HAL
open science

DEVIATION OF ERGODIC AVERAGES FOR SUBSTITUTION DYNAMICAL SYSTEMS WITH EIGENVALUES OF MODULUS ONE

Xavier Bressaud, Alexander I. Bufetov, Pascal Hubert

To cite this version:

Xavier Bressaud, Alexander I. Bufetov, Pascal Hubert. DEVIATION OF ERGODIC AVERAGES FOR SUBSTITUTION DYNAMICAL SYSTEMS WITH EIGENVALUES OF MODULUS ONE. Proceedings of the London Mathematical Society, 2014. hal-01256152

HAL Id: hal-01256152

https://hal.science/hal-01256152v1

Submitted on 14 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1106.2666v1 [math.DS] 14 Jun 2011

DEVIATION OF ERGODIC AVERAGES FOR SUBSTITUTION DYNAMICAL SYSTEMS WITH EIGENVALUES OF MODULUS ONE

XAVIER BRESSAUD, ALEXANDER I. BUFETOV, AND PASCAL HUBERT

Abstract. Deviation of ergodic sums is studied for substitution dynamical systems with a matrix that admits eigenvalues of modulus 1. The functions γ we consider are the corresponding eigenfunctions. In Theorem 1.1 we prove that the limit inferior of the ergodic sums $(n, \gamma(x_0) + \ldots + \gamma(x_{n-1}))_{n\in\mathbb{N}}$ is bounded for every point x in the phase space. In Theorem 1.2, we prove existence of limit distributions along certain exponential subsequences of times for substitutions of constant length. Under additional assumptions, we prove that ergodic integrals satisfy the Central Limit Theorem (Theorem 1.3, Theorem 1.9).

Contents

1. Introduction 2
1.1. Limit inferior of ergodic sums 2
1.2. Limit distributions for substitutions of constant length 3
1.3. Historical remarks. 9
1.4. Organization of the paper 9
1.5. Acknowledgements. 9
2. Background 10
2.1. Words and sequences 10
2.2. Substitutions 10
2.3. Vershik automorphisms and suspension flows 11
3. Proof of Theorem 1.1 13
4. Proof of Corollaries 1.5 and 1.6 15
4.1. Proof of Corollary 1.5 15
4.2. Proof of Corollary 1.6 16
4.3. Proof of Proposition 1.7 18
5. Asymptotic behaviour of ergodic sums 18

Date: 14 June 2011.
1991 *Mathematics Subject Classification.* Primary: 37B10; Secondary: 37E05.
Key words and phrases. substitutive systems, ergodic sums, interval exchange transformations, Markov approximation.

5.1. Ergodic means 18
5.2. Decomposition of unstable leaves. 19
5.3. A family of automata 20
5.4. Ergodic sums and automata. 21
6. Markov approximation and the proof of Theorem 1.2 23
6.1. Markov chains 23
6.2. Proof of Theorem 1.2 24
7. Proof of Theorem 1.9 26
7.1. Assumptions on the Markov chain 26
7.2. Choice of a full-measure set of admissible sequences 27
7.3. A Central Limit Theorem for non homogeneous Markov chains 27
7.4. Proof of Theorem 1.9 29
7.5. Standard assumptions 30
7.6. Proof of Proposition 7.4 31
7.7. The variance 36
8. Specific cases and examples 39
8.1. Synchronization 39
8.2. Strongly non synchronizable substitutions on a 2 letters alphabet 41
8.3. Examples 42
References 45

1. Introduction

Let σ be a primtive substitution over a finite alphabet $\mathcal{A}$, let M_σ be the matrix substitution, and let X_σ be the corresponding subshift in the space of bi-infinite sequences (the precise definitions are recalled in Section 2). The aim of this paper is to study the asymptotic behaviour of ergodic sums for the minimal dynamical system (X_σ, T) in the (non-hyperbolic) case when the matrix M_σ has an eigenvalue of modulus one. For a function $f : X_\sigma \to \mathbb{R}$, all $x \in X_\sigma$ and all $n \in \mathbb{N}$, set

$$S_n f(x) = \sum_{k=0}^{n-1} f(T^k x).$$

In this paper, we shall only consider functions f depending on the first coordinate of the symbolic sequence. In what follows, we will identify such a function f with the corresponding vector in $\mathbb{C}^{\#\mathcal{A}}$.

1.1. **Limit inferior of ergodic sums.**

Theorem 1.1. *Let σ be a primitive substitution with matrix M_σ. Assume that M_σ has an eigenvalue θ_2 of modulus 1 and let $\Gamma = (\gamma(a), a \in \mathcal{A})^t$ be a right eigenvector associated to θ_2. There exists a constant C (depending only on σ and γ) such that, for any $x \in X_\sigma$, we have*

$$\liminf_{n\to\infty} |S_n\gamma(x)| < C \text{ and } \liminf_{n\to\infty} |S_n\gamma(T^{-n}x)| < C. \tag{1.1}$$

Recall that, since T is minimal, by the Gottschalk-Hedlund Theorem, the function γ is a coboundary if and only if the ergodic sums $S_n\gamma(x)$ remain bounded, as $n \to \infty$, for some and, consequently, for every $x \in X_\sigma$. Adamczewski [Ad] proved that, if θ_2 is not a root of unity, then for almost all x, we have

$$\limsup_{n\to\infty} |S_n\gamma(x)| = \infty. \tag{1.2}$$

If θ_2 is a root of unity, then Adamczewski [Ad] gives a necessary and sufficient combinatorial condition ensuring that γ is a coboundary. For *almost all* points the conclusion of Theorem 1.1 is true for any ergodic transformation. It is important for us to have the result for all points. Observe that there exist infinitely many primitive substitutions with a hyperbolic matrix with the following property (see [BHM]). If the function Γ is in the unstable space of the matrix or equivalently if there exists x such that

$$\limsup_{n\to\infty} |S_n\gamma(x)| = \infty$$

then there exists x such that

$$\liminf_{n\to\infty} |S_n\gamma(x)| = \infty \text{ and } \liminf_{n\to\infty} |S_n\gamma(T^{-n}x)| = \infty$$

Below, we give further applications of Theorem 1.1 to affine interval exchange transformations (Corollaries 1.5 and 1.6).

1.2. **Limit distributions for substitutions of constant length.** Informally, the mechanism of the Central Limit Theorem for substitutions with eigenvalue 1 can be formulated as follows. A long orbit of a substitution dynamical system admits a representation as a disjoint union of segments of different exponential scales. Since the eigenvalue is equal to 1, these segments give roughly the same contribution. By the Markov property, these segments are asymptotically independent, whence the Central Limit theorem.

We are only able to make this argument precise under significant additional restrictions.

Recall that a substitution σ has constant length d if the image of each letter has length d. For instance, this condition implies that the maximal eigenvalue of the matrix M_σ is d. Furthermore, we assume that 1

is an eigenvalue and consider the eigenvector γ, $\gamma M_\sigma = \gamma$. For a given $t \in [0,1]$, we consider ergodic sums $S_{\lfloor d^n t \rfloor}(\gamma)$ as a sequence of random variables where the initial point is distributed according to the unique shift invariant measure on X_σ. The behaviour of these sequences significantly depends on the expansion of t in basis d; namely on the sequence $\tau_t = (\tau_n)_{n\in\mathbb{N}} \in \{0,\dots,d-1\}^{\mathbb{N}}$ such that $t = \sum_{n\leq 1} \tau_n d^n$. We first consider the case when this expansion is eventually periodic.

Theorem 1.2. *Let σ be a primitive substitution of constant length d. Assume that M_σ has eigenvalue 1 with eigenvector γ. Let $t \in [0,1]$ and assume that the expansion $\tau_t = (\tau_n)_{n\in\mathbb{N}}$ of t in basis d is eventually periodic. Then,*

$$\frac{1}{\sqrt{n}} S_{\lfloor d^n t \rfloor}(\gamma) \xrightarrow{distr} \nu$$

where ν is a probability measure with generalized density

$$p_0\delta_0 + \sum_{k=1}^{\ell} \frac{p_k}{\sqrt{2\pi}\sigma_k} e^{-x^2/(2\sigma_k^2)}, \ \sum_{k=0}^{\ell} p_k = 1.$$

Moreover, the limit is trivial if and only if γ is a coboundary.

The nature of the limit law depends on the expansion of t in basis d. In Section 8 we give examples showing that the limit distribution need not be Gaussian and can have atoms.
In Section 7 we provide a precise algorithmically verifiable sufficient condition on the morphism σ under which the limit distribution is indeed Gaussian for almost all t. We thus arrive at the following theorem.

Theorem 1.3. *There exist infinitely many primitive substitutions σ of constant length with eigenvalue 1 and eigenvector γ such that, for Lebesgue almost every $t \in [0,1]$, the following holds:*

(1) *there exist positive constants α_1, α_2 such that the variance $V_n = Var(S_{\lfloor d^n t \rfloor}(\gamma))$ satisfies the inequality*

$$\alpha_1 n^{4/5} \leq V_n \leq \alpha_2 n.$$

(2) *The sequence*

$$\frac{1}{\sqrt{V_n}} S_{\lfloor d^n t \rfloor}(\gamma)$$

converges in distribution to the normal law $N(0,1)$.

Our argument is an adaptation to our context of R. L. Dobrushin's central limit theorem [Dob] for non homogeneous Markov chains. We follow the martingale approach of S. Sethuraman and S.R.S. Varadhan [SV].

Conjecture 1.4.

(1) *For any primitive substitution of constant length with eigenvalue 1, there exists t such that the limit distribution of*

$$\frac{1}{\sqrt{V_n}} S_{\lfloor d^n t \rfloor}(\gamma)$$

is not Gaussian.

(2) *For any primitive substitution of constant length with eigenvalue 1, for almost every t,*

$$\frac{1}{\sqrt{V_n}} S_{\lfloor d^n t \rfloor}(\gamma)$$

converges to the normal law $N(0,1)$.

1.2.1. *Affine interval exchange transformations.* Our study was partly motivated by questions arising in the study of so-called Affine Interval Exchange Transformations. An interval exchange transformation is defined by the length $\lambda = (\lambda_1, \ldots, \lambda_r)$ of the intervals $I_1, \ldots, I_r$ and a permutation π of $\{1, \ldots, r\}$. It is denoted by $T_{(\lambda,\pi)}$. An affine interval exchange transformation is a piecewise affine bijective map of the interval. As for interval exchange transformation it is defined by the length of the intervals, a permutation, and the slopes of the map (constant on each interval). This additional information is encoded in a vector $(w_1, \ldots, w_r)$ with $w_i > 0$ for $i = 1, \ldots, r$. Given these data, one can determine $\{a_i, i = 1, \ldots, r\}$ such that for all $i = 1, \ldots, r$ and all $x \in I_i$,

$$T(x) = w_i x + a_i.$$

To an affine interval exchange transformation T, we assign the corresponding log-slope vector

$$\vec{\Gamma} = (\log w_1, \ldots, \log w_r).$$

A natural idea to analyze this class of dynamical systems has been to ask whether they are semi-conjugate to some interval exchange transformation (as a circle diffeomorphism is semi-conjugate to a rotation). A subsequent question is to ask if such a semi-conjugacy is a conjugacy. Indeed, it is not always the case : affine interval exchange transformation may have wandering intervals. It is well-known that the problem of conjugacy reduces to estimates on the ergodic sums of interval exchange transformations (see for instance [CG], [C], [LM], [BHM], [MMY2]). Together with this observation, our result yields new classes of examples corresponding to self-similar interval exchange transformation with eigenvalues of modulus one.

Let T be an affine interval exchange transformation. One can assign to T its *Rauzy expansion* in the same way as for usual interval exchange transformations [MMY2]. Let us consider the case where the Rauzy expansion of T is periodic with positive Rauzy renormalization matrix R. In this case T admits a semi-conjugacy onto the underlying self-similar interval exchange transformation T_0. Let $(\lambda_1, \dots \lambda_n)$ be the lengths of T_0 and as before, let $(w_1, \dots, w_n)$ be the slopes of T. Camelier and Gutierrez [CG] showed furthermore that

$$\sum_i^n \lambda_i \log w_i = 0. \tag{1.3}$$

Conversely, given a self-similar interval exchange transformation T_0 and a vector of slopes $(w_1, \dots, w_n)$ satisfying (1.3) then Camelier and Gutierrez showed that there exists an affine interval exchange transformation T with slopes $(w_1, \dots, w_n)$ semi-conjugate to T_0.
We next consider the question of conjugation. Consider the decomposition of $\mathbb{C}^n$ into characteristic spaces of the matrix ${}^tR$:

$$\mathbb{C}^n = E^u \oplus E^c \oplus E^s.$$

Here E^u corresponds to eigenvalues with absolute value greater than 1, E^s to eigenvalues with absolute value less than 1, E^c to eigenvalues on the unit circle. If $\vec{\Gamma} \in E^s$, then Camelier and Gutierrez [CG] showed that T is conjugate to T_0. If $\vec{\Gamma} \in E^u$, Camelier and Gutierrez [CG] and Cobo [C] gave examples in which T has wandering intervals. In [BHM] , it is showed that if $\vec{\Gamma} \in E^u$ and Galois conjugate to the Perron-Frobenius vector of ${}^tR$ then there is no conjugacy. Marmi-Moussa-Yoccoz [MMY2] obtained the same conclusion for generic interval exchange transformation ([MMY2]) using the positivity of the Lyapunov exponent of the Kontsevich-Zorich cocycle (proved by Forni [Fo], Avila-Viana [AV]).
Theorem 1.1 implies

Corollary 1.5. *Let $T_0 = T_{(\lambda,\pi)}$ be a self-similar interval exchange transformation and R the associated matrix obtained by Rauzy induction. Assume that R has an eigenvalue θ of modulus 1.*

- *If $\theta = \pm 1$, let Γ be an eigenvector associated to θ for ${}^tR$*
- *If $\theta = e^{i\phi}$ belongs to $\mathbb{C} \setminus \mathbb{R}$, let V_ϕ (resp. $V_{-\phi}$) be an eigenvector associated to θ (resp. to $e^{-i\phi}$) for ${}^tR$. Let Γ be a vector with real entries belonging to the vector space generated by V_ϕ and $V_{-\phi}$*

If an affine interval exchange transformation with vector of the logarithm of the slopes is Γ is semi-conjugate to T_0 then it is also topologically conjugated to T_0.

This corollary applies to many examples:

Corollary 1.6. *There exist infinitely many self similar interval exchange transformation $T_{(\lambda,\pi)}$ on $r \geq 4$ intervals such that every affine interval exchange transformation that is semi-conjugate to $T_{(\lambda,\pi)}$ is conjugate to $T_{(\lambda,\pi)}$.*

We recall [Ve2] that the Veech group $\Gamma(X)$ of a translation surface X is a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$ obtained as the image by derivation of the group of affine diffeomorphisms of the translation surface X. The surface X is a Veech surface if its Veech group is a lattice in $\mathrm{SL}_2(\mathbb{R})$. An element of Γ is the derivative of a pseudo-Anosov diffeomorphism if and only if it is a hyperbolic element. A Veech surface is primitive if it is not a cover of any other translation surface. Assume that X is a primitive Veech surface of genus 2. The group $\Gamma(X)$ is then defined over a quadratic field K ([Mc]). We have the following proposition:

Proposition 1.7. *On each primitive Veech surface in genus 2, there exists a pseudo-Anosov diffeomorphism whose dilatation has two conjugates of modulus one.*

Remark. In other words, the dilatation of our pseudo-Anosov diffeomorphism is a Salem number.
To a pseudo-Anosov diffeomorphism given by Proposition 1.7, assign an interval exchange transformation using Veech's zippered rectangle construction. The resulting self-similar interval exchange transformation T_0 then has the desired property: any affine interval exchange transformation T semi-conjugate to T_0 is in fact conjugate to T_0. Indeed the unstable space is spanned by the Perron-Frobenius vector and by the result of Camelier-Gutierrez (see equation (1.3)), the vector of slopes cannot belong to the unstable space. If the vector of slopes belongs to the stable space, then again results of Camelier and Gutierrez imply the existence of a conjugacy. If the vector of slopes belongs to the central space then Corollary 1.5 applies and again yields the existence of a conjugacy. Others examples can be obtained using the square tiled surfaces with degenerate Lyapunov spectrum constructed by Forni [Fo2] and Forni-Matheus-Zorich [FMZ].

1.2.2. *Symbolic flows and deviation of ergodic averages.* To prove Theorem 1.2 we introduce a family of Markov chains which provide approximations of the ergodic sums. To give more specific statements

—but without entering now into details of the proof— let us consider, for each $N \in \mathbb{N}$ and $\tau_0 \in \{0, \dots, d^N\}$, the Markov chain X^{N,τ_0} defined on the state space $\mathcal{A} \times \mathcal{A}^2 \times \{0, \dots, d^N\}$ having probability transitions:

$$p_{(a,V,m)\to(b,W,k)} = \begin{cases} d^{-N} & \text{if } \sigma^N(a) = PbS \text{ with } |P| = m \\ & \text{and } \sigma^N(V) = P'WS' \text{ with } |P'| = m + \tau_0 \\ 0 & \text{otherwise.} \end{cases}$$

Consider the map $G_{\tau_0}(a, V, m) = \gamma(S) + \gamma(P')$; we say that G_{τ_0} is a coboundary if it remains bounded on all paths of the Markov chain X^{N,τ_0}.

Corollary 1.8. *Assume that the chain X^{N,τ_0} only possesses one recurrent component and set $t_0 = \sum_{n\geq 1} \tau_0 d^{-nN}$. If $S_n\gamma$ is unbounded, then, as $n \to \infty$, the sequence of random variables $S_{\lfloor d^n t_0 \rfloor}\gamma/\sqrt{n}$ converges in distribution to a normal law with positive variance.*

We prove by analyzing examples that the irreducibility assumption does not always hold and that the behaviour of the fluctuations at exponential scale can be very complicated. For instance, we exhibit examples of substitutions σ with an invariant vector γ that is not a coboundary such that $S_{d^n t}(\gamma)$ is bounded for some values of t. The combinatorial properties of the graphs of the corresponding Markov chains remain very unclear. Nevertheless, using results on non stationary Markov chains, we can prove that if there is t_0 with a periodic expansion such that the associated Markov chain is aperiodic with positive variance, then fluctuations are normal for Lebesgue almost every $t \in [0, 1]$. More specifically, we prove the following Theorem.

Theorem 1.9. *Let σ be a primitive substitution of constant length with eigenvalue 1. Let γ be an eigenvector associated with the eigenvalue 1. Assume that there exists N and τ_0 such that the Markov chain X^{N,τ_0} is recurrent aperiodic and the function G_{τ_0} is not a coboundary. Then, for Lebesgue almost every $t \in [0, 1]$, the following holds:*

(1) *there exist positive constants α_1, α_2 such that the variance $V_n = Var(S_{\lfloor d^n t \rfloor}(\gamma))$ satisfies the inequality*

$$\alpha_1 n^{4/5} \leq V_n \leq \alpha_2 n.$$

(2) *The sequence*

$$\frac{1}{\sqrt{V_n}} S_{\lfloor d^n t \rfloor}(\gamma)$$

converges in distribution to the normal law $N(0, 1)$.

The sequence V_n is the variance of a non stationary Markov chain. The assumption (positivity of the variance for a specific τ_0) means that for the homogeneous Markov chain associated to $t_0 = \sum_{n\geq 1} \tau_0 d^{-nN}$, the approximation of the ergodic sums of γ on this chain is not a coboundary. It is not very explicit but it is possible to check it on examples. We prove that the hypothesis of Theorem 1.9 is satisfied in many cases.

1.3. **Historical remarks.** For deviations of ergodic averages of generic interval exchange transformations see [Zo], [Fo]. For holonomy flows of pseudo-Anosov diffeomorphisms for which the second eigenvalue θ_2 of the corresponding action in homology is real and satisfies the inequality $\theta_2 > 1$, limit theorems are obtained in [Bu2]. In this case, limit distributions have compact support.
Let σ be a primitive substitution over a finite alphabet $\mathcal{A}$ and γ be a function from $\mathcal{A}$ to $\mathbb{C}$. Let $X_\sigma \subseteq \mathcal{A}^{\mathbb{Z}}$ be the subshift defined from σ and x be a point in X_σ. The stepped line associated to x is the piecewise affine curve $\ell(x)$ in $\mathbb{R} \times \mathbb{C}$ with vertices $(n, \gamma(x_0) + \ldots + \gamma(x_{n-1}))_{n\in\mathbb{N}}$ and $(n, \gamma(x_{-n}) + \ldots + \gamma(x_{-1}))_{n\geq 1}$. Stepped lines were studied by many authors (see for instance [ArIt] and [ABB] and references there). They are closely related to the famous Rauzy fractals and to the fractal curves studied by Dumont and Thomas in [DT1], [DT2]. Using a different language, the result of Adamczewski [Ad] about discrepancy of substitutive systems describes the behaviour of

$$\limsup_{n\to\infty}(|\gamma(x_0) + \ldots + \gamma(x_n)|) \tag{1.4}$$

in terms of the eigenvalues of the matrix M_σ of σ.

1.4. **Organization of the paper.** Section 2 sets our notation on words, sequences, substitutions and suspension flows. Theorem 1.1 is proved in Section 3 while its corollaries 1.5 and 1.6 are proved in Section 4. In Section 5, we prove preliminary results about approximation of ergodic sums by Markov chains. More precisely, we construct an explicit family of finite automata which allow us to code the ergodic sums as sums of nonhomogeneous Markov random variables. The main difficulty is that the underlying Markov chain may fail to be irreducible (examples are given in Section 8). In Section 6 we prove Theorem 1.2. In Section 7 we prove Theorem 1.9. In Section 8 we discuss various examples including those proving Theorem 1.3 (see Remark 8.7).

1.5. **Acknowledgements.** We thank S. Geninska for mentioning Beardon's theorem [Be]. X. B. was partially supported by project ANR JCJC: LAM while this work was in progress. A. I. B. is an Alfred

P. Sloan Research Fellow. He is supported in part by grant MK-4893.2010.1 of the President of the Russian Federation, by the Programme on Mathematical Control Theory of the Presidium of the Russian Academy of Sciences, by the Programme 2.1.1/5328 of the Russian Ministry of Education and Research, by the Edgar Odell Lovett Fund at Rice University, by the NSF under grant DMS 0604386, and by the RFBR-CNRS grant 10-01-93115. P. H. was partially supported by project blanc ANR: ANR-06-BLAN-0038 while this work was in progress.

2. Background

2.1. **Words and sequences.** Let $\mathcal{A}$ be a finite set. One calls it an *alphabet* and its elements *symbols*. A *word* is a finite sequence of symbols in $\mathcal{A}$, $w = w_0 \dots w_{\ell-1}$. The length of w is denoted $|w| = \ell$. One also defines the empty word ε. The set of words in the alphabet $\mathcal{A}$ is denoted $\mathcal{A}^*$ and $\mathcal{A}^+ = \mathcal{A}^* \setminus \{\varepsilon\}$. We will need to consider words indexed by integer numbers, that is, $w = w_{-m} \dots w_{-1}.w_0 \dots w_\ell$ where $\ell, m \in \mathbb{N}$ and the dot separates negative and non-negative coordinates. The set of one-sided infinite sequences $x = (x_i)_{i\in\mathbb{N}}$ in $\mathcal{A}$ is denoted by $\mathcal{A}^{\mathbb{N}}$. Analogously, $\mathcal{A}^{\mathbb{Z}}$ is the set of two-sided infinite sequences $x = (x_i)_{i\in\mathbb{Z}}$.
Given a sequence x in $\mathcal{A}^+$, $\mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ one denotes $x[i,j]$ the subword of x appearing between indexes i and j. Similarly one defines $x(-\infty, i]$ and $x[i, \infty)$. Let $w = w_{-m} \dots w_{-1}.w_0 \dots w_\ell$ be a word on $\mathcal{A}$. One defines the cylinder set $[w]$ as $\{x \in \mathcal{A}^{\mathbb{Z}} : x[-m, \ell] = w\}$.
The shift map $T : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ or $T : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$ is given by $T(x) = (x_{i+1})_{i\in\mathbb{N}}$ for $x = (x_i)_{i\in\mathbb{N}}$. A subshift is any shift invariant and closed (for the product topology) subset of $\mathcal{A}^{\mathbb{Z}}$ or $\mathcal{A}^{\mathbb{N}}$. A subshift is minimal if all of its orbits by the shift are dense.
In what follows we will use the shift map in several contexts, often in restriction to a subshift. To simplify notation we keep the symbol T all the time.

2.2. **Substitutions.** We refer to [Qu] and [F] and references therein for the general theory of substitutions.
A *substitution* is a map $\sigma : \mathcal{A} \to \mathcal{A}^+$. It naturally extends to $\mathcal{A}^+$, $\mathcal{A}^{\mathbb{N}}$ and $\mathcal{A}^{\mathbb{Z}}$; for $x = (x_i)_{i\in\mathbb{Z}} \in \mathcal{A}^{\mathbb{Z}}$ the extension (which is a morphism of monoid) is given by

$$\sigma(x) = \dots \sigma(x_{-2})\sigma(x_{-1}).\sigma(x_0)\sigma(x_1) \dots$$

where the central dot separates negative and non-negative coordinates of x. A further natural convention is that the image of the empty word ε is ε.

Let M be the matrix with indices in $\mathcal{A}$ such that M_{ab} is the number of times letter b appears in $\sigma(a)$ for any $a, b \in \mathcal{A}$. The substitution is *primitive* if there is $N > 0$ such that for any $a \in \mathcal{A}$, $\sigma^N(a)$ contains any other letter of $\mathcal{A}$ (here σ^N means N consecutive iterations of σ). Under primitivity one can assume without loss of generality that $M > 0$.

A substitution σ is of *constant length* d if, for all $a \in \mathcal{A}$, $|\sigma(a)| = d$.

Let $X_\sigma \subseteq \mathcal{A}^{\mathbb{Z}}$ be the subshift defined from σ. That is, $x \in X_\sigma$ if and only if any subword of x is a subword of $\sigma^N(a)$ for some $N \in \mathbb{N}$ and $a \in \mathcal{A}$.

Assume σ is primitive. Given a point $x \in X_\sigma$ there exists a unique sequence $(p_i, c_i, s_i)_{i\in\mathbb{N}} \in (\mathcal{A}^* \times \mathcal{A} \times \mathcal{A}^*)^{\mathbb{N}}$ such that for each $i \in \mathbb{N}$: $\sigma(c_{i+1}) = p_i c_i s_i$ and

$$\ldots \sigma^3(p_3)\sigma^2(p_2)\sigma^1(p_1)p_0.c_0 s_0 \sigma^1(s_1)\sigma^2(s_2)\sigma^3(s_3)\ldots$$

is the central part of x, where the dot separates negative and non-negative coordinates. This sequence is called the prefix-suffix decomposition of x (see for instance [CS]).

If only finitely many suffixes s_i are nonempty, then there exists $a \in \mathcal{A}$ and non-negative integers ℓ and q such that

$$x^+ = x[0, \infty) = c_0 s_0 \sigma^1(s_1) \ldots \sigma^\ell(s_\ell) \lim_{n\to\infty} \sigma^{nq}(a)$$

Analogously, if only finitely many p_i are non empty, then

$$x^- = x(-\infty, -1] = (\lim_{n\to\infty} \sigma^{np}(b))\sigma^m(p_m) \ldots \sigma^1(p_1)p_0$$

for some $b \in \mathcal{A}$ and non-negative integers p and m.

2.3. **Vershik automorphisms and suspension flows.** In this subsection, we recall the construction of Vershik's automorphisms [Ver], [VL] and their continuous analogues [Ito], [Bu]. We refer to Section 2, 3, 5 of [Bu] for details and further references. Given an oriented graph Γ with m vertices, let $\mathcal{E}(\Gamma)$ be the set of edges of Γ. For $e \in \mathcal{E}(\Gamma)$ we denote $I(e)$ its initial vertex and $F(e)$ its terminal vertex. To the graph Γ we assign a non-negative $m \times m$ non-negative matrix $A(\Gamma)$ by the formula

$$A = A(\Gamma)_{i,j} = \sharp\{e \in \mathcal{E}(\Gamma) \ : \ I(e) = i, F(e) = j\} \tag{2.1}$$

We assume that A is a primitive matrix (in other words, that the corresponding topological Markov chain is irreducible and aperiodic).

We define the Markov compactum:

$$Y = \{y = y_1 \dots y_n \cdots : y_n \in \mathcal{E}(\Gamma), F(y_{n+1}) = I(y_n)\}$$

The shift on Y is denoted by $\mathfrak{S}$. Assume that there is an order on the set of edges starting from a given vertex. This partial order extends to a partial order on Y: we write $y < y'$ if there exists $l \in \mathbb{N}$ such that $y_l < y'_l$ and $y_n = y'_n$ for $n > l$. The Vershik automorphism T^Y is the map from Y to itself defined by

$$T^Y y = \min_{y' > y} y'.$$

As A is primitive, there is a unique probability measure invariant under T^Y denoted by μ_Y.
We now define a suspension flow over (Y, T^Y) in the following way: let H be the Perron-Frobenius eigenvector of A. Then h_t is the special flow over (Y, T^Y) with roof function $\tau(y) = h_{I(y_1)}$. The phase space of the flow is

$$Y(\tau) = \{(y, t) : y \in Y, 0 \leq t < \tau(y)\}.$$

The measure μ_Y induces a probability measure ν_Γ on Y_τ. The vector H is normalized in such a way that the space $Y(\tau)$ have total measure 1.
For each $e \in \mathcal{E}(\Gamma)$, the set

$$\{(y, t) : y \in Y, y_1 = e,\ 0 \leq t < h_{I(y_1)}\}$$

is called a *rectangle* in what follows.
The space $X = \{x = \dots x_{-n} \dots x_n \cdots : x_n \in \mathcal{E}(\Gamma), F(x_{n+1}) = I(x_n)\}$ is the natural extension of (Y, S). The space X endowed with the Parry measure (the measure of maximal entropy) is canonically isomorphic to $(Y(\tau), \nu_\Gamma)$.
Following Livshits [Liv], we now connect Vershik's automorphisms with substitution dynamical systems. Consider the alphabet $\mathcal{A} = \{1, \dots, m\}$ as the set of vertices of Γ. For all $a \in \mathcal{A}$, we denote $\sigma(a)$ the sequence $\{F(e) : I(e) = a\}$ ordered with the partial order on $\{e : I(e) = a\}$. The dynamical system (X_σ, T) is a topological factor of the Vershik automorphism (Y, T^Y). The semi-conjugacy is given by the prefix-suffix decomposition. Almost every $u \in X_\sigma$ can be written in the form

$$u = \cdots \sigma^n(P_n) \cdots \sigma(P_1) P_0 . a_0 S_0 \sigma(P_1) \cdots \sigma^n(P_n) \cdots ,$$

where, for all $n \in \mathbb{N}$, $\sigma(a_{n+1}) = P_n a_n S_n$. Thus, it is clear that such point correspond to the unique path in Y such that, for all $n \geq 0$, $a_n = F(y_{n+1})$ (and $= I(y_n)$ for $n > 0$) and y_{n+1} has exactly $|P_n|$ predecessors in the partial order around $I(y_{n+1})$. We recall that the

semi-conjugacy is not always a topological conjugacy because there may be multiple writings but it is a measurable conjugacy.
When σ has constant length, the roof function of the suspension flow is constant because ${}^t(1,\dots,1)$ is an eigenvector for A associated to the maximal eigenvalue.
We observe that the matrix M_σ of the substitution is connected to the matrix A by the relation $M_\sigma = A^T$.

3. Proof of Theorem 1.1

Without loss of generality, we assume that M is a positive matrix. Let $x \in X_\sigma$, θ be an eigenvalue of M of modulus one and $\Gamma = (\gamma(a), a \in \mathcal{A})$ an eigenvector associated to θ. We give a proof for

$$\liminf_{n\to\infty} |(\gamma(x_0) + \ldots + \gamma(x_n))| < C.$$

The other estimate is quite similar to obtain.
The prefix-suffix sequence $(p_i, c_i, s_i)_{i\in\mathbb{N}}$ of x is a path in a *finite* automaton (the prefix-suffix automaton, see [CS]). Therefore, the prefixes and suffixes belong to a finite set $\mathcal{P} \times \mathcal{S}$ depending only on σ.

First Case We first assume that infinitely many suffixes are non empty in the prefix-suffix decomposition of x. Let k be a positive integer, by the prefix-suffix decomposition, we have

$$\begin{aligned} x = \ldots \sigma^{k+1}(p_{k+1})\sigma^n(p_n)\sigma^{k-1}(p_{k-1})\ldots\sigma^1(p_1)p_0.c_0s_0\sigma^1(s_1) \\ \ldots\sigma^{k-1}(s_{k-1})\sigma^k(s_k)\sigma^{k+1}(s_{k+1})\ldots \end{aligned}$$

We denote by $\hat{P}_k$ the word $\sigma^{k-1}(p_{k-1})\ldots\sigma^1(p_1)p_0$ and by $\hat{S}_k$ the word $c_0s_0\sigma^1(s_1)\ldots\sigma^{k-1}(s_{k-1})$. Thus we have

$$x = \ldots\sigma^{k+1}(p_{k+1})\sigma^k(p_k)\hat{P}_k.\hat{S}_k\sigma^k(s_k)\sigma^{k+1}(s_{k+1})\ldots$$

and

$$\sigma^k(c_k) = \hat{P}_k.\hat{S}_k.$$

Assume that s_{k+1} is non empty. Then $\sigma(s_{k+1})$ contains c_k because every letter appears in the image of every letter. We consider the first occurrence of c_k in $\sigma(s_{k+1})$ and get

$$\sigma(s_{k+1}) = \Pi_k c_k \Sigma_k,$$

where Π_k and Σ_k are words of bounded length (the bound only depends on σ). Consequently $\hat{S}_k\sigma^k(s_k)\sigma^k(\Pi_k)\sigma^k(c_k)\sigma^k(\Sigma_k)$ is a prefix of $x^+ = x_0x_1\ldots x_n\ldots$. This word is equal to $\hat{S}_k\sigma^k(s_k)\sigma^k(\Pi_k)\hat{P}_k\hat{S}_k\sigma^k(\Sigma_k)$. Thus, the word $W_k = \hat{S}_k\sigma^k(s_k)\sigma^k(\Pi_k)\hat{P}_k$ is a prefix of x^+.

We extend the function γ to the monoid $\mathcal{A}^*$ by the natural manner.

$$\gamma : \begin{array}{lcl} \mathcal{A}^* & \to & \mathbb{C} \\ w = w_1 \dots w_t & \mapsto & \displaystyle\sum_{i=1}^{t} \gamma(w_i). \end{array}$$

It is clear that the extension of γ is a morphism of monoid: if v and w are two words $\gamma(vw) = \gamma(v) + \gamma(w)$.
The key lemma is the following

Lemma 3.1. *The sequence $(\gamma(W_k))_{k\in\mathbb{N}}$ is bounded by a constant that only depends on σ and γ.*

Proof. By the hypothesis on θ and γ, for every $w \in \mathcal{A}^*$ and $m \in \mathbb{N}$,

$$\gamma(\sigma^m(w)) = \theta^m \gamma(w).$$

This implies that

$$|\gamma(\sigma^m(w))| = |\gamma(w)|.$$

We use this remark to calculate $\gamma(W_k) = \gamma(\hat{S}_k) + \theta^k\gamma(s_k) + \theta^k\gamma(\Pi_k) + \gamma(\hat{P}_k) = \gamma(\hat{S}_k\hat{P}_k) + \theta^k\gamma(s_k) + \theta^k\gamma(\Pi_k)$.
Thus, $\gamma(W_k) = \gamma(\hat{S}_k\Pi_k) + \theta^k(\gamma(s_k) + \gamma(\Pi_k)) = \gamma(\sigma^k(c_k)) + \theta^k(\gamma(s_k) + \gamma(\Pi_k)) = \theta^k(\gamma(c_k) + \gamma(s_k) + \gamma(\Pi_k))$.
This yields that

$$|\gamma(W_k)| \leq |\gamma(c_k)| + |\gamma(s_k)| + |\gamma(\Pi_k)|.$$

The length of the words c_k, s_k, Π_k are bounded independently on x. Thus, each factor on the right hand side of the inequality is bounded and the sequence $(|\gamma(W_k)|)_{k\in\mathbb{N}}$ is bounded independently on x.

□

Now we end the proof of Theorem 1.1 in the first case. We already proved that $(W_k)_{k\in\mathbb{N}}$ is a sequence of prefixes of x^+. By hypothesis, the length of W_k tends to infinity (because infinitely many suffixes s_n are non empty). By Lemma 3.1 the sequence $(|\gamma(W_k)|)_{k\in\mathbb{N}}$ is bounded independently on x. This proves Theorem 1.1 in the first case.

Second Case If the suffixes of x are eventually empty, it means that x^+ belongs to the negative orbit of a periodic point for σ. The proof is similar to the previous one. There exists $a \in \mathcal{A}$ and non-negative integers ℓ and q such that

$$x^+ = c_0 s_0 \sigma^1(s_1) \dots \sigma^\ell(s_\ell) \lim_{n\to\infty} \sigma^{nq}(a).$$

Let $S_\ell = c_0 s_0 \sigma^1(s_1) \dots \sigma^\ell(s_\ell)$; by construction of the prefix-suffix automaton, there exists a letter c and a finite word P_ℓ such that $\sigma^{\ell+1}(c) = P_\ell S_\ell$, thus

$$x = \dots P_\ell . S_\ell \lim_{n\to\infty} \sigma^{nq}(a).$$

Since the letter c appears in the word $\sigma(a)$, there exists words of bounded length Π and Σ such that $\sigma(a) = \Pi c \Sigma$. Thus for each $n > \ell+2$, $S_\ell \sigma^{nq}(a) = S_\ell \sigma^{nq-\ell-2} \sigma^{\ell+1}(\Pi c \Sigma)$ is a prefix of x^+. Moreover

$$S_\ell \sigma^{nq-\ell-2}(\sigma^{\ell+1}(\Pi)) \sigma^{nq-\ell-2}(\sigma^{\ell+1}(c)) = S_\ell \sigma^{nq-\ell-2}(\sigma^{\ell+1}(\Pi)) \sigma^{nq-\ell-2}(P_\ell S_\ell).$$

Set $W_n = S_\ell \sigma^{nq-\ell-2}(\sigma^{\ell+1}(\Pi)) \sigma^{nq-\ell-2}(P_\ell)$; it is a prefix of x^+.
Now $\gamma(W_n) = \gamma(S_\ell) + \gamma(\Pi) + \gamma(P_\ell) = \gamma(\sigma^{\ell+1}(c)) + \gamma(\Pi) = \gamma(c) + \gamma(\Pi)$. The absolute value of this quantity is bounded. The bound only depends on σ and γ since c is a letter and Π belongs to a finite set (prefixes of $\sigma(a)$).

4. Proof of Corollaries 1.5 and 1.6

4.1. **Proof of Corollary 1.5.** Let $T = T_{(\lambda,\pi)}$ be a self-similar interval exchange transformation on r intervals and $\Gamma = (\gamma_1, \dots, \gamma_r)$ a vector orthogonal to λ. The set of affine interval exchange transformations semi-conjugated to $T_{(\lambda,\pi)}$ with slopes $w = (e_1^\gamma, \dots, e_r^\gamma)$ is never empty (see [CG], [C]); this set is denoted by $S_{aff}(T, w)$.
Cobo gave an if and only if condition ensuring the existence of f in $S_{aff}(T, w)$ which is not conjugated to T (see [C] page 392–393). This condition is the following:
there exists a point $x \in X_\sigma$ such that

$$\sum_{n\geq 1} e^{\gamma(x_{-n}\dots x_{-1})} + \sum_{n\geq 1} e^{-\gamma(x_0\dots x_{n-1})} < \infty. \tag{4.1}$$

Otherwise every element of $S_{aff}(T, w)$ is conjugated to T by a homeomorphism.
Theorem 1.1 says that, if Γ is an eigenvector associated to an eigenvalue of modulus one, then for every $x \in X_\sigma$ the sequence $(\gamma(x_0 \dots x_{n-1}))_{n\in\mathbb{N}}$ possesses a bounded subsequence. Therefore, if $\theta = \pm 1$,

$$\sum_{n\geq 1} e^{\gamma(x_{-n}\dots x_{-1})} + \sum_{n\geq 1} e^{-\gamma(x_0\dots x_{n-1})}$$

diverges for every x $\in X_\sigma$.
If $\theta \in \mathbb{C} \setminus \mathbb{R}$, by assumption, there exist complex numbers a and b such that $\Gamma = aV_\phi + bV_{-\phi}$. We apply Theorem 1.1 to the sequences $(V_\phi(x_0 \dots x_{n-1}))_{n\in\mathbb{N}}$ and $(V_{-\phi}(x_0 \dots x_{n-1}))_{n\in\mathbb{N}}$. It is clear in the proof of Theorem 1.1 that one can find a subsequence (n_k) such that $(V_\phi(x_0 \dots x_{n_k-1}))_{k\in\mathbb{N}}$ and $(V_{-\phi}(x_0 \dots x_{n_k-1}))_{k\in\mathbb{N}}$ are simultaneously bounded.

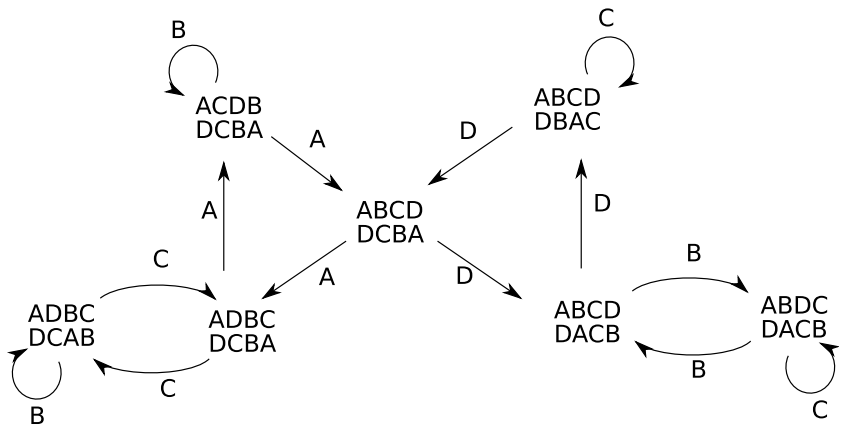

FIGURE 1. Rauzy diagram.

Thus equation (4.1) cannot be fulfilled which completes the proof of Corollary 1.5.

4.2. **Proof of Corollary 1.6.** From Theorem 1.1, it is enough to give an infinite family of self-similar interval exchanges transformations on 4 intervals that have eigenvalues of modulus 1. The examples are obtained in the most simple Rauzy class (see figure 1).
We recall that a Rauzy class is the closure of a given pair of permutations under Rauzy induction and that self similar interval exchange transformations correspond to loops in the Rauzy diagram. Marmi-Moussa-Yoccoz gave an if and only if condition ensuring that a loop is realized by a self similar interval exchange transformation. In our situation, the Marmi-Moussa-Yoccoz criterion says that a loop is admissible if all the labels A, B, C, D appear along the loop (see [MMY1]). This condition will be fulfilled by the family of examples studied here.
Every arrow of the diagram induces a substitution. Given an interval exchange transformation T on an interval I, let T' be its image under Rauzy induction acting on I'. The substitution is defined by following the images of elements of the partition I' in the partition I under T until they come back to I'. The product of these substitutions gives the substitution corresponding to the loop or to the self similar interval exchange transformation associated to the loop. Thus a loop is labelled by a finite word on the alphabet $\{A, B, C, D\}$.
The symmetric permutation is

$$\begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$$

Proposition 4.1. *For every $n \geq 1$, the matrix M_n of the path starting from the symmetric permutation labelled by the loop $DBBDC^nDAAA$ has the following properties:*

(1) *The Perron-Frobenius eigenvalue $\theta_1^{(n)}$ is an algebraic integer of degree 4.*
(2) *It has two conjugate of modulus 1, $\theta_2^{(n)} = \overline{\theta_3^{(n)}}$ (it is a Salem number).*
(3) *If $\theta_2^{(n)} = e^{2i\pi\alpha_n}$, then α_n is an irrational number.*

Proof. The substitution on the alphabet $\{1, 2, 3, 4\}$ is

$$\sigma_n = \begin{cases} 1 & \mapsto & 14 \\ 2 & \mapsto & 14224 \\ 3 & \mapsto & 14(23)^{n+1}24 \\ 4 & \mapsto & 14(23)^n 24. \end{cases}$$

The corresponding matrix is

$$M_n = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & n+2 & n+1 \\ 0 & 0 & n+1 & n \\ 1 & 2 & 2 & 2 \end{pmatrix}.$$

The characteristic polynomial of M_n is

$$P_n = X^4 - (6+n)X^3 + (10+n)X^2 - (6+n)X + 1.$$

By the general theory of interval exchange transformations (see [Yo] [Zo2]), P_n is a reciprocal polynomial. Thus, the roots $\theta_1, \theta_2, \theta_3, \theta_4$ of P_n satisfy $\theta_4 = 1/\theta_1$, $\theta_3 = 1/\theta_2$ and $\theta_1 > |\theta_2| \geq |\theta_2| > \theta_4$ (as n is given, the dependence on n is omitted in the sequel).
We first prove that P_n has two complex roots of modulus 1. Let $t = \theta_1 + 1/\theta_1$ and $s = \theta_2 + 1/\theta_2$. We remark that θ_2 and θ_3 are complex conjugate if and only if $-2 \leq s \leq 2$. In fact, if θ_2 is real and larger than 1, $s > 2$; if θ_2 has modulus 1, $\theta_2 = e^{2i\pi\alpha}$ and $s = 2\cos(2\pi\alpha)$.
A simple calculation yields the equations $t + s = 6 + n$ and $ts = 8 + n$. Thus, $s = \frac{6+n-\sqrt{n^2+8n+4}}{2}$. We easily check that this quantity belongs to $(-2, 2)$ independently on n. Thus θ_2 has modulus 1.
Now, we prove that P_n is irreducible. Otherwise, θ_2 is ± 1 or an algebraic integer of degree 2. A direct computation proves that ± 1 are not roots of P_n. Moreover, the only irreducible polynomials of degree 2, with roots of modulus 1 are the cyclotomic polynomials $X^2 + X + 1$, $X^2 + 1$, $X^2 - X + 1$ (the polynomial has the form $X^2 - sX + 1$ with $-2 < s < 2$). To prove that P_n is irreducible, we check that it is not a multiple of $X^2 + X + 1$, $X^2 + 1$, $X^2 - X + 1$.

We end the proof by showing that α is irrational. If α is rational it means that θ_2 is a root of unity. This is impossible since $\theta_1 > 1$ is a conjugate of θ_2.

$\square$

4.3. **Proof of Proposition 1.7.** Denote by τ the non-trivial Galois automorphism of K and by $\Gamma(X)$ the Veech group of X. Given $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, its Galois conjugate is $\tau(A) = \begin{pmatrix} \tau(a) & \tau(b) \\ \tau(c) & \tau(d) \end{pmatrix}$. Denote by Γ' the image of $\Gamma(X)$ under Galois conjugacy. Let ϕ be an affine diffeomorphism of X and A the corresponding element of the Veech group, its action on $H_1(X, \mathbb{R})$ is by $\begin{pmatrix} A & 0 \\ 0 & \tau(A) \end{pmatrix}$ in a suitable basis.
The group Γ' is a non discrete and non elementary subgroup of $\mathrm{SL}_2(\mathbb{R})$. By a result of Beardon [Be], it contains an elliptic element of infinite order. Let $A' = \tau(A'')$ be this element. If A'' is hyperbolic, then Proposition 1.7 follows. We now show that A'' cannot be elliptic nor parabolic. As $\Gamma(X)$ is discrete, every elliptic element of $\Gamma(X)$ is of finite order. But $A' = \tau(A'')$ is an element of infinite order. Consequently, A'' is not an elliptic element. Now the image of a parabolic element under Galois conjugacy is again parabolic (since an element of $\mathrm{SL}_2(\mathbb{R})$ is parabolic if and only if the absolute value of its trace is 2), so A'' cannot be parabolic. Therefore A'' is indeed hyperbolic and is the derivative of a pseudo-Anosov diffeomorphism with the desired properties.

5. Asymptotic behaviour of ergodic sums

5.1. **Ergodic means.** We are interested in the distribution of the ergodic integral

$$\int_0^{\lambda_1^n t} f \circ h_s(x) ds$$

considered as a random variable on the probability space (X, ν_Γ) and normalized to have variance 1.
The first observation is that for all $n \in \mathbb{Z}$, the law of x is the same as the law of $\mathfrak{S}^n(x)$ by $\mathfrak{S}$-invariance of the measure ν_Γ (we mean: the law of Id is the same as the law of $\mathfrak{S}^n$). Hence the distribution of the ergodic integral is the same as that of the sequence

$$\int_0^{\lambda_1^n t} f \circ h_s(\mathfrak{S}^{-n}(x)) ds.$$

Let us state this result more formally:

Lemma 5.1. *For any measurable function $\varphi : Y \to \mathbb{R}$, all $z \in \mathbb{R}$ and all $m \in \mathbb{Z}$,*

$$\nu(\{x \in Y : \varphi(x) \leq z\}) = \nu(\{x \in Y : \varphi(\mathfrak{S}^m x) \leq z\}).$$

From here till the end of the paper σ is a substitution of constant length d, thus $\lambda_1 = d$. We choose a function f constant on rectangles (defined in Subsection 2.3). We are interested in the case when this function "corresponds" to the coordinates of an eigenvector associated with eigenvalue 1, i.e. an invariant vector (ie $f(\sigma(a)) = f(a)$ for every letter a).

5.2. **Decomposition of unstable leaves.** Pieces of unstable leaves on which we compute the ergodic averages cross transversally the "rectangles" of the partition. To such a piece of unstable leave we associate a symbolic sequence given by the names of the successive rectangles it does cross. The self similar structure implies that these symbolic sequences are words in the language of the substitution σ. More precisely, such a piece of leaf starts inside a rectangle (the initial rectangle) then crosses transversally a given number of rectangles, to end up inside a last rectangle (the final rectangle). Here, we try to describe the symbolic sequence using the self-similar structure. We have to take care of the initial and final rectangles which are not completely crossed. For (important) technical reasons we distinguish not only the final but the last two rectangles.

Let $x \in Y$ and $t \in [1; \infty)$. We consider a piece γ_t^x of unstable leave determined by its starting point x and its length t. The sequence $(x_n)_{n\in\mathbb{Z}}$ determines a bi-infinite sequence of prefixes/center/suffixes (P_n^x, c_n^x, S_n^x) satisfying for all $n \in \mathbb{Z}$, $F(x_n) = c_n^x$ and $\sigma(c_{n+1}^x) = P_n^x c_n^x S_n^x$. We observe that around x the symbolic sequence (sequence of codes of rectangles seen along the unstable leave through x) is (note that this decomposition stops if all suffixes are empty)

$$\omega^x = \cdots \sigma^n(P_n^x) \cdots \sigma(P_1^x) P_0^x . c_0^x S_0^x \sigma(S_1^x) \cdots \sigma^n(S_n^x) \cdots .$$

We compute the distance from x to the boundary of the rectangle in which it is by $t^- = \sum_{n<0} d^n |S_n^x|$.
We decompose (in basis d) : $t = \sum_{n\leq M} d^n \tau_n$ and let $k_t = \lfloor t \rfloor - 1 = \sum_{n=0}^{M} d^n \tau_n - 1$. We observe that $\gamma_t(x)$ crosses at least k_t rectangles. It does cross one more only if $\sum_{n<0} d^n |P_n^x| + \sum_{n<0} d^n \tau_n > 1$ (which happens only if there is H such that for all $h < H$, $|P_{-h}^x| + \tau_{-h} = d-1$, and $|P_{-H}^x| + \tau_{-H} \geq d$).
We denote $K = K_t$ the greater integer such that $d^K < k_t$. We observe that $\ell_x = |S_0^x \sigma(S_1^x) \cdots \sigma^K(S_K^x)| < k_t$. We also observe that

there is a letter W such that $|S_0^x\sigma(S_1^x)\cdots\sigma^K(S_K^x)\sigma^{K+1}(W)| > k_t$ and that $S_0^x\sigma(S_1^x)\cdots\sigma^K(S_K^x)\sigma^{K+1}(W)$ is a factor of ω^x. We decompose the prefix of length $k_t - \ell_x$ of $\sigma^{K+1}(W)$ into prefixes in such way that $\sum_{i=0}^K d^i|S_i^x| + \sum_{i=0}^K d^i|P_i^t| = k$ and $\prod_{i=0}^K \sigma^i(S_i^x)\prod_{i=K}^0 \sigma^i(P_i^t) = \omega_1^x\cdots\omega_{k_t}^x$.
Then, we set $t^+ = t - k - t^-$. We have to check if whether $t^+ < 1$ or $1 \le t^+ < 2$. We use the identity $f(\sigma(a)) = f(a)$ to get, if $t^+ < 1$, the following equation:

$$\int_0^t f\circ h_s(x)ds = f(a_0)t^- + \sum_{n=1}^K f(S_n^x) + \sum_{n=K}^1 f(P_n^t) + t^+ f(a_k), \tag{5.1}$$

and, if $1 \le t^+ < 2$,
(5.2)

$$\int_0^t f\circ h_s(x)ds = f(a_0)t^- + \sum_{n=1}^K f(S_n^x) + \sum_{n=K}^1 f(P_n^{x,t}) + f(a_k) + (t^+-1)f(a_{k+1}).$$

It is important to notice that except for the boundary terms, the decomposition depends on x through $(x_n)_{n\ge 0}$ and of t through $\lfloor t\rfloor$. The other data tell about what happens more precisely in the boundary terms. We get

$$\left|\int_0^t f\circ h_s(x)ds - \left(\sum_{n=1}^K f(S_n^x) + \sum_{n=K}^1 f(P_n^t)\right)\right| \le 3||f||_\infty. \tag{5.3}$$

If we step to $d^n t$ and shift the starting point to $\sigma^{-n}(x)$, we will obtain the same decomposition and something more precise on boundary terms. Hence, to get recursively to the next step we only have to keep track of c_0^x and the word $a_{k_t}a_{k_t+1}$. To encode this information, we define a family of automata which is described in the next paragraph.

5.3. **A family of automata.** The following family of automata associated with the substitution σ will allow us approximate the ergodic averages in a Markovian way.

For all $W \in \mathcal{A}$ and all integer $1 \le m \le d$, we decompose $\sigma(W) = P^{W,m}c^{W,m}S^{W,m}$ with $P^{W,m} \in \mathcal{A}^{m-1}$, $c^{W,m} \in \mathcal{A}$ and $S^{W,m} \in \mathcal{A}^{d-m}$.
For all $\tau \in \{0,\dots,d-1\}$ we define an automaton $\mathcal{A}_\tau$. States are couples $(a,W) \in \mathcal{A}\times\mathcal{A}^2$ and edges are labelled by a real number v and an integer m. For all $v \in \mathbb{R}$, $m \in \{1,\dots,d\}$ and integers and $a, b, \in \mathcal{A}$, $V, W \in \mathcal{A}^2$ we put a labelled arrow

$$(a,V) \overset{v,m}{\to} (b,W)$$

if $b = c^{a,m}$, $W = c^{V,m+\tau}c^{V,m+\tau+1}$ and $v = \Pi_1(e_{S^{a,m}} + e_{PV,m+\tau})$. Here the word $\sigma(V)$ has length $2d$, thus it has a prefix of length $m+\tau$. Observe that in particular, $\sigma(a) = P^{a,m}bS^{a,m}$ so that the outgoing degree d of any vertex is the same as the outgoing degree of the corresponding vertex in the prefix/suffix automaton. Note that for $\tau = 0$ we can forget the second letter of V and W; indeed since $m \leq d$, $c^{V,m}$ is determined by the first letter of V.
Examples of such automata are described in Section 8. The readers should refer to this section to understand concrete examples.

5.4. **Ergodic sums and automata.** We fix $t \in [1; d)$ and write $t = \sum_{n\leq 0} \tau_n d^n$ (note that $\tau_0 > 0$). We let W be a word of length d i.e. $W \in \mathcal{A}^d$ and choose $x \in Y$ such that $\omega^x \in [W]$, i.e. $\omega_0^x \cdots \omega_{d-1}^x = W$ or again $F(x_0)F((T_Y x)_0) \cdots F((T_Y^{d-1}x)_0) = W$. The sequence $(x_n)_{n\in\mathbb{Z}}$ determines a bi-infinite sequence of prefixes/centres/suffixes (P_n^x, c_n^x, S_n^x) satisfying for all $n \in \mathbb{Z}$, $F(x_n) = c_n^x$ and $\sigma(c_{n+1}^x) = P_n^x c_n^x S_n^x$.
We observe that around x the symbolic sequence (sequence of codes of rectangles seen along the unstable leave through x) is (except if all prefixes are empty)

$$\omega^x = \cdots \sigma^n(P_n^x) \cdots \sigma(P_1^x)P_0^x.c_0^x S_0^x \sigma(S_1^x) \cdots \sigma^n(S_n^x) \cdots .$$

We note that $c_0^x = W_0$. In the simpler case, W is the prefix of length d of $c_0^x S_0^x \sigma(S_1^x)$; or of $c_0^x S_0^x \sigma^l(S_l^x)$ where S_l^x is the first non empty suffix after S_0^x. If all suffixes are empty, then we must look at $T_Y^k(x)$. Anyhow, we set $a^0 := c_0^x = W_0$ and $V^0 := W_{\tau_0}W_{\tau_0+1}$. We also set $U^0 = W_1 \cdots W_{\tau_0-1}$ (may be empty, if $\tau_0 = 1$).

We construct recursively a sequence $(a^n, V^n, U^n)_{n\geq 0}$. Assume (a^n, V^n, U^n) determined for some n. Let then m be the length $m = |P_{-(n+1)}^x| + 1 (= d - |S_{-(n+1)}^x|)$. We set

$$\begin{cases} a^{n+1} & := & c_{-(n+1)}^x = c^{a^n,m} \\ V^{n+1} & := & V_{\tau_{-n}+m}^n V_{\tau_{-n}+m+1}^n = c^{V^n,\tau_{-n}+m}c^{V^n,\tau_{-n}+m+1} \\ U^{n+1} & := & S_{-(n+1)}^x \sigma(U^n)V_1^n \cdots V_{\tau_{-n}+m-1}^n. \end{cases}$$

Remark 5.2. *This sequence (a^n, V^n) is obtained by the family of automata described in the previous section. The sequence of automata depends on the digits of t.*

Since f depends only on the first coordinate, for brevity we write $f(W)$ for the sum $f(W_1) + \cdots + f(W_k)$ if $W = W_1 \cdots W_k$. We set $t_-^n = d^n \sum_{k<-n} d^k |S_{-k}^x|$ and $t_+^n = d^n \sum_{k<-n} d^k(\tau_k + |P_{-k}^x|)$ and observe that $0 < t_+^n < 2$.

Lemma 5.3. *If $t_+^n < 1$, then*

$$\int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds = f(a_n)t_-^n + f(U^n) + t_+^n f(V_1^n).$$

If $t_+^n \geq 1$, then

$$\int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds = f(a_n)t_-^n + f(U^n) + f(V_1^n) + (t_+^n - 1)f(V_2^n).$$

In particular, we have

$$\left| \int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds - f(U^n) \right| \leq 3||f||_\infty.$$

Proof. The symbolic decomposition around $\mathfrak{S}^m(x)$ is

$$\omega^{\mathfrak{S}^m x} = \cdots \sigma^n(P_{m+n}^x) \cdots \sigma(P_{m+1}^x) P_m^x . c_m^x S_m^x \sigma(S_{m+1}^x) \cdots \sigma^n(S_{m+n}^x) \cdots .$$

Hence for $m = -n$, recalling that $\omega^x \in [W]$, we obtain (looking at the right hand part)

$$\omega_+^{\mathfrak{S}^{-n}x} = c_{-n}^x S_{-n}^x \sigma(S_{-n+1}^x) \cdots \sigma^{n-1}(S_{-1}^x) \sigma^n(W_1 \cdots W_d) \cdots$$

By construction, U^n is a prefix of this word (shifted once). It remains to compute its length. We observe that $|U^0| = \tau_0 - 1$. But $|U^{n+1}| = |S_{-(n+1)}^x| + d|U^n| + \tau_{-(n+1)} + m - 1 = d|U^n| + d + \tau_{-(n+1)}$. Hence if we assume $|U^n| = d^n(\sum_{k=-n}^0 \tau_k d^k) - 1$, we obtain $|U^{n+1}| = d^{n+1}(\sum_{k=-n}^0 \tau_k d^k) - d + d + \tau_{-(n+1)} = d^{n+1} \sum_{k=-(n+1)}^0 \tau_k d^k - 1 (= \lfloor d^{n+1} t \rfloor - 1)$. In view of the definitions of t_+^n and t_-^n, combining this discussion with equations (5.1) and (5.2), we get the statement of the lemma. □

To pass from limit theorems for the Markov chain to those for the ergodic integral, we need to combine Lemma 5.3 with the following simple observation.

Proposition 5.4. *Let $X_1^{(n)}$ and $X_2^{(n)}$, $n \in \mathbb{N}$, be two sequences of random variables on a probability space $(\Omega, \mathbb{P})$ satisfying the conditions:*

1. *there exists a constant K such that for all n we have the inequality $|X_1^{(n)} - X_2^{(n)}| \leq K$ almost surely;*
2. *$Var(X_2^{(n)}) \to \infty$ as $n \to \infty$.*

Then

$$\lim_{n\to\infty} \frac{Var(X_2^{(n)})}{Var(X_1^{(n)})} = 1.$$

Proof. First, note that we only need to consider the case when $\mathbb{E}X_1^{(n)} = \mathbb{E}X_2^{(n)} = 0$, since if $X_1^{(n)}$ and $X_2^{(n)}$ satisfy the assumptions of the Proposition, then so do $X_1^{(n)} - \mathbb{E}X_1^{(n)}$ and $X_2^{(n)} - \mathbb{E}X_2^{(n)}$. If $\mathbb{E}X_1^{(n)} = \mathbb{E}X_2^{(n)} = 0$, then we have

$$|\mathbb{E}(X_1^{(n)})^2 - \mathbb{E}(X_2^{(n)})^2| \leq K\left(K + \mathbb{E}|X_1^{(n)}|\right) \leq K\left(K + \sqrt{Var(X_2^{(n)})}\right)$$

by the Cauchy-Buniakovsky-Schwarz inequality, and the proof is complete. □

6. Markov approximation and the proof of Theorem 1.2

The following construction plays a basic role in the sequel.

6.1. **Markov chains.** Let $t \in \mathbb{R}_+^*$. Let n_0 be the unique integer such that $d^{n_0}t \in [1; d)$ and set $t = \sum_{k \leq n_0} \tau_k d^k$. For now, we work with $n_0 = 1$. Let $(\Omega, \mathcal{F}, P)$ be a probability space. We construct a random variable I on Ω valued in $\mathcal{A}^d$ with distribution

$$P(I = W) = \nu(\{x \in Y : \omega_0^x \cdots \omega_d^x = W\}).$$

We define a map $g_0^t : \mathcal{A}^d \to \mathbb{R}$ by $g_0(W) = f(W_1 \cdots W_{k_t - 1})$. Let now $(m_n)_{n \geq 1}$ be a sequence of iid random variables uniformly distributed in $\{1, \ldots, d\}$ independent of I. We define a process valued in $\mathcal{E} = \mathcal{A} \times \mathcal{A}^2$ by setting $X_0 = (I_1, I_k I_{k+1})$ and recursively $X_{n+1} = (a^{n+1}, V^{n+1})$ to be the end vertex of the edge labelled m_n of the automaton $A_{\tau_{-n}}$ starting from $X_n = (a^n, V^n)$. Finally, we set

$$g_n^t(X_n, m_n) = f(S^{a^n, m_n}) + f(P^{V^n, m_n + \tau_{-n}}). \tag{6.1}$$

Lemma 6.1. *The sequence $(X_n)_{n \geq 0}$ is a Markov chain.*
We have the following commutative diagram

$$\begin{array}{ccc} \mathcal{E}^{\mathbb{N}} & \xrightarrow{\mathfrak{S}_{\mathfrak{t}}} & \mathcal{E}^{\mathbb{N}} \\ \downarrow\pi & & \downarrow\pi \\ Y & \xrightarrow{\mathfrak{S}} & Y \end{array}$$

where $\mathfrak{S}_{\mathfrak{t}}$ is the shift on the Markov chain. The probability measure P invariant by the Markov chain projects to ν.

Proof. $(X_n)_{n \geq 0}$ is a Markov chain since (a_{n+1}, V_{n+1}) only depends on (a_n, V_n). Moreover, there is a projection π between infinite paths of (X_n) and Y. This map consists in forgetting the second variable V^n. Formally, the map is defined by : $m_n = |P_n^x| + 1$. To check that the image measure is ν under the projection, it suffices to observe that $\sum_{V \in \mathcal{A}^2} P(X_0 = (a, V)) = \nu(\{x \in Y : \omega_0^x = a\})$ which is obvious. □

We claim that

Lemma 6.2.

$$f(U^n) = g_0(X_0) + \sum_{k=1}^{n} g_k(X_k, m_k).$$

Proof. It is a straightforward computation. Check that $f(U^0) = g_0(X_0)$. Then write recursively $U^{n+1} = S^x_{-(n+1)}\sigma(U^n)V^n_1 \cdots V^n_{\tau_{-n}+m-1}$. Hence $f(U^{n+1}) - f(U^n) = f(S^x_{-(n+1)}) + f(V^n_1 \cdots V^n_{\tau_{-n}+m-1})$. □

Now, we assume that t has a periodic expansion thus the Markov chain is homogeneous. To simplify notations, we will forget the dependence in t. Up to a change of basis, we assume that the expansion of t has period 1, all the digits are equal, i.e. for all k, $\tau_k = \tau$.

6.2. **Proof of Theorem 1.2.**
The Markov chain (X_n) is endowed with an initial probability measure μ_0 defined on $\mathcal{A} \times \mathcal{A}^2$ and a transition matrix Q.

Lemma 6.3. *Let $B = \{V \in \mathcal{A}^2, \nu([B]) > 0\}$. The Markov chain (V_n) is recurrent on $\mathcal{A} \times B$.*

Proof. The proof is straightforward since the substitution is primitive. □

Observe that the processes (a_n) and (V_n) are also Markov chains. Let $\tilde{\mu}$ be an invariant measure for the Markov chain (X_n).

Lemma 6.4. *The initial measure $\tilde{\mu_0}$ of the Markov chain (X_n) projects on the first coordinates on the invariant measure for the Markov chain (a_n) and on the last coordinate on the invariant measure for the Markov chain (V_n).*

Proof. The initial measure satisfies:

$$\mu_0(a, m, V) = P(I_1 = a, I_{k_t I_{k_t+1}} = V)P(m_0 = m).$$

It projects on $\mu^1_0(a) = \nu([a])$ and on $\mu^2_0(V) = \nu([V])$. These measures are obviously invariant under the Markov chains. □

An important lemma is the following:

Lemma 6.5. *For every integer n and for every probability measure $\tilde{\mu}$ invariant by the chain (X_n), we have $E[g_n(X_n)] = 0$ and $E_{\tilde{\mu}}[g_0(X_0)] = 0$.*

Proof. We recall that $E_\nu[\int_0^{d^n t} f\circ h_s(x)ds] = 0$ since $E[f] = 0$ by hypothesis. This does not *a priori* imply the conclusion of the Lemma since the Markov chain (X_n) is not always recurrent. We must exclude the possibility the averages on different ergodic components compensate each other.
By Lemma 6.2, there exists a constant K such that, independently on n,

$$E[\sum_{k=1}^{n} g_k(X_k, m_k)] < K. \tag{6.2}$$

Since expectation is a linear operator, and in view of the definition of g_n (recall equation (6.1)), we can decompose

$$E[g_n(X_n, m)] = E[g_n^1(a_n, m)] + E[g_n^2(V_n, m)]. \tag{6.3}$$

Let $\tilde{\mu}$ be any invariant measure under (X_n). We have:

$$E_{\tilde{\mu}}[g_0(X_0, m)] = E_{\tilde{\mu}}[g_n(X_n, m)] = E_{\tilde{\mu}}[g_n^1(a_n, m)] + E_{\tilde{\mu}}[g_n^2(V_n, m)].$$

By Lemma 6.4 and equation (6.3) this yields

$$E_{\tilde{\mu}}[g_0(X_0, m)] = E[g_n^1(a_n, m)] + E[g_n^2(V_n, m)] = E[g_n(X_n, m)].$$

Therefore, $E[g_n(X_n, m)]$ is independent on n. It follows from (6.2) that $E[g_n(X_n, m)] = 0$ and $E_{\tilde{\mu}}[g_0(X_0] = 0$.

□

We just proved that for every recurrent class the mean of the invariant measure supported by this class is equal to zero. Thus the central limit theorem for stationary Markov chains implies the random variable $\frac{1}{\sqrt{n}}\sum_{k=1}^{n} g_k(X_k, m_k)$ has a limiting distribution. The contribution of each recurrent class is a normal law.

Remark 6.6. Observe that, as a consequence, for every positive ε,

$$\frac{1}{n^{\frac{1}{2}+\varepsilon}} \sum_{k=1}^{n} g_k(X_k, m_k)$$

converges in distribution to 0.

To finish the proof of Theorem 1.2, it is enough to show:

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \int_0^{d^n t} f \circ h_s(x)ds \overset{(d)}{=} \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g_k(X_k, m_k). \tag{6.4}$$

In other words :

Proposition 6.7. *For all bounded continous* $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\lim_{n\to\infty} \int_Y \varphi\left(\frac{1}{\sqrt{n}} \int_0^{d^n t} f \circ h_s ds\right) \nu = \lim_{n\to\infty} E\left(\varphi\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n g_k(X_k, m_k)\right)\right).$$

Proof. By Lemmas 5.3 and 6.2, the quantities $\int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds$ and $\sum_{k=1}^n g_k(X_k, m_k)$ are equal up to a bounded term. As we divide by $\sqrt{n}$ a bounded term is negligible. Thus the limit distribution of the sequences

$$\frac{1}{\sqrt{n}} \int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds \text{ and } \frac{1}{\sqrt{n}} \sum_{k=1}^n g_k(X_k, m_k)$$

are the same.
By Lemma 5.1, the law of

$$\int_0^{d^n t} f \circ h_s(x)ds$$

is equal to the law of

$$\int_0^{d^n t} f \circ h_s(\mathfrak{S}^{-n}x)ds.$$

This proves the proposition. □

Together with Proposition 5.4, this implies Theorem 1.2.

7. Proof of Theorem 1.9

7.1. **Assumptions on the Markov chain.** In the beginning of Section 5, to a substitution σ we assigned a family of automata $\{A_i, i \in \{1, \dots, \}$ in such a way that to any $t \in [0, 1]$ with decomposition in basis d (set $\mathcal{A} = \{0, \dots, d-1\}$) given by $t = \sum_{n>0} \tau_n d^{-n}$ we associate a sequence of automata in the family, $A_{\tau_0}, A_{\tau_1}, \dots$. To such t we also associate a sequence of functions (g_k^t) defined on the state space E of the underlying Markov chain ($g_k^t = g_{\tau_k}$). When t has a periodic expansion in basis d, the sequence of automata is periodic and hence the associated Markov chain is a standard homogeneous Markov chain.
Here we assume that there is such a t_0 for which the Markov chain is *aperiodic and has positive variance*. Let us first explain the meaning of this assumption. We stress that it is not only an assumption on the Markov chain but also on the functions (g_k^t). Let $\chi = \chi_1 \cdots \chi_p$ denote a word in $\mathcal{A}^p$ such that $t_0 = \sum_{n<0}(\sum_{i=1}^p \chi_i d^{-i})d^{-np}$. First of all, we assume that the matrix of the Markov chain associated with t_0 is primitive (without loss of generality, we can assume that the matrix has strictly positive entries). The positive variance asumption means

that the sequence of functions g^{t_0} does not yield a coboundary. Hence there are two cycles on which it takes two different values.
In slightly different words, it implies that there are two paths u and v in E^{p+1} (with $u_0 = v_0$ and $u_p = v_p$) and some $\epsilon > 0$ such that

$$\left| \sum_{k=1}^{p} (g_{\chi_k}(u_k)) - \sum_{k=1}^{p} (g_{\chi_k}(v_k)) \right| \geq \epsilon > 0, \tag{7.1}$$

We claim that under this assumption, for Lebesgue almost all t,

$$\frac{\sum_{k=0}^{n-1} (g_k^t(X_k) - E[g_k^t(X_k)])}{\sqrt{V(\sum_{k=0}^{n-1} g_k^t(X_k)))}} \Rightarrow N(0,1). \tag{7.2}$$

As we shall see more precisely in Section 7.4, by Lemma 5.3 and Lemma 6.2, convergence (7.2) implies Theorem 1.9.

7.2. **Choice of a full-measure set of admissible sequences.** Observe that the Lebesgue measure on $[0,1]$ maps onto the uniform product measure on $\mathcal{A}^{\mathbb{N}}$ by the map sending a real number to its expansion in basis d. Hence, with probability 1, the expansion of t contains infinitely many occurrences of $\kappa = \chi^3$ (we stress that we ask for three successive occurrences of χ).
Denote (q_n) the sequence of occurrences of κ in the expansion of t and $m_n = q_n + 3p$ (recall $p = |\chi| = |\kappa|/3$; fix $m_0 = 0$). The sequence $(m_n - m_{n-1})$ is a sequence of independent identically distributed random variables with geometric law. We claim that it does not grow too fast.
We use Borel-Cantelli Lemma to say that the inequality $B_n = \{m_n - m_{n-1} > n^{\frac{1}{4}}\}$ almost surely holds only for a finite number of $n \in \mathbb{N}$ since $\sum_{n\geq 1} \mathbb{P}(B_n) = \sum_{n\geq 1} c\,\theta^{n^{1/4}} < \infty$. Let $J \subset [0,1]$ be the set of probability 1 on which the expansion of t contains infinitely many occurrences of κ and B_n is true a finite number of times. Observe that for all $t \in J$, there is C (depending on t) such that for all $n \in \mathbb{N}$,

$$m_n - m_{n-1} < C n^{\frac{1}{4}}. \tag{7.3}$$

7.3. **A Central Limit Theorem for non homogeneous Markov chains.**

7.3.1. *The Dobrushin Theorem.* We recall Dobrushin's central limit theorem for non-homogeneous Markov chains (see [Dob]). We follow the exposition by Sethuraman and Varhadan [SV]. We only state the theorem for Markov chains defined on a finite state space, since that result is sufficient for our purposes.

For each $n \geq 1$, let $\{X_i^{(n)} : 1 \leq i \leq n\}$ be n observations of a non-homogeneous Markov chain on a finite state space E with transition matrices $\{\pi_{i,i+1}^{(n)} : 1 \leq i \leq n-1\}$ and initial distribution $\mu^{(n)}$.
Recall that, for discrete Markov chains, the Dobrushin's ergodic coefficient of a transition matrix π is defined as $\alpha(\pi) = 1 - \delta(\pi)$ where

$$\delta(\pi) = \sup_{a,b \in E; c \in E} |\pi(a,c) - \pi(b,c)|.$$

Recall the well-known inequality:

$$\alpha(\pi) \geq \min_{a,c \in E} \pi(a,c). \tag{7.4}$$

Moreover if η is a continuous function its oscillation is $Osc(\eta) = \sup_{\omega,\omega'} |\eta(\omega) - \eta(\omega')|$. If η is a function defined on the states of a Markov chain, we have the property : $Osc(\pi\eta) \leq \delta(\pi) Osc(\eta)$.
Let $\alpha_n = \min_{1 \leq i \leq n-1} \alpha(\pi_{i,i+1}^{(n)})$. Let $\{f_i^{(n)} : 1 \leq i \leq n\}$ be real valued functions on E such that there exists some finite constants C_n with $\sup_{1 \leq i \leq n} \sup_{x \in X} |f_i^{(n)}(X_i^{(n)})| \leq C_n$. Define, for $n \geq 1$, the sum

$$S_n = \sum_{i=1}^{n} f_i^{(n)}(X_i^{(n)}).$$

Theorem 7.1 (Theorem 1.1 in [SV])**.** *If*

$$\lim_{n \to \infty} C_n^2 \alpha_n^{-3} \left[\sum_{i=1}^{n} V(f_i^{(n)}(X_i^{(n)})) \right]^{-1} = 0,$$

then, we have the standard normal convergence

$$\frac{S_n - E[S_n]}{\sqrt{V(S_n)}} \Rightarrow N(0,1).$$

7.3.2. *Decomposition for mixing.* Let now $t \in J$ be fixed. Most quantities depend on t but, from now on, we omit the sub/superscript. Let $(\tau_n)_{n \geq 0}$ be the expansion of t. We denote p_n the transition matrix associated with the automaton A_{τ_n} and $p_{n,m} = p_n \cdots p_{m-1}$ the transition matrix between times n and m. For all n consider the word $W_n = \tau_{m_{n-1}} \cdots \tau_{m_n - 1}$. This yields a decomposition of our sequence in words ending with $\kappa = \chi\chi\chi$. We define the (non homogeneous) Markov chain $Y_n = X_{m_{n-1}}$. We denote by π_n its transition matrices, $\pi_n = p_{m_{n-1}, m_n}$. The function we want to control is a function of $(X_{m_{n-1}}, \ldots, X_{m_n - 1})$. Let us set $\tilde{Y}_n = (X_{m_{n-1}+1}, \ldots, X_{m_n - 1})$. Observe that, conditionally to Y_n and Y_{n+1}, $\tilde{Y}_n$ is independent of $\{Y_i, i \in \mathbb{N} \setminus \{n, n+1\}\}$ (and of

$\{\tilde{Y}_i, i \in \mathbb{N} \setminus \{n\}\})$. We introduce

$$f_n(Y_n, \tilde{Y}_n) = \sum_{k=m_{n-1}}^{m_n-1} g_k(X_k) - \mathbb{E}[\sum_{k=m_{n-1}}^{m_n-1} g_k(X_k)].$$

Remark 7.2. Ideally we would like to use Dobrushin's result about sums of observables of a non stationary Markov chain. The difficulty is that the functions do not depend only on the Markov chain itself. We could represent this quantity as a function of Y_n, Y_{n+1} and another random variable independent of the chain. Still it is not enough to apply the theorem in its standard versions. The usual way to avoid this problem would be to build a new chain joining coordinates by pairs $((Y_n, Y_{n+1}))$; but this construction breaks down the lower bound for the ergodic coefficient. This does not change the result fundamentally but we have to explain how the proof has to be adapted.

We state the main results of this section as follows. We denote $F_n = f_n(Y_n, \tilde{Y}_n)$ and $S_n = \sum_{k=1}^n F_k$. The positive variance assumption allows to show that, for all $t \in J$,

Lemma 7.3. *There is $\gamma > 0$ such that*

$$V(S_n) \geq \gamma n. \tag{7.5}$$

The proof is given in Section 7.7. Relying on this result, we show that

Proposition 7.4. *For all $t \in J$,*

$$\frac{S_n}{\sqrt{V(S_n)}} \Rightarrow N(0,1)$$

The remainder of this section is devoted to the proof of Proposition 7.4. We are going to adapt Dobrushin's Central Limit Theorem for inhomogeneous Markov chain (following the exposition by Sethuraman and Varhadan [SV]) to our special case. But we first show how it yields Theorem 1.9.

7.4. **Proof of Theorem 1.9.** We are now in position to prove Theorem 1.9. Let us fix $t \in J$.

Lemma 5.3 and Lemma 6.2 show that we can apply Proposition 5.4, to get equivalence of the variances V_n introduced in the statement of Theorem 1.9 and $V(\sum_{k=0}^{n-1} g_k^t(X_k)))$. Since the g_k^t belong to a finite family of bounded functions, we easily get an upper bound for the variances: $V(\sum_{k=0}^{n-1} g_k^t(X_k)) \leq \sup_{k\geq 0} ||g_k^t||_\infty n \leq \alpha_2 n$. For the lower bound, we have to be more careful. On the one hand, Lemma 7.3 shows that there is $\gamma > 0$ such that $V(S_n) \geq \gamma n$. On the other

hand, we use the relationship (7.3) between n and m_n to observe that $m_n \leq n.C.n^{1/4} \leq C.n^{5/4}$. Hence $V(\sum_{k=0}^{m_n-1} g_k^t(X_k))) \geq \gamma(m_n/C)^{4/5}$ and (for $m_n \leq m < m_{n+1}$), $V(\sum_{k=0}^{m-1} g_k^t(X_k))) \geq \gamma(m_n/C)^{4/5} \geq \gamma((m - m^{1/4})/C)^{4/5} \geq \alpha_1 n^{4/5}$. This proves statement (i).

We use again (7.3) and boundedness of the family (g_k^t) to check that for $m_n \leq m < m_{n+1}$, $|\sum_{k=0}^{m-1} g_k^t(X_k))) - S_n| \leq$ const. $n^{1/4}$. In view of the variance lower bound, Proposition 7.4 yields

$$\frac{\sum_{k=0}^{n-1}(g_k^t(X_k) - E[g_k^t(X_k)])}{\sqrt{V(\sum_{k=0}^{n-1} g_k^t(X_k)))}} \Rightarrow N(0,1). \tag{7.6}$$

Lemma 5.3 and Lemma 6.2 show that (7.6) yields statement (ii) of Theorem 1.9.

We conclude recalling that J has full Lebesgue measure. □

7.5. **Standard assumptions.** Even though our situation is very similar, we cannot apply directly Dobrushin's Theorem because our sum runs on observables that are not functions of the state of the (mixing) Markov chain. Nonetheless, our argument follows the general outline of [SV]. We use the same interplay between boundedness, rate of mixing and variance. More specifically, in our case, the ergodic coefficient is constant (bounded away from 0 ; so we need just to ensure that $C_n^2/V(S_n) \to 0$).

7.5.1. *Upper bound.* The upper bound

$$C_n = \sup_{i \leq n} \sup_{z,\tilde{z}} f_i(z,\tilde{z}) \leq \sup_{i \leq n}(||g_i||_\infty (m_{i+1} - m_i)) \leq C_1.n^{1/4}.$$

In other words,

$$C_n = \sup_{i \leq n} F_i \leq C_1 n^{1/4}.$$

7.5.2. *Ergodic coefficient.* The ergodicity coefficient is bounded from below uniformly because of the presence of the word κ at the end which guarantees some mixing. Let us be more specific.
By inequality (7.4) since all the words W_n end with the word κ (associated to a (power of a) primitive matrix), p_{q_n,m_n} is uniformly bounded from below (by a constant α), and we have

$$\pi_n(a,b) \geq \sum_{c \in A} p_{m_{n-1},q_n}(a,c) p_{q_n,m_n}(c,b) \geq \sum_{c \in A} p_{m_{n-1},q_n}(a,c)\, \alpha \geq \alpha.$$

From which we deduce that $\alpha(\pi_n) \geq \alpha$ uniformly in n. This holds for the Markov chain $(Y_n)_{n \geq 0}$.

Remark 7.5. We stress a technical difference with [SV]. The point is that, as in [SV], we have $Osc(E[f(Y_j, \tilde{Y}_j)|Y_i]) \leq \alpha^{j-i} Osc(f(Y_j))$; but that we will need a bound on $Osc(E[f_j(Y_j, \tilde{Y}_j)|Y_i, \tilde{Y}_i])$ where $\tilde{Y}_i$ depends on what happens between m_{i-1} and $m_i - 1$. Since we need only asymptotic decay, i.e. for $j - i$ large, we will simply shift by 1 using that

$$Osc(E[f_j(Y_j, \tilde{Y}_j)|Y_i, \tilde{Y}_i]) \leq Osc(E[f_j(Y_j, \tilde{Y}_j)|Y_{i+1}]).$$

7.5.3. *Variance.* The variance of each elementary contribution is uniformly bounded from below. Furthermore the sum of the variances of the first n terms is increasing rapidly enough with n. In the usual setting the last statement follows from the first one using assumption 2. In our case, it seems more convenient to prove directly the second statement. That is the object of Lemma 7.3 which claims that there is $\gamma > 0$ such that $V(S_n) \geq \gamma n$ and whose proof is postponned to Section 7.7.

7.5.4. *Outline of the argument.* The crucial point for the proof (following [SV]) is the interplay between these three quantities. Roughly speaking the variance must grow fast enough to kill unboundedness and lack of ergodicity, as expressed by equation (1.3) in [SV]. In our case the ergodic coefficient is independent on n so it is enough to have the variance growing faster than the upper bound. To draw the parallel let us write :

$$\lim_{n\to\infty} C_n^2 \alpha_n^{-3} V_n^{-1} = \lim_{n\to\infty} C^2 n^{1/2} \rho^{-3} \frac{1}{\gamma n} = 0.$$

But we stress again that the situation is slightly more tricky because the observable we are interested in is not a simple function of the states of the Markov chain for which the assumptions are fullfilled (f_n depends on Y_n, but also on $\tilde{Y}_n$). That is why we treat the variance separately. For the remainder of the proof, we can not just quote [SV] but it appears that we can follow exactly the same lines. To construct the Martingale approximant, we take $\tilde{Y}_n$ into account. To prove that the assumptions of the main Martingale differences CLT arguments are fulfilled, we use the same ideas based on the control of the oscillations of the conditionnal expectations of $f_n(Y_n, \tilde{Y}_n)$ which rely essentially on the summability of the long range correlations and on the sublinearity of the upper bound (together with the linear growth of the variance).

7.6. **Proof of Proposition 7.4.** The central idea is to use the following standard CLT (implied by Corollary 3.1 in [HH]) for martingale

differences : if $M_k^{(n)} = \sum_{l=2}^k \xi_l^{(n)}$ is a martingale with respect to a filtration $\mathcal{F}_k$ and if,

$$\max_{1\leq i\leq n} ||\xi_i^{(n)}||_\infty \to 0, \quad \text{and,} \tag{7.7}$$

$$\sum_{i=1}^n E[(\xi_i^{(n)})^2]|\mathcal{F}_{i-1}] \to 1 \text{ in } L^2, \tag{7.8}$$

then,

$$M_n^{(n)} \Rightarrow N(0,1).$$

The quantities whose fluctuations we study are not a martingale mainly because increments are not independent for a Markov chain in that they depend on the present state. It is rather standard to bypass this difficulty by introducing

$$Z_k^{(n)} = \sum_{i=k}^n E[f_i(Y_i, \tilde{Y}_i) \,|\, Y_k, \tilde{Y}_k]. \tag{7.9}$$

Then considering the scaled differences

$$\xi_k^{(n)} = \frac{1}{\sqrt{V(S_n)}} \left(Z_k^{(n)} - E[Z_k^{(n)} \,|\, Y_{k-1}, \tilde{Y}_{k-1}] \right),$$

we define the (family of) processes $M_k^{(n)} = \sum_{i=2}^k \xi_i^{(n)}$ and observe that the processes $(M_k^{(n)})_{1\leq k\leq n}$ are martingales with respect to the filtrations $\{\mathcal{F}_k, 1 \leq k \leq n\}$, where $\mathcal{F}_k = \sigma\{Y_l, \tilde{Y}_l : 1 \leq l \leq k\}$. Indeed, since $\xi_k^{(n)}$ is $\mathcal{F}_k$-measurable, for $k \leq n-1$,

$$\begin{aligned}
E[M_{k+1}^{(n)}|\mathcal{F}_k] &= \sum_{i=2}^k \xi_i^{(n)} + E[\xi_{k+1}^{(n)}|\mathcal{F}_k] \\
&= M_k^{(n)} + (\sqrt{V(S_n)})^{-1}(E\left[Z_{k+1}^{(n)} - E[Z_{k+1}^{(n)} \,|\, Y_k, \tilde{Y}_k] \,|\, \mathcal{F}_k\right]) \\
&= M_k^{(n)} + (\sqrt{V(S_n)})^{-1}(E[Z_{k+1}^{(n)}|\mathcal{F}_k] - E[Z_{k+1}^{(n)} \,|\, Y_k, \tilde{Y}_k]|\mathcal{F}_k]) \\
&= M_k^{(n)}.
\end{aligned}$$

Observing that

$$\begin{aligned} E[Z_{k+1}^{(n)} \,|\, Y_k, \tilde{Y}_k] &= E[\sum_{i=k+1}^{n} E[f_i(Y_i, \tilde{Y}_i) \,|\, Y_{k+1}, \tilde{Y}_{k+1}] \,|\, Y_k, \tilde{Y}_k] \\ &= \sum_{i=k+1}^{n} E[E[f_i(Y_i, \tilde{Y}_i) \,|\, Y_{k+1}, \tilde{Y}_{k+1}] \,|\, Y_k, \tilde{Y}_k] \\ &= \sum_{i=k+1}^{n} E[f_i(Y_i, \tilde{Y}_i) \,|\, Y_k, \tilde{Y}_k], \end{aligned}$$

we can put (7.9) the other way round to obtain, for $k \leq n-1$,

$$f_k(Y_k, \tilde{Y}_k) = Z_k^{(n)} - E[Z_{k+1}^{(n)} \,|\, Y_k, \tilde{Y}_k],$$

for $k = n$, $f_n(Y_n, \tilde{Y}_n) = Z_n^{(n)}$. We obtain, for the sums we are interested in:

$$S_n = \sum_{k=2}^{n} \left(Z_k^{(n)} - E[Z_k^{(n)} \,|\, Y_{k-1}, \tilde{Y}_{k-1}] \right) + Z_1^{(n)},$$

and for their variances (using the orthogonality of the martingale increments):

$$V(S_n) = \sum_{k=2}^{n} V\left(Z_k^{(n)} - E[Z_k^{(n)} \,|\, Y_{k-1}, \tilde{Y}_{k-1}] \right) + V(Z_1^{(n)}. \tag{7.10}$$

Hence, we can approximate $S_n/\sqrt{V(S_n)}$ by $M_n^{(n)}$ (see Step 5.) and use the above argument to conclude, provided we check both assumptions (7.7) and (7.8) (as well as approximation).

To do so, we will follow the proof of [SV].

Step 1. We prove inequalities (playing the role of Lemma 3.1 in [SV]). We will make repeated use of the property : $Osc(\pi\eta) \leq \delta(\pi) Osc(\eta)$. Let $1 \leq i < j \leq n$. As $||F_j^{(n)}||_\infty \leq C_n$, its oscillation $Osc(F_j) \leq 2C_n$. From the control we have on the mixing coefficient (see Remark 7.5),

$$Osc(E[F_j^{(n)}|\mathcal{F}_i]) \leq Osc(F_j^{(n)})\delta(\pi_{i+1,j}) \leq 2C_n(1-\alpha)^{j-i-1}.$$

For centered random variables, $||||_\infty$ is bounded by $2Osc()$, so that, for all $1 \leq i < j \leq n$,

$$||E[F_j|\mathcal{F}_i||_\infty \leq 2Osc(E[F_j^{(n)}|\mathcal{F}_i]) \leq \text{const. } n^{1/4}(1-\alpha)^{j-i}. \tag{7.11}$$

The same arguments also yield the similar

$$Osc(E[F_j^2|\mathcal{F}_i]) \leq \text{const. } n^{1/2}(1-\alpha)^{j-i}. \tag{7.12}$$

Step 2. Putting together inequality (7.11) and the variance lower bound Lemma 7.3 (corresponding to Proposition 3.2 [SV]), we obtain an analog of Lemma 3.2 [SV].
Using inequality (7.11) and $EF_i = 0$, we see that

$$||Z_k^{(n)}||_\infty \leq \sum_{i=k}^{n} ||E[F_i^{(n)}|\mathcal{F}_k]||_\infty \leq \text{const. } n^{1/4} \sum_{i=k}^{n} (1-\alpha)^{i-k} \leq 2C_n/\alpha.$$

Then by the variance lower bound (Lemma 7.3),

$$\sup_{1\leq k\leq n} \frac{||Z_k^{(n)}||_\infty}{\sqrt{V(S_n)}} \leq \text{const. } \frac{n^{1/4}}{\alpha\sqrt{\gamma n}}, \tag{7.13}$$

which is, under our assumptions $o(1)$.

This will yield the assymptotic equivalence of $S_n/\sqrt{V(S_n)}$ and $M_n^{(n)}$ as well as negligibility (7.7) of the differences $\xi_k^{(n)}$.

Step 3. Now let us set $H_i = E[(\xi_i^{(n)})^2]|\mathcal{F}_{i-1}]$. Before all, we observe that

$$E[\sum_{i=1}^{n} H_i] = \sum_{i=1}^{n} E[E[(\xi_i^{(n)})^2]|\mathcal{F}_{i-1}]] = \sum_{i=1}^{n} E[(\xi_i^{(n)})^2] = 1 + o(1)$$

in view of (7.10) and (7.13). In the next two steps, we are going to prove that $\sum_{i=1}^{n} H_i \to 1$ in L^2. This first step is rather general and correspond to Lemma 3.3 in [SV]. Observe that, since $E[\sum_{i=1}^{n} H_i] = 1 + o(1)$, the L^2-norm

$$E[(\sum_{i=1}^{n} H_i - 1)^2] = E[(\sum_{i=1}^{n} H_i)^2] - 2E[\sum_{i=1}^{n} H_i] + 1 = E[(\sum_{i=1}^{n} H_i)^2] - 1 + o(1)$$

tends to 0 if and only if $E[(\sum_{i=1}^{n} H_i)^2] \to 1$. Following the proof of Lemma 3.3 in [SV] we write

$$E[(\sum_{i=1}^{n} H_i)^2] = \sum_{i=1}^{n} E[H_i^2] + 2\sum_{i=1}^{n-1} E[H_i(\sum_{j=i+1}^{n} H_j)],$$

and,

$$(\sum_{i=1}^{n} E[H_i])^2 = \sum_{i=1}^{n} E[H_i^2] + 2\sum_{i=1}^{n-1} E[H_i]E[(\sum_{j=i+1}^{n} H_j)].$$

Taking the difference and recalling $E[\sum_{i=1}^{n} H_i] = 1 + o(1)$, we deduce that

$$E[(\sum_{i=1}^{n} H_i)^2] - 1 = 2\sum_{i=1}^{n-1} E[H_i \left(\sum_{j=i+1}^{n} H_j - E[\sum_{j=i+1}^{n} H_j] \right)] + o(1).$$

We observe that, since H_i is measurable with respect to $\mathcal{F}_{i-1}$,

$$E[H_i \sum_{j=i+1}^{n} H_j] = E[H_i E[\sum_{j=i+1}^{n} H_j|\mathcal{F}_{i-1}]].$$

Hence,

$$\begin{aligned} |E[(\sum_{i=1}^{n} H_i)^2] - 1| &\leq \sum_{i=1}^{n-1} E[H_i \left(E[\sum_{j=i+1}^{n} H_j|\mathcal{F}_{i-1}] - E[\sum_{j=i+1}^{n} H_j] \right)] + o(1) \\ &\leq \sup_{1\leq i\leq n-1} Osc(E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) \left(\sum_{i=1}^{n-1} E[H_i] \right) + o(1). \end{aligned}$$

Step 4. Hence it remains only to prove

$$\sup_{1\leq i\leq n-1} Osc(E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) = o(1). \tag{7.14}$$

First we write

$$\begin{aligned} E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) &= E[\sum_{j=i+1}^{n} E[(\xi_j^{(n)})^2]|\mathcal{F}_{j-1}]|\mathcal{F}_{i-1}]) \\ &= E[\sum_{j=i+1}^{n} (\xi_j^{(n)})^2|\mathcal{F}_{i-1}] \end{aligned}$$

and use the orthogonality property (martingale) of the increments ($E[\xi_r^{(n)}\xi_s^{(n)}|\mathcal{F}_u] = 0$ provided $r > s > u$) :

$$E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) = E[(\sum_{j=i+1}^{n} \xi_j^{(n)})^2|\mathcal{F}_{i-1}].$$

Then we write all this in terms of Z and F. We recall that for $k \leq n-1$, $F_k = Z_k^{(n)} - E[Z_{k+1}^{(n)}|\mathcal{F}_k]$, while $F_n = Z_n^{(n)}$.

$$\begin{aligned} \sum_{j=i+1}^{n} \xi_j^{(n)} &= \sum_{j=i+1}^{n} Z_j^{(n)} - E[Z_j^{(n)}|\mathcal{F}_{j-1}] \\ &= V(S_n)^{-1}(\sum_{j=i+1}^{n-1} F_j + E[Z_{j+1}^{(n)}|\mathcal{F}_j] - E[Z_j^{(n)}|\mathcal{F}_{j-1}]) + F_n - E[Z_n^{(n)}|\mathcal{F}_{n-1}]) \\ &= V(S_n)^{-1}(\sum_{j=i+1}^{n} F_j - E[Z_{i+1}^{(n)}|\mathcal{F}_i]). \end{aligned}$$

Hence,

$$\begin{aligned} E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) &= V(S_n)^{-1}E[(\sum_{j=i+1}^{n} F_j - E[Z_{i+1}^{(n)}|\mathcal{F}_i])^2|\mathcal{F}_{i-1}] \\ &= V(S_n)^{-1}E[(\sum_{j=i+1}^{n} F_j)^2|\mathcal{F}_{i-1}] - V(S_n)^{-1}E[E[Z_{i+1}^{(n)}|\mathcal{F}_i]^2|\mathcal{F}_{i-1}]. \end{aligned}$$

The last term is bounded by $\sup_{2\leq i\leq n-1} V(S_n)^{-1}||Z_{i+1}^{(n)}||_\infty^2 = o(1)$; hence its oscillation is also uniformly $o(1)$. For the first term, we write

$$Osc(V(S_n)^{-1}E[(\sum_{j=i+1}^{n} F_j)^2|\mathcal{F}_{i-1}]) \leq V(S_n)^{-1} \sum_{i+1\leq j,m\leq n} Osc(E[F_jF_m|\mathcal{F}_{i-1}]).$$

But, in the same spirit as for the proof of inequality (7.11), we have, for all $1 \leq l < i < j \leq n$,

$$\begin{aligned} Osc(E\left[F_i^{(n)}E[F_j^{(n)}|\mathcal{F}_i]|\mathcal{F}_l\right]) &\leq (1-\alpha)^{i-l-1}Osc(F_i^{(n)}E[F_j^{(n)}|\mathcal{F}_i]) \\ &\leq (1-\alpha)^{i-l-1}\Big(Osc(F_i^{(n)})||E[F_j^{(n)}|\mathcal{F}_i]||_\infty \\ &\qquad + ||F_i^{(n)}||_\infty Osc(E[F_j^{(n)}|\mathcal{F}_i])\Big) \\ &\leq \text{const. } n^{1/2}(1-\alpha)^{i-l}(1-\alpha)^{j-i}. \end{aligned}$$

It follows that, for all $i \leq n$,

$$Osc(E[(\sum_{j=i+1}^{n} H_j)|\mathcal{F}_{i-1}]) \leq \text{const. } V(S_n)^{-1}n^{1/2}/\alpha^2 = o(1).$$

Step 5. Conclusion. On the one hand, Step 2. shows that almost surely,

$$\left|M_n^{(n)} - \frac{S_n}{\sqrt{V(S_n)}}\right| \leq \frac{||Z_1^{(n)}||_\infty}{\sqrt{V(S_n)}} = o(1).$$

On the other hand, Steps 1. and 2. show that (7.7) holds while Steps 3. and 4. show that (7.8) holds so that $M_n^{(n)} \Rightarrow N(0,1)$. This concludes the proof of Proposition 7.4 □

7.7. **The variance.** We end the proof by giving a direct proof for Lemma 7.3 about the lower bound of the variance.

Step 1. To start with, we claim that there is $\gamma > 0$ such that, for all $i \geq 1$,

$$E\left[\left(f_i(Y_i,\tilde{Y}_i) - E\left[f_i(Y_i,\tilde{Y}_i)|Y_i,Y_{i+1}\right]\right)^2 \,|\, Y_i,Y_{i+1}\right] > \gamma. \tag{7.15}$$

To simplify we write $F_i = f_i(Y_i,\tilde{Y}_i)$ and $G_i = F_i - E[F_i|Y_i,Y_{i+1}]$ so that (7.15) becomes :

$$E[G_i^2 \,|\, Y_i,Y_{i+1}] > \epsilon.$$

To prove the claim, we fix $i \geq 1$. We set $m = m_{i-1}$, $M = m_i$, $p = |\chi|$, $q = M_i - 3p$, $r = M - 2p$, $s = M - p$, and $\ell = M - m$. We stress that with these notations, $\tau_q \cdots \tau_{r-1} = \tau_r \cdots \tau_{s-1} = \tau_s \cdots \tau_{M-1} = \chi$ and recall that the matrix associated with χ is primitive.

For $w \in E^{p+1}$, we set $\mathcal{E}_w = \{X_r = w_0, \ldots, X_s = w_p\}$. Elementary considerations about Markov chains and primitivity of the matrices associated to χ yield existence of $\epsilon_1 > 0$ and $\epsilon_2 > 0$ such that, for all $w \in E^{p+1}$, $P(X_r = w_0, X_s = w_p \,|\, Y_i,Y_{i+1}) > \epsilon_1$, $P(\mathcal{E}_w \,|\, X_r = w_0, X_s = w_p) > \epsilon_2$.

We also observe that (7.1) shows that that there are two paths u and v in E^{p+1} (with $u_0 = v_0$ and $u_p = v_p$) such that

$$\left|\sum_{k=r}^{s-1}(g_k(u_k)) - \sum_{k=r}^{s}(g_k(v_k))\right| \geq \epsilon > 0.$$

We observe that $E[G_i^2 \,|\, Y_i,Y_{i+1}] \geq E[\mathbb{1}_{\mathcal{E}_u \cup \mathcal{E}_v} G_i^2 \,|\, Y_i,Y_{i+1}]$, so that, since $\mathcal{E}_u \cap \mathcal{E}_v = \emptyset$, $E[G_i^2 \,|\, Y_i,Y_{i+1}] \geq E[\mathbb{1}_{\mathcal{E}_u} G_i^2 \,|\, Y_i,Y_{i+1}] + E[\mathbb{1}_{\mathcal{E}_v} G_i^2 \,|\, Y_i,Y_{i+1}]$.

By definition,

$$E[G_i^2 \,|\, Y_i,Y_{i+1}] = \sum_{\tilde{z} \in A^\ell} G_i(z_1 \cdots z_\ell)^2 P(\tilde{Y}_i = \tilde{z}|Y_i,Y_{i+1}).$$

In the same language,

$$E[\mathbb{1}_{\mathcal{E}_u} G_i^2 \,|\, Y_i,Y_{i+1}] \geq \sum_{w \in E^r, z \in E^p} G_i(wuz)^2 P(X_m \cdots X_M = wuz|Y_i,Y_{i+1}).$$

Writing $P(X_m \cdots X_M = wuz|Y_i,Y_{i+1}) = P(\mathcal{E}_u|X_r = u_0, X_s = u_p) P(X_m \cdots X_r = wu_0, X_s \cdots X_M = u_p z|Y_i,Y_{i+1})$, we obtain

$$\begin{aligned} E[G_i^2 \,|\, Y_i,Y_{i+1}] \;\; \geq \;\; & \sum_{w \in E^r, z \in E^p} (G_i(wuz)^2 + G_i(wvz)^2) \\ & \times \epsilon_2 \times P(X_m \cdots X_r = wu_0, X_s \cdots X_M = u_p z|Y_i,Y_{i+1}). \end{aligned}$$

For $w \in E^r$ and $z \in E^p$, observe that

$$|G_i(wuz) - G_i(wvz)| = \left| \sum_{k=r}^{s-1} (g_k(u_k)) - \sum_{k=r}^{s-1} (g_k(v_k)) \right| \geq \epsilon,$$

so that $|G_i(wuz)|^2 + |G_i(wvz)|^2 \geq \epsilon^2$ (since $a^2 + b^2 \geq a^2 + b^2 - 2|a||b| = (|a| - |b|)^2$). Hence

$$\begin{aligned} E[G_i^2 \,|\, Y_i, Y_{i+1}] &\geq \sum_{w \in E^r, z \in E^p} (\epsilon^2/4)\epsilon_2 P(X_m \cdots X_r = wu_0, X_s \cdots X_M = u_p z | Y_i, Y_{i+1}) \\ &\geq (\epsilon^2/4)\epsilon_2 P(X_r = u_0, X_s = u_p | Y_i, Y_{i+1}) \\ &\geq \epsilon^2 \epsilon_2 \epsilon_1 / 4, \end{aligned}$$

and the claim is proved with $\gamma = \epsilon^2 \epsilon_2 \epsilon_1 / 4$.

Step 2. To fulfill the proof of Lemma 7.3, we prove that

$$V(\sum_{i=0}^{n-1} f_i(Y_i, \tilde{Y}_i)) \geq \sum_{i=0}^{n-1} E\left[\left(f_i(Y_i, \tilde{Y}_i) - E\left[f_i(Y_i, \tilde{Y}_i) | Y_i, Y_{i+1} \right] \right)^2 \right],$$

or, in other words,

$$V(\sum_{i=0}^{n-1} F_i) \geq \sum_{i=0}^{n-1} E\left[G_i^2 \right].$$

We recall that $EF_i = 0$.

$$\begin{aligned} V(\sum_i F_i) &= E\left[(\sum_i F_i)^2 \right] \\ &= E\left[E\left[(\sum_i F_i)^2 \,|\, Y_0, \ldots, Y_n \right] \right] \\ &= E\left[E\left[\left(\sum_i (F_i - E[F_i | Y_i, Y_{i+1}]) + \sum_i E[F_i | Y_i, Y_{i+1}] \right)^2 \,|\, Y_0, \ldots, Y_n \right] \right] \\ &= E\left[E\left[\left(\sum_i G_i + \sum_i E[F_i | Y_i, Y_{i+1}] \right)^2 \,|\, Y_0, \ldots, Y_n \right] \right] \\ &= E\left[E\left[\left(\sum_i G_i \right)^2 + H + \left(\sum_i E[F_i | Y_i, Y_{i+1}] \right)^2 \,|\, Y_0, \ldots, Y_n \right] \right], \end{aligned}$$

where the product term H is

$$\begin{aligned} H &= E\left[\left(\sum_i G_i\right)\left(\sum_i E[F_i|Y_i,Y_{i+1}]\right) \mid Y_0,\dots,Y_n\right] \\ &= E\left[\sum_i G_i \mid Y_0,\dots,Y_n\right]\left(\sum_i E[F_i|Y_i,Y_{i+1}]\right), \end{aligned}$$

and, for all i,

$$\begin{aligned} E\left[G_i \mid Y_0,\dots,Y_n\right] &= E\left[F_i - E[F_i|Y_i,Y_{i+1}] \mid Y_0,\dots,Y_n\right] \\ &= E\left[F_i \mid Y_0,\dots,Y_n\right] - E[F_i|Y_i,Y_{i+1}] \\ &= 0. \end{aligned}$$

To conclude, we observe that, since for all $i<j\le k$, $E\left[G_iG_k \mid Y_j\right]=0$ (by independence of the past and the future conditionally to the present for the Markov chain (X_n)),

$$\begin{aligned} E\left[E\left[\left(\sum_i G_i\right)^2 \mid Y_0,\dots,Y_n\right]\right] &= E\left[E\left[\sum_i (G_i)^2 \mid Y_0,\dots,Y_n\right]\right] \\ &= \sum_i E\left[E\left[(G_i))^2 \mid Y_0,\dots,Y_n\right]\right] \\ &= \sum_i E\left[E\left[(G_i))^2 \mid Y_i,Y_{i+1}\right]\right] \\ &\ge E\left[\sum_i \gamma\right] \\ &\ge n\gamma, \end{aligned}$$

where the last inequality is indebt to (7.15). □

8. Specific cases and examples

To get an idea of the whole picture it seems useful to study what happens in specific cases, making further assumptions on the substitution or/and on the possible choices of t. We also give concrete examples at the end of this section.

8.1. **Synchronization.** We consider the case $t=1$ and more generally t integer or say $d^{n_0}t$ integer for some n_0, i.e. the expansion of t is finite. To study this simple case, we observe that (at least after a finite number of steps), we can work with a *simplified automaton* with state space $\mathcal{A}\times\mathcal{A}$ because when $\tau_k=0$ we never need to know the second letter of V since $m_k+\tau_k\le d$.

Lemma 8.1. *When the expansion of t is finite, the subset of states $\{(a,a)\ ;\ a\in\mathcal{A}\}$ is a recurent class of the (simplified) Markov chain. Moreover, on this class, the cocycle g is a coboundary. Instead of normal fluctuations the limit law of $\lim_{n\to\infty}\frac{1}{\sqrt{n}}\int_0^{d^n t} f\circ h_s(x)ds$ is a Dirac mass at 0.*

Proof. We assume that $\tau_k=0$ for all k large enough (larger than n_0). Thus, the Markov chain is stationary. In the automaton associated to $\tau=0$, it is clear that a couple (a,a) is followed by a couple (b,b) since the positions of both letters are the same in $\sigma(a)$. Therefore the set $\{(a,a)\ ;\ a\in\mathcal{A}\}$ is a recurrent class.
If $\mathfrak{S}^{-n_0}x$ belongs to a and $h_{d^{n_0}t}(\mathfrak{S}^{-n_0}x)$ is in the same letter a then the integral along the unstable leave does not depend on the exact position of $\mathfrak{S}^{-n_0}x$ in a. Indeed shift of the piece of leaf along does not change the value of the integral (what is added from one side is subtracted from the other side). Moreover, since the value of $\tau_k=0$ is zero, if we shift $\mathfrak{S}^{-n_0}x$ at the beginning of a the point $\mathfrak{S}^{-n_0}x$ to $h_{d^{n_0}t}(\mathfrak{S}^{-n_0}x)$ is also at the beginning of a. Thus the leaf from $\mathfrak{S}^{-n_0}x$ to $h_{d^{n_0}t}(\mathfrak{S}^{-n_0}x)$ covers exactly a union of blocks of that scale. Thus, if we refine for all $n>n_0$, $\int_0^{d^n t} f(h_s(\sigma^{-n}x))ds=\int_0^{d^{n_0}t} f(h_s(\sigma^{-n_0}x))ds$. The value along the Markov chain can change (since it is only an approximation of the integrals, at each scale) but it must keep at bounded distance of the value of the integral. We claim that the sum of the values along any cycle of the recurrent class $\{(a,a)\ ;\ a\in\mathcal{A}\}$ is zero. In fact, given a cycle of class, if the sum of the values of g along this cycle were non zero, we could iterate this cycle a large number of time and get arbitrarily large value for g. Thus g is a coboundary. The limit law of of $\lim_{n\to\infty}\frac{1}{\sqrt{n}}\int_0^{d^n t} f\circ h_s(x)ds$ is therefore a Dirac mass at 0 since $\int_0^{d^n t} f\circ h_s(x)ds$ is bounded and we divide by $\sqrt{n}$. □

Definition 8.2. *We say that a letter is* synchronizable *if it appears at the same position in the image of two distinct letters. We say that a substitution σ is* synchronizable *if for every couple of letters (b,c) there is a letter a that appears at the same position in $\sigma(b)$ and $\sigma(c)$. A substitution is* strongly non synchronizable *if it has no synchronizable letter.*

Proposition 8.3. *If a substitution σ is synchronizable, then the (simplified) Markov chain has a unique recurrent class, the set $\{(a,a)\ ;\ a\in\mathcal{A}\}$. On this class, the cocycle g is a coboundary. Therefore the limit law of $\lim_{n\to\infty}\frac{1}{\sqrt{n}}\int_0^{d^n t} f\circ h_s(x)ds$ is a Dirac mass at 0 when the expansion of t in base d is finite.*

Proof. By Lemma 8.1, it is enough to prove that the simplified automaton has only one recurrent class. At each step of the Markov chain there is a positive probability to reach the recurrent class $\{(a,a)\ ;\ a \in \mathcal{A}\}$. Therefore the only recurrent class is the class $\{(a,a)\ ;\ a \in \mathcal{A}\}$. □

Lemma 8.4. *If σ has no synchronizable letters, then the recurent class $\{(a,a)\ ;\ a \in \mathcal{A}\}$ of the (simplified) Markov chain is disconnected of the remainder of the graph. Hence it is reached only if reached at initial time.*

Proof. If b and c are different letters of the alphabet $\mathcal{A}$, there is no arrow from the (b,c) to $\{(a,a)\ ;\ a \in \mathcal{A}\}$. Otherwise there would be a synchronizable letter. Therefore the class $\{(a,a)\ ;\ a \in \mathcal{A}\}$ is disconnected from the remainder of the graph. □

In the last paragraph, Figure 2 gives an example of a synchronizable substitution on a 3 letters alphabet. Figure 3 gives an example of a substitution on a 2 letters alphabet which is non synchronizable.

8.2. **Strongly non synchronizable substitutions on a 2 letters alphabet.** The following propositions give a recipe to construct strongly non synchronizable substitutions on a 2 letters alphabet. We also study the ergodic sums for these substitutions for some values of t.

Proposition 8.5. *If $\mathcal{A} = \{a,b\}$ a non synchronizable substitution σ of length d on $\mathcal{A}$ with eigenvalue 1 is uniquely determined by the image of a ($\sigma(b)$ is obtained from $\sigma(a)$ by exchanging a by b and vice versa). Moreover the matrix M of σ has the form $M = \begin{pmatrix} k+1 & k \\ k & k+1 \end{pmatrix}$ where $k+1$ is the number of a in $\sigma(a)$ and $d = 2k+1$.*

Proof. First of all, if σ is strongly non synchronizable, for every j in $\{1,\dots,d\}$, the letter number j in $\sigma(a)$ is different from the letter number j in $\sigma(b)$. Therefore one of these letter is an a, the other one is a b. Thus $\sigma(b)$ obtained from $\sigma(a)$ is equal to $s(\sigma(a))$ where s is the substitution $s(a) = b$, $s(b) = a$. Let ℓ be the number of a in $\sigma(a)$ and k the number of b in $\sigma(a)$. We have $k+\ell = d$. By the previous discussion, the matrix M has the form $M = \begin{pmatrix} \ell & k \\ k & \ell \end{pmatrix}$. Since the matrix M has an eigenvalue 1, $\ell = k+1$. □

Proposition 8.6. *Let σ be a substitution of length d on a 2 letters alphabet $\mathcal{A} = \{a,b\}$. We assume that σ is strongly non synchronizable and that every word of length 3 appears in $\sigma(a)$. Then the graph associated to the automaton $\tau = 1$ is strongly connected and aperiodic.*

Moreover g is not a coboundary on this graph. Thus the limit law for $\lim_{n\to\infty}\frac{1}{\sqrt{n}}\int_0^{d^n t} f\circ h_s(x)ds$ is a normal law when the digits of t are ultimately constant equal to 1.

Remark 8.7. *This gives a family of examples that satisfy the hypothesis of Theorem 1.9 and hence a proof of Theorem 1.3.*

Proof. To prove connectedness of the graph, we first prove that there is an arrow from (a, aa) to any other vertex. Let $(\alpha, \beta\gamma)$ be a triple of letters in $\mathcal{A}$. The word $\alpha\beta\gamma$ appears in $\sigma(a)$. We consider the word $\sigma(aa)$. In this word we see $\alpha W\alpha\beta\gamma$ where the length of W is equal to d. Thus the length between α and $\beta\gamma$ is $d+1$. By definition of the automaton, this means that we have an arrow from (a, aa) to $(\alpha, \beta\gamma)$. Now, let us prove that from any triple $(\alpha, \beta\gamma)$ there is an arrow to (a, aa). If $\alpha = a$ and $\beta = a$ the word aaa appears somewhere in $\sigma(a)$. Thus, by definition of the automaton with $\tau = 1$, there is an arrow from $(a, a\gamma)$ to (a, aa). If $\alpha = a$ and $\beta = b$, the word abb appears in $\sigma(a)$. At the same position baa appears in $\sigma(b)$ since σ is not strongly synchronizable. Thus there is an arrow from $(a, b\gamma)$ to (a, aa). If $\alpha = b$, $\beta = a$, the same reasonning holds. If $\alpha = \beta = b$, then aaa appears in $\sigma(b)$. Thus, by the same argument, there is an arrow from $(b, b\gamma)$ to (a, aa). This proves that the graph is strongly connected.
In fact, we also proved that it is aperiodic since there is an arrow from (a, a) to itself (thus a cycle of length 1).
We prove now that g is not a coboundary. We show that the value of g on some arrow from (a, aa) to itself is different from 0 which is enough. Let aaa be a subword of $\sigma(a)$. We have $\sigma(aa) = PaaaSPaaaS$ where the length of SP is $d-3$. We consider the arrow from (a, aa) to itself corresponding to $m = |P| + 1$. The label of this arrow is the number v_{aaSPa}. The word $aaSPa$ is a reordering of $PaaaS$. Thus, the numbers v_{aaSPa} and v_{PaaaS} coincide. But $v_{PaaaS} = v_{\sigma(a)} = v_a$. One can check that an eigenvector associated to the eigenvalue 1 of M is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Thus v_a is different from 0.
We now apply Theorem 1.2. As there is only one recurrent class which is aperiodic, the limit distribution is a normal law. □

8.3. **Examples.** In this paragraph, we construct explicit examples with various properties:
Figure 2 represents the automatom for

$$\sigma : \begin{cases} 1 & \to & 12 \\ 2 & \to & 13 \\ 3 & \to & 23 \end{cases}$$

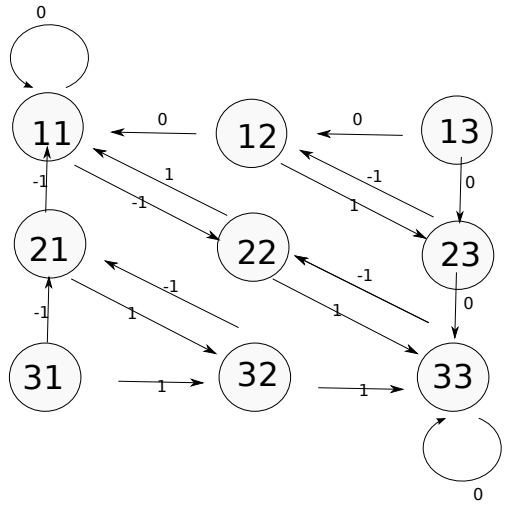

FIGURE 2.

and $\tau = 0$. The synchronization property is satisfied. The graph has one recurrent component but is not strongly connected (there are transient components). The eigenfunction f is not a coboundary nevertheless, when the 2-adic expansion of t is finite, the cocycle g is a coboundary.

Figure 3 represents the automaton for

$$\sigma : \begin{cases} 1 & \to & 112 \\ 2 & \to & 221 \end{cases}$$

and $\tau = 0$. The graph has 2 recurrent components: the component $\{(a, a)\ ;\ a \in \mathcal{A}\}$ where the cocycle is a coboundary, another component where the Markov chain has non zero variance. Thus, the limiting distribution is a superposition of a Dirac at 0 and a normal distribution. This means that starting from a set of positive measure there is no fluctuation at the scale $\frac{1}{\sqrt{n}}$, on another set of positive measure there are fluctuations at the same scale.

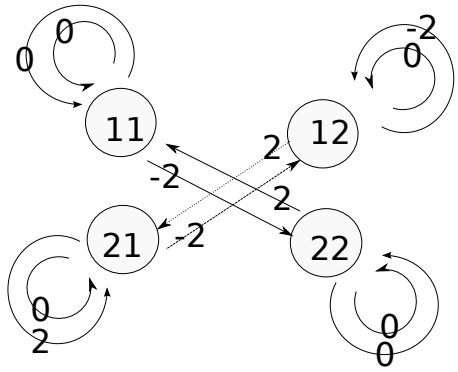

FIGURE 3.

Figure 4 represents the automaton for

$$\sigma : \begin{cases} 1 & \to & 112 \\ 2 & \to & 221 \end{cases}$$

and $\tau = 1$. The graph is strongly connected, aperiodic, the variance is positive thus the limiting distribution is a normal distribution.
Figure 5 represents the automaton for

$$\sigma : \begin{cases} 1 & \to & 112 \\ 2 & \to & 221 \end{cases}$$

and $\tau = 2$.
The graph contains two connected components. On one component the cocycle is a coboundary. We have no explanation for this phenomenon. On the other one, there is a positive variance.

References

[Ad] B. Adamczewski, Symbolic discrepancy and self-similar dynamics, Ann. Inst. Fourier (Grenoble) 54 (2004), 2201–2234.

[ArIt] P. Arnoux, S. Ito, Pisot substitutions and Rauzy fractals, Journées Montoises d'Informatique Théorique (Marne-la-Valle, 2000). Bull. Belg. Math. Soc. Simon Stevin 8 (2001), no. 2, 181–207.

[ABB] P. Arnoux, J. Bernat, X. Bressaud, *Geometrical models for substitutions*, Experimental Math. to appear.

[AV] A. Avila, M. Viana, *Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture* Acta Math. 198 (2007), no. 1, 1–56.

[Be] A. Beardon, The geometry of discrete groups, Graduate Texts in Mathematics, 91. Springer-Verlag, New York, 1983. xii+337 pp.

[BHM] X. Bressaud, P. Hubert, A. Maass, *Persistence of wandering intervals in self-similar affine interval exchange transformations*, Ergodic Theory and Dynamical Systems 30 (2010), no. 3, 665686.

[Bu] A. I. Bufetov, *Finitely additive measures on the asymptotic foliations of a Markov compactum*, www.arxiv.org, (2009).

[Bu2] A. I. Bufetov, *Limit theorems for translation flows* , www.arxiv.org, (2010).

[CS] V. Canterini, A. Siegel, *Automate des préfixes-suffixes associé une substitution primitive*, [Prefix-suffix automaton associated with a primitive substitution] J. Théor. Nombres Bordeaux 13 (2001), no. 2, 353–369.

[CG] R. Camelier, C. Gutierrez, *Affine interval exchange transformations with wandering intervals*, Ergodic Theory and Dynamical Systems 17, (1997), 1315-1338.

[C] M. Cobo, *Piece-wise affine maps conjugate to interval exchanges*, Ergodic Theory and Dynamical Systems 22, (2002), 375-407.

[Dob] R. Dobrushin, *Central limit theorems for non-stationary Markov chains I, II*, Theory of Probab. and its Appl. 1, 65–80, 329–383 (1956).

[DT1] J.M. Dumont, A. Thomas, *Digital sum moments and substitutions*, Acta Arith. 64 (1993), 205–225.

[DT2] J.M. Dumont, A. Thomas, *Digital sum problems and substitutions on a finite alphabet*, J. Number Theory 39 (1991), no. 3, 351–366.

[F] P. Fogg, Substitutions in Dynamics, Arithmetics and Combinatorics, Lecture Notes in Mathematics, 1794, Springer-Verlag, 2002.

[Fo] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) 155 (2002), no. 1, 1–103.

[Fo2] G. Forni, *On the Lyapunov exponents of the Kontsevich-Zorich cocycle*, Handbook of dynamical systems. Vol. 1B, 549?580, Elsevier B. V., Amsterdam, 2006.

[FMZ] G. Forni, C. Matheus, A. Zorich *Square-tiled cyclic covers*, preprint 2010.

[HH] P. Hall, C. C. Heyde, *Martingale limit theory and its application*,Academic Press, New-York (1980)

[Ito] S. Ito, A construction of transversal flows for maximal Markov automorphisms, Tokyo J. Math. 1 (1978), no. 2, 305–324.

[L] G. Levitt, *La décomposition dynamique et la différentiabilité des feuilletages des surfaces*, Ann. Inst. Fourier 37, (1987), 85-116.

[Liv] A. N. Livshits, Sufficient conditions for weak mixing of substitutions and of stationary adic transformations (Russian) Mat. Zametki 44 (1988), no.

6, 785–793, 862; translation in Math. Notes 44 (1988), no. 5-6, 920–925 (1989).

[LM] I. Liousse, H. Marzougui, Echanges d'intervalles affines conjugués à des linéaires [Conjugation between affine and linear interval exchanges] Ergodic Theory Dynam. Systems 22 (2002), 2, 535–554.

[Mc] C. McMullen, *Billiards and Teichmüller curves on Hilbert modular surfaces*, J. Amer. Math. Soc. **16**, no. 4 (2003) 857–885

[MMY1] S. Marmi, P. Moussa, J.C. Yoccoz, The cohomological equation for Roth-type interval exchange maps, J. Amer. Math. Soc. 18 (2005), no. 4, 823–872.

[MMY2] S. Marmi, P. Moussa, J.C. Yoccoz, Affine interval exchange maps with a wandering interval, Proc. Lond. Math. Soc. (3) 100 (2010), no. 3, 639669.

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

[Ra] G. Rauzy, *Echanges d'intervalles et transformations induites*, Acta Arith. 34, (1979), 315–328.

[SV] S. Sethuraman and S.R.S. Varadhan , *A Martingale Proof of Dobrushin's Theorem for Non-Homogeneous Markov Chains*, Electronic Journal of Probability 10, (2005), no 36, 1221–1235.

[Th] W. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. A.M.S. **19**, (1988) 417–431.

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

[Ve2] W. A. Veech, *Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards*, Invent. Math. 97 (1989), no. 3, 553583.

[Ver] A. M. Vershik, *The adic realizations of the ergodic actions with the homeomorphisms of the Markov compact and the ordered Bratteli diagrams* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 223 (1995), Teor. Predstav. Din. Sistemy, Kombin. i Algoritm. Metody. I, 120–126, 338; translation in J. Math. Sci. (New York) 87 (1997), no. 6, 4054–4058.

[VL] A. M. Vershik, A. N. Livshits, *Adic models of ergodic transformations, spectral theory, substitutions, and related topics. Representation theory and dynamical systems*, 185–204, Adv. Soviet Math., 9, Amer. Math. Soc., Providence, RI, 1992.

[Yo] J. C. Yoccoz, *Continuous fraction algorithms for interval exchange maps: an introduction*, in "Frontiers in Number Theory, Physics and Geometry, volume I. On Random matrices, Zeta Functions and Dynamical Systems", P. Cartier, B. Julia, P. Moussa, P. Vanhove (Editors), Springer Verlag, Berlin 2006, 403–437.

[Zo] A. Zorich, *Finite Gauss measure on the space of interval exchange transformations, Lyapunov exponents*, Annales de l'Institut Fourier 46:2, (1996), 325–370.

[Zo1] A. Zorich, Deviation for interval exchange transformations, Ergodic Theory Dynam. Systems 17 (1997), no. 6, 1477–1499.

[Zo2] A. Zorich, *Flat surfaces*, in "Frontiers in Number Theory, Physics and Geometry, volume I. On Random matrices, Zeta Functions and Dynamical Systems", P. Cartier, B. Julia, P. Moussa, P. Vanhove (Editors), Springer Verlag, Berlin 2006, 439–585.

Université Paul Sabatier, Institut de Mathématiques de Toulouse, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France
E-mail address: bressaud@math.univ-toulouse.fr

Rice University, Houston, TX, USA and the Steklov Institute of Mathematics, Moscow, Russian Federation
E-mail address: aib1@rice.edu

Laboratoire Analyse, Topologie et Probabilités, Case cour A, Faculté des Sciences de Saint-Jerôme, Avenue Escadrille Normandie-Niemen, 13397 Marseille Cedex 20, France.
E-mail address: hubert@cmi.univ-mrs.fr

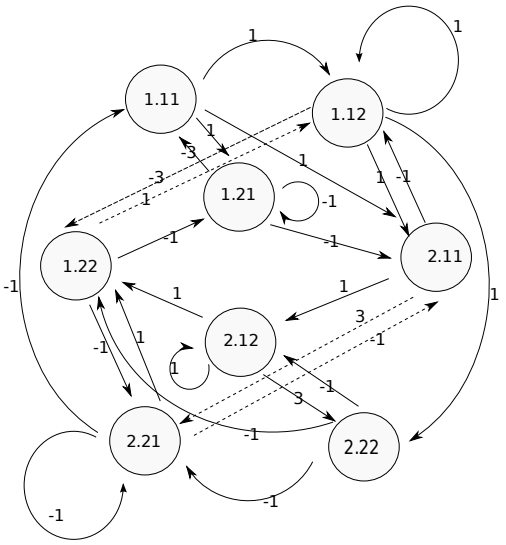

Figure 4.

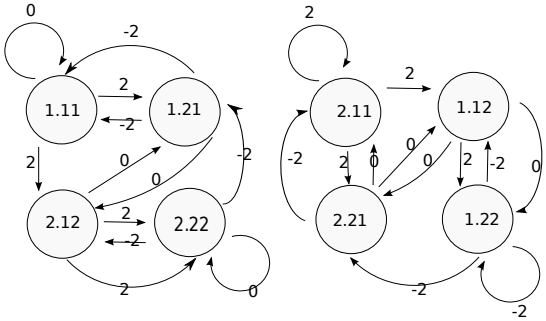

FIGURE 5.