesgue measure on Λ_r, and μ the Lebesgue measure on $[0,1)$. It is proved in (Keane 1975) that for an *irreducible* permutation π ($\pi\{1,\ldots k\} \neq \{1,\ldots k\}$ for every $k < r$) the i.d.o.c. property is implied by the *total irrationality* of the λ_i (the λ_i have no rational relation except $\lambda_1 + \cdots + \lambda_r = 1$), and thus for m-almost every $\lambda \in \Lambda_r$, $T_{\lambda,\pi}$ satisfies it, and hence is minimal, or equivalently the $x(\xi)$ are uniformly recurrent.

It follows from Theorem 7.3.3, and it is stated in (Keane 1975), that

Proposition 7.4.6 *For three-interval exchange maps the i.d.o.c. condition implies unique ergodicity (hence uniform frequencies for the natural codings).*

It was conjectured by M. Keane (Keane 1975) that the i.d.o.c. condition implies unique ergodicity for every r. After this conjecture was disproved (see Section 7.5.2.1), a weaker result was considered as a question by the same author (Keane 1977) and proved independently by H. Masur (Masur 1982) and W. Veech (Veech 1982).

Theorem 7.4.7 *For a given irreducible π, $T_{\lambda,\pi}$ is uniquely ergodic (or equivalently the $x(\xi)$ have uniform frequencies) for m-almost every $\lambda \in \Lambda_r$.*

The proofs of W. Veech and H. Masur use deep geometrical methods; but a later proof of M. Boshernitzan uses mainly combinatorial methods; it is published in (Boshernitzan 1985) but can be simplified (and made completely combinatorial) by using the criteria of (Boshernitzan 1984) (Boshernitzan 1992) described in Section 7.3. Thus we give here this simplified proof, with an updated vocabulary.

Proposition 7.4.8 (Boshernitzan 1985) *Let $U_{n,\varepsilon}$ be the set of $\lambda \in \Lambda_r$ such that $e_n(T_{\lambda,\pi}) \leq \frac{\varepsilon}{n}$ (see Section 7.3.3). If $0 < \varepsilon < \frac{1}{r}$, $m(U_{n,\varepsilon}) \leq 3r^3\varepsilon$.*

Proof Let $G_n(\lambda)$ denote the Rauzy graphs (see Definition 7.3.5) of length n of any infinite word $x(\xi)$ in the natural coding of $T_{\lambda,\pi}$. As the complexity of any $x(\xi)$ is $(r-1)n+1$, $G_n(\lambda)$ has at most $3r-3$ branches.

The weight functions of Definition 7.2.5 can be carried over to the Rauzy graphs: ψ is a *weight function on a graph* if it is positive on each vertex, the sum of its values on vertices is 1, and it can be extended to the edges such that for every vertex w

$$\psi(w) = \sum_{\text{incoming edges}} \psi(e) = \sum_{\text{outgoing edges}} \psi(e).$$

And the function ψ_λ, defined on the vertices of $G_n(\lambda)$ by associating with the vertex $w_1 \cdots w_n$ the measure of the cylinder, $\mu[w_1 \cdots w_n]$, is a weight function on the graph $G_n(\lambda)$; the weight of an edge $w_1 \cdots w_{n+1}$ is also $\mu[w_1 \cdots w_{n+1}]$.

We fix now a Rauzy graph G of length n; let $\Lambda(G)$ be the set of $\lambda \in \Lambda_r$ such that $G_n(\lambda) = G$. For a given word $w = w_1 \cdots w_n$, $\psi_\lambda(w_1)$ is just λ_{w_1}; for all $\lambda \in \Lambda(G)$, all the Rauzy graphs $G_i(\lambda)$, $1 \leq i \leq n$, are fixed, and when we look at the defining equalities of the weight function ψ_λ on $G_i(\lambda)$, we see that the measures of cylinders of length $i+1$ are computed by explicit formulas from those of length i; thus the numbers $\psi_\lambda(w_1 \cdots w_i)$, $1 < i \leq n$, can be computed inductively; they depend linearly on λ. Because $T_{\lambda,\pi}$ preserves the measure μ, $\psi_\lambda(w_1 \cdots w_n) = \psi_\lambda(w'_1 \cdots w'_n)$ if $w_1 \cdots w_n$ and

$w'_1 \cdots w'_n$ are on the same branch of G; hence, for fixed λ, $\psi_\lambda(w_1 \cdots w_n)$ takes $1 \leq t \leq 3r-3$ values, which we denote by $\varphi_1(\lambda), \dots, \varphi_t(\lambda)$; the φ_j are linear functionals, $e_n(T_{\lambda,\pi})$ is just the smallest of the $\varphi_j(\lambda)$, $1 \leq j \leq t$. Furthermore, again through the defining equalities of the successive weight functions on $G_i(\lambda)$, we can retrieve λ from the values $\psi_\lambda(w)$ on all the vertices of G; thus $\Lambda(G)$ is a convex set and every weight function ψ on G yields a $\lambda \in \Lambda(G)$ such that $\psi_\lambda = \psi$.

We want to estimate the measure of $\{\lambda \in \Lambda(G) \mid \varphi_i(\lambda) \leq \frac{\varepsilon}{n}\}$; for this, we use a general result for which we refer the reader to (Boshernitzan 1985), Corollary 7.4: *If φ is the restriction of a linear functional to a convex set K of dimension d, taking values between 0 and A, then, if V denotes the volume,*

$$V(\varphi^{-1}[0,B)) \leq \frac{dB}{A} V(K).$$

We apply it with $K = \Lambda(G)$, restricting ourselves to those with $m(\Lambda(G)) > 0$, $\varphi = \varphi_i$, $B = \frac{\varepsilon}{n}$; the dimension is $r-1$, the volume is the Lebesgue measure; we need an estimate on A; for this, we claim that *for each vertex s of G, there exists a weight function such that $\psi(s) \geq \frac{1}{rn}$.* To do this, we choose a $\lambda \in \Lambda(G)$ such that $T_{\lambda,\pi}$ is minimal, which is possible as $m(\Lambda(G)) > 0$; this implies that G is strongly connected and thus we can find a loop $s = s_0 \to \cdots \to s_k \to s_0$ in G; by taking it of minimal length, we ensure it has no repetition. Then we define ψ' to be $\frac{1}{k+1}$ on the s_i and 0 on the other vertices; ψ' is not a weight function as it may be 0 on some vertices, but $\psi = (1-\delta)\psi' + \delta\psi_\lambda$ is a weight function, and as $k \leq (r-1)n+1$ we can choose δ such that our claim is proved.

Thus we have $A \geq \frac{1}{rn}$, and thus, for all G with $m(\Lambda(G)) > 0$ and hence for all G,

$$m(\{\lambda \in \Lambda(G) \mid \varphi_i(\lambda) \leq \frac{\varepsilon}{n}\}) \leq (r-1)r\varepsilon m(\lambda(G)).$$

As $t \leq 3r-3$,

$$m(\{\lambda \in \Lambda(G) \mid \min_{1 \leq i \leq t} \varphi_i(\lambda) \leq \frac{\varepsilon}{n}\}) \leq 3(r-1)^2 r\varepsilon m(\lambda(G)),$$

which implies the proposition. □

Proof of Theorem 7.4.7 If $0 < \varepsilon < \frac{1}{r}$ and $n \geq 1$, we put $V_{n,\varepsilon} = \Lambda_r \setminus U_{n,\varepsilon}$, and $V_\varepsilon = \cap_{N \geq 1} \cup_{n > N} V_{n,\varepsilon} \cap \{\lambda \mid T_{\lambda,\pi}$ is i.d.o.c.$\}$.

If λ is in V_ε, there are infinitely many n such that $e_n(T_{\lambda,\pi}, \mu) \geq \frac{\varepsilon}{n}$, hence $ne_n(T_{\lambda,\pi}, \mu) \not\to 0$ when $n \to +\infty$, and $T_{\lambda,\pi}$ is uniquely ergodic by Theorem 7.3.8. Thus $m(\{\lambda \mid T_{\lambda,\pi}$ is uniquely ergodic$\})$ is at least $m(V_\varepsilon) \geq 1 - 3r^3\varepsilon)$, and thus is one as ε is arbitrary. □

The above proof does not use any of the geometrical properties of interval exchange maps.

Explicit examples of infinite words with uniform frequencies coming from coding of four-interval exchanges can be deduced from (Keane 1977) or found in (Ferenczi and Zamboni 2008) or (Cheung and Masur 2006). Examples for higher number of intervals can be deduced from (Sataev 1975), see Proposition 7.5.3 below.

7.5 Counter-examples

7.5.1 The full shift

As we have seen in Section 7.4, *non-uniformly recurrent* words and *words of positive entropy* are generally expected not to have uniform frequencies, unless they have been built for the specific purpose of having them.

Typical examples falling into both these categories are the words with full complexity $p(n) = k^n$ on any finite alphabet A of cardinality k, such as the *Champernowne word* $011011100101110\cdots$, built by concatenating the expansions in base 2 of $0, 1, 2, \ldots, n, \ldots$: they do not have uniform frequencies, and their associated symbolic system is the *full shift* $A^{\mathbb{N}}$.

Proposition 7.5.1 *The full shift has uncountably many ergodic invariant measures.*

Proof Take a probability vector $\pi = (\pi_1, \ldots, \pi_k)$ and assign to cylinders the measure $\mu_\pi([w_1 \cdots w_n]) = \pi_{w_1} \cdots \pi_{w_n}$. These measures are ergodic (see Exercise 7.17). □

The system $(A^{\mathbb{N}}, \mathcal{B}, \mu_\pi, S)$ is then called a (one-sided) *Bernoulli shift*. But there are lots of other ergodic invariant measures for the full shift, for examples the Dirac measure on each periodic orbit (see Exercise 7.11), or the measures arising from the uniquely ergodic examples given in Section 7.4 (that can be considered as invariant measures on the full shift defined on the same alphabet).

The non-trivial *subshifts of finite type* (where we consider all the infinite words in which a prescribed set of finite words does not occur) and the *sofic systems* (which constitute the closure of the subshifts of finite type for a natural notion of homomorphism), defined in (Weiss 1973) behave like the full shift in having positive entropy, with no uniform recurrence or frequencies. See Section 1.6 and 2.3.1 for definitions.

7.5.2 Interval exchange maps again

7.5.2.1 Veech's counter-examples

Systems which are minimal and not uniquely ergodic are not so easy to build; the famous examples of H. Furstenberg (Furstenberg 1961) are defined on multi-dimensional tori, and do not give rise naturally to infinite words.

Then came the examples of W. Veech (Veech 1969), which use ideas of (Furstenberg 1961) together with very involved arithmetic considerations:

Theorem 7.5.2 (Veech 1969) *For two irrationals α and β, let $f = -2\chi_{[0,\beta)} + 1$ and T be the map on $\mathbb{R}/\mathbb{Z} \times \{-1, 1\}$ defined by*

$$T(\xi, e) = (\xi + \alpha, f(\xi)e).$$

If β is not of the form $p\alpha + q$ with p and q integers (independence condition)*, T is minimal.*

If α and β satisfy the coboundary condition

$$ef(\xi)g(\xi) = g(\xi + \alpha),$$

for every $\xi \in [0, 1)$, some measurable function g and number $e = \pm 1$, T is not uniquely ergodic.

For each α with unbounded partial quotients, there exist uncountably many numbers β such that both these conditions are satisfied.

This result does not appear as such in the paper (Veech 1969). Indeed, it is written under a partly symbolic form with T replaced by the map $\overline{T}(y, e) = (S(y), y_0.e)$ on $X_x \times \{-1, 1\}$, where $x = (x_n)_n$ is defined as follows: we fix $\xi \in [0, 1)$, and $x_n = -1$ if $\xi + n\alpha$ (modulo 1) falls into $[0, \beta)$, $x_n = +1$ otherwise; hence, $X_x \subset \{-1, 1\}^{\mathbb{N}}$. It is then straightforward, though not written in (Veech 1969), that, for α and β satisfying the conditions, the iterates of $(x, 1)$ give an infinite word on the alphabet $\{(-1, -1), (1, -1), (-1, 1), (1, 1)\}$ which is uniformly recurrent and does not have uniform frequencies, and these may be considered as the first infinite words with these properties.

If now we replace $\mathbb{R}/\mathbb{Z} \times \{-1\}$ with $[0, 1)$ and $\mathbb{R}/\mathbb{Z} \times \{1\}$ with $[1, 2)$, and normalise by 2, we see that T is also a *five-interval exchange map* defined (if for example $\beta < 1-\alpha$) by $\lambda = \frac{1}{2}(\beta, 1-\alpha-\beta, \alpha+\beta, 1-\alpha-\beta, \alpha)$ and the permutation $1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 5, 4 \mapsto 1, 5 \mapsto 4$. Under the independence condition, T satisfies the i.d.o.c. condition; and its natural coding on a five-letter alphabet is another uniformly recurrent infinite word, without uniform frequencies if α and β satisfy the coboundary condition.

We can look at another map T', *induced* (see Definition 7.5.4 below) on $[0, \beta)$ by the rotation of angle α. The map T' is a three-interval exchange map, satisfying the i.d.o.c. condition if α and β satisfy the independence condition. It is an implicit consequence of (Veech 1969), stated in (Veech 1984), that the coboundary condition is equivalent to an unexpected property, namely that T' has -1 as an eigenvalue, meaning that there exists g in $L^2([0,1))$ with $g \circ T' = -g$. A direct proof of this property is given in (Ferenczi, Holton, and Zamboni 2004).

We can also look at the map T'', which is an *exduction* on $[0, \beta)$ of the rotation of angle α: namely $T''\xi = \xi + 1$ for $0 \leq \xi < \beta$, and for $\beta \leq \xi < 1 + \beta$, $T''\xi$ is the representative of $\xi + \alpha$ modulo 1 which falls into $[0, 1)$. It is also a three-interval exchange map, defined after normalisation (if for example $\beta < 1 - \alpha$) by $\lambda = \frac{1}{1+\beta}(\beta, 1 - \alpha - \beta, \alpha + \beta)$ and the permutation $1 \mapsto 3, 2 \mapsto 2, 3 \mapsto 1$. It can be viewed as a dual version of T' (as the rotation is an induced map of T'') and shares the same properties. The map T'' was introduced in (Keynes and Newton 1976) in order to exhibit T''^2 as a non-uniquely ergodic five-interval exchange map.

The coboundary condition is studied further by Y. Cheung (Cheung 2003) where, for fixed β, estimates are given for the Hausdorff dimension of the set of α for which α and β satisfy it; the non-unique ergodicity of T is also seen as the non-ergodicity of some directions for a *billiard flow.*

A nice generalisation of Veech's result appeared a few years later in a paper of E. Sataev. The lack of communication between West and East at that time explains that Sataev apparently did not know the paper (Veech 1969) and that in turn (Sataev 1975) was widely ignored thereafter.

Proposition 7.5.3 (Sataev 1975) *For any integer $r \geq 2$ and any integer $1 \leq k \leq r$, there exist an irrational α, $r-1$ disjoint intervals $I_j \subseteq [0, 1)$ and $r - 1$ different permutations π_j of $\{1, \ldots, r\}$, $1 \leq j \leq r - 1$, such that U is minimal and has exactly k ergodic invariant probability measures, where U is the map on $\mathbb{T}^1 \times \{1, \ldots, r\}$ defined by $U(\xi, e) = (\xi + \alpha, h(\xi)e)$, and $h(\xi)$ is the permutation π_j when ξ is in I_j and the identity elsewhere.*

This gives exchanges of at least r^2 and at most $2r^2$ intervals which have a prescribed number $1 \leq k \leq r$ of ergodic invariant probability measures.

7.5.2.2 Keane's counter-examples

Then M. Keane (Keane 1977) lowered the number of intervals required for a counter-example to four, which is optimal in view of Proposition 7.4.6. But his paper uses very different techniques, and there appear for the first time two ideas which were to be named and systematically studied later:

one is the *induction*, a different form of which will give the *Rauzy induction* and is the starting point of the geometrical methods mentioned in Section 7.4.4; the other one is the use of *matrices for adic systems*, which will be developed in the next section.

We recall that

Definition 7.5.4 If T is map from a set X to itself, and A a subset of X, the *induced map* of T on A is defined by $T_A(z) = T^{r_A(z)}z$, with $r_A(z) = \min\{n > 0 \mid T^n z \in A\}$, for all $z \in A$ for which $r_A(z)$ is finite.

The induction idea works as follows: M. Keane takes a four-interval exchange for the permutation (in our notation, see Section 7.4.4 above) $1 \to 4, 2 \to 2, 3 \to 1, 4 \to 3$, denoted by π, with a probability vector λ as yet unknown.

Some inequalities on the λ_i ensure that the induced map of $T_{\lambda,\pi}$ on the fourth interval Δ_4 is well defined and is another four-interval exchange map which, after renormalisation and a renumbering of the intervals which reverses their order, can be defined by the permutation π and a vector λ' such that $\lambda = \mathbf{A}_{m,p}\lambda'$, where m and p are integers and $\mathbf{A}_{m,p}$ is the matrix

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ m-1 & m & 0 & 0 \\ p & p & p-1 & p \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

This induction process is then iterated.

Proposition 7.5.5 (Keane 1977) *For every infinite sequence of matrices* $\mathbf{A}_{m_k,p_k}$*, if* P *is the positive cone in* $\mathbb{R}^4$*, the set* $\cap_{k\in\mathbb{N}}\mathbf{A}_{m_1,p_1}\cdots\mathbf{A}_{m_k,p_k}P$ *is non-empty.*

Let E *be this set normalised by* $\lambda_1 + \cdots + \lambda_4 = 1$*. For every* $\lambda \in E$*, the four-interval exchange* $T_{\lambda,\pi}$ *is such that, if we iterate* k *times the induction on the fourth interval, after renormalising, and reversing the order if* k *is even, we get the four-interval exchange* $T_{\lambda'_{(k)},\pi}$*, with* $\lambda = \mathbf{A}_{m_1,p_1}\cdots\mathbf{A}_{m_k,p_k}\lambda'_{(k)}$*.*

The matrix part of (Keane 1977) will be stated in greater generality as Proposition 7.5.10; it shows that under mild conditions on the m_k and p_k, T is not uniquely ergodic. Minimality for this example can be realised through the i.d.o.c. condition, and also with the stronger requirement of total irrationality, which was not satisfied by Veech's examples.

Further examples of non-uniquely ergodic four-interval exchanges can be found in (Marmi, Moussa, and Yoccoz 2005) and

(Ferenczi and Zamboni 2008), and a generalisation of Keane's examples to r intervals has been done in (Yoccoz 2005) for any $r \geq 4$. In all cases, by natural coding we get uniformly recurrent words without uniform frequencies.

7.5.3 Adic words and languages

7.5.3.1 Adic systems in the symbolic framework

The *Bratteli–Vershik dynamical systems*, initially called *adic systems* by A. Vershik , are defined and studied at length in Chapter 6 of the present book. They do not fit into the framework of symbolic dynamics. Indeed, they are defined on what looks like a space of infinite words, but not as a shift, and some of them are not topologically isomorphic to a symbolic system since they are not *expansive* (see Section 6.5.3). For many of them, however, there is a standard way to code them into a well-defined subshift, see the discussion after Proposition 7.5.8, and we take this coded form as a (somewhat pedestrian) definition of what we call a *symbolic adic system*. More precisely we define an *adic infinite word* as a word whose language is generated by a finite number of families of words, build by recursive concatenation rules:

Definition 7.5.6 An infinite word x is *adic* if there exist finite words $B_{n,1}, \ldots, B_{n,k_n}$, for $n \in N$, such that $L(x)$ is the set of words w for which there exist n and i such that w is a factor of $B_{n,i}$, with the additional conditions

(i) $B_{0,i} = i$, $1 \leq i \leq k_0$,

(ii) for each $1 \leq i \leq k_n$, there exist an integer $t(n,i) > 0$, and $t(n,i)$ integers $1 \leq k_s(n,i) \leq k_{n-1}$ such that
$$B_{n,i} = \Pi_{s=1}^{t(n,i)} B_{n-1,k_s(n,i)},$$

(iii) for every p there exists $N(p)$ such that for every word w in $L(x)$ of length at least $N(p)$ and every decomposition
$$w = U\Pi_{j=1}^{r} B_{p,l_j} V$$
where U is a (proper, possibly empty) suffix of some B_{p,l_0} and V is a (proper, possibly empty) prefix of some $B_{p,l_{r+1}}$, U and the l_i, $1 \leq i \leq r$, depend only on w.

The *nth matrix* $\mathbf{M}_n(x)$ of the adic word is the matrix which has on its ith line, $1 \leq i \leq k_{n-1}$, and jth column, $1 \leq j \leq k_n$, the number of $1 \leq s \leq t(n,j)$ such that $k_s(n,j) = i$.

The notation B is for *block* which is more used by ergodicians than *word*, and the matrix counts the number of $B_{n-1,i}$ which appear in the defining formula for $B_{n,j}$.

The third condition in the definition is called a condition of *recognisability*; though it is cumbersome to write, it is generally easy to check.

We have already seen examples of adic words in Section 7.4.3, with those of the Grillenberger words which are built explicitly with blocks $B_{n,i}$. As has been stated, the examples of Section 7.5.2.2 fall into this category: it is an easy consequence of Proposition 7.5.5 that (with the notations of Section 7.5.2.2)

Proposition 7.5.7 *For any λ in E, the natural codings of $T_{\lambda,\pi}$ are adic words with $k_n = 4$ for all n and $\mathbf{M}_n = \mathbf{A}_{m_n,p_n}$.*

More generally, natural codings of interval exchange maps are adic words, with explicit constructions being given for some permutations in (Ferenczi, Holton, and Zamboni 2003) (Ferenczi and Zamboni 2009) (Ferenczi and Zamboni 2008); another construction can be deduced from Section 6.5.6.

Recent general results on Bratteli–Vershik systems compute all their invariant probability measures, see for example (Fisher 2009), (Bezuglyi, Kwiatkowski, Medynets, et al. 2009), and Section 6.8 with further references. We give here a simple particular case of these results, adapted to the needs of the present section, together with a sketch of a self-contained proof which does not need the whole Bratteli–Vershik machinery, but still may be skipped by readers who are more interested in word combinatorics than in dynamical systems.

Proposition 7.5.8 *Let x be an adic word with matrices $\mathbf{M}_n$, such that $k_n = k$ and $\det \mathbf{M}_n \neq 0$ for every n, and*

$$\lim_{n\to+\infty} \min\{\Sigma_{a\in C}a \mid C \textit{ column of } \mathbf{M}_1 \cdots \mathbf{M}_n\} = +\infty.$$

If P is the positive cone in $\mathbb{R}^k$, each point μ in the set $\cap_{n\in\mathbb{N}}\mathbf{M}_1 \cdots \mathbf{M}_n P$, normalised by $\mu_1 + \cdots + \mu_k = 1$, defines an invariant probability measure on (X_x, S) such that $\mu[i] = \mu_i$; every invariant probability measure on (X_x, S) is of that form, and at most k of them are ergodic.

Proof The recognisability condition ensures that for every n, every infinite word y in X_x admits a unique infinite decomposition $UB_{n,l_1}B_{n,l_2}\cdots$; for $1 \le i \le k$ we define the set $F_{n,i}$ to be the set of $y \in X_x$ for which the suffix U is empty and $l_1 = i$. We check that for each given n, the $S^j F_{n,i}$, $1 \le i \le k$,

$0 \le j \le |B_{n,i}| - 1$ form a partition of X_x (which is indeed a *Kakutani–Rokhlin partition*, see Definition 6.4.1); these partitions are increasing (the atoms of the $(n+1)$th partition are subsets of atoms of the nth partition), and, except possibly for a countable number of points, two points belonging to the same atom of the nth partition for every n are the same (these infinite words coincide on arbitrarily long initial segments, as the condition on the columns of $M_1 \cdots M_n$ ensures that the $|B_{n,i}|$ go to infinity with n); the system (X_x, S) is indeed a Bratteli–Vershik dynamical system and moreover it is of *finite rank*, see (Ferenczi 1997). This is enough to ensure that a measure on X_x is determined by its values on the atoms of these partitions, thus, if it is S-invariant, by its values on the $F_{n,i}$.

It follows from the definitions that $F_{n-1,i}$ is a union of images by S of the $F_{n,j}$, $1 \le j \le k$, with an iterate of $F_{n,j}$ appearing in $F_{n-1,i}$ whenever some $k_s(n,i)$ is equal to j in the defining decomposition of $B_{n,i}$ thus, if $\rho_n = (\mu(F_{n,1}), \ldots, \mu(F_{n,k}))$, we get $\rho_{n-1} = M_n \rho_n$. Thus the measure μ is completely determined by the vector ρ_0, and the space of such vectors is of dimension at most k; we can define such a measure if, and only if, all the vectors ρ_n have positive coordinates; thus, after normalising, we get the claimed result. □

Let us mention that if we describe a Bratteli–Vershik dynamical system through a sequence of Kakutani–Rokhlin partitions as in Theorem 6.4.3, this gives immediately a symbolic system as in Definition 7.5.6, by putting the letter l at the jth place of the word $B_{n,i}$ whenever (with the notation of Theorem 6.4.3) $T^j B_i(n)$ falls into $B_i(0)$. Unfortunately, the symbolic system we get is not always isomorphic to the system defined from the KR-partitions, as happens when we start from the dyadic odometer (see Section 6.5.1), which is aperiodic but gives rise to an infinite word defined by $k_n = 1$, $B_0 = 0$, $B_{n+1} = B_n B_n$, which is just the periodic infinite word $0000 \cdots$; the isomorphism does however work whenever the words we get satisfy the recognisability condition, which is generally the case - though obviously not for the B_n of the dyadic odometer.

7.5.3.2 Some families of examples

For an adic word, minimality can be ensured by mild properties of *primitivity* of the matrices:

Proposition 7.5.9 *Let x be an adic word with matrices $\mathbf{M}_n$, such that, for every n, there exists $m \ge n$ such that all the entries of $\mathbf{M}_n \cdots \mathbf{M}_m$ are positive; then x is uniformly recurrent.*

Proof Exercise 7.20. □

This gives a cornucopia of uniformly recurrent words without uniform frequencies, by ensuring there is more than one normalised element in $\cap_{n\in\mathbb{N}} M_1 \cdots M_n P$. And first we can state the result which is implicitly proved in the matrix part of (Keane 1977):

Proposition 7.5.10 *An adic word with matrices* $\mathbf{M}_n = \mathbf{A}_{m_n,p_n}$*, with* $p_1 \geq 9$ *and* $3(p_n + 1) \leq m_n \leq \frac{1}{2}(p_{n+1} + 1)$ *for all* n*, does not have uniform frequencies.*

Thus the non-unique ergodicity of Keane's examples comes only from the adic structure of the system and Proposition 7.5.8, which is proved in (Keane 1977) in this particular case (the general case uses basically the same reasoning). And we see from Propositions 7.5.7 and 7.5.8 that there is a duality between the lengths of the intervals and the values of the invariant measures on them: with the notations of Section 7.5.2.2, for any $\lambda \in E$, every invariant probability measure on $([0, 1), T_{\lambda,\pi})$ is defined from a vector $\mu \in E$ by giving measure μ_i to the ith interval (it follows from the proofs of both Theorem 7.4.7 and Proposition 7.5.8 that the measures of the four initial intervals determine the measure μ completely). Under the above conditions on the (m_k, p_k), E is not reduced to a point but is a segment, whose two endpoints give the two ergodic invariant measures (note that here the adic structure predicts at most four ergodic invariant measures; but we have seen in Section 7.4.4 that, by Katok's result (Katok 1973), this can be reduced to two for a four-interval exchange). If we choose λ to be in the interior of this segment, these two ergodic measures are absolutely continuous with respect to the Lebesgue measure but different from it; if we choose λ to be an endpoint, one ergodic measure is the Lebesgue measure and the other one is singular; a recent work of J. Chaika (Chaika 2008) has proved that this singular measure can have a support of arbitrarily small Hausdorff dimension.

The non-unique ergodicity for the examples of (Ferenczi and Zamboni 2008) comes also directly from their adic definition.

The paper (Ferenczi, Fisher, and Talet 2009) is devoted to the building of examples with the lowest possible k_n.

Proposition 7.5.11 (Ferenczi, Fisher, and Talet 2009) *Any adic words with the following matrices* $\mathbf{M}_n$ *are uniformly recurrent without uniform frequencies:*

(i) $\begin{pmatrix} q_n & 1 \\ 1 & q_n \end{pmatrix}$ *if* $\Sigma_{n=0}^{+\infty} \frac{1}{q_n} < 1$.

(ii) $\begin{pmatrix} q_n & r_n \\ s_n & t_n \end{pmatrix}$ *if* $q_0 t_0 > r_0 s_0$, $\Sigma_{n=0}^{+\infty} \frac{q_n r_{n+1}}{r_n t_{n+1}} < +\infty$, $\Sigma_{n=0}^{+\infty} \frac{b_n s_{n+1}}{a_n q_{n+1}} < +\infty$, *where* (a_n, b_n) *is the first line of the matrix* $\mathbf{M}_0 \cdots \mathbf{M}_n$.

(iii) $\begin{pmatrix} q_n & 1 & 0 \\ 1 & q_n & 1 \\ 0 & q_n & 1 \end{pmatrix}$ *if* $\Sigma_{n=0}^{+\infty} \frac{1}{q_n} < 1$.

(iv) $\begin{pmatrix} q_n & 0 & q_n - 1 \\ r_n & r_n - 1 & r_n \\ 1 & 1 & 1 \end{pmatrix}$ *if* $r_0 \geq 6$ *and* $3r_n + 1 \leq 2q_n \leq r_{n+1}$ *for all* $n > 0$.

The first examples provide probably the simplest words which may be built with uniform recurrence but without uniform frequencies. The last family of examples is an abstract version of Keane's examples of Section 7.5.2.2 with $k = 3$; note that they are not natural codings of three-interval exchanges, because of Proposition 7.4.6.

In the above examples, uniform recurrence and uniform frequencies are ensured by sufficient conditions on the matrix only; however, in general uniform recurrence and uniform frequencies depend on the actual recursion formulas giving the words $B_{n,i}$, see Exercise 7.18.

7.5.3.3 Complexity and the Cassaigne–Kaboré word

The complexity of an adic word depends on the actual recursion formulas; let us mention that if $k_n = k$ for all n the complexity is sub-exponential (the topological entropy is 0), but examples can be built with $\limsup \frac{p_x(n)}{f(n)} = +\infty$ for any given f with subexponential growth: such examples are built in Proposition 3 of (Ferenczi 1996) under a slightly different form; with the notations of that paper, those built for bounded K can be defined as adic words with $k_n = K + 1$ by putting $B_{n,i} = B_n 1^{i-1}$, $1 \leq i \leq K + 1$.

At the other end of the spectrum J. Cassaigne and I. Kaboré (Cassaigne and Kaboré 2009) have investigated the words without uniform frequencies with the lowest possible complexity function.

Proposition 7.5.12 (Cassaigne and Kaboré 2009) *There exists a uniformly recurrent adic word without uniform frequencies such that* $p_x(n) \leq 3n + 1$ *and* $\liminf_{n \to +\infty} \frac{p_x(n)}{n} = 2$.

This tends to show that the bounds in Theorems 7.3.2 and 7.3.3 are optimal.

The example itself uses sequences of numbers $0 < l_n < m_n < p_n$, with l_n tending to infinity, $\frac{p_n}{m_n}$ growing fast enough, and $\frac{m_n}{l_n}$ faster enough; for example $l_n = 2^{4+2^{n+1}}$, $m_n = 2^{2^{n+3}}$, $p_n = 2^{5.2^{n+1}}$. Then, we define $B_{0,1} = 1$, $B_{0,2} = 2$, $B_{n+1,1} = B_{n,1}^{m_n} B_{n,2}^{l_n}$ and $B_{n+1,2} = B_{n,1}^{m_n} B_{n,2}^{p_n}$.

The infinite word x can be defined as the limit of the $B_{n,1}$ when $n \to +\infty$. Its complexity can be computed by using bispecial factors (see Section 4.5): we find that $p_x(n+1) - p_x(n)$ is 2 or 3, the value 2 being taken on intervals $[s_n, t_n]$ with $t_n \geq l_n s_n$, which proves the assertions. The uniform recurrence comes from Proposition 7.5.9, and the absence of uniform frequencies can be proved either by checking we are in the second class of examples in Proposition 7.5.11, or by imitating the simpler proof for the first class of examples in Proposition 7.5.11, or directly, see Exercise 7.19 below.

7.5.3.4 Earlier examples

X. Bressaud (unpublished) used the same formulas as J. Cassaigne and I. Kaboré but with l_n replaced by 1, m_n growing fast and p_n much faster than m_n; his word is also uniformly recurrent and without uniform frequencies.

A. Frid (CANT 2006, unpublished) suggested the following example: let x be the infinite word that is the limit of the finite words B_n defined by $B_0 = 1$ and $B_{n+1} = B_n \sigma(B_n)^{k_n}$, where $\sigma(w)$ denotes the word obtained by exchanging 2 and 1 in w.

If k_n grows sufficiently fast, then A. Frid showed that x is uniformly recurrent and does not have frequencies. More precisely (X_x, S) admits exactly two ergodic invariant measures. Indeed, this is an adic word if we put $B_{n,1} = B_n$, $B_{n,2} = \sigma(B_n)$; it looks different from the previous examples as the higher entries in the matrix are not on the diagonal, but x is also the limit of $B_{2n,1}$, and if we look only at the $B_{2n,1}$ and $B_{2n,2}$ we get an adic word whose matrix is of the same type as those in Proposition 7.5.11.

7.5.3.5 The Pascal-adic language

The reference for this whole section is (Méla and Petersen 2005). As we noticed in the proof of Proposition 7.5.8, all the adic examples of Section 7.5.3, which are built with a constant k_n, could also be described with the older notion of *finite rank* (Ferenczi 1997), and the adic terminology may seem to be just a more fashionable presentation; however, this terminology comes into its own when k_n is unbounded, and the following very interesting example could not have been defined within the framework of earlier notions. Note that here the dynamical system is not defined by one infinite word, but by a language; but this slight generalisation does not change anything to the properties and techniques involved.

Definition 7.5.13 The *Pascal-adic language* is the set L of words w for which there exist n and i such that w is a factor of $B_{n,i}$, where the $B_{n,i}$, $1 \leq i \leq n+2$, are defined by $B_{0,i} = i$, $1 \leq i \leq 2$; for every n, $B_{n,1} = B_{n-1,1}$, $B_{n,n+2} = B_{n-1,n+1}$ and for $2 \leq j \leq n+1$

$$B_{n,j} = B_{n-1,j-1}B_{n-1,j}.$$

The *Pascal-adic system* is the subshift (X_L, S), where X_L is the set of all $x \in \{1,2\}^{\mathbb{N}}$ such that $L(x) \subseteq L$ (see Proposition 7.1.2).

This language is related to the Pascal triangle, as the length of $B_{n,i}$ is the binomial coefficient $\binom{n+1}{i-1}$, and the blocks are built in the same way as these coefficients are built along the Pascal triangle.

The system (X_L, S) is not minimal, as $B_{n,2} = 1^n2$ for every $n \geq 1$ so $1^\omega \in X_L$. The following result states that it has infinitely many ergodic invariant probability measures.

Theorem 7.5.14 (Méla and Petersen 2005) *For every word w in L, for every real number $0 \leq \alpha \leq 1$, there exists $f_\alpha(w)$ such that for any sequence $k_n \to +\infty$ with $\lim_{n\to+\infty} \frac{k_n}{n} = \alpha$,*

$$\lim_{n\to+\infty} \frac{|B_{n,k_n}|_w}{|B_{n,k_n}|} = f_\alpha(w).$$

Moreover, all the ergodic invariant probability measures on (X_L, S) are the measures μ_α given on the cylinders by $\mu_\alpha([w]) = f_\alpha(w)$.

Thus, this is a very interesting intermediate case where the invariant measures are known (and there are not too many of them), and there are what we could call directional frequencies. As for the complexity, it satisfies $\lim_{n\to+\infty} \frac{p(n)}{n^3} = \frac{1}{6}$.

7.6 Further afield

In spite of the Jewett–Krieger theorem, uniquely ergodic systems may be considered to represent a small and particularly well-behaved class of dynamical systems; however, even more complicated systems can present properties which generalise directly the notion of unique ergodicity: instead of hoping that there will be one invariant measure, we know some explicit invariant measures, and what we want to prove is that these constitute the whole set of invariant measures. For example, a measure-theoretic system (X, T, μ) has *minimal self-joinings* if every ergodic measure on $X \times X$, invariant by $T \times T$, and whose marginals on X are μ, is either the product measure $\mu \times \mu$, or a diagonal measure defined by $\nu(A \times B) = \mu(A \cap T^k B)$ for some k. This notion was defined by

D. J. Rudolph in (Rudolph 1979). The standard example of a system with minimal self-joinings is *Chacon's map*, the symbolic system associated with the Chacon word of Section 1.4. Moreover, the (mostly combinatorial) techniques used in (del Junco, Rahe, and Swanson 1980) to prove this result were generalised in a famous series of papers by M. Ratner, starting from the joinings of *horocycle flows* (Ratner 1983) and culminating in the proof of *Raghunathan's conjecture* on unipotent groups (Ratner 1991).

The results on interval-exchange maps can generally be interpreted into results on more geometrical systems; in particular, Boshernitzan's ne_n criterion is not only a tool in proving Theorem 7.4.7 above, but also a way of knowing that a given interval exchange is uniquely ergodic; thus it is proved by W. Veech (Veech 1999) to imply an earlier criterion of H. Masur (Masur 1992) for the unique ergodicity of some foliations on Riemann surfaces of genus at least two equipped with a holomorphic 1-form. Then he uses this criterion to precise the knowledge of invariant measures for another famous flow, the Teichmüller geodesic flow.

7.7 Exercises

Section 7.1

Exercise 7.1 Prove that an infinite word x is eventually periodic if, and only if, X_x is finite. Prove that x is periodic if, and only if, X_x is finite and $S|_{X_x} : X_x \to X_x$ is onto.

Exercise 7.2 Consider the words $x = 0101101^3 01^4 01^5 \cdots$ and $y = 0100110^3 1^3 0^4 1^4 \cdots$. Describe the subshifts X_x and X_y and the subshifts they contain. Which ones are minimal? Make a picture describing the action of S on X_x and X_y.

Exercise 7.3 Prove that the set of cylinders of a subshift (X, S) is equal to the set of closed balls (any centre, any radius) of X for the distance defined in 1.2.10. Prove that it is also equal to the set of open balls of X for this distance. Prove that any element of a ball is a centre of it.

Exercise 7.4 Let (X, S) be a subshift. Prove that

$$X = \bigcap_{n \in \mathbb{N}} \bigcup_{u \in L_n(X)} [u],$$

where $[u]$ is considered as a cylinder in the dynamical system $A^{\mathbb{N}}$.

Exercise 7.5 Using exercise 1.8, prove directly that the Chacon word has uniform frequencies.
Hint. Show that $f_w(x) = \lim_{n\to+\infty} \frac{|b_n|_w}{|b_n|}$.

Section 7.2

Exercise 7.6 Rephrase the definition of a weight function in terms of Rauzy graphs.

Exercise 7.7 Give an example of a non-minimal non-uniquely ergodic subshift (X, S) such that any infinite word x in X has uniform frequencies.

Exercise 7.8 Let (X, S) be a subshift, and n be a positive integer. Assume that there exist n elements $x_1, \dots, x_n$ of X and n elements $w_1, \dots, w_n$ of $L(X)$ such that, for any $i, j \le n$ satisfying $i \ne j$, we have $f_{w_i}(x_j) > 0$ and $f_{w_i}(x_i) = 0$. Prove that (X, S) admits at least n ergodic invariant measures.

Exercise 7.9 Let (X, S) be a subshift, such that for any letter $a \in A$, there exists a word $x \in X$ such that $\limsup_{n\to\infty} \frac{|x_0\cdots x_n|_a}{n} > 1/2$. Prove that (X, S) admits at least $\mathrm{Card}(A)$ ergodic invariant measures.

Exercise 7.10 Let $(a_n)_{n\in\mathbb{N}}$ be an integer sequence which grows sufficiently fast. Let x be the word $0^{a_0}1^{a_1}0^{a_2}1^{a_3}0^{a_4}1^{a_5}0^{a_6}1^{a_7}\cdots$. Show that x does not have frequencies, but all of the minimal subshifts contained in X_x are uniquely ergodic.

Section 7.3

Exercise 7.11 For any element v of A^+, let us define the measure $\mu_v = \frac{1}{|v|}\sum_{k=0}^{|v|-1} \delta_{S^k(v^\omega)}$. Prove that the set $\{\mu_v \mid v \in A^+\}$ is a subset of $\mathcal{E}(A^{\mathbb{N}}, S)$ which is dense in $\mathcal{M}(A^{\mathbb{N}}, S)$. In particular, $\overline{\mathcal{E}(A^{\mathbb{N}}, S)} = \mathcal{M}(A^{\mathbb{N}}, S)$.
Hint. To approximate an invariant measure μ in $\mathcal{M}(A^{\mathbb{N}}, S)$, let n be a big integer, let $\{u_1, \dots, u_k\}$ be an enumeration of $L_n(A^{\mathbb{N}}) = A^n$, and consider the word $v = u_1^{p_1} \cdots u_k^{p_k}$, where p_i/q is a rational approximation of $\mu([u_i])$.

Section 7.4

Exercise 7.12 An infinite word x is said to be *episturmian* if $L(x)$ is closed under reversal and has at most one right special factor of each length (Droubay, Justin, and Pirillo 2001). Prove that an episturmian infinite word has uniform frequencies.

Exercise 7.13 Let x be an infinite word on A. A finite word $q \in A^*$ is said to be a *quasiperiod* of x if x is covered by the occurrences of q (in particular, q is a prefix of x). An infinite word x is said to be *multi-scale quasiperiodic* if it admits infinitely many quasiperiods (Marcus and Monteil 2006). Prove that a multi-scale quasiperiodic infinite word has uniform frequencies.

Exercise 7.14 An infinite word x is said to be *linearly recurrent* if $\exists K \geq 0,\ \forall u \in L(x),\ \forall v \in L_{K|u|}(x),\ |v|_u \geq 1$ (Durand 2000) (see also Section 6.5.3). Prove that a linearly recurrent infinite word has uniform frequencies. Prove that a fixed point of a primitive substitution is linearly recurrent. Prove that an infinite word x is linearly recurrent if, and only if, there exists μ in $\mathcal{M}(X_x, S)$ such that $\liminf_{n\to\infty} ne_n(\mu) > 0$.
Hint. If $v = v_0 \cdots v_n$ is a long element of $L(x)$ having no occurrence of u, assume that $vu \in L(x)$ and compute the measure of $\cup_{i\leq n}[v_i v_{i+1} \cdots v_n u]$ (this trick is due to M. Boshernitzan).

Exercise 7.15 Give an example of a minimal subshift of linear complexity whose invariant measures are not determined by their values on the cylinders of length 1.
Hint. Consider the substitution $0 \mapsto 0011$, $1 \mapsto 0101$.

Exercise 7.16 Give an example of an infinite word x such that, for any k, $f_{0^k}(x) = 1$ and such that X_x has infinitely many ergodic invariant measures. Is it possible to construct a uniformly recurrent example?

Section 7.5

Exercise 7.17 Let $\mu = \mu_\pi$ be the Borel measure defined on the full-shift $(A^{\mathbb{N}}, S)$ as in the proof of Proposition 7.5.1. Prove that, for any two finite words u and v and for any $n \geq |u|$, we have $\mu([u] \cap S^{-n}([v])) = \mu([u])\mu([v])$. Since μ is outer-regular, we know that for any Borel subset B of $A^{\mathbb{N}}$ and for any $\varepsilon > 0$, there exist some finite words $w_1, \ldots, w_\ell$ such that B is included in the disjoint union $\sqcup_{i=1}^{\ell}[w_i]$ and such that $\mu(\sqcup_{i=1}^{\ell}[w_i]) \leq \mu(B) + \varepsilon$. Prove that, for any two Borel subsets B and C of $A^{\mathbb{N}}$, we have $\lim_{n\to\infty} \mu(B \cap S^{-n}(C)) = \mu(B)\mu(C)$ (the measure-theoretic dynamical systems satisfying this property are said to be *strongly mixing*). Prove that, for any S-invariant Borel subset B of $A^{\mathbb{N}}$, we have $\mu(B) = \mu(B)^2$. Conclude that the system $(A^{\mathbb{N}}, \mathcal{B}, \mu_\pi, S)$ is ergodic.

Exercise 7.18 Let x denote the fixed point of the substitution $a \mapsto aaab$ $b \mapsto b$ which begins with the letter a. Show that X_x is neither minimal nor uniquely ergodic. Show that both x and Chacon's word can be described in the adic framework with the same (constant) sequence of matrices.

Exercise 7.19 Let $B_{n,1}$ and $B_{n,2}$ be as in Section 7.5.3.3. Show that $1 \leq \frac{|B_{n,2}|}{|B_{n,1}|} \leq \frac{p_{n-1}}{l_{n-1}}$. Show that $\frac{|B_{n+1,2}|_2}{|B_{n+1,2}|} \geq \frac{p_n}{m_n+p_n} \geq \frac{|B_{n,2}|_2}{|B_{n,2}|}$. Show that $\frac{|B_{n+1,1}|_1}{|B_{n+1,1}|} \geq \frac{m_n}{m_n+l_n\frac{p_{n-1}}{l_{n-1}}} \frac{|B_{n,1}|_1}{|B_{n,1}|}$. Use infinite products to give lower bounds for $\frac{|B_{n,2}|_2}{|B_{n,2}|}$ and $\frac{|B_{n,1}|_1}{|B_{n,1}|}$. Show that $\frac{|B_{n,2}|_2}{|B_{n,2}|} + \frac{|B_{n,1}|_1}{|B_{n,1}|} \geq \frac{3}{2}$ for good choices of l_n, m_n, p_n. Conclude that this contradicts the existence of uniform frequencies.

Exercise 7.20 Prove Proposition 7.5.9.

7.8 Note: Dictionary between word combinatorics and symbolic dynamics

combinatorics on words	symbolic dynamics
infinite word x	subshift (X_x, S)
factorial and extendable language L	subshift (X, S)
finite word w	cylinder $[w]$
uniform recurrence	minimality
weight function	invariant measure
uniform frequencies	unique ergodicity
positive uniform frequencies	strict ergodicity

References

[Bezuglyi, Kwiatkowski, Medynets, et al. 2009] Bezuglyi, S., Kwiatkowski, J., Medynets, K., and Solomyak, B. Invariant measures on stationary Bratteli diagrams. *Ergod. Th. & Dynam. Sys.* To appear.

[Boshernitzan 1984] Boshernitzan, M. A unique ergodicity of minimal symbolic flows with linear block growth. *J. Analyse Math.* **44**, (1984) 77–96.

[Boshernitzan 1985] Boshernitzan, M. A condition for minimal interval exchange maps to be uniquely ergodic. *Duke Math. J.* **52**(3), (1985) 723–752.

[Boshernitzan 1992] Boshernitzan, M. A condition for unique ergodicity of minimal symbolic flows. *Ergod. Th. & Dynam. Sys.* **12**(3), (1992) 425–428.

[de Bruijn 1946] de Bruijn, N. G. A combinatorial problem. *Proc. Konin. Neder. Akad. Wet.* **49**, (1946) 758–764.

[Cassaigne and Kaboré 2009] Cassaigne, J. and Kaboré, I. Un mot sans fréquences de faible complexité (provisional title), 2009. In preparation.

[Chaika 2008] Chaika, J. Hausdorff dimension for ergodic measures of interval exchange transformations. *J. Mod. Dyn.* **2**(3), (2008) 457–464.

[Cheung 2003] Cheung, Y. Hausdorff dimension of the set of nonergodic directions. *Ann. of Math. (2)* **158**(2), (2003) 661–678. With an appendix by M. Boshernitzan.

[Cheung and Masur 2006] Cheung, Y. and Masur, H. A divergent Teichmüller geodesic with uniquely ergodic vertical foliation. *Israel J. Math.* **152**, (2006) 1–15.

[Denker, Grillenberger, and Sigmund 1976] Denker, M., Grillenberger, C., and Sigmund, K. *Ergodic theory on compact spaces.* Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin, 1976.

[Droubay, Justin, and Pirillo 2001] Droubay, X., Justin, J., and Pirillo, G. Epi-Sturmian words and some constructions of de Luca and Rauzy. *Theoret. Comput. Sci.* **255**(1-2), (2001) 539–553.

[Durand 2000] Durand, F. Linearly recurrent subshifts have a finite number of non-periodic subshift factors. *Ergod. Th. & Dynam. Sys.* **20**, (2000) 1061–1078.

[Ferenczi 1995] Ferenczi, S. Les transformations de Chacon: combinatoire, structure géométrique, lien avec les systèmes de complexité $2n + 1$. *Bull. Soc. Math. France* **123**, (1995) 271–292.

[Ferenczi 1996] Ferenczi, S. Rank and symbolic complexity. *Ergod. Th. & Dynam. Sys.* **16**, (1996) 663–682.

[Ferenczi 1997] Ferenczi, S. Systems of finite rank. *Colloq. Math.* **73**(1), (1997) 35–65.

[Ferenczi, Fisher, and Talet 2009] Ferenczi, S., Fisher, A., and Talet, M. Some (non)-uniquely ergodic adic transformations. *J. Analyse Math.* To appear.

[Ferenczi, Holton, and Zamboni 2003] Ferenczi, S., Holton, C., and Zamboni, L. Q. Structure of three-interval exchange transformations. II. A combinatorial description of the trajectories. *J. Anal. Math.* **89**, (2003) 239–276.

[Ferenczi, Holton, and Zamboni 2004] Ferenczi, S., Holton, C., and Zamboni, L. Q. Structure of three-interval exchange transformations III: ergodic and spectral properties. *J. Anal. Math.* **93**, (2004) 103–138.

[Ferenczi and Zamboni 2008] Ferenczi, S. and Zamboni, L. Q. Eigenvalues and simplicity for interval exchange transformations, 2008. Preprint.

[Ferenczi and Zamboni 2009] Ferenczi, S. and Zamboni, L. Q. Structure of k-interval exchange transformations: induction, trajectories, and distance theorems. *J. Analyse Math.* To appear.

[Fisher 2009] Fisher, A. M. Nonstationary mixing and the unique ergodicity of adic transformations. *Stoch. Dyn.* To appear.

[Furstenberg 1961] Furstenberg, H. Strict ergodicity and transformation of the torus. *Amer. J. Math.* **83**, (1961) 573–601.

[Grillenberger 1972] Grillenberger, C. Construction of strictly ergodic systems I. Given entropy. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **25**, (1972/73) 323–334.

[Hahn and Katznelson 1967] Hahn, F. and Katznelson, Y. On the entropy of uniquely ergodic transformations. *Trans. Amer. Math. Soc.* **126**, (1967) 335–360.

[Jewett 1969] Jewett, R. The prevalence of uniquely ergodic systems. *J. Math. Mech.* **19**, (1969/70) 717–729.

[del Junco, Rahe, and Swanson 1980] del Junco, A., Rahe, M., and Swanson, L. Chacon's automorphism has minimal self-joinings. *J. Analyse Math.* **37**, (1980) 276–284.

[Katok 1973] Katok, A. B. Invariant measures of flows on orientable surfaces. *Dokl. Akad. Nauk SSSR* **211**, (1973) 775–778.

[Keane 1975] Keane, M. S. Interval exchange transformations. *Math. Zeitschrift* **141**, (1975) 25–31.

[Keane 1977] Keane, M. Non-ergodic interval exchange transformations. *Israel J. Math.* **26**(2), (1977) 188–196.

[Keynes and Newton 1976] Keynes, H. B. and Newton, D. A "minimal", non-uniquely ergodic interval exchange transformation. *Math. Z.* **148**(2), (1976) 101–105.

[Krieger 1972] Krieger, W. On unique ergodicity. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pp. 327–346. Univ. California Press, Berkeley, Calif., 1972.

[Marcus and Monteil 2006] Marcus, S. and Monteil, T. Quasiperiodic infinite words : multi-scale case and dynamical properties, 2006. Preprint.

[Marmi, Moussa, and Yoccoz 2005] Marmi, S., Moussa, P., and Yoccoz, J.-C. The cohomological equation for Roth-type interval exchange maps. *J. Amer. Math. Soc.* **18**(4), (2005) 823–872 (electronic).

[Masur 1982] Masur, H. Interval exchange transformations and measured foliations. *Ann. of Math. (2)* **115**(1), (1982) 169–200.

[Masur 1992] Masur, H. Hausdorff dimension of the set of nonergodic foliations of a quadratic differential. *Duke Math. J.* **66**(3), (1992) 387–442.

[Méla and Petersen 2005] Méla, X. and Petersen, K. Dynamical properties of the Pascal adic transformation. *Ergodic Theory Dynam. Systems* **25**(1), (2005) 227–256.

[Monteil 2005] Monteil, T. *Illumination dans les billards polygonaux et dynamique symbolique.* Ph.D. thesis, Institut de Mathématiques de Luminy, Université de la Méditerranée, 2005. Chapter 5.

[Monteil 2009] Monteil, T. An upper bound for the number of ergodic invariant measures of a minimal subshift with linear complexity, 2009. In preparation.

[Oseledec 1966] Oseledec, V. I. The spectrum of ergodic automorphisms. *Dokl. Akad. Nauk SSSR* **168**, (1966) 1009–1011.

[Oxtoby 1952] Oxtoby, J. C. Ergodic sets. *Bull. Amer. Math. Soc.* **58**, (1952) 116–136.

[Pytheas Fogg 2002] Pytheas Fogg, N. *Substitutions in Dynamics, Arithmetics and Combinatorics*, vol. 1794 of *Lecture Notes in Mathematics.* Springer-Verlag, 2002. Ed. by V. Berthé and S. Ferenczi and C. Mauduit and A. Siegel.

[Queffélec 1987] Queffélec, M. *Substitution Dynamical Systems — Spectral Analysis*, vol. 1294 of *Lecture Notes in Mathematics.* Springer-Verlag, 1987.

[Ratner 1983] Ratner, M. Horocycle flows, joinings and rigidity of products. *Ann. of Math. (2)* **118**(2), (1983) 277–313.

[Ratner 1991] Ratner, M. On Raghunathan's measure conjecture. *Ann. of Math. (2)* **134**(3), (1991) 545–607.

[Rauzy 1983] Rauzy, G. Suites à termes dans un alphabet fini. In *Séminaire de Théorie des Nombres de Bordeaux*, pp. 25.01–25.16. 1982/83.

[Rosenthal 1988] Rosenthal, A. Strictly ergodic models for noninvertible transformations. *Israel J. Math.* **64**(1), (1988) 57–72.

[Rudin 1987] Rudin, W. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edn., 1987.

[Rudin 1991] Rudin, W. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, second edn., 1991.

[Rudolph 1979] Rudolph, D. J. An example of a measure preserving map with minimal self-joinings, and applications. *J. Analyse Math.* **35**, (1979) 97–122.

[Sataev 1975] Sataev, E. A. The number of invariant measures for flows on orientable surfaces. *Izv. Akad. Nauk SSSR Ser. Mat.* **39**(4), (1975) 860–878. English translation in *Mathematics of the USSR-Izvestiya* **9**(4) (1975), 813–830.

[Veech 1969] Veech, W. A. Strict ergodicity in zero dimensional dynamical systems and the Kronecker-Weyl theorem mod 2. *Trans. Amer. Math. Soc.* **140**, (1969) 1–33.

[Veech 1982] Veech, W. A. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math. (2)* **115**(1), (1982) 201–242.

[Veech 1984] Veech, W. A. The metric theory of interval exchange transformations. I. Generic spectral properties. *Amer. J. Math.* **106**(6), (1984) 1331–1359.

[Veech 1999] Veech, W. A. Measures supported on the set of uniquely ergodic directions of an arbitrary holomorphic 1-form. *Ergodic Theory Dynam. Systems* **19**(4), (1999) 1093–1109.

[Weiss 1973] Weiss, B. Subshifts of finite type and sofic systems. *Monatsh. Math.* **77**, (1973) 462–474.

[Yoccoz 2005] Yoccoz, J.-C. Échanges d'intervalles, 2005. Cours du Collège de France.

[Yoccoz 2006] Yoccoz, J.-C. Continued fraction algorithms for interval exchange maps: an introduction. In *Frontiers in number theory, physics, and geometry. I*, pp. 401–435. Springer, Berlin, 2006.

Notation Index

General Index

Contents