SYMBOLIC DISCREPANCY AND SELF-SIMILAR DYNAMICS

BORIS ADAMCZEWSKI †

Résumé anglais :

ABSTRACT. We consider subshifts arising from primitive substitutions, which are known to be uniquely ergodic dynamical systems. In order to precise this point, we introduce a symbolic notion of discrepancy. We show how the distribution of such a subshift is in part ruled by the spectrum of the incidence matrices associated with the underlying substitution. We also give some applications of these results in connection with the spectral study of substitutive dynamical systems.

Résumé français :

ABSTRACT. Nous considérons des systèmes dynamiques naturellement associés aux substitutions primitives et connus pour être uniquement ergodiques. Afin d'étudier plus précisément cette propriété, nous introduisons différentes notions de discrépance symbolique. Nous montrons comment les propriétés de répartition d'un tel système sont en partie déterminées par les matrices d'incidences associées à la substitution sous-jacente. Nous donnons également certaines applications de ces résultats à l'étude spectrale des systèmes dynamiques substitutifs.

Titre français : Discrépance symbolique et dynamiques auto-similaires.

Key words: discrepancy,substitutions, subshifts, bouded remainder sets, self-similar dynamics.

Mots clés : discrépance, substitutions, sous-shifts, ensembles à restes bornés, dynamiques auto-similaires.

2000 AMS classification: 11K38, 37A30, 37A45, 37B10, 68R15.

1. Introduction

In this paper, we introduce two functions of discrepancy, one associated with symbolic sequences and the other with subshifts (respectively defined in (1) and (2)). We mainly deal with the asymptotic behaviour of these two functions, focusing on sequences obtained as fixed points of primitive substitutions and on subshifts arising from them. Such sequences naturally appear as soon as one studies dynamical systems with a self-similar structure (that is, the induced system on some subset is topologically conjugate to the original one). This is in particular the case for one-dimensional toral quadratic rotations (see for instance [1]) and interval exchanges with parameters lying in the same quadratic field [3]. This work is motivated by questions arising at once from Diophantine approximation and ergodic theory and shares some links with [2, 22, 23, 8, 15].

† LRI, CNRS UMR 8623, Bât 490 Université Paris-Sud 91405 Orsay Cedex France. e-mail: adamczew@lri.fr.

We first recall the definition of discrepancy on a finite set. Let $\mathcal{A}$ be a finite set. Endowed with the discrete topology, $\mathcal{A}$ is a compact set. Let us consider a probability measure μ on $\mathcal{A}$. A sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ which takes its values in $\mathcal{A}$ is said uniformly distributed with respect to the measure μ if:

$$\forall\, a \in \mathcal{A}, \quad \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \chi_{\{a\}}(u_k) = \mu(a),$$

where $\chi_{\{a\}}$ denotes the characteristic function of the set $\{a\}$. Then, we define the discrepancy function $\Delta_N(\mu, \mathbf{u})$ of the sequence $\mathbf{u}$ by:

$$\Delta_N(\mu, \mathbf{u}) = \max_{a\in\mathcal{A}} \left| \sum_{k=0}^{N-1} \chi_{\{a\}}(u_k) - N\mu(a) \right|. \tag{1}$$

These definitions, given for a finite set, come directly from the more classical notions of uniform distribution modulo 1 and discrepancy for real sequences. Some generalizations to topological, compact or quasi-compact groups could also be found in the literature. Two important references on this subject are the books of L. Kuipers and H. Niederreiter [17] and of M. Drmota and R. F. Tichy [7].

With any symbolic sequence $\mathbf{u}$ defined over a finite set, one can also associate the dynamical system $(\overline{\mathcal{O}(\mathbf{u})}, T)$, where $\overline{\mathcal{O}(\mathbf{u})}$ is the closure of the orbit of $\mathbf{u}$ under the classical shift transformation T. This leads us to ask the following questions: what could be a (natural) version of discrepancy on such a symbolic space? How far is such a system from having an ideal distribution?

Let $\mathbf{u}$ be a symbolic sequence, $\mathcal{X} = (\overline{\mathcal{O}(\mathbf{u})}, T)$ the subshift arising from $\mathbf{u}$, and μ a shift invariant measure for $\mathbf{u}$. Then, we define the discrepancy function of the dynamical system $\mathcal{X}$ (with respect to μ) by:

$$D_N(\mathcal{X}) = \sup_{V\in\overline{\mathcal{O}(\mathbf{u})}} \sup_{w\in\mathcal{L}(\mathbf{u})} \left| \sum_{k=0}^{N-1} \chi_{[w]}(T^k(V)) - N\mu([w]) \right|, \tag{2}$$

where $[w]$ denotes the cylinder associated with the word w. This definition is really close to the classical one given in the case of the torus, since we just have replaced intervals by cylinders. Moreover, if (X, T, μ) denotes a uniquely ergodic dynamical system and f a continuous function on X, then

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k \to \int f d\mu,$$

uniformly. The discrepancy function introduced above measures the uniform speed of convergence of Birkhoff'sums with a restriction to a class of very regular continuous functions (i.e., characteristic functions of cylinders). Since cylinders form a natural basis of the topology for the space $\overline{\mathcal{O}(\mathbf{u})}$, we can interpret $D_N(\mathcal{X})$ as a measure of unique ergodicity for subshifts having this property. Note that it can be easily deduced from a classical result generally attributed to Curtis, Lyndon and Hedlund (see for instance [18]), that the order of magnitude of the functions $\Delta_N(\mathbf{u})$ and $D_N(\mathcal{X})$ are topological invariants (i.e., preserved by a topological isomorphism between two subshifts).

This article is organized as follows. In Section 2, we present our results (mainly, Theorem 1 and Theorem 3). We discuss spectral properties of primitive substitutive subshifts in Section 3. The main tool for proving Theorem 1 is introduced in Section 4. Finally, Section 5 and Section 6 are respectively devoted to the proof of Theorem 1 and Theorem 3.

2. Definitions and main results

2.1. **Symbolic sequences.** A finite and nonempty set $\mathcal{A}$ is called alphabet. The length of a finite word ω, denoted by $|\omega|$, is the number of letters it is built from. The empty word, ε, is the unique word of length 0. We denote by $\mathcal{A}^*$ the set of finite words on $\mathcal{A}$ and by $\mathcal{A}^{\mathbb{N}}$ the set of sequences on $\mathcal{A}$. Let $\mathbf{u} = (u_k)_{k\in\mathbb{N}}$ be a symbolic sequence defined over the alphabet $\mathcal{A}$. A factor of $\mathbf{u}$ is a finite word of the form $u_i u_{i+1} \ldots u_j$, $0 \le i \le j$. We denote by $\mathcal{L}(\mathbf{u})$ the set of all the factors of the sequence $\mathbf{u}$, $\mathcal{L}(\mathbf{u})$ is called the language of $\mathbf{u}$. If ω is a finite word and a a letter, then $|w|_a$ is the number of occurrences of the letter a in ω.

2.2. **Substitutions.** Endowed with concatenation, the set $\mathcal{A}^*$ is a free monoïd with unit element ε. A map from $\mathcal{A}$ to $\mathcal{A}^* \setminus \{\varepsilon\}$ can be extended by concatenation to an endomorphism of the free monoïd $\mathcal{A}^*$ and then to a map from $\mathcal{A}^{\mathbb{N}}$ to itself. A *substitution* σ on the alphabet $\mathcal{A}$ is such a morphism satisfying

(i) There exists $a \in \mathcal{A}$ such that a is the first letter of $\sigma(a)$,
(ii) For all $b \in \mathcal{A}$, $\lim_{n\to+\infty} |\sigma^n(b)| = +\infty$.

Then, $(\sigma^n(aa\ldots))_{n\in\mathbb{N}}$ converges in $\mathcal{A}^{\mathbb{N}}$, endowed with the product of the discrete topologies on $\mathcal{A}$, to a sequence $\mathbf{u}$. This sequence is a *fixed point* of σ, i.e., $\sigma(\mathbf{u}) = \mathbf{u}$. Given a substitution σ defined on $\mathcal{A} = \{1, 2, \ldots, d\}$, the matrix $M_\sigma = (|\sigma(j)|_i)_{(i,j)\in\mathcal{A}^2}$ is called the *incidence matrix* associated with σ. A substitution is *primitive* if there exists a power of its incidence matrix for which all the entries are positive. To a fixed point of a primitive substitution $\mathbf{u}$ we associate a *natural probability measure* μ on $\mathcal{A}$, the measure of a letter $a \in \mathcal{A}$ being given by the value of its frequency in $\mathbf{u}$, that is, $\mu(a) = \lim\limits_{n\to\infty} \dfrac{|u_0u_1\ldots u_{n-1}|_a}{n}$ (the existence of frequencies is for instance proved in [21]).

2.3. **Spectrum of a primitive substitution.** The Perron-Frobenius theorem implies that the incidence matrix of a primitive substitution admits a simple real eigenvalue greater than the modulus of all the others eigenvalues (see for instance [21]). This eigenvalue is greater than one and is called the *Perron* eigenvalue of the substitution. We order the spectrum S_{M_σ} of the incidence matrix associated with a primitive substitution σ as follows:

$$S_{M_\sigma} = \{\theta_i,\ 2 \le i \le d'\} \cup \{\theta_1 = \theta\},$$

where θ is the Perron eigenvalue of σ, d' is the number of distinct eigenvalues and

$$\forall\, 2 \le i, k \le d',\ \ i < k \Rightarrow \begin{cases} |\theta_i| > |\theta_k|, \\ \text{or } |\theta_i| = |\theta_k| \text{ and } \alpha_i \ge \alpha_k, \end{cases} \tag{3}$$

where α_j denotes the multiplicity of the eigenvalue θ_j in the minimal polynomial of M_σ. Furthermore, if $|\theta_i| = |\theta_k| = 1$, $\alpha_i = \alpha_k$, θ_i is not a root of unity and θ_k is a root of unity, then $i < k$. With these conditions, the quantities θ, $|\theta_2|$ and α_2 are well-defined (whereas several choice for the value of θ_2 can sometimes be done).

Given a primitive substitution σ, one can define its substitution of order 2, denoted by σ_2 (see Section 6 for a precise construction). It is proved in [21] that such a substitution is primitive too and shares the same Perron eigenvalue. Following (3), we can thus define the quantities $|\theta_{2,2}|$ and $\alpha_{2,2}$, associated with the spectrum of M_{σ_2}.

2.4. **Landau symbols.** Let f and g be two real positive functions. We recall the definition of some Landau symbols:

$$f = O(g) \quad \text{if} \quad \exists C > O, \text{ such that } f(x) < Cg(x),\ \forall x \in \mathbb{R}_+,$$

$$f = \Omega(g) \quad \text{if} \quad \limsup_{x \to +\infty} \frac{f(x)}{g(x)} > 0.$$

The following notation will be used in the most of our results. We will write

$$f = (O \cap \Omega)(g),$$

if both $f = O(g)$ and $f = \Omega(g)$. Though such a relation does of course not imply that $f \sim g$, we will consider, with a slight abuse of language, that the order of magnitude of the function f is g, as soon as the relation $f = (O \cap \Omega)(g)$ is satisfied.

2.5. **Main results.** Following these definitions, a fixed point of a primitive substitution is uniformly distributed with respect to its natural probability measure. Our main result is a precise estimate of the discrepancy for such sequences. We show how it is in part ruled by the spectrum of the incidence matrix associated with the substitution. We prove that in most of cases (more precisely for cases (i), (ii) and (iii) of Theorem 1) the order of magnitude of the discrepancy just depends on the incidence matrix. In these cases, the asymptotic behaviour of the discrepancy is thus not modified by any permutation of the letters in the definition of the substitution. However, Theorem 1 could not be reduced to a result on primitive matrices since case (iv) really depends on the substitution. For instance, the two following substitutions σ and τ respectively defined by:

$$\begin{array}{l} 1 \longmapsto 14 \\ 2 \longmapsto 31 \\ 3 \longmapsto 34 \\ 4 \longmapsto 31 \end{array} \quad \text{and} \quad \begin{array}{l} 1 \longmapsto 12 \\ 2 \longmapsto 13 \\ 3 \longmapsto 34 \\ 4 \longmapsto 13 \end{array}$$

share the same incidence matrix but have a different discrepancy. More precisely, the fixed point of σ (generated by the letter 1) has a bounded discrepancy function, whereas the order of magnitude of the discrepancy associated with the one of τ is $\log N$. Since we can associate to a primitive substitition a natural probability measure μ as in section 2.2, we will write in the following $\Delta_N(\mathbf{u})$ instead of $\Delta_N(\mu, \mathbf{u})$ for the discrepancy of a fixed point $\mathbf{u}$ of a primitive substitution.

Theorem 1. *Let $\mathbf{u} = (u_k)_{k \geq 0}$ be a fixed point of a primitive substitution σ defined over the alphabet $\mathcal{A}$, θ, θ_2 and α_2 defined as in Section 2.3, μ the natural probability measure associated with $\mathbf{u}$ (defined as in Section 2.2) and*

$$\Delta_N(\mathbf{u}) = \max_{a \in \mathcal{A}} \left| \sum_{k=0}^{N-1} \chi_{\{a\}}(u_k) - N\mu(a) \right|,$$

where $\chi_{\{a\}}$ denotes the characteristic function of the set a. Then, the following holds:

(i)	*if* $|\theta_2| < 1$, *then*	$\Delta_N(\mathbf{u})$ *is bounded* ,
(ii)	*if* $|\theta_2| > 1$, *then*	$\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_2} N^{(\log_\theta |\theta_2|)}\right)$,
(iii)	*if* $|\theta_2| = 1$ *and* θ_2 *is not a root of unity, then*	$\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{(\alpha_2+1)}\right)$,
(iv)	*if* $|\theta_2| = 1$ *and* θ_2 *is a root of unity, then*	
	either $A_{\sigma,\mathbf{u}} \neq 0$ *and*	$\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{(\alpha_2+1)}\right)$,
	or $A_{\sigma,\mathbf{u}} = 0$ *and*	$\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_2}\right)$,

where the complex number $A_{\sigma,\mathbf{u}}$ *(which just depends on the pair* $(\sigma, \mathbf{u})$*) is defined in Section 5.4 and could explicitly be computed.*

We derive from Theorem 1 the following characterization of fixed points of primitive substitutions having an almost perfect distribution (that is, with a bounded discrepancy).

Corollary 2. *A fixed point* $\mathbf{u}$ *of a primitive substitution* σ *has a bounded discrepancy function* $\Delta_N(\mathbf{u})$ *if and only if one of the following holds:*

(i) $|\theta_2| < 1$,
(ii) $|\theta_2| = 1$, $\alpha_2 = 0$, θ_2 *is a root of unity and* $A_{\sigma,\mathbf{u}} = 0$.

Next, we consider dynamical systems arising from primitive substitutions, which are well-known to be uniquely ergodic subshifts (see [19]). We prove an analogous of Theorem 1, for the discrepancy of these systems, with respect to their unique invariant measure. The main interest of Theorem 3 is that it gives a uniform information with respect to all the factors, whereas Theorem 1 just deals with letters. Theorem 3 is obtained *via* Theorem 1 together with the use of the notion of derived sequences introduced in [10]. However, Theorem 3 is far from being a simple by-product of Theorem 1.

Theorem 3. *Let* $\mathbf{u}$ *be a fixed point of a primitive substitution* σ, θ, $\theta_{2,2}$ *and* $\alpha_{2,2}$ *defined as in Section 2.3,* $\mathcal{X} = (\overline{\mathcal{O}(\mathbf{u})}, T, \mu)$ *the dynamical system arising from* $\mathbf{u}$*, and*

$$D_N(\mathcal{X}) = \sup_{V \in \overline{\mathcal{O}(\mathbf{u})}} \sup_{w \in \mathcal{L}(\mathbf{u})} \left| \sum_{k=0}^{N-1} \chi_{[w]}(T^k(V)) - N\mu([w]) \right| .$$

Then, the following holds:

(i) *if* $|\theta_{2,2}| < 1$, *then* $D_N(\mathcal{X})$ *is bounded* ,

(ii) *if* $|\theta_{2,2}| > 1$, *then* $D_N(\mathcal{X}) = (O \cap \Omega)\left((\log N)^{\alpha_{2,2}} N^{(\log_\theta |\theta_{2,2}|)}\right)$,

(iii) *if* $|\theta_{2,2}| = 1$ *and* $\theta_{2,2}$ *is not a root of unity, then* $D_N(\mathcal{X}) = (O \cap \Omega)\left((\log N)^{(\alpha_{2,2}+1)}\right)$,

(iv) *if* $|\theta_{2,2}| = 1$ *and* $\theta_{2,2}$ *is a root of unity, then* $D_N(\mathcal{X}) = O\left((\log N)^{(\alpha_{2,2}+1)}\right)$,

and $D_N(\mathcal{X}) = \Omega\left((\log N)^{\alpha_{2,2}}\right)$.

In particular, it follows that such dynamical systems are *uniformly well distributed* with respect to their unique ergodic measure, that is:

Corollary 4. *Let* $\mathbf{u}$ *be a fixed point of a primitive substitution* σ, $\mathcal{X} = (\overline{\mathcal{O}(\mathbf{u})}, T, \mu)$ *the dynamical system arising from* $\mathbf{u}$, *then we have that*

$$\sup_{w \in \mathcal{L}(\mathbf{u})} \left| \frac{1}{N} \sum_{k=0}^{N-1} \chi_{[w]}(T^k(V)) - \mu([w]) \right|$$

tends to 0 *uniformly in* $V \in \overline{\mathcal{O}(\mathbf{u})}$.

Our method provides explicit constants for all the bounds given in Theorem 1 and 3. In case (iv) of Theorem 1, it is even possible to compute the quantity

$$\limsup_{N \to \infty} \frac{\Delta_N(\mathbf{u})}{(\log N)^{(\alpha_2+1)}}.$$

Most of the arguments to prove it are exposed in [2]. The quantities $\theta_{2,2}$ and $\alpha_{2,2}$ could probably be replaced respectively by θ_2 and α_2 in Theorem 3, but at this point this can just be proved in the case where $|\theta_2| > 1$. Theorem 1 and 3 are respectively proved in Section 5 and 6.

3. An application to spectral theory

We first discuss how our results can be used to obtain a spectral information for primitive substitutive subshits and thus for their measure-theoretic isomorphic dynamical systems.

Let $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system. For any $B \in \mathcal{B}$ and $x \in X$, we consider

$$\Delta_N(T; B; x) = \left| \sum_{k=0}^{N-1} \left(\chi_B(T^k(x)) - \mu(B)\right) \right|.$$

By Birkhoff's ergodic theorem, for any $B \in \mathcal{B}$ and almost all $x \in X$,

$$\frac{1}{N} \Delta_N(T; B; x) \to 0,$$

when N tends to infinity. A set B is called a *bounded remainder set* for T if $\Delta_N(T; B; x)$ is bounded on a set of measure one. For an irrational rotation on the one-dimensional torus, H. Kesten (more precisely, the "if" part is due to E. Hecke in 1922) gives the following characterization of bounded remainder sets which are intervals.

Theorem 5 (Kesten [16]). *Let T be a translation on $\mathbb{T}$ by an irrational α. Then, an interval I is a bounded remainder set for T if and only if its length belongs to* $\mathbb{Z}\alpha \mod 1$.

In 1973, H. Fürstenberg, H. Keynes and L. Shapiro (see also K. Petersen [20] and G. Halász [14]) proved the following strong generalisation of Kesten's theorem making relevant the notion of discrepancy in ergodic theory.

Theorem 6 (Fürstenberg and al. [13]). *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system. If a subset B of $\mathcal{B}$ is a bounded remainder set for the transformation T, then $e^{2\pi i\mu(B)}$ is an eigenvalue for T. Moreover, if $e^{2\pi i r}$ is an eigenvalue for T, there exists a bounded remainder set $B \in \mathcal{B}$ such that $\mu(B) = r$.*

In the symbolic framework, intervals could naturally be replaced by cylinders. Therefore, we can translate the previous problematic by: given a subshift, does there exist cylinders which are bounded remainder sets? Our study answers partially this question in the case of primitive substitutive subshifts. In particular, we obtain that, for a substitution satisfying the condition $\theta_{2,2} < 1$ (for instance the Fibonacci or Tribonacci subshifts), all the cylinders are bounded remainder sets. Theorem 3 proves even that in this case there exists a uniform bound.

In view of Theorem 6, we can easily translate such results to provide a simple condition of non-weak mixing for primitive substitutive subshifts.

Theorem 7. *Let $\mathbf{u}$ be a fixed point of a primitive substitution σ and $\mathcal{X}$ be the associated subshift. If $\Delta_N(\mathbf{u})$ is bounded, that is, if one of the following holds:*

(i) $|\theta_2| < 1$,
(ii) $|\theta_2| = 1$, $\alpha_2 = 0$, *θ_2 is a root of unity and* $A_{\sigma,\mathbf{u}} = 0$,

then, $\mathcal{X}$ is not weakly mixing.

In particular, we recover that the subshift arising from a Pisot type substitution (that is, the Perron eigenvalue is a Pisot number and the characteristic polynomial is irreducible) is never weakly mixing. This result, mentioned in [24], derives from one of B. Solomyak [26] but we do not require in Theorem 7 the irreducibility of the characteristic polynomial of the incidence matrix, contrary to the assumption done in [26]. Moreover, since the sets considered are cylinders, one expects to be able to compute their measures. Corollary 2 and Theorem 3 provide thus a concrete way to obtain eigenvalues for a primitive substitutive subshift. In the case of the Fibonacci and Tribonacci subshifts, all the eigenvalues can be found by proving that any cylinder corresponds to a bounded remainder set and using the fact that the eigenvalues form an additive group. More precisely, we can deduce from our study that the discrepancy function $\mathcal{D}_N(\mathcal{X})$ is bounded for these two subshifts. Such a result is probably also true for all Pisot type substitutions.

4. Notation and preliminary results

In this section, we introduce the main tool that we will have to use for our study together with some preliminary results.

4.1. **The $S_{\mathbf{u}}^f(N)$ functions.** In order to estimate the discrepancy of a symbolic sequence $\mathbf{u}$, it is useful to associate a "weight" with each letter of $\mathcal{A}$. Then, the study of the discrepancy takes the following formulation: what is the number of occurences of each letter in a given prefix of $\mathbf{u}$? We already used this idea in [2] for particular sequences, and we propose now to give a more general statement of this fact. The properties of $(S_{\mathbf{u}}^f(N))_{N\in\mathbb{N}^*}$ were investigated, in particular for some sequences related to the distribution of digits in arithmetical sequences, in [4, 5, 6, 8, 9, 23].

Definition 8. Let $\mathbf{u} = u_0u_1\dots u_n\dots$ be a symbolic sequence defined over the alphabet $\mathcal{A} = \{1, 2, \dots, d\}$. If $f = (f(i))_{i\in\mathcal{A}} \in \mathbb{C}^d$ and $N \in \mathbb{N}^*$, then we define:

$$S_{\mathbf{u}}^f(N) = \sum_{i=1}^{d} |u_0u_1\dots u_{N-1}|_i f(i).$$

Just as, if $\omega \in \mathcal{A}^*$, we define:

$$S^f(\omega) = \sum_{i=1}^{d} |\omega|_i f(i).$$

Definition 9. Let $\mathbf{u} = u_0u_1\dots u_n\dots$ be a sequence defined over the alphabet $\mathcal{A} = \{1, 2, \dots, d\}$ and such that each letter a of $\mathcal{A}$ admits a positive frequency $\mu(a)$ in $\mathbf{u}$. Let $\Lambda = (\mu(i))_{i\in\mathcal{A}} \in \mathbb{C}^d$ denote the frequencies vector of $\mathbf{u}$. Then, for $1 \le i \le d-1$, we introduce the vectors $f_i \in \mathbb{C}^d$, defined by

$$f_i(j) = \begin{cases} 1 & \text{if } j = i, \\ \frac{\mu(i)}{\mu(i)-1} & \text{else.} \end{cases} \tag{4}$$

Proposition 10. *The two following assertions are equivalent:*

(i) $\Delta_N(\mathbf{u}) = O(g(N))$ *(respectively $o(g(N))$),*

(ii) $\forall f = (f(i))_{i\in\mathcal{A}} \in \mathbb{C}^d$, $f \perp \Lambda$, $S_{\mathbf{u}}^f(N) = O(g(N))$ *(respectively $o(g(N))$).*

In (i), the constant in the O just depends on $\mathbf{u}$ and in (ii), it depends on $\mathbf{u}$ and f.

Proof. Since

$$\Delta_N(\mathbf{u}) = \max_{a\in\mathcal{A}} \left| \sum_{k=0}^{N-1} \chi_{\{a\}}(u_k) - N\mu(a) \right| = \max_{a\in\mathcal{A}} \left| |u_0u_1\dots u_{N-1}|_a - N\mu(a) \right|,$$

the previous definition implies

$$\Delta_N(\mathbf{u}) = \max_{i=1,2,\dots,d-1} (1-\mu(i))|S_{\mathbf{u}}^{f_i}(N)| \tag{5}$$

and the fact that the f_i form a basis of the vectorial space $\Lambda^\perp$ ends the proof. □

Proposition 11. *The two following assertions are equivalent:*

(i) $\Delta_N(\mathbf{u}) = \Omega(g(N))$,

(ii) $\exists\, f = (f(i))_{i\in\mathcal{A}} \in \mathbb{C}^d$, $f \perp \Lambda$, $S_{\mathbf{u}}^f(N) = \Omega(g(N))$.

Proof. It comes directly from (5). □

For any word $w \in \mathcal{A}^*$, let us introduce the vector $L(w) = (|w|_i)_{i\in\mathcal{A}}$. Then, for a substitution σ devined over $\mathcal{A}$ we have

$$L(\sigma(w)) = M_\sigma(L(w)), \tag{6}$$

where M_σ denotes the incidence matrix of σ. For fixed i and j in $\mathcal{A}$, the sequence $(|\sigma^n(j)|_i)_{n\in\mathbb{N}}$ satisfies a linear reccurence whose coefficients are those of the minimal polynomial of M_σ. Following Section 2.3, there exist complex numbers $\lambda_{i,j}^{k,l}$ and $\lambda_{i,j}$ such that for every $n \in \mathbb{N}$,

$$|\sigma^n(j)|_i = \lambda_{i,j}\theta^n + \sum_{k=2}^{d'} \left(\sum_{l=0}^{\alpha_k} \lambda_{i,j}^{k,l} n^l \theta_k^n \right). \tag{7}$$

Let us notice that Equation (6) and (7) imply that, for each letter j, the vector $(\lambda_{i,j})_{i\in\mathcal{A}}$ is an eigenvector of M_σ associated with its Perron eigenvalue θ. There thus exists a complex number ε_j such that $\lambda_{i,j} = \varepsilon_j \mu(i)$. Then, for any vector $f = (f(i))_{i=1,2,\ldots,d} \in \mathbb{C}^d$ lying in the orthogonal vectorial space spanned by μ, it follows:

$$\begin{aligned} S^f(\sigma^n(j)) &= \textstyle\sum_{i=1}^d |\sigma^n(j)|_i f(i) \\ &= \underbrace{\left(\sum_{i=1}^d \lambda_{i,j} f(i) \right)}_{0} \theta^n + \textstyle\sum_{k=2}^{d'} \left(\sum_{l=0}^{\alpha_k} \left(\sum_{i=1}^d \lambda_{i,j}^{k,l} f(i) \right) n^l \theta_k^n \right). \end{aligned} \tag{8}$$

Since for $k > 2$, $|\theta_2| \geq |\theta_k|$,

$$S^f(\sigma^n(j)) = O\left(n^{\alpha_2}|\theta_2|^n\right),$$

where the constant in the O just depends on j, if we assume f and σ fixed. Then, for every word $\omega \in \mathcal{A}^*$

$$S^f(\sigma^n(\omega)) = O\left(n^{\alpha_2}|\theta_2|^n\right), \tag{9}$$

where the constant in the O just depends on ω, if we assume f and σ fixed. In order to make the following more friendly readable, we introduce the following notation:

$$F_{f,k,l}(\omega) = \sum_{j=1}^{m} \left(\sum_{i=1}^{d} \lambda_{i,\omega_j}^{k,l} f(i) \right),$$

where $\omega = \omega_1\omega_2\ldots\omega_m$ is a word defined over $\mathcal{A}$ and f a vector lying in the orthogonal vectorial space spanned by μ. It thus implies

$$S^f(\sigma^n(w)) = \sum_{k=2}^{d'} \sum_{l=0}^{\alpha_k} F_{f,k,l}(\omega) n^l \theta_k^n. \tag{10}$$

4.2. **Numeration systems associated with substitutions.** We present now the key tool for proving Theorem 1. These generalized numeration systems associated with substitutions have been introduced simultaneously by J-M. Dumont and A. Thomas [8], and G. Rauzy [23].

Definition 12. Let σ be a substitution and let us assume that $\mathbf{u}$ is a fixed point for σ generated by the letter 1. The subset of $\mathcal{A}^*$ composed by the proper prefixes of the image by σ of the letters will be denote by $Pref_\sigma$. The *prefix automaton* associated with the pair $(\sigma, \mathbf{u})$ is defined in the following way:

– $\mathcal{A}$ is the set of states,

– $Pref_\sigma$ is the set of labels,

– there is a transition from the state i to the state j labelled by the (possibly empty) word m if mj is a prefix of $\sigma(i)$.

Definition 13. An *admissible labelled path* $\mathcal{C}$ in the prefix automaton associated with a pair $(\sigma, \mathbf{u})$ will be denoted by

$$((i_0, i_1, E_0), (i_1, i_2, E_1), \ldots, (i_{n-1}, i_n, E_{n-1})),$$

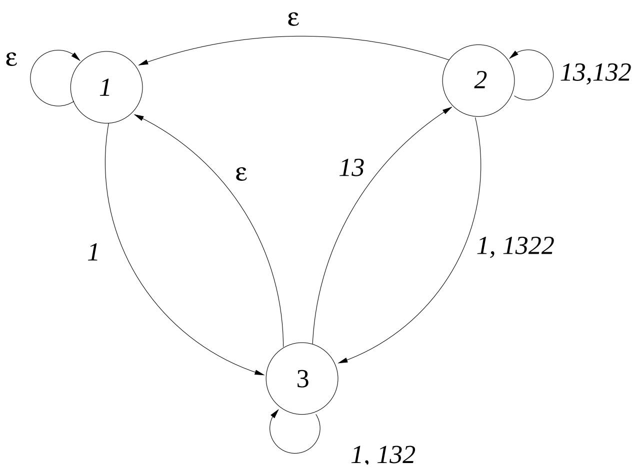

FIGURE 1. Example of a prefix automaton in the case of the substitution $1 \longmapsto 13$, $2 \longmapsto 13223$, $3 \longmapsto 123$.

$i_j \in \mathcal{A}$ for $0 \leq j \leq n$, $E_j \in Pref_\sigma$ for $0 \leq j \leq n-1$. The positive integer n is the length of the path. The set composed by the admissible labelled paths of length n will be denoted by $\mathcal{C}_\sigma^n$.

The main theorem concerning the prefix automaton is the following.

Theorem 14 (Dumont and Thomas [8], Rauzy [23]). *Let* $\mathbf{u}$ *be a fixed point generated by the letter* 1 *of a substitution* σ*, then we have:*

(i) for every positive integer N*, there exists a unique admissible path of length* n_N *in the prefix automaton associated with the pair* $(\sigma, \mathbf{u})$*, starting from* 1 *and labelled by the sequence* $(E_0, E_1 \ldots E_{n_N})$*, such that* $\mathbf{u}_N = \sigma^{n_N}(E_0)\sigma^{n_N-1}(E_1)\ldots E_{n_N}$*, where* $\mathbf{u}_N$ *denotes the prefix of* $\mathbf{u}$ *of the length* N *and such that* $E_0 \neq \varepsilon$*.*

(ii) Conversely, to any such a path, there corresponds a unique prefix of $\mathbf{u}$*, given by the above formula.*

(iii) Moreover, $|\sigma^{n_N}(1)| \leq N < |\sigma^{n_N+1}(1)|$*.*

The remaining of the paper is now devoted to the proofs of the main results.

5. Proof of Theorem 1

This section is devoted to the proof of Theorem 1. Most of the upper and lower bounds are respectively stated in Section 5.1 and 5.2. The essential difficulties appear when $|\theta_2| = 1$ and this case is treated in Section 5.3.

5.1. **First upper bounds.** We apply Theorem 14 in order to obtain upper bounds for the discrepancy of fixed point of primitive substitutions.

Proposition 15. *Let* $\mathbf{u}$ *be a fixed point of a primitive substitution, then:*

$$\begin{aligned}
&\text{if } |\theta_2| < 1, \text{ then} && \Delta_N(\mathbf{u}) \text{ is bounded,} \\
&\text{if } |\theta_2| > 1, \text{ then} && \Delta_N(\mathbf{u}) = O\left((\log N^{\alpha_2}) N^{\frac{\log|\theta_2|}{\log\theta}}\right), \\
&\text{if } |\theta_2| = 1, \text{ then} && \Delta_N(\mathbf{u}) = O\left(\log N^{\alpha_2+1}\right),
\end{aligned}$$

where the constants in the O *just depend on* $\mathbf{u}$.

Proof. Equality (9) together with the fact that the words E_i in Theorem 14 lie in a finite set imply:

$$S_{\mathbf{u}}^f(N) = O\left(\sum_{k=0}^{n_N} k^{\alpha_2}|\theta_2|^k\right),$$

and thus

$$S_{\mathbf{u}}^f(N) = O\left(n_N^{\alpha_2}\sum_{k=0}^{n_N} |\theta_2|^k\right), \tag{11}$$

where n_N is defined as in Theorem 14. One can notice that the assertion (iii) of Theorem 14 implies in view of Equality (7) that

$$\exists\, C > O,\ \exists\, C' > 0, \text{ such that } \ C\theta^{n_N} < N < C'\theta^{n_N}.$$

We thus deduce

$$\left| n_N - \frac{\log N}{\log\theta}\right| = O(1). \tag{12}$$

Then, we have to distinguish three cases, depending on the modulus of the eigenvalue θ_2, and using (11), it follows:

$$\begin{aligned}
&\text{if } |\theta_2| < 1, \text{ then} && S_{\mathbf{u}}^f(N) = O(1),\\
&\text{if } |\theta_2| > 1, \text{ then} && S_{\mathbf{u}}^f(N) = O\left((\log N)^{\alpha_2} N^{\frac{\log|\theta_2|}{\log\theta}}\right),\\
&\text{if } |\theta_2| = 1, \text{ then} && S_{\mathbf{u}}^f(N) = O\left((\log N)^{\alpha_2+1}\right),
\end{aligned}$$

where the constants in the O just depend on $\mathbf{u}$ and f. Proposition 10 allows us to conclude. □

5.2. **First lower bounds.** We want now to show the pertinency of Proposition 15. Hence, we are going to construct a sequence of prefixes of $\mathbf{u}$ with the worst possible distribution. Following Proposition 11, we know that it is sufficient to exhibit a vector $f \in \mathbb{C}^d$, $f \perp \mu$, and an increasing sequence of integers, $(N_k)_{k\in\mathbb{N}}$, such that the sequence $\left(|S_{\mathbf{u}}^f(N_k)|\right)_{k\in\mathbb{N}}$ takes "high" values, in a sense that we will have, of course, to make clear.

Let us first recall the following classical result of linear algebra.

Lemma 16. *Let M be a $d \times d$ complex matrix and let us denote by $\{(\theta_i, \alpha_i),\ 1 \le i \le d'\}$ the spectrum of M, where the θ_i mean the distinct eigenvalues of M and the α_i their multiplicity in the minimal polynomial of M. Let r be a positive integer and θ a non-zero eigenvalue of M^r, then the multiplicity of θ in the minimal polynomial of M^r is equal to the maximum of $\{\alpha_i$ such that $\theta_i^r = \theta,\ 1 \le i \le d'\}$.*

If we apply the previous lemma to the incidence matrix associated with the primitive substitution σ, we obtain the following.

Corollary 17. *Let k be a positive integer. Let us denote by $\{\theta_i',\ 2 \le i \le d''\} \cup \{\theta'\}$ the spectrum of σ^k, so that the θ_i' are ordered as in (2.3). Then, the following holds:*

$$\theta' = \theta^k,\ |\theta_2'| = |\theta_2|^k, \ \textit{and}\ \alpha_2' = \alpha_2.$$

In view of Corollary 17, we can thus freely consider any power of σ whithout changing the conditions which appear in Theorem 1. We are now ready to prove the following.

Proposition 18. *Let* $\mathbf{u}$ *be a fixed point of a primitive substitution, then:*

$$\text{if } |\theta_2| \geq 1, \text{ then } \Delta_N(\mathbf{u}) = \Omega\left((\log N^{\alpha_2}) N^{\frac{\log|\theta_2|}{\log\theta}}\right).$$

Proof. If j is a fixed element of $\mathcal{A}$ and $2 \leq k \leq d'$, then the vectors $\left(\lambda_{i,j}^{k,\alpha_k}\right)_{i\in\mathcal{A}}$ (defined in (7)) are eigenvectors associated with the eigenvalue θ_k (or eventually zero vectors). That follows directly from Equation (6) and (7), and from the fact that the sequences $\left(n^k\theta_i^n\right)_{n\in\mathbb{N}}$ form a free familly of the vectorial space spanned by complex sequences. Moreover, there exists at least one letter j_0 such that the vector $\left(\lambda_{i,j_0}^{2,\alpha_2}\right)_{i\in\mathcal{A}}$ is a non-zero vector, because otherwise this would provide a polynomial P of degree less than the one of the minimal polynomial associated with M and such that $P(M) = 0$. Let us consider the following vectorial subspace of $\mathbb{C}^d$:

$$E = \left\langle \{\mu\} \cup \left\{ \left(\lambda_{i,j}^{k,\alpha_k}\right)_{i\in\mathcal{A}},\ 2 < k \leq d',\ j \in \mathcal{A} \right\} \right\rangle.$$

Since $\left(\lambda_{i,j_0}^{2,\alpha_2}\right)_{i\in\mathcal{A}}$ is a non-zero vector, it is an eigenvector associated with the eigenvalue θ_2 and it thus does not lie in E. Hence,

$$E^{\perp} \not\subseteq \left\langle \left(\lambda_{i,j_0}^{2,\alpha_2}\right)_{i\in\mathcal{A}} \right\rangle^{\perp}.$$

There thus exists a vector $f_0 = (f_0(i))_{i\in\mathcal{A}} \in \mathbb{C}^d$ such that $f_0 \in E^{\perp}$ and $f_0 \not\perp \left(\lambda_{i,j_0}^{2,\alpha_2}\right)_{i\in\mathcal{A}}$.

Because of the primitivity of σ and Corollary 17, we can assume without restriction that all the entries of the incidence matrix associated with σ are greater or equal than two, which implies:

- the prefix automaton associated with σ is strongly connected, that is to say, for any pair $(i,j) \in \mathcal{A}^2$, the path (i,j) is admissible.
- for any pair $(i,j) \in \mathcal{A}^2$, the letter j has at least one occurence in a proper prefix of $\sigma(i)$.

In particular, there exists a proper prefix of $\sigma(1)$ in which the letter j_0 occurs. Let us denote by $w_1w_2\ldots w_{r-1}w_r$, $w_r = j_0$, such a proper prefix. Then, we distinguish two cases:
either

$$\sum_{j=1}^{r-1}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{2,\alpha_2} f_0(i)\right) \neq 0,$$

or

$$\sum_{j=1}^{r}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{2,\alpha_2} f_0(i)\right) \neq 0,$$

with $w_r = j_0$. In any case, there exists $m \in \{r-1, r\}$ satisfying

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{2,\alpha_2} f_0(i)\right) \neq 0.$$

We set $w = w_1\ldots w_m$. We have

$$S^{f_0}(\sigma^n(\omega)) = \underbrace{\sum_{j=1}^{m}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{2,\alpha_2} f_0(i)\right)}_{\neq 0} n^{\alpha_2}\theta_2^n$$

$$+\sum_{j=1}^{m}\left(\sum_{k=2}^{d'}\left(\sum_{l=0}^{\alpha_k-1}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{k,l}f_0(i)\right)n^l\theta_k^n\right)\right)$$
$$=\underbrace{\sum_{j=1}^{m}\left(\sum_{i=1}^{d}\lambda_{i,w_j}^{2,\alpha_2}f_0(i)\right)}_{\neq 0}n^{\alpha_2}\theta_2^n+o\left(n^{\alpha_2}\theta_2^n\right),$$

where the constant in the o just depends on ω (for a fixed pair (σ, f_0)). This implies that $S^{f_0}(\sigma^n(\omega)) = \Omega\left((n^{\alpha_2}\theta_2^n\right)$. Since $\mathbf{u}$ begins with 1, ω is a prefix of $\mathbf{u}$ and $\sigma^n(\omega)$ too. For $k \in \mathbb{N}$, let $N_k = |\sigma^k(\omega)|$. Then,

$$\left|k - \frac{\log N_k}{\log \theta}\right| = O(1),$$

and there exists a positive constant C such that:

$$S_{\mathbf{u}}^{f_0}(N_k) > C\left((\log N_k)^{\alpha_2} N_k^{\frac{\log|\theta_2|}{\log\theta}}\right).$$

It follows that $S_{\mathbf{u}}^{f_0}(N) = \Omega\left((\log N)^{\alpha_2} N^{\frac{\log|\theta_2|}{\log\theta}}\right)$, which ends the proof in view of Proposition 11. □

5.3. **The critical case.** Proposition 15 and 18 give the right order of magnitude of $\Delta_N(\mathbf{u})$ if $|\theta_2| \neq 1$. When $|\theta_2| = 1$, we just obtain that the extreme irregularities lie between $(\log N)^{\alpha_2}$ and $(\log N)^{\alpha_2+1}$. In particular, we are *a priori* not able to say if $\Delta_N(\mathbf{u})$ is bounded or not when $|\theta_2| = 1$ and $\alpha_2 = 0$. The following section is precisely devoted to the understanding of these critical cases. We will show that the knowledge of the incidence matrix associated with the substitution is not always sufficient to solve this problem. However, we provide an algorithmic way of answering it in the contentious cases.

5.3.1. **Case *(iii)* of Theorem 1.** The following proposition states that the discrepancy is maximal (in view of Proposition 15) when θ_2 is not a root of unity.

Proposition 19. *Let $\mathbf{u}$ be a fixed point of a primitive substitution. If $|\theta_2| = 1$ and θ_2 is not a root of unity, then*

$$\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_2+1}\right).$$

Proof. Let $\mathbf{u}$ be a fixed point of a primitive substitution σ and let us assume that $|\theta_2| = 1$ and θ_2 is not a root of unity. There thus exists a real γ, $\gamma \notin \mathbb{Q}$, such that $e^{i\gamma} = \theta_2$.

We recall that the primitivity of σ allows us to assume without restriction that all the entries of M are greater than or equal to two (see Corollary 17). As in the first part of the proof of Proposition 18, we can prove the existence of:

- a vector $f_0 \in \mathbb{C}^d$ and a letter j_0 such that $f_0 \in E^{\perp}$ and $f_0 \not\perp \left(\lambda_{i,j_0}^{2,\alpha_2}\right)_{i\in\mathcal{A}}$, where $E = \left\langle \{\mu\} \cup \left\{\left(\lambda_{i,j}^{k,\alpha_k}\right)_{i\in\mathcal{A}},\ 2 < k \leq d',\ j \in \mathcal{A}\right\}\right\rangle$,
- a proper prefix of $\sigma(1)$, w, such that $F_{f_0,2,\alpha_2}(w) \neq 0$ (see (10) for a definition of $F_{f_0,2,\alpha_2}(w) \neq 0$).

This implies that

$$S^{f_0}(\sigma^n(\omega)) = \underbrace{F_{f_0,2,\alpha_2}(w)}_{\neq 0}\, n^{\alpha_2} e^{in\gamma} + O(n^{\alpha_2-1}),$$

where the constant in the O just depends on w.

Let us consider a positive integer N. Following Theorem 14, there exists a unique admissible path in the prefix automaton associated with the pair $(\sigma,\mathbf{u})$, beginning with 1 and labelled by $(E_0,E_1,\ldots,E_{n_N})$, $E_0\neq\varepsilon$ such that:

$$\mathbf{u}_N=\sigma^{n_N}(E_0)\sigma^{n_N-1}(E_1)\ldots E_{n_N}.$$

We thus have

$$S_{\mathbf{u}}^{f_0}(N)=\sum_{k=0}^{n_N}S_{\mathbf{u}}^{f_0}(\sigma^k(E_{n_N-k})).$$

The fact that the prefixes E_i lie in a finite set $(Pref_\sigma)$ implies

$$S_{\mathbf{u}}^{f_0}(N)=\sum_{k=0}^{n_N}F_{f_0,2,\alpha_2}(E_{n_N-k})k^{\alpha_2}e^{ik\gamma}+O\left((n_N)^{\alpha_2}\right),$$

where the constant in the O just depend on $\mathbf{u}$. There thus exists a positive number C independent of N, such that:

$$\left|S_{\mathbf{u}}^{f_0}(N)\right|>\left|\sum_{k=0}^{n_N}F_{f_0,2,\alpha_2}(E_{n_N-k})k^{\alpha_2}e^{ik\gamma}\right|-C(n_N)^{\alpha_2}. \tag{13}$$

The next step consists in exhibiting a sequence $(\mathcal{C}_j)_{j\in\mathbb{N}}$ of arbitrarily large admissible paths in the prefix automaton associated with the pair $(\sigma,\mathbf{u})$, starting from the state 1, labelled by $(E_0,E_1,\ldots,E_{n_j})$, $E_0\neq\varepsilon$ and such that:

$$|S_{\mathbf{u}}^{f_0}(N_j)|>\left|\sum_{k=0}^{n_j}F_{f_0,2,\alpha_2}(E_{n_j-k})k^{\alpha_2}e^{ik\gamma}\right|-C(n_j)^{\alpha_2}>M(n_j)^{(\alpha_2+1)},$$

where the integers N_j are given by

$$N_j=\sum_{k=0}^{n_j}|\sigma^k(E_{n_j-k})|.$$

Then, Equation (13) and Proposition 11 will allow us to conclude that $\Delta_N(\mathbf{u})=\Omega((\log N)^{\alpha_2+1})$ because the assertion (iii) in Theorem 14 ensures that

$$\left|n_j-\frac{\log N_j}{\log\theta}\right|=O(1).$$

Since w is a proper prefix of $\sigma(1)$, there exists a letter j_0 such that wj_0 is a prefix of $\sigma(1)$. The labelled path $(1,j_0,w)$ is thus admissible. Let us denote by $(j_k)_{k\in\mathbb{N}}$ the sequence of states defined as follows: j_1 is the first letter of $\sigma(j_0)$ and more generally let j_{k+1} be the first letter of $\sigma(j_k)$. Hence, for every positive integer k, the labelled path $((1,j_0,w),(j_0,j_1,\varepsilon),(j_1,j_2,\varepsilon),\ldots,(j_k,j_{k+1},\varepsilon))$ is admissible. The set $\mathcal{A}$ being finite, we can find two positive integers k_0 and k_1, $k_1>k_0$, such that $j_{k_0}=j_{k_1}$. Then, we have to distinguish two cases:

- either $j_{k_0}=1$,
- or $j_{k_0}\neq 1$.

We first assume that $j_{k_0}=1$, as it is represented on Figure 2.

Let us consider a real δ, $\pi>\delta>0$ and V a subset of $\mathbb{C}$ defined by:

$$V=\left\{e^{i\lambda},\ \ -\delta<\lambda<\delta\right\}.$$

Then, there exists a positive number c such that for any positive integer N:

$$(v_1,v_2,\ldots,v_N)\in V^N\Longrightarrow\left|\sum_{i=1}^{N}v_i\right|>cN. \tag{14}$$

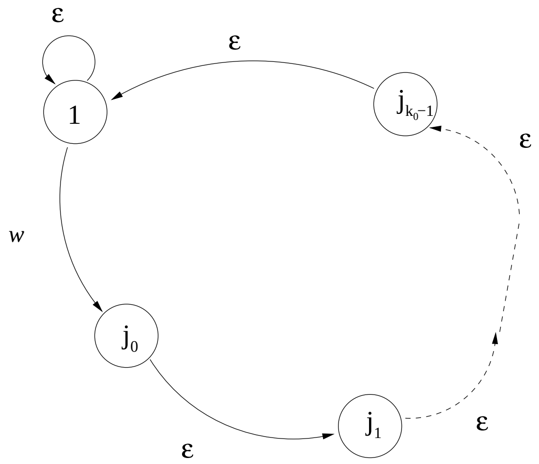

FIGURE 2. Case $j_{k_0} = 1$.

Let α be an irrational number and I an interval of the one-dimensional torus $\mathbb{T}$, then there exists a positive integer m such that:

$$\forall\, x \in \mathbb{T},\ \exists n \in \mathbb{N},\ n \leq m,\ x + n\alpha \in I, \tag{15}$$

as it is for instance proved in [25]. The irrationality of γ ensures, in view of (15), the existence of a sequence of integers $(n_k)_{k\in\mathbb{N}}$ and of an integer m satisfying

$$\forall\, j \in \mathbb{N},\ e^{in_j\gamma} \in V,\ n_0 = 0 \text{ and } m > n_j - n_{j-1} > k_0.$$

In particular, Inequality (14) implies

$$\left|\sum_{j=1}^{N} (n_j)^{\alpha_2} e^{in_j\gamma}\right| > cN^{(\alpha_2+1)}. \tag{16}$$

For any positive integer j, let us denote by $\mathcal{C}_j$ the labelled path of length n_j defined in the following way:

$$\Bigg((1,j,w),(j,j_1,\varepsilon),\ldots,(j_{k_0-1},1,\varepsilon),\underbrace{(1,1,\varepsilon),(1,1,\varepsilon),\ldots,(1,1,\varepsilon)}_{(n_N-n_{N-1})-k_0 \text{ times}},$$

$$(1,j,w),(j,j_1,\varepsilon),\ldots,(j_{k_0-1},1,\varepsilon),\underbrace{(1,1,\varepsilon),(1,1,\varepsilon),\ldots,(1,1,\varepsilon)}_{(n_{N-1}-n_{N-2})-k_0 \text{ times}},$$

$$\vdots \qquad\qquad \vdots$$

$$\vdots \qquad\qquad \vdots$$

$$\vdots \qquad\qquad \vdots$$

$$(1,j,w),(j,j_1,\varepsilon),\ldots,(j_{k_0-1},1,\varepsilon),\underbrace{(1,1,\varepsilon),(1,1,\varepsilon),\ldots,(1,1,\varepsilon)}_{(n_1-n_0)-k_0 \text{ times}},(1,j,w)\Bigg).$$

This labelled path is admissible in the prefix automaton associated with the pair $(\sigma, \mathbf{u})$. Moreover $F_{f_0,2,\alpha_2}(\varepsilon) = 0$ and we can deduce from Equation (13) and assertion (ii) in Theorem 14 the existence of an integer N_j satisfying

$$\left|S_{\mathbf{u}}^{f_0}(N_j)\right| \geq \left|\sum_{k=1}^{j} F_{f_0,2,\alpha_2}(w)(n_k)^{\alpha_2} e^{in_k\gamma}\right| - C(n_j)^{\alpha_2}.$$

Therefore, Inequality (16) implies that

$$\left|S_{\mathbf{u}}^{f_0}(N_j)\right| \geq \frac{c(n_j)^{(\alpha_2+1)}}{|F_{f_0,2,\alpha_2}(w)|} - C(n_j)^{\alpha_2}$$

and since (iii) in Theorem 14 ensures that

$$\left|n_j - \frac{\log N_j}{\log \theta}\right| = O(1),$$

we have

$$\left|S_{\mathbf{u}}^{f_0}(N_j)\right| > M(\log(N_j))^{(\alpha_2+1)},$$

where M is positive and does not depend on j. Hence, the sequence of labelled paths $(\mathcal{C}_j)_{j\in\mathbb{N}}$ provides an increasing sequence of integers $(N_j)_{j\in\mathbb{N}}$ such that

$$\Delta_{N_j}(\mathbf{u}) > D(\log(N_j))^{(\alpha_2+1)}.$$

Since $\Delta_N(\mathbf{u}) = O((\log N)^{\alpha_2+1})$ has already been proved in Proposition 15, this ends the proof in the case $j_{k_0} = 1$.

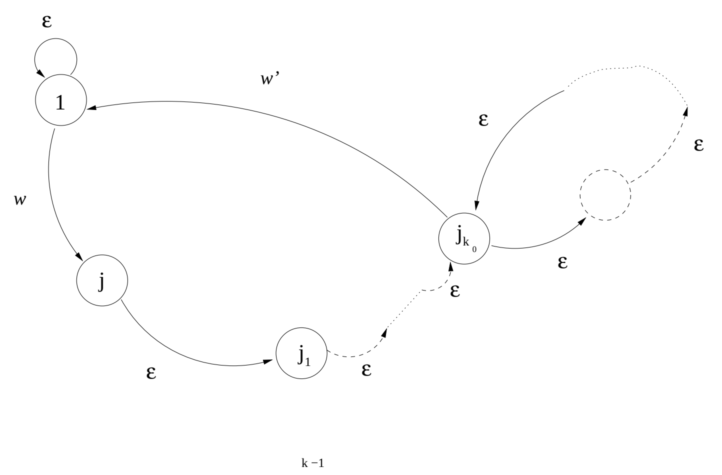

FIGURE 3. Case $j_{k_0} \neq 1$.

The case $j_{k_0} \neq 1$ (see Figure 3), is similar but a little bit more technical. It can be dealt by using, instead of (15), the fact that for any irrational α any interval I of $\mathbb{T}$ there exists a positive integer m satisfying

$$\forall\, x \in \mathbb{T},\ \exists\,(n,l) \in \mathbb{N}^2,\ n \leq m,\ l \leq m,\ \ x + n\alpha \in I \text{ and } x + l((k_1 - k_0)\alpha) \in I.$$

□

5.4. **Case *(iv)* of Theorem 1.** We begin this section with a definition of the complex number $A_{\sigma,\mathbf{u}}$ used in Theorem 1 and Corollary 2. The meaning of $A_{\sigma,\mathbf{u}}$ is strongly connected with the notion, introduced in [2], of elementary loops in the prefix automaton.

Definition 20. Let σ be a substitution and let us suppose that $\mathbf{u}$ is a fixed point for σ generated by the letter 1. We call elementary loop any admissible labelled path $((i_0, i_1, E_0), \ldots, (i_{n-1}, i_n, E_{n-1}))$ in the prefix automaton associated with the pair $(\sigma, \mathbf{u})$, satisfying the following conditions:

– $i_0 = i_n$,

– $\forall\ 0 \le j < k < n,\ i_k \neq i_j$.

We will denote by $\mathcal{E}l(\sigma, \mathbf{u})$ the set composed by all the elementary loops in the prefix automaton associated with the pair $(\sigma, \mathbf{u})$.

Remark 21. Since $\mathcal{A}$ and $Pref_\sigma$ are finite sets, $\mathcal{E}l(\sigma, \mathbf{u})$ is finite too.

Let σ be a primitive substitution defined over the alphabet $\mathcal{A} = \{1, 2, \ldots, d\}$ and such that θ_2 is a root of unity. This implies (see Section 2.3) that all the eigenvalues of M_σ whose modulus equals one and whose multiplicity equals α_2 are roots of unity. Then, let n_0 be the l.c.m. of the orders of these eigenvalues (considered as roots of unity).

Let $\mathbf{u}$ be a fixed point for σ generated by the letter 1. The sequence $\mathbf{u}$ is thus also a fixed point generated by the letter 1 for σ^{n_0}. For any admissible labelled path $\mathcal{C} = ((i_0, i_1, E_0), \ldots, (i_{n-1}, i_n, E_{n-1}))$, in the prefix automaton associated with the pair $(\sigma^{n_0}, \mathbf{u})$, we introduce:

$$F_{f,2,\alpha_2}(\mathcal{C}) = \sum_{j=0}^{n-1} F_{f,2,\alpha_2}(E_j), \tag{17}$$

where $F_{f,2,\alpha_2}(E_j)$ is defined following Equality (10). For any vector $f \in \mathbb{C}^d$, we introduce the quantity

$$A^f_{\sigma^{n_0},\mathbf{u}} = \max\left\{|F_{f,2,\alpha_2}(B)|,\ B \in \mathcal{E}l(\sigma^{n_0}, \mathbf{u})\right\}.$$

Then, we can define a complex number, denoted by $A_{\sigma,\mathbf{u}}$, just depending on the pair $(\sigma, \mathbf{u})$, by:

$$A_{\sigma,\mathbf{u}} = \max\left\{A^{f_j}_{\sigma^{n_0},\mathbf{u}},\ 1 \le j < d\right\}, \tag{18}$$

where vectors f_i are defined as in (4).

We are now ready to state the following proposition.

Proposition 22. *Let* $\mathbf{u}$ *be a fixed point of the primitive substitution* σ*. If* θ_2 *is a root of unity, then*

- *either* $A_{\sigma,\mathbf{u}} \neq 0$ *and* $\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{(\alpha_2+1)}\right)$,
- *or* $A_{\sigma,\mathbf{u}} = 0$ *and* $\Delta_N(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_2}\right)$.

In order to prove Proposition 22, we first need the two following results.

Lemma 23. *Let C be a positive number. Then, for any sequence of complexe numbers $(a_k)_{k\in\mathbb{N}}$ satisfying $|\sum_{k=0}^m a_k| \le C$ for any positive integer m, we have*

$$\sum_{k=0}^{n} a_k k^l = O(n^l),$$

the constant in the O just depending on C.

Proof. It comes directly from a classical Abel summation. □

Lemma 24. *Let $\mathbf{u}$ be a fixed point of a primitive substitution and f be a vector in $\mathbb{C}^d$ such that $f \perp \mu$. Then, there exists a constant $C > 0$ such that for any admissible labelled path $\mathcal{C}$ in the prefix automaton associated with the pair $(\sigma, \mathbf{u})$, one can find $(B_1, B_2, \ldots, B_k) \in (\mathcal{E}l)^k$, k possibly equal to zero, satisfying*

$$\left| F_{f,2,\alpha-2}(\mathcal{C}) - \sum_{i=1}^{k} F_{f,2,\alpha_2}(B_i) \right| \le C,$$

where $F_{f,2,\alpha_2}(B_i)$ is defined following Equality (17).

Proof. Let us reason by induction on the length n of the path. Let us consider $C = \max\left\{F_{f,2,\alpha_2}(\mathcal{C}),\ \mathcal{C} \in \mathcal{C}_\sigma^k,\ k \le d\right\}$. If $\mathcal{C}$ means an admissible path of length smaller than or equal to d, then by definition of C we have

$$|F_{f,2,\alpha_2}(\mathcal{C}) - 0| \le C,$$

which shows that the proposition is satisfied for $n \le d$.

Now, let $n \in \mathbb{N}$, $n > d$, and let us assume that the proposition is satisfied for any admissible path of length k, $k < n$. If $\mathcal{C} = ((i_0, i_1, E_0), \ldots, (i_{n-1}, i_n, E_{n-1}))$ is an admissible labelled path of length n, then there exists $(l, h) \in \{0, 1, 2, \ldots, n\}^2$, $l < h$, such that $i_l = i_h$ since the cardinality of $\mathcal{A}$ is equal to d. Let us denote by h' the minimum of $\{m,\ m > l$, such that $i_m = i_l\}$. It follows that $(i_l, \ldots, i_{h'})$ is an elementary loop and $((i_0, i_1, E_0), \ldots, (i_l, i_{h'+1}, E_h), \ldots, (i_{n-1}, i_n, E_n))$ is an admissible labelled path of length smaller than or equal to n. Thus, the induction hypothesis implies that there exists $(B_1, B_2, \ldots, B_k) \in (\mathcal{E}l)^k$, k eventually equal to zero, such that:

$$\left| F_{f,2,\alpha_2}\left((i_0, i_1, E_0), \ldots, (i_l, i_{h'+1}, E_h), \ldots, (i_{n-1}, i_n, E_n)\right) - \sum_{i=1}^{k} F_{f,2,\alpha_2}(B_i) \right| \le C.$$

But

$$\begin{aligned} F_{f,2,\alpha_2}(\mathcal{C}) = \quad & F_{f,2,\alpha_2}\left((i_0, i_1, E_0), \ldots, (i_l, i_{h'+1}, E_{h'}), \ldots, (i_{n-1}, i_n, E_{n-1})\right) \\ & + F_{f,2,\alpha_2}\left((i_l, i_{l+1}, E_l), \ldots, (i_{h'-1}, i_{h'}, E_{h'-1})\right), \end{aligned}$$

and thus if $B_{k+1} = ((i_l, i_{l+1}, E_l), \ldots, (i_{h'-1}, i_{h'}, E_{h'-1}))$, it follows that

$$\left| F_{f,2,\alpha_2}(\mathcal{C}) - \sum_{i=1}^{k+1} F_{f,2,\alpha_2}(B_i) \right| \le C,$$

concluding the proof. □

Proof of Proposition 22. We can assume without restriction that $\theta_2 = 1$.

We first assume that $A_{\sigma,\mathbf{u}} = 0$. Let f be a vector in $\mathbb{C}^d$ such that $f \perp \mu$ and N be a positive integer. Following Theorem 14, there exists $(E_0, E_1, \ldots, E_{n_N}) \in (Pref_\sigma)^{n_N}$ such that:

$$S^f_{\mathbf{u}}(N) = \sum_{k=0}^{n_N} S^f_{\mathbf{u}}(\sigma^k(E_{n_N-k})).$$

The fact that the E_i lie in the finite set $Pref_\sigma$ implies that

$$S^f_{\mathbf{u}}(N) = \sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k})k^{\alpha_2} + O\left((n_N)^{\alpha_2}\right),$$

where the constant in the O just depends on $\mathbf{u}$. We thus have to show that

$$\sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k})k^{\alpha_2} = O\left((n_N)^{\alpha_2}\right).$$

Lemma 24 implies the existence of a positive C and $(B_1, B_2, \ldots, B_k) \in (\mathcal{E}l(\sigma, \mathbf{u}))^k$ such that

$$\left|\sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k}) - \sum_{i=1}^{k_N} F_{f,2,\alpha_2}(B_i)\right| \leq C,$$

since by definition

$$\sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k}) = F_{f,2,\alpha_2}(\mathcal{C}).$$

But if $A_{\sigma,\mathbf{u}} = 0$, then $F_{f,2,\alpha_2}(B) = 0$ for any $B \in \mathcal{E}l(\sigma, \mathbf{u})$, which gives

$$\left|\sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k})\right| \leq C.$$

Since C does not depend on the path $\mathcal{C}$ and thus on the integer N, Lemma 23 implies that

$$\sum_{k=0}^{n_N} F_{f,2,\alpha_2}(E_{n_N-k})k^{\alpha_2} = O\left((n_N)^{\alpha_2}\right),$$

where the constant in the O does not depend on the choice of N. Following Proposition 18, we obtain that $\Delta_N(\mathbf{u}) = (O \cap \Omega)((\log N)^{\alpha_2})$ because $n_N = \log N + O(1)$.

Now, let us assume $A_{\sigma,\mathbf{u}} \neq 0$. There thus exists a vector f_j, defined as in (4), such that $A^{f_j}_{\sigma,\mathbf{u}} \neq 0$. Let us denote by $B = ((i_0, i_1, E_0), \ldots, (i_{p-1}, i_p, E_{P-1}))$ an element of $\mathcal{E}l(\sigma, \mathbf{u})$ satisfying

$$|F_{f_j,2,\alpha_2}(B)| = A^{f_j}_{\sigma,\mathbf{u}}.$$

Since the prefix automaton is strongly connected, there exists an admissible labelled path starting from the state 1 and ending in i_0. Let $\mathcal{C}_0 = ((a_0, a_1, E'_0), (a_1, a_2, E'_1), \ldots, (a_{l-1}, i_0, E'_{l-1}))$ be such a path, with $E'_0 \neq \varepsilon$ and $a_0 = 1$. For every positive integer k, we introduce the following labelled path:

$$\Big((a_0, a_1, E'_0), (a_1, a_2, E'_1), \ldots, (a_{l-1}, i_0, E'_{l-1}),\ \underbrace{(i_o, i_1, E_0), \ldots, (i_{p-1}, i_p, E_{p-1})}_{\text{iterated } k \text{ times}}\Big).$$

This path of length $l + kp$ is thus admissible, begins with 1 and satisfies $E_0 \neq \varepsilon$. Following the proposition (ii) in Theorem 14, there exists a positive integer N_k such that:

$$S^{f_j}_{\mathbf{u}}(N_k) = \sum_{m=0}^{kp} F_{f_j,2,\alpha_2}(E_{(m \bmod p)})m^{\alpha_2} + O(1).$$

Moreover, following Equality (12), we have

$$kp \sim \frac{\log(N_k)}{\log\theta},$$

which implies that

$$\lim_{k\to\infty} \frac{\left|S_{\mathbf{u}}^{f_j}(N_k)\right|}{(\log N_k)^{(\alpha_2+1)}} = \frac{A_{\mathbf{u}}^{f_j}}{\log\theta} > 0.$$

Finally, $\left|S_{\mathbf{u}}^{f_j}(N)\right| = \Omega((\log N)^{(\alpha_2+1)})$, which ends the proof in view of Proposition 11 and 18. □

Proof of Theorem 1. It comes directly from Proposition 15, 18, 19 and 22. □

6. Proof of Theorem 3

The goal of this section is to prove Theorem 3. We proceed as follows; we first generalize Theorem 1 (see Proposition 28) and next we have to study the discrepancy of the derivative sequences associated with a fixed point of a primitive substitution (see Lemma 31 and 32) in order to found uniform bounds in Proposition 28. Then a finitude argument due to [12] (Theorem 30) will allow us to conclude the proof.

Discrepancy functions associated with a symbolic sequence. We give here a generalization of Theorem 1. The discrepancy function associated with a symbolic sequence $\mathbf{u}$ measures the speed of convergence of the vector

$$\left(\frac{|u_0u_1\dots u_{N-1}|_a}{N}\right)_{a\in\mathcal{A}}$$

towards the frequencies vector of the sequence $\mathbf{u}$. We want now to introduce a similar notion but with words playing the role of letters. Let $\mathbf{u}$ be a symbolic sequence defined over the alphabet $\mathcal{A}$. Then, we can define, for any positive integer n, a discrepancy function of order n for $\mathbf{u}$, in the following way:

$$\Delta_N^{(n)}(\mathbf{u}) = \max_{w\in\mathcal{L}_n(\mathbf{u})} \left|\sum_{k=0}^{N-1} \chi_{[w]}(T^k(\mathbf{u})) - N\mu(w)\right|,$$

where $|u|_w$ denotes the number of occurrences of the word w in the word u. We obtain in particular $\Delta_N^{(1)}(\mathbf{u}) = \Delta_N(\mathbf{u})$. In view of the previous study, it is quite natural to ask if we can estimate the growth order of these discrepancy functions in the case of fixed points of primitive substitutions. In particular, is it possible to obtain such an information in terms of the incidence matrix associated with the substitution?

In order to answer this question, we recall now a useful construction which can be found in [21]. Let σ be a primitive substitution defined over the alphabet $\mathcal{A}$ and $\mathbf{u}$ an associated fixed point. For any positive integer l, $\mathcal{A}_l$ denotes the alphabet $\{1, 2, \dots, P_{\mathbf{u}}(l)\}$, where $P_{\mathbf{u}}$ is the complexity function of $\mathbf{u}$. We can thus consider a map Θ_l from $\mathcal{L}_l(\mathbf{u})$ to $\mathcal{A}_l$ wich associates with each factor of length l its order of occurrence in $\mathbf{u}$. If i denotes a letter of the alphabet $\mathcal{A}_l$, we can conversely associate with i a unique word $\Theta_l^{-1}(i) = w_0w_1\dots w_{l-1} \in \mathcal{L}_l(\mathbf{u})$ since Θ_l is one-to-one. If

$$\sigma(\Theta^{-1}(i)) = \sigma(w_0w_1\dots w_{l-1}) = y_0y_1\cdots y_{|\sigma(w_0)|-1}y_{|\sigma(w_0)|}\cdots y_{|\sigma(\Theta^{-1}(i)|-1},$$

then, we define the substitution of order l for σ by:

$$\sigma_l(i) = \Theta\left((y_0y_1\dots y_{l-1})(y_1y_2\dots y_l)\dots(y_{|\sigma(y_0)|-1}\cdots y_{|\sigma(y_0)|+l-2})\right). \tag{19}$$

So defined, $|\sigma_l(i)| = |\sigma(w_0)|$.

We recall now some results about the previous construction.

Proposition 25 (Queffélec [21]). *For every positive integer l, the substitution of order l for a subtitution σ admits the sequence $\mathbf{u}_l = \sigma_l^{\infty}(1)$ as a fixed point. Moreover, if $\mathbf{u} = u_0u_1 \dots u_n \dots$ means $\sigma^{\infty}(1)$, then the sequence $\Theta_l^{-1}(\mathbf{u}_l)$ is composed by all the factors of length l of $\mathbf{u}$ without repetition and in the same order as in $\mathbf{u}$, that is to say,*

$$\Theta_l^{-1}(\mathbf{u}_l) = (u_0u_1 \dots u_{l-1})(u_1u_2 \dots u_l) \dots (u_nu_{n+1} \dots u_{n+l-1}) \dots .$$

We can already notice that if $\mathbf{u}_l = u_0^{(l)}u_1^{(l)} \dots u_n^{(l)} \dots$, then

$$|u_0^{(l)}u_1^{(l)} \dots u_n^{(l)}|_i = |u_0u_1 \dots u_{n+l-1}|_{\Theta_l^{-1}(i)}. \tag{20}$$

This implies in particular the following corollary.

Corollary 26. *The order of magnitude of the discrepancy function of order l for $\mathbf{u}$ is the same as that of the discrepancy function (of order one) of $\mathbf{u}_l$.*

Proposition 27 (Queffélec [21]). *If σ is a primitive substitution then for every positive integer l, the substitution σ_l is primitive too and its incidence matrix M_l has the same Perron eigenvalue as the one of σ.*

The eigenvalues of M_l, $l \geq 2$, are those of M_2 with perhaps in addition the eigenvalue 0. Moreover, if P_2 is the minimal polynomial of M_2, then there exists an integer m such that $P_l = P_2X^m$, where P_l means the minimal polynomial of M_l.

Following Equation (2.3), we can note

$$S_{M_{\sigma_l}} = \{\theta_{l,i},\ 2 \leq 2 \leq d_l\} \cup \{\theta_{l,1} = \theta_l\}$$

the spectrum of the incidence matrix associated with σ_l. Proposition 27 implies that $\theta_l = \theta$, $\theta_{l,2} = \theta_{2,2}$ and $\alpha_{l,2} = \alpha_{2,2}$, where $\alpha_{l,2}$ means the multiplicity of $\theta_{l,2}$ in the minimal polynomial of M_l. In view of Corollary 26 and Proposition 27, we can state the following result.

Proposition 28. *Let $\mathbf{u}$ be a fixed point of a primitive substitution σ. Then, we have for every integer $l \geq 2$:*

(i)	*if $|\theta_{2,2}| < 1$, then*	$\Delta_N^{(l)}(\mathbf{u})$ *is bounded* ,
(ii)	*if $|\theta_{2,2}| > 1$, then*	$\Delta_N^{(l)}(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_{2,2}} N^{(\log_\theta |\theta_{2,2}|)}\right)$,
(iii)	*if $|\theta_{2,2}| = 1$ and $\theta_{2,2}$ is not a root of unity, then*	$\Delta_N^{(l)}(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{(\alpha_{2,2}+1)}\right)$,
(iv)	*if $|\theta_{2,2}| = 1$ and $\theta_{2,2}$ is a root of unity, then*	$\Delta_N^{(l)}(\mathbf{u}) = O\left((\log N)^{(\alpha_{2,2}+1)}\right)$,
	and	$\Delta_N^{(l)}(\mathbf{u})B = \Omega\left((\log N)^{\alpha_{2,2}}\right)$.

Moreover, in the case where $\theta_{2,2}$ is a root of unity, then:

- *either* $\forall\, l \geq 2$, $\Delta_N^{(l)}(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_{2,2}}\right)$,
- *or there exists an integer $m \geq 2$ such that,*
  for $l < m$, $\Delta_N^{(l)}(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_{2,2}}\right)$,
  and for $l \geq m$, $\Delta_N^{(l)}(\mathbf{u}) = (O \cap \Omega)\left((\log N)^{\alpha_{2,2}+1}\right)$.

For proving Proposition 28, we need the following lemma.

Lemma 29. *Let l be a positive integer, $\mathbf{u}$ an infinite sequence defined over the alphabet $\mathcal{A}$ and suppose that there exists a function f such that $\Delta_N^{(l)}(\mathbf{u}) = \Omega(f(N))$. Then, we have $\Delta_N^{(l+1)}(\mathbf{u}) = \Omega(f(N))$.*

Lemma 29 points out the fact that the order of magnitude of the discrepancy functions $\Delta_N^{(l)}(\mathbf{u})$ associated with a symbolic sequence $\mathbf{u}$ could not decrease with respect to l.

Proof of Proposition 28. Equality (i), (ii), (iii) and (iv) come directly from Corollary 26, Proposition 27 and Theorem 1. Then, the last point of Proposition 28 is a consequence of Lemma 29. □

Discrepancy for derivative sequences. The different constants obtained for Δ_N^l in Proposition 28 depend *a priori* on l. To state Theorem 3 the next step is to prove that one can found uniform constants in Proposition 28. Our approach consists in exhibiting some connections between the discrepancy of a linearly reccurent sequence and the one of its derivative sequences and then to use a finitude argument given in [12] (Theorem 30). Then, the last step will be to show that one can deal with any sequence $\mathbf{v} \in \overline{\mathcal{O}(\mathbf{u})}$ insted of the sequence $\mathbf{u}$ itself.

We recall the main definitions and results that are given in [12] concerning the notion of return words. Let $\mathbf{u}$ be a uniformly recurrent sequence over the alphabet $\mathcal{A}$ and let u be a nonempty factor of $\mathbf{u}$. A *return word* to u of $\mathbf{u}$ is a factor $u_{[i,k-1]}$ $(= u_i u_{i+1} \dots u_{k-1})$ of $\mathbf{u}$ such that i and k are two consecutive occurrences of u. If j denotes the first occurrence of u in $\mathbf{u}$, the sequence $T^j(\mathbf{u})$ can be written in a unique way as a concatenation of return words to u. Let $\mathcal{R}_{\mathbf{u},u}$ be the set of return words to u in $\mathbf{u}$. Then $T^j(\mathbf{u}) = \omega_0\omega_1 \dots \omega_i \dots$, where $\omega_i \in \mathcal{R}_{\mathbf{u},u}$. The fact that $\mathbf{u}$ is uniformly recurrent implies that $\mathcal{R}_{\mathbf{u},u}$ is a finite set. We can therefore consider a bijective map $\Phi_{\mathbf{u},u}$ from $\mathcal{R}_{\mathbf{u},u}$ to the finite set $\{1, 2, \dots, \mathrm{Card}(\mathcal{R}_{\mathbf{u},u})\} = \mathcal{A}_{\mathbf{u},u}$, where, for definiteness, the return words are ordered according to their first occurrence (i.e., $\Phi_{\mathbf{u},u}^{-1}(1)$ is the first return word ω_0, $\Phi_{\mathbf{u},u}^{-1}(2)$ is the first ω_i which is different from ω_0, and so on). The *derivative sequence* of $\mathbf{u}$ on u is the sequence with values in the alphabet $\mathcal{A}_u$ given by

$$\mathcal{D}_u(\mathbf{u}) = \Phi_{\mathbf{u},u}(\omega_0)\Phi_{\mathbf{u},u}(\omega_1) \dots \Phi_{\mathbf{u},u}(\omega_i) \dots .$$

To such a sequence we can associate a morphism $\Theta_{\mathbf{u},u}$ from $\mathcal{A}_{\mathbf{u},u}$ to $\mathcal{A}^*$ defined by:

$$\Theta_{\mathbf{u},u}(i) = \omega_i.$$

We obtain $\Theta_{\mathbf{u},u}(\mathcal{D}_u(\mathbf{u})) = T^j(\mathbf{u})$. The morphism $\Theta_{\mathbf{u},u}$ is called the return morphism to u of $\mathbf{u}$.

An important point is that one can characterize primitive substitutive sequences in terms of derivative sequences.

Theorem 30 (Durand [12])**.** *Let $\mathbf{u}$ be a uniformly recurrent sequence. Then, the set $\mathcal{D}er(\mathbf{u})$ of its derivative sequences, is finite if and only if $\mathbf{u}$ is a primitive substitutive sequence.*

Now, if $w_1, w_2, \dots, w_k$ denote some factors of $\mathbf{u}$, we define the quantities:

$$\Delta_N(\mathbf{u}; w_i) = \left| \sum_{k=0}^{N-1} \chi_{[w_i]}(T^k(\mathbf{u})) - N\mu([w]) \right|$$

and

$$\Delta_N(\mathbf{u}; w_1, w_2, \dots, w_k) = \max_{1 \le i \le k} \Delta_N(\mathbf{u}; w_i).$$

Let $\mathbf{u}$ be a K-linearly reccurent sequence, w be a factor of $\mathbf{u}$ and $\mathbf{v}$ be the derivative sequence of $\mathbf{u}$ on w. Then, the sequence $\mathbf{v}$ is also linearly recurrent (with another constant) and this implies that any letter in $\mathbf{v}$ has a frequency. It thus exists a natural probability measure ν for $\mathbf{v}$ (in the sense of Section 2.2). We will thus write $\Delta_n(\mathbf{v})$ instead of $\Delta_n(\nu, \mathbf{v})$ for the discrepancy of the sequence $\mathbf{v}$ with respect to ν.

Then, we can state the following.

Lemma 31. *Let $\mathbf{u}$ be a K-linearly reccurent sequence, w be a factor of $\mathbf{u}$ and $\mathbf{v}$ be the derivative sequence of $\mathbf{u}$ on w. Then, there exists a constant C such that:*

$$\forall m \in \mathbb{N}^*,\, \exists n \in \mathbb{N}^*,\, n < Km, \quad \Delta_m(\mathbf{v}) < C\Delta_n(\mathbf{u}; w_1, w_2, \ldots, w_r),$$

where $w_1, w_2, \ldots, w_r$ denote the return words of $\mathbf{u}$ to w.

Lemma 32. *Let $\mathbf{v}$ be a derivative sequence of a K-LR sequence $\mathbf{u}$. Then, there exists a constant $C_{\mathbf{v}}$ such that for any factor w of $\mathbf{u}$ satisfying $\mathcal{D}_w(\mathbf{u}) = \mathbf{v}$ the following holds:*

$$\forall n \in \mathbb{N}^*,\, \exists m \in \mathbb{N}^*,\, m < n, \quad \Delta_n(\mathbf{u}; w) < C_{\mathbf{v}}\Delta_m(\mathbf{v}).$$

Proof of Lemma 31. Let w be a factor of the sequence $\mathbf{u}$ and $\mathbf{v} = D_w(\mathbf{u})$ be the derivative sequence of $\mathbf{u}$ on w. Let j denote the first occurrence of W in $\mathbf{u}$ and r the cardinality of the set $\mathcal{R}_{\mathbf{u},w}$. The invariant measures associated with $\mathbf{u}$ and $\mathbf{v}$ are respectivly denoted by μ and ν (we refer the reader to [11] for a proof of the unique ergodicity of a linearly recurrent subshift). We will write Θ instead of $\Theta_{\mathbf{u},w}$.

We thus obtain that $\Theta(\mathbf{v}) = T^j(\mathbf{u})$, where T denotes the usual shift. With any integer m, we can associate an integer n such that:

$$\Theta(v_0 v_1 \ldots v_m) = u_j u_{j+1} \ldots u_n$$

and

$$n - j = \sum_{k=1}^{r} |v_0 v_1 \ldots v_m|_k |w_k|.$$

Since $\mathbf{u}$ is K-LR, we have that $j < Km$. Moreover, we have by definition of a return word:

$$|v_0 v_1 \ldots v_m|_i = \sum_{k=0}^{n} \chi_{[w_i]}(T^k(\mathbf{u})).$$

Then

$$\frac{\sum_{k=0}^{n} \chi_{[w_i]}(T^k(\mathbf{u}))}{n} = \frac{|v_0 v_1 \ldots v_m|_i}{\sum_{k=1}^{r} |v_0 v_1 \ldots v_m|_k |w_k|},$$

which implies

$$\mu([w_i]) = \frac{\nu([i])}{\sum_{k=1}^{r} \nu([k])|w_k|},$$

for $i \in \{1, 2, \ldots r\}$. It thus follows that for $i \in \{1, 2, \ldots r\}$:

$$\sum_{k=0}^{n} \chi_{[w_i]}(T^k(\mathbf{u})) - n\mu([w_i]) = |v_0 v_1 \ldots v_m|_i - \frac{\left(\sum_{k=1}^{r} |v_0 v_1 \ldots v_m|_k |w_k|\right) \nu([i])}{\sum_{k=1}^{r} \nu([k])|w_k|}$$

$$= |v_0 v_1 \ldots v_m|_i - m\nu([i]) + \nu([i]) \left(m - \frac{\sum_{k=1}^{r} |v_0 v_1 \ldots v_m|_k |w_k|}{\sum_{k=1}^{r} \nu([k])|w_k|} \right)$$

$$= |v_0 v_1 \ldots v_m|_i - m\nu([i]) + \frac{\nu([i])}{\sum_{k=1}^{r} \nu([k])|w_k|} \left(\sum_{k=1}^{r} |w_k|(m\nu([k]) - |v_0 v_1 \ldots v_m|_k \right).$$

Let us now consider the following $d \times d$ real matrix:

$$M = \begin{pmatrix} 1 + \frac{\nu([1])}{h} & \frac{\nu([1])}{h} & \ldots\ldots\ldots\ldots & \frac{\nu([1])}{h} \\ \frac{\nu([2])}{h} & 1 + \frac{\nu([2])}{h} & \ldots\ldots\ldots\ldots & \frac{\nu([2])}{h} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\nu([r])}{h} & \frac{\nu([r])}{h} & \ldots\ldots\ldots\ldots & 1 + \frac{\nu([r])}{h} \end{pmatrix},$$

where $h = \sum_{k=1}^{r} \nu([k])|w_k|$. It thus follows that:

$$\begin{pmatrix} |u_0 u_1 \ldots u_n|_{w_1} - n\mu(w_1) \\ \vdots \\ |u_0 u_1 \ldots u_n|_{w_r} - n\mu(w_r) \end{pmatrix} = M \begin{pmatrix} |w_1|(m\nu([1]) - |v_0 v_1 \ldots v_m|_1) \\ \vdots \\ |w_d|(m\nu([r]) - |v_0 v_1 \ldots v_m|_r) \end{pmatrix}.$$

By a simple computation one can check that M is invertible if and only if

$$\sum_{l=1}^{r} \left(\frac{\nu([l])}{\sum_{k=1}^{r} \nu([k])|w_k|} \right) = \frac{1}{\sum_{k=1}^{r} \nu([k])|w_k|} \neq -1.$$

This implies the existence, for every $1 \leq j \leq r$, of coefficents $\left(a_1^{(j)}, a_2^{(j)}, \ldots, a_r^{(j)}\right) \in \mathbb{R}^r$ such that for any integer m there exists an integer $n < Km$ satisfying

$$|v_0 v_1 \ldots v_m|_j - m\nu([j]) = \sum_{k=1}^{r} a_k^{(j)} \left(|u_0 u_1 \ldots u_n|_{w_k} - n\mu([w_k])\right)$$

and therefore the existence of a positive number C satisfying

$$\Delta_m(\mathbf{v}) < C\Delta_n(\mathbf{u}; w_1, w_2, \ldots, w_r),$$

concluding the proof. □

Proof of Lemma 32. We keep here the notation of the previous proof. Let us introduce the vector $f_w = (f_w(i))_{i\in\mathbb{R}^r}$ defined by $f_w(i) = 1 - |w_i|\mu([w])$. By definition of a derivative sequence, we are allowed to claim that

$$\sum_{k=0}^{n-1} \chi_{[w]}(T^k(\mathbf{u})) - n\mu([w]) = S_{\mathbf{v}}^{f_w}(m) - j\mu([w]),$$

where j denotes the first occurence of w in the sequence $\mathbf{u}$ and m and n are defined as above (we recall in particular that $m < n < Km$). By definition of the measure μ, we have

$$\sum_{k=0}^{n-1} \chi_{[w]}(T^k(\mathbf{u})) - n\mu([w]) = o(n),$$

and thus

$$S_{\mathbf{v}}^{f_w}(m) = o(m).$$

In view of Equality (8), we obtain that $f_w \perp \mu'$, where $\mu' = (\nu([i]))_{i=1,\ldots,r}$ is the frequencies vector of the sequence $\mathbf{v}$. As we have already noticed, the familly of vectors (f_k), $1 \le k \le r-1$, defined by

$$f_k(i) = \begin{cases} 1 & \text{if } i = k, \\ \frac{\nu([k])}{\nu([k])-1} & \text{else,} \end{cases}$$

is a basis of the vectorial subspace $(\mu')^{\perp}$. There thus exist coefficients $(b_1, \ldots, b_{r-1}) \in \mathbb{R}^r$ such that $f_w = \sum_{k=1}^{r-1} b_k f_k$. Since $\mathbf{u}$ is a K-LR, $|w|\mu([w]) < K$ (see [11]). It follows that $0 < |w_i|\mu([w]) < K^2$ and thus $\|f_w\|_\infty < K^2$. The fact that $S_{\mathbf{v}}^{f_k}(m) = (1-\nu([k]))\,(|v_0v_1\ldots v_m|_k - m\nu([k]))$ implies $\left|S_{\mathbf{v}}^{f_k}(m)\right| < \Delta_m(\mathbf{v})$. Moreover, by linearity we have $S_{\mathbf{v}}^{f_w}(m) = \sum_{k=1}^{r-1} b_k S_{\mathbf{v}}^{f_k}(m)$ and it thus follows:

$$\Delta_n(\mathbf{u}; w) < \sum_{k=1}^{r-1} b_k \Delta_m(\mathbf{v}) + j\mu([w]) < \sum_{k=1}^{r-1} b_k \Delta_m(\mathbf{v}) + K.$$

Since $\|f_w\|_\infty < K^2$, there exists a constant $C_{\mathbf{v}}$, which does not depend on the choice of w, and such that $\Delta_n(\mathbf{u}; w) < C_{\mathbf{v}}\Delta_m(\mathbf{v})$, hence the proof. □

Proof of Theorem 3. Let $\mathbf{u}$ be a fixed point of a primitive substitution σ and $\mathcal{X} = (\overline{\mathcal{O}(\mathbf{u})}, T, \mu)$ the dynamical system arising from $\mathbf{u}$. In order to fix the ideas, we assume that $|\theta_{2,2}| > 1$ and we note $f(N) = (\log N)^{\alpha_{2,2}} N^{(\log_\theta |\theta_{2,2}|)}$. Then, Proposition 28 implies that $D_N(\mathcal{X}) = \Omega(f(N))$ and that for any word $w \in \mathcal{L}(\mathbf{u})$, $\Delta_N(\mathbf{u}; w) = O(f(N))$. It remains to prove that $D_N(\mathcal{X}) = O(f(N))$. If $w_1, w_2, \ldots, w_d$, denote the return words of $\mathbf{u}$ to w, we thus deduce that there exists $C_1 > 0$ such that

$$\Delta_N(\mathbf{u}; w_1, w_2, \ldots, w_d) = \max_{1\le i\le d} \Delta_N(\mathbf{u}; w_i) < C_1 f(N).$$

Now let $\mathbf{v}$ be the derivative sequence of $\mathbf{u}$ on w. Lemma 31 thus implies the existence of $C_2 > 0$ satisfying $\Delta_N(\mathbf{v}) < C_2 f(N)$, because f is a sublinear function (that is, $\forall(x,y)\ f(x+y) \le f(x) + f(y)$) and hence

$$\sup_{\mathbf{v}\in\mathcal{D}er(\mathbf{u})} \Delta_N(\mathbf{v}) < C_3 f(N),$$

with $C_3 > 0$ since, following Theorem 30, $\mathcal{D}er(\mathbf{u})$ is a finite set. Therefore, Lemma 32 implies

$$\sup_{w \in \mathcal{L}(\mathbf{u})} \Delta_N(\mathbf{u}; w) < C_4 f(N),$$

for some $C_4 > 0$.

Let $\mathbf{w} = (w_k)_{k \geq 0} \in \overline{\mathcal{O}(\mathbf{u})}$ and j be the first occurrence of the word $w_0 \dots w_{N-1}$ in $\mathbf{u}$ (such an occurrence always exists by minimality). Since $\mathbf{u}$ is linearly reccurent, there exists $K > 0$ (just depending on $\mathbf{u}$), such that $j < KN$. Then, we have

$$\Delta_N(\mathbf{w}; w) \leq \Delta_j(\mathbf{u}; w) + \Delta_{j+N}(\mathbf{u}; w) < C_4 f(KN) + C_4 f((K+1)N) < C_5 f(N),$$

where C_5 neither depends on $\mathbf{w}$ nor on w, since f is increasing and sublinear. Finally, we obtain that $D_N(\mathcal{X}) = O(f(N))$, which achieves the proof since the other cases could be dealt exactly in the same way. □

Acknowledgments. I would like to thank Valérie Berthé, Nicolas Bédaride, Pierre Liardet and Christian Mauduit for useful discussions and comments during the preparation of this paper. The author is also grateful to the anonymous referee for valuable suggestions.

References

[1] B. Adamczewski. Codages de rotations et phénomènes d'autosimilarité. *J. Théor. Nombres Bordeaux*, 14:351–386, 2002.

[2] B. Adamczewski. Répartitions des suites $(n\alpha)_{n \in \mathbb{N}}$ et substitutions. *Acta Arith.*, 112:1–22, 2004.

[3] M. D. Boshernitzan and C. R. Carroll. An extension of Lagrange's theorem to interval exchange transformations over quadratic fields. *J. Anal. Math.*, 72:21–44, 1997.

[4] J. Brillhart, P. Erdős, and P. Morton. On sums of Rudin-Shapiro coefficients. II. *Pacific J. Math.*, 107:39–69, 1983.

[5] J. Coquet. A summation formula related to the binary digits. *Invent. Math.*, 73:107–115, 1983.

[6] F. M. Dekking. On the distribution of digits in arithmetic sequences. In *Seminar on number theory, 1982–1983 (Talence, 1982/1983)*, pages Exp. No. 32, 12. Univ. Bordeaux I, Talence, 1983.

[7] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications.* Springer-Verlag, Berlin, 1997.

[8] J.-M. Dumont and A. Thomas. Systèmes de numération et fonctions fractales relatifs aux substitutions. *Theoret. Comput. Sci.*, 65:153–169, 1989.

[9] J.-M. Dumont and A. Thomas. Digital sum problems and substitutions on a finite alphabet. *J. Number Theory*, 39:351–366, 1991.

[10] F. Durand. A characterization of substitutive sequences using return words. *Discrete Math.*, 179:89–101, 1998.

[11] F. Durand. Linearly recurrent subshifts have a finite number of non-periodic subshift factors. *Ergodic Theory Dynam. Systems*, 20:1061–1078, 2000.

[12] F. Durand. Combinatorial and dynamical study of substitutions around the theorem of cobham. *Dynamics and Randomness, Nonlinear Phenomena and Complex Systems*, Kluwer Acad. Pub, 53 – 94, 2002.

[13] H. Furstenberg, H. Keynes, and L. Shapiro. Prime flows in topological dynamics. *Israel J. Math.*, 14:26–38, 1973.

[14] G. Halász. Remarks on the remainder in Birkhoff's ergodic theorem. *Acta Math. Acad. Sci. Hungar.*, 28:389–395, 1976.

[15] C. Holton and L. Q. Zamboni. Geometric realizations of substitutions. *Bull. Soc. Math. France*, 126:149–179, 1998.

[16] H. Kesten. On a conjecture of Erdős and Szüsz related to uniform distribution mod 1. *Acta Arith.*, 12:193–212, 1966/1967.

[17] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[18] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding.* Cambridge University Press, Cambridge, 1995.

[19] P. Michel. Stricte ergodicité d'ensembles minimaux de substitution. *C. R. Acad. Sci. Paris Sér. A*, 278:811–813, 1974.

[20] K. Petersen. On a series of cosecants related to a problem in ergodic theory. *Compositio Math.*, 26:313–317, 1973.

[21] M. Queffélec. *Substitution dynamical systems—spectral analysis.* Lecture Notes in Mathematics **1294**, Springer-Verlag, Berlin, 1987.

[22] G. Rauzy. Nombres algébriques et substitutions. *Bull. Soc. Math. France*, 110:147–178, 1982.

[23] G. Rauzy. Sequences defined by iterated morphisms. In *Sequences (Naples/Positano, 1988)*, pages 275–286. Springer, New York, 1990.

[24] A. Siegel. Représentation géométrique, combinatoire et arithmétique des systèmes substitutifs de type Pisot. Thèse de doctorat de l'Université de la Méditerranée, 2000.

[25] N. B. Slater. Gaps and steps for the sequence $n\theta$ mod 1. *Proc. Cambridge Philos. Soc.*, 63:1115–1123, 1967.

[26] B. Solomyak. On the spectral theory of adic transformations. In *Representation theory and dynamical systems*, pages 217–230. Amer. Math. Soc., Providence, RI, 1992.