Journal of Integer Sequences, Vol. 16 (2013),
Article 13.2.6

Do the Properties of an S-adic Representation Determine Factor Complexity?

Fabien Durand
LAMFA, CNRS UMR 7352
Université de Picardie Jules Verne
UFR des Sciences
33, rue Saint-Leu
80039 Amiens Cedex 1
France
fabien.durand@u-picardie.fr

Julien Leroy
Department of Mathematics
University of Liège
Grande Traverse 12 (B37)
B-4000 Liège, Belgium
and
LAMFA, CNRS UMR 7352
Université de Picardie Jules Verne
UFR des Sciences
33, rue Saint-Leu
80039 Amiens Cedex 1
France
julien.leroy@uni.lu

Gwenaël Richomme
Université Paul-Valéry Montpellier 3
UFR IV, Dpt MIAp, Case J11
Route de Mende
34199 Montpellier Cedex 5
France
and
LIRMM (CNRS, Univ. Montpellier 2) - UMR 5506 - CC 477
161 rue Ada
34095 Montpellier Cedex 5
France
gwenael.richomme@lirmm.fr

Abstract

The S-adic conjecture postulates the existence of a condition C such that a sequence has linear complexity if and only if it is an S-adic sequence satisfying C for some finite set S of morphisms. We present an overview of the factor complexity of S-adic sequences and we give some examples that either illustrate some interesting properties, or that are counterexamples to what might seem to be a “good” condition C.

1 Introduction

A common tool in the study of a sequence (or an infinite word) $\mathbf{w}$ over a finite alphabet A is the complexity function $p_{\mathbf{w}}$ (or simply p) that counts the number of factors of each length n occurring in the sequence, i.e., $p_{\mathbf{w}}(n) = \mathrm{Card}(\{u \in A^* \mid |u| = n \text{ and } \exists x \in A^*, \mathbf{y} \in A^{\mathbb{N}} : \mathbf{w} = xu\mathbf{y}\})$ (see [BR10, Chap. 4] for a survey on complexity function). The set of factors of length n of $\mathbf{w}$ is denoted by $\mathrm{Fac}_n(\mathbf{w})$ and $\mathrm{Fac}(\mathbf{w}) = \bigcup_{n\in\mathbb{N}} \mathrm{Fac}_n(\mathbf{w})$. The complexity function $p_{\mathbf{w}}$ is clearly bounded by $(\mathrm{Card}(A))^n, n \in \mathbb{N}$, but not every function can be a complexity function. As an example, it is well known (see [MH38]) that either the sequence is ultimately periodic (and then $p_{\mathbf{w}}(n)$ is ultimately constant), or its complexity function grows at least like $n+1$. Non-periodic sequences with minimal complexity $p_{\mathbf{w}}(n) = n+1$ for all n exist. They are called *Sturmian sequences* and a large literature is devoted to them (see [Lot02, Chap. 2] and [Fog02, Chap. 6] for surveys on these sequences).

There is a huge literature about sequences with low complexity. Indeed, such sequences occur, for instance, in combinatorics [All94, Cas96, Cas97, GJ09, Kos98, Rot94], dynamical systems [AR91, Bos84, Bos92, Fer95, Fer99], number theory [AB07, AB11, AZ98, FM97], and geometry [CS01, HZ98, Rau82]. In this paper, by “linear complexity” we mean that the number of factors of length n is $O(n)$, and by “low complexity” we usually mean linear complexity.

Moreover, many well-known families of sequences can also be indefinitely desubstituted with a finite number of morphisms. Formally, an *S-adic sequence* is defined as follows. Let $\mathbf{w}$ be a sequence over a finite alphabet A. If S is a set of morphisms (possibly infinite), an *S-adic representation* of $\mathbf{w}$ is given by a sequence $(\sigma_n : A_{n+1}^* \to A_n^*)_{n\in\mathbb{N}}$ of morphisms in S and a sequence $(a_n)_{n\in\mathbb{N}}$ of letters, $a_i \in A_i$ for all i, such that[1] $A_0 = A$, $\mathbf{w} = \lim_{n\to+\infty} \sigma_0\sigma_1\cdots\sigma_n(a_{n+1}^\omega)$ and $\lim_{n\to+\infty} |\sigma_0\sigma_1\cdots\sigma_n(a_{n+1})| = +\infty$, where a_{n+1}^ω is the sequence only composed of occurrences of a_{n+1}. The sequence $(\sigma_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is the *directive word* of the representation. In the sequel, we will say that a sequence $\mathbf{w}$ is *S-adic* if there exists a set S of morphisms such that $\mathbf{w}$ admits an S-adic representation.

The notion of S-adicity first appeared when Ferenczi [Fer96] proved that the set of factors of any uniformly recurrent sequence with linear complexity admits an S-adic representation, with S finite, satisfying $\lim_{n\to+\infty} \min_a \in A_{n+1} |\sigma_0\sigma_1\cdots\sigma_n(a)| = +\infty$. The present work found its motivation in that paper. Indeed, an open problem is to determine the link between being an S-adic sequence and having linear complexity (see [AR91, Fer96, Ler12b]). This problem is called the *S-adic conjecture*:

[1]The topology over $A^{\mathbb{N}}$ is the classical product topology of the discrete topology over A.

Conjecture 1.1 (S-adic conjecture)**.** There exists a condition C such that a sequence has linear complexity if and only if it is an S-adic sequence satisfying Condition C for some finite set S of morphisms.

It is clear that we cannot avoid considering a particular condition since there exist some pure substitutive sequences with quadratic complexity.

In this paper, we present an overview of the factor complexity (aka "subword complexity") of S-adic sequences and we give some examples that either illustrate some interesting properties, or that are counterexamples to what might seem to be a "good" condition C.

In what follows, we consider that alphabets are finite subsets of $\mathbb{N}$ and if $\sigma : A^* \to B^*$ is a morphism with $A = \{0, 1, \ldots, k\}$, we write $\sigma = [\sigma(0), \ldots, \sigma(k)]$. The following example is classical when considering S-adic sequences.

Example 1.2. Let us define the four morphisms R_0, R_1, L_0 and L_1 over $\{0, 1\}$ by $R_0 = [0, 10]$, $R_1 = [01, 1]$, $L_0 = [0, 01]$ and $L_1 = [10, 1]$. Since the work of Morse and Hedlund [MH40], it is well known that for any Sturmian sequence $\mathbf{w}$, there is a sequence $(k_n)_{n\in\mathbb{N}}$ of integers such that

$$\mathbf{w} = \lim_{n\to+\infty} L_0^{k_0} R_0^{k_1} L_1^{k_2} R_1^{k_3} L_0^{k_4} R_0^{k_5} \cdots L_1^{k_{4n+2}} R_1^{k_{4n+3}}(0^\omega). \tag{1}$$

With, for all $n \geq 1$, the additional assumptions "$k_{2n} + k_{2n+1} \geq 1$" and "if $k_{2n} = 0$, then $k_{2n-1} \neq 0$", the sequence $(k_n)_{n\in\mathbb{N}}$ is unique [BHZ06, LR07].

It is important to notice that, when we talk about an S-adic sequence, the corresponding directive word $(\sigma_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is always implicit (even when it is not unique). Indeed, for a given set S of morphisms, we will see that two distinct S-adic sequences can have different properties depending on their respective directive words.

Observe that another classical tool to study sequences is topological entropy ($h = \limsup_{n\to+\infty} \frac{\log(p(n))}{n}$), also used in dynamical systems. However, as we are interested in very low complexity (subexponential) sequences, it is of no help.

The paper is organized as follows. Section 2 aims to present an overview of the complexity functions of morphic and S-adic sequences. Section 3 explores some necessary and some sufficient conditions on directive words that might lead to a correct statement of the S-adic conjecture. These conditions are illustrated through examples and counterexamples and expressed in terms of growth rate, return words, and composition rules of the morphisms and repetitions. A key tool to prove or disprove the linearity of the complexity of many examples is the deep result of Cassaigne characterizing the linearity of the complexity in terms of special factors (Theorem 3.7 on page 10). Section 4 goes further than the S-adic conjecture and proposes a new problem about the possible complexities of everywhere-growing S-adic sequences.

2 Comparison between morphic and S-adic sequences

The aim of this section is to compare morphic sequences with S-adic sequences. In particular, we show that the factor complexity of morphic sequences is rather restricted and can depend on combinatorial criteria, although this is not the case for S-adic sequences.

2.1 Morphic and pure morphic sequences

Purely morphic sequences correspond to S-adic sequences with $\mathrm{Card}(S) = 1$. If $S = \{\sigma\}$, we then have $\sigma_0\sigma_1\sigma_2\cdots = \sigma\sigma\sigma\cdots = \sigma^\omega$. In that case, the complexity functions that can occur were completely determined by Pansiot in [Pan84]. Indeed, he proved that for pure morphic sequences $\mathbf{w} = \sigma^\omega(a) = \lim_{n\to+\infty}\sigma^n(a^\omega)$ with σ *non-erasing* (i.e., $\sigma(b)$ is not the empty word for all letters b), the complexity function $p_{\mathbf{w}}(n)$ can have only five asymptotic behaviors that are[2] $\Theta(1)$, $\Theta(n)$, $\Theta(n\log n)$, $\Theta(n\log\log n)$ and $\Theta(n^2)$. Moreover, when the sequence $\mathbf{w}$ is aperiodic, Morse and Hedlund proved that its complexity function cannot be $\Theta(1)$ (see [MH38]). Pansiot proved that the class of complexity of the sequence only depends on the growth rate of the length of the images.

Definition 2.1. Recall that a morphism $\sigma : A^* \to A^*$ is said to be *everywhere-growing* if it does not admit *bounded letter*, i.e., letter b such that $\lim_{n\to+\infty}|\sigma^n(b)| < +\infty$. We let $A_{\mathfrak{B},\sigma}$ (or $A_{\mathfrak{B}}$ when no confusion is possible) denote the set of bounded letters of σ. By opposition, a non-bounded letter is called a *growing* letter. Since for all letters a, we have $|\sigma^n(a)| \in \Theta(n^{\alpha_a}\beta_a^n)$ for some α_a in $\mathbb{N}$ and $\beta_a \geq 1$ (see [RS80]), any everywhere-growing morphism satisfies exactly one of the following three definitions:

1. a morphism $\sigma : A^* \to A^*$ is *quasi-uniform* if there exists $\beta \geq 1$ such that for all letters $a \in A$, $|\sigma^n(a)| \in \Theta(\beta^n)$;

2. a morphism $\sigma : A^* \to A^*$ is *polynomially diverging* if there exists $\beta > 1$ and a function $\alpha : A \to \mathbb{N}$, $\alpha \neq 0$, such that for all letters $a \in A$, $|\sigma^n(a)| \in \Theta(n^{\alpha(a)}\beta^n)$;

3. a morphism $\sigma : A^* \to A^*$ is *exponentially diverging* if there exist $a_1, a_2 \in A$, $\alpha_1, \alpha_2 \in \mathbb{N}$ and $\beta_1, \beta_2 > 1$ with $\beta_1 \neq \beta_2$ such that for each $i \in \{1, 2\}$, $|\sigma^n(a_i)| \in \Theta(n^{\alpha_i}\beta_i^n)$.

Theorem 2.2 (Pansiot [Pan84])**.** *Let $\mathbf{w} = \sigma^\omega(a)$ be a pure morphic sequence with σ non-erasing.*

1. If σ is everywhere-growing and

i. quasi-uniform, then[3] $p_{\mathbf{w}}(n) \in \mathcal{O}(n)$;

ii. polynomially diverging, then $p_{\mathbf{w}}(n) \in \Theta(n\log\log n)$;

iii. exponentially diverging, then $p_{\mathbf{w}}(n) \in \Theta(n\log n)$.

2. If σ is not everywhere-growing and if there are infinitely many factors of $\mathbf{w}$ in $A_{\mathfrak{B}}^*$, then $p_{\mathbf{w}}(n) = \Theta(n^2)$.

3. If σ is not everywhere-growing and if there are only finitely many factors of $\mathbf{w}$ in $A_{\mathfrak{B}}^*$, then there exists a pure morphic sequence $\tau^\omega(b)$ with $\tau : B^* \to B^*$ everywhere-growing and a non-erasing morphism $\lambda : B^* \to A^*$ such that $\mathbf{w} = \lambda(\tau^\omega(b))$. In this case, we have $p_{\mathbf{w}}(n) \in \Theta(p_{\tau^\omega(b)}(n))$.

[2] $f(n) \in \Theta(g(n))$ if $\exists C_1, C_2 > 0, n_0\ \forall n \geq n_0\ |C_1 g(n)| \leq |f(n)| \leq |C_2 g(n)|$.

[3] $f(n) \in \mathcal{O}(g(n))$ if $\exists C > 0, n_0\ \forall n \geq n_0\ |f(n)| \leq |Cg(n)|$.

Unfortunately, Theorem 2.2 only holds for non-erasing morphisms. However, the following result states that when the morphism is erasing, the obtained pure morphic sequence is a *morphic sequence* (i.e., an image under a morphism of a pure morphic sequence) with non-erasing morphisms. The result is due to Cobham [Cob68] and was recovered later by Pansiot [Pan83]. It can also be found in Cassaigne and Nicolas's survey [CN03].

Theorem 2.3 (Cobham [Cob68] and Pansiot [Pan83])**.** *If* $\mathbf{w}$ *is a morphic sequence, it is the image under a letter-to-letter morphism of a pure morphic word* $\sigma^\omega(a)$ *with* σ *a non-erasing morphism.*

Theorems 2.2 and 2.3 show that to compute the complexity function of a pure morphic sequence, it is sometimes necessary to view it as a morphic sequence. It is therefore natural to be interested in the complexity function of such sequences. By definition, any pure morphic sequence is morphic. The converse has been known to be false since at least 1980, when Berstel proved that the Arshon word is morphic but not pure morphic [Ber80]. Other examples can be found in the literature; for instance, the result by Séébold showing that the unique (up to letter permutation) binary overlap-free word which is a fixed point of a morphism is the Thue-Morse word [Séé85]).

Moreover, it is not only the case that the class of morphic sequences strictly contains the class of pure morphic sequences, but also the asymptotic behavior of their complexity functions are different. Indeed, Example 2.4 shows that complexity classes given by Pansiot are no longer sufficient.

Example 2.4 (Deviatov [Dev08])**.** Let $\mathbf{w}$ be the morphic sequence $\tau(\sigma^\omega(0))$ where σ and τ are defined by

$$\sigma : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 12 \\ 2 \mapsto 23 \\ 3 \mapsto 3 \end{cases} \quad \text{and} \quad \tau : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \\ 2 \mapsto 2 \\ 3 \mapsto 2 \end{cases}$$

We have $p_{\mathbf{w}} \in \Theta(n\sqrt{n})$.

Other examples can be found in [Pan85]. Indeed, for all $k \geq 1$, Pansiot explicitly built a morphic sequence $\mathbf{w}$ whose complexity function satisfies $p_{\mathbf{w}}(n) \in \Theta(n\sqrt[k]{n})$. However, these behaviors $\Theta(n\sqrt[k]{n})$ seem to be the only new behaviors with respect to pure morphic sequences. Indeed, in [Dev08] Deviatov proved the next result and conjectured an equivalent result of Pansiot's theorem (Theorem 2.2) for morphic sequences.

Theorem 2.5 (Deviatov [Dev08])**.** *Let* $\mathbf{w}$ *be a morphic sequence. Then, either* $p_{\mathbf{w}}(n) \in \Theta(n^{1+\frac{1}{k}})$ *for some* $k \in \mathbb{N}^*$*, or* $p_{\mathbf{w}}(n) \in \mathcal{O}(n \log n)$*.*

Conjecture 2.6 (Deviatov [Dev08])**.** The complexity function of any morphic sequence only adopts one of the following asymptotic behaviors: $\Theta(1)$, $\Theta(n)$, $\Theta(n \log\log n)$, $\Theta(n \log n)$, $\Theta(n^{1+\frac{1}{k}})$ for some $k \in \mathbb{N}$.

In particular, Theorem 2.5 implies that the highest complexity that one can get is the same for morphic sequences and for pure morphic sequences. This can be explained by the following result.

Proposition 2.7 (Cassaigne and Nicolas [CN03]). *Let* $\mathbf{w}$ *be a one-sided sequence over* A *and* $\sigma : A^* \to B^*$ *be a non-erasing morphism. If* $M = \max_{a \in A} |\sigma(a)|$, *for all* n *we have* $p_{\sigma(\mathbf{w})}(n) \leq M p_{\mathbf{w}}(n)$. *Moreover, if* $\mathbf{w}$ *is pure morphic and* σ *is injective, then* $p_{\sigma(\mathbf{w})}(n) \in \Theta(p_{\mathbf{w}}(n))$.

It has been pointed out by a referee that Proposition 2.7 could be improved by replacing the injectivity hypothesis with the following one: *there exists* K *such that for all* $u \in \mathrm{Fac}(\sigma(\mathbf{w}))$, *we have* $\mathrm{Card}\left(\sigma^{-1}\left(\{u\}\right) \cap L(\mathbf{w})\right) \leq K$. Furthermore, the comparison between the respective complexity of $\mathbf{w}$ and $\sigma(\mathbf{w})$ seems to depend on the lack of injectivity of σ.

The previous discussion shows that the factor complexity of morphic sequences is rather constrained. To conclude this section, we give some examples of how some additional combinatorial criteria can restrict it even further.

A well-known fact is that if a pure morphic sequence is *k-power-free*, i.e., it does not contains any factor of the form u^k, then its factor complexity grows at least linearly and at most like $n \log n$ (see [ER83]). We will consider a similar criterion for S-adic sequences in Section 3.5.

Another such criterion is the *uniform recurrence* of the sequence, i.e., any factor occurs infinitely often and with bounded gaps in $\mathbf{w}$. For morphic sequences, this implies that the complexity is linear (see [NP09]). Actually, the uniform recurrence of a morphic sequence is even equivalent to its *linear recurrence*, i.e., any factor u occurs infinitely often and with gaps bounded by $K|u|$ (see [Dur98, Dur]). Furthermore, if $\mathbf{w}$ is a morphic and uniformly recurrent sequence over A, then $\mathbf{w}$ is a morphic sequence $\tau(\sigma^\omega(a))$ with σ a primitive morphism (this result was already proved in [Dur98] in the particular case of morphic sequences $\psi(\varphi^\omega(a))$ with ψ non-erasing and φ everywhere-growing). The following result also provides an algorithm to check whether a pure morphic sequence is uniformly recurrent.

Theorem 2.8 (Damanik and Lenz [DL06]). *A pure morphic sequence* $\mathbf{w} = \sigma^\omega(a)$ *is uniformly recurrent if and only if there is a growing letter* $b \in A$ *that occurs with bounded gaps in* $\mathbf{w}$ *and such that for all letters* $c \in A$ *there is a power* σ^k *such that* c *occurs in* $\sigma^k(b)$.

2.2 S-adic sequences

The previous section shows that the factor complexity of morphic sequences is rather restricted (especially for pure morphic sequences). In particular, the class of uniformly recurrent morphic sequences is strictly contained in the class of morphic sequences with linear complexity. In this section, we show that, for S-adic sequences, the state of affairs is quite different. The first important result is the following.

Proposition 2.9 (Cassaigne [Fog11]). *Let* A *be an alphabet and* $l \notin A$. *There exists a finite set* S *of morphisms over* $A' = A \cup \{l\}$ *such that any sequence over* A *is* S*-adic.*

In particular, this implies that any function which is the complexity function of some sequence is also the complexity function of some S-adic sequence. (although for morphic sequences, the complexity is at most quadratic). Moreover, the following proposition implies that the set of possible asymptotic behaviors for the complexity function of S-adic sequences is uncountable (although it is finite for pure morphic sequences and countable for morphic ones since the set of morphic sequences is countable). See also [MM10] for another approach to build sequences whose complexity is closed to a given function.

Proposition 2.10 (Cassaigne [Cas03]). *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a function such that*

i. $\lim_{t\to+\infty} \frac{f(t)}{\log t} = +\infty$;

ii. f *is differentiable, except possibly at* 0;

iii. $\lim_{t\to+\infty} f'(t)t^\beta = 0$ *for some* $\beta > 0$;

iv. f' *is decreasing.*

Then there exists a uniformly recurrent sequence $\mathbf{w}$ *over* $\{0,1\}$ *such that*[4] $\log(p_{\mathbf{w}}(n)) \sim f(n)$.

In particular, the function $f(n)$ in the previous proposition can be taken equal to n^α for any α with $0 < \alpha < 1$.

Another big difference is that the class of uniformly recurrent S-adic sequences with linear complexity forms a very small subset of the class of uniformly recurrent S-adic sequences. Indeed, recall that the *topological entropy* of a sequence over an alphabet A is the real number h with $0 \le h \le \log(\mathrm{Card}(A))$ defined by

$$h = \limsup_{n\to\infty} \frac{\log(p(n))}{n}.$$

A uniformly recurrent sequence $\mathbf{w}$ over an alphabet A with at least two letters $a, b \in A$ cannot have maximal complexity ($p_{\mathbf{w}}(n) = \mathrm{Card}(A)^n$): since all powers a^n occur in $\mathbf{w}$, there are unbounded gaps between two successive occurrences of b in $\mathbf{w}$. However, together with Proposition 2.9, the following result shows that, except the maximal one $p_{\mathbf{w}}(n) = \mathrm{Card}(A)^n$, any high complexity can be attained by uniformly recurrent S-adic sequences.

Theorem 2.11 (Grillenberger [Gri73]). *Let A be an alphabet with $d = \mathrm{Card}(A) \ge 2$ and $h \in [0, \log(d)[$. There exists a uniformly recurrent one-sided sequence $\mathbf{w}$ over A with topological entropy h.*

The proof of previous theorem is furthermore constructive and Ferenczi and Monteil noticed [BR10, Chap. 7] that this construction is S-adic with an infinite set S.

The following result provides an S-adic characterization of uniformly recurrent sequences. A sequence is said to be *primitive S-adic (with constant s_0)* if $s_0 \in \mathbb{N}$ is such that for all $r \in \mathbb{N}$, all letters in A_r occur in all images $\sigma_r \cdots \sigma_{r+s_0}(a)$, $a \in A_{r+s_0+1}$. It is said to be *weakly primitive S-adic* if for all $r \in \mathbb{N}$, there exists $s > 0$ such that all letters in A_r occur in all

[4] $f(n) \sim g(n)$ if $\forall \varepsilon > 0\ \exists n_0\ \forall n > n_0\ |f(n)/g(n) - 1| < \varepsilon$.

images $\sigma_r \cdots \sigma_{r+s}(a)$, $a \in A_{r+s+1}$. Any primitive S-adic sequence is weakly primitive and Durand [Dur03] proved that any weakly primitive S-adic sequence is uniformly recurrent. Following a construction based on return words (see Section 3.4 for the definition), one can prove that the converse is true. Recall that a sequence is said to be *proper S-adic* if all morphisms in its directive word are *proper*, i.e., for all n there are letters $a, b \in A_n$ such that $\sigma_n(A_{n+1}) \subset aA_n^*b$. Then we have

Theorem 2.12 (Durand [Dur03], Leroy [Ler12a]). *A sequence $\mathbf{w}$ is uniformly recurrent if and only if it is weakly primitive S-adic.*

Moreover, the directive word can be chosen to be proper and if $\mathbf{w}$ does not have linear complexity, then S is infinite.

The proof is similar to the proof of the following result, which gives an S-adic characterization of linearly recurrent sequences.

Theorem 2.13 (Durand [Dur03]). *A sequence $\mathbf{w}$ is linearly recurrent if and only if it is primitive and proper S-adic with* $\mathrm{Card}(S) < +\infty$.

The next example shows that the linear recurrence property is an even stronger condition than primitivity for S-adic sequences (contrary to what holds in the morphic case).

Example 2.14 (Durand [Dur03]). Let $S = \{\sigma, \tau\}$ where σ and τ are defined by

$$\sigma : \begin{cases} 0 \mapsto 021 \\ 1 \mapsto 101 \\ 2 \mapsto 212 \end{cases} \quad \text{and} \quad \tau : \begin{cases} 0 \mapsto 012 \\ 1 \mapsto 021 \\ 2 \mapsto 002 \end{cases}$$

The sequence

$$\mathbf{w} = \lim_{n \to +\infty} \sigma\tau\sigma^2\tau \cdots \sigma^n\tau(0^\omega)$$

is primitive S-adic but not linearly recurrent.

3 S-adicity and linear complexity

The aim of this section is to explore some ideas that one might have about the S-adic conjecture. First, we present some sufficient conditions for S-adic sequences to have linear complexity, and we show that we cannot make them weaker. Then, we give some counterexamples to some conditions that one might naturally believe to be sufficient to get linear complexity.

3.1 The growth rate of the length

Durand [Dur00, Dur03] gave some sufficient conditions for an S-adic sequence to have linear complexity. The main condition is the one given by the following result and is a generalization of what exists for pure morphic sequences (see Theorem 2.2). The other conditions are simply consequences of it.

Proposition 3.1 (Durand [Dur03]). *Let* $\mathbf{w}$ *be an* S*-adic sequence with* $\mathrm{Card}(S) < +\infty$ *and whose directive word is* $(\sigma_n)_{n\in\mathbb{N}}$ *with* $\sigma_n : A_{n+1}^* \to A_n^*$ *and* A_0 *the alphabet of* $\mathbf{w}$*. If there is a constant* D *such that for all* n*,*

$$\max_{a,b\in A_{n+1}} \frac{|\sigma_0 \cdots \sigma_n(a)|}{|\sigma_0 \cdots \sigma_n(b)|} \leq D, \tag{2}$$

then $p_{\mathbf{w}}(n) \leq D \max_{\sigma_n \in S, a\in A_{n+1}} |\sigma_n(a)|(\mathrm{Card}(A))^2 n$ *with* $A = \cup_{n\in\mathbb{N}} A_n$*.*

Corollary 3.2 (Durand [Dur03]). *If* $\mathbf{w}$ *is* S*-adic with* $\mathrm{Card}(S) < \infty$ *and all morphisms in* S *are uniform, then* $p_{\mathbf{w}}(n) \leq l(\mathrm{Card}(A))^2 n$ *with* $A = \cup_{n\in\mathbb{N}} A_n$ *and* $l = \max\limits_{\sigma\in S, a\in A(\sigma)} |\sigma(a)|$*.*

Proposition 3.3 (Durand [Dur00]). *If* $\mathbf{w}$ *is a primitive* S*-adic sequence with* $\mathrm{Card}(S) < +\infty$ *and constant* s_0 *directed by* $(\sigma_n)_{n\in\mathbb{N}}$ *with* $\sigma_n : A_{n+1}^* \to A_n^*$ *and* A_0 *the alphabet of* $\mathbf{w}$*, then there exists a constant* D *such that for all non-negative integers* r*,*

$$\max_{a,b\in A_{r+s_0+1}} \frac{|\sigma_r \cdots \sigma_{r+s_0}(a)|}{|\sigma_r \cdots \sigma_{r+s_0}(b)|} \leq D.$$

Corollary 3.4. *Let* S *be a set of non-erasing morphisms and* $\tau \in S$ *be strongly primitive (i.e., for all letters* a *of* $\tau(A)$*, the letter* a *occurs in all images* $\tau(b)$ *for* $b \in A$*). Any* S*-adic sequence for which* τ *occurs infinitely often with bounded gaps in the directive word is uniformly recurrent and has linear complexity.*

The everywhere-growing property can be naturally transposed to S-adic sequences. But, contrary to what holds in the pure morphic case, under that assumption, the condition given by Equation (2) is not equivalent to linear complexity. Indeed, even some Sturmian sequences do not satisfy it (those with unbounded coefficients $(k_n)_{n\in\mathbb{N}}$ in Example 1.2). One might therefore try to make that condition a little bit weaker.

We can observe in Example 1.2 that Condition (2) is still satisfied infinitely often. Indeed, let us consider the directive word $(\tau_n)_{n\in\mathbb{N}}$ defined by $\tau_0 = L_0^{k_0} R_0^{k_1} L_1$, $\tau_{2n} = L_0^{k_{4n}-1} R_0^{k_{4n+1}} L_1$ for $n \geq 1$ and $\tau_{2n+1} = L_1^{k_{4n+2}-1} R_1^{k_{4n+3}} L_0$ for $n \geq 0$. For all n, there exist integers i and j such that either $\tau_n = [0^i 1 0^{j+1}, 0^i 1 0^j]$ or $\tau_n = [1^i 0 1^j, 1^i 0 1^{j+1}]$. With these morphisms, we have

$$\tau_0 \tau_1 \cdots \tau_n \cdots = L_0^{k_0} R_0^{k_1} L_1^{k_2} R_1^{k_3} \cdots L_1^{k_{4n+2}} R_1^{k_{4n+3}} \cdots \tag{3}$$

and there is a constant K such that for all n, $\max_{a,b\in A_{n+1}} \frac{|\tau_0 \cdots \tau_n(a)|}{|\tau_0 \cdots \tau_n(b)|} \leq K$. The sequence $(\tau_n)_{n\in\mathbb{N}}$ is called a *contraction* of the directive word $(L_0^{k_0} R_0^{k_1} \cdots)$. Observe that the set $S = \{\tau_n \mid n \in \mathbb{N}\}$ of morphisms might be infinite (when $(k_n)_{n\in\mathbb{N}}$ is unbounded). Consequently, it may be interesting to work either with infinite sets of morphisms or with contractions.

However, Example 3.5 below shows that Proposition 3.1 is no longer true when $\mathrm{Card}(S) = \infty$. Indeed, if we consider the contraction $(\sigma_n)_{n\in\mathbb{N}}$ of the directive word of Proposition 3.6 defined for all $n \geq 0$ by

$$\sigma_n = \gamma^{k_n} \mu, \tag{4}$$

we have $|\sigma_0 \cdots \sigma_n(0)| = |\sigma_0 \cdots \sigma_n(1)|$ for all n, although the complexity is no longer linear when the sequence $(k_n)_{n\in\mathbb{N}}$ is unbounded.

Example 3.5. Let μ be the *Thue-Morse morphism* $[01, 10]$ and let γ be the morphism $[001, 1]$. From Theorem 2.2 we know that the non-uniformly recurrent sequence

$$\gamma^\omega(0) = 0010011^2 0010011^3 0010011^2 0010011^4 \cdots$$

has quadratic complexity.

Proposition 3.6. *Let $(k_n)_{n\in\mathbb{N}}$ be a sequence of non-negative integers. The sequence*

$$\mathbf{w}_{\gamma,\mu} = \lim_{n\to+\infty} \gamma^{k_0}\mu\gamma^{k_1}\mu\gamma^{k_2}\mu\cdots\gamma^{k_n}\mu(0^\omega)$$

is uniformly recurrent. Moreover, $\mathbf{w}_{\gamma,\mu}$ has at most linear complexity if and only if the sequence $(k_n)_{n\in\mathbb{N}}$ is bounded. Finally, for all n we have

$$|\gamma^{k_0}\mu\gamma^{k_1}\mu\gamma^{k_2}\mu\cdots\gamma^{k_n}\mu(0)| = |\gamma^{k_0}\mu\gamma^{k_1}\mu\gamma^{k_2}\mu\cdots\gamma^{k_n}\mu(1)|,$$

and denoting

$$\ell_n = |\gamma^{k_0}\mu\gamma^{k_1}\mu\gamma^{k_2}\mu\cdots\gamma^{k_n}\mu(0)|,$$

we have

$$p_{\mathbf{w}_{\gamma,\mu}}(\ell_n) \leq 3\ell_n.$$

Before proving the result, recall that a *right* (resp., *left*) *special factor* in a language $L \subset A^*$ is a factor such that there are at least two letters $a, b \in A$ for which ua and ub (resp., au and bu) belong to L. A *bispecial factor* is a factor which is both left and right special. This definition can be extended to words and sequences $\mathbf{w}$ by replacing L by the set of factors $\mathrm{Fac}(\mathbf{w})$. Let us recall the following result.

Theorem 3.7 (Cassaigne [Cas96])**.** *A sequence has linear complexity if and only if there is a constant K such that for all n, the number of right (resp., left) special factors of length n is less than K.*

Proof of Proposition 3.6. First, as μ occurs infinitely often in the directive word, $\mathbf{w}_{\gamma,\mu}$ is weakly primitive $\{\gamma, \mu\}$-adic and so, by Theorem 2.12, it is uniformly recurrent.

Now let us study the complexity depending on the sequence $(k_n)_{n\in\mathbb{N}}$. The case of a bounded sequence is a direct consequence of Corollary 3.4. Hence, let us assume that the sequence $(k_n)_{n\in\mathbb{N}}$ is unbounded and let us show that the complexity is not at most linear. Due to Theorem 3.7, we only have to prove that the number of right special factors of length n of $\mathbf{w}_{\gamma,\mu}$ is unbounded.

As already mentioned in Example 3.5, the fixed point $\gamma^\omega(0)$ has quadratic complexity. Consequently the number of right special factors of $\gamma^\omega(0)$ of a length n is unbounded.

We let the reader verify that bispecial factors of $\gamma^\omega(0)$ are the words ε, 0, 1 and the words $\gamma(1v)$ with v bispecial, and that other right special factors of $\gamma^\omega(0)$ are the suffixes of words $\gamma(v)$ for v right special. Thus, by induction, one can state that all right special factors of length n of $\gamma^\omega(0)$ occur in $\gamma^{n+1}(0)$.

Now let us show that if u is a right special factor in $\gamma^{k_{n+1}}(0)$, then $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(u)$ is a right special factor of $\mathbf{w}$ of length $|u|2^q$ with $q = \sum_{i=0}^{n}(k_i + 1)$. Indeed, as $\mu(0)$ and $\gamma(0)$

start with 0 and $\mu(1)$ and $\gamma(1)$ start with 1, the image of u is still a right special factor. Moreover, $\mu(u)$ contains exactly $|u|$ occurrences of the letter 0 and $|u|$ occurrences of the letter 1, and both γ and μ map a word with the same number of 0 and 1 to a word of double length with the same number of 0 and 1. Hence $|\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(u)| = |u|2^q$ with q defined as previously. Now, if u and v are two distinct right special factors of length m of $\gamma^\omega(0)$, then $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(u)$ and $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(v)$ are two distinct special factors of length $m2^q$ of $\mathbf{w}$. As the number of right special factors of a given length of $\gamma^\omega(0)$ is unbounded, and as the sequence $(k_n)_{n\in\mathbb{N}}$ is unbounded, the number of right special factors of a given length of $\mathbf{w}$ is also unbounded, which concludes the first part of the proof.

The last step is to show that, for all integers ℓ_n, we have $p_{\mathbf{w}_{\gamma,\mu}}(\ell_n) \leq 3\ell_n$. For all non-negative integers n, we already know that

$$|\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(0)| = |\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(1)| = \ell_n = 2^q.$$

with q as defined previously by $\sum_{i=0}^{n}(k_i+1)$. Consequently, all factors u of length ℓ_n are factors of $|\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(v)|$ for some word v of length 2. As there are only 4 possible binary words of length 2 and as there are no more than ℓ_n+1 distinct factors of length ℓ_n in a word of length $2\ell_n$, we obtain $p_{\mathbf{w}_{\gamma,\mu}}(\ell_n) \leq 4\ell_n + 4$. However, it is easily deduced from the following equations that the sets of factors of length ℓ_n that respectively occur in $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(00)$ and in $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(11)$ are equal.

$$\begin{aligned}
\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(00) &= \gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_{n-1}}\mu\left(\gamma^{k_n}(0)1\gamma^{k_n}(0)1\right)\\
\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(11) &= \gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_{n-1}}\mu\left(1\gamma^{k_n}(0)1\gamma^{k_n}(0)\right)
\end{aligned}$$

Consequently, we obtain $p_{\mathbf{w}_{\gamma,\mu}}(\ell_n) \leq 3\ell_n + 3$. But, among the $3\ell_n+3$ words, both words $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(0)$ and $\gamma^{k_0}\mu\gamma^{k_1}\mu\cdots\gamma^{k_n}\mu(1)$ have been counted several times (twice for one of them and four times for the other). Hence $p_{\mathbf{w}_{\gamma,\mu}}(\ell_n) \leq 3\ell_n - 1$. □

3.2 The condition C of the conjecture might not only concern the set S

Since it seems hard to make the condition of Equation (2) weaker, another idea is to determine new sufficient conditions that are independent of it. The first attempt in this direction was proposed by Boshernitzan. He asked whether the following holds: *if S contains only morphisms for which the fixed points (if it exists) of their powers have linear complexity, then any S-adic sequence has linear complexity.*

But Boshernitzan eventually provided the following counterexample to that statement. Since we did not find any detailed proof of it in the literature, we provide it below.

Example 3.8. Let γ and E be the morphisms over $\{0,1\}$ respectively defined by $[001,1]$ and $[1,0]$. Observe that both morphisms γE and $E\gamma$ are *primitive*. Consequently, they admit a power whose fixed points have linear complexity. We consider the sequence

$$\mathbf{w}_{\gamma,E} = \lim_{n\to+\infty} \gamma E\gamma^2 E\gamma^3 E\cdots\gamma^{n-1}E\gamma^n(0^\omega).$$

Proposition 3.9 (Boshernitzan)**.** *The sequence $\mathbf{w}_{\gamma,E}$ is S-adic for $S = \{\gamma E, E\gamma\}$. Further, it is uniformly recurrent and does not have linear complexity.*

Proof. First, by definition, $\mathbf{w}_{\gamma,E}$ is indeed S-adic for $S = \{\gamma E, E\gamma\}$.

Next, the composition $\gamma E\gamma$ is strongly primitive and occurs infinitely often in the directive word of $\mathbf{w}_{\gamma,E}$. It is therefore a consequence of Theorem 2.12 that $\mathbf{w}_{\gamma,E}$ is uniformly recurrent.

To prove that $\mathbf{w}_{\gamma,E}$ does not have linear complexity, by Theorem 3.7, it is sufficient to prove that the number of its right special factors of length n is unbounded. For this purpose, let us introduce notation. For all $k \in \mathbb{N}^*$, let us define the morphism $\Gamma_k = \gamma E\gamma^2 E \cdots \gamma^{k-1}E\gamma^k E$ and, for all $k \in \mathbb{N}$, the sequence

$$\mathbf{w}^{(k)} = \lim_{n\to+\infty} \gamma^{k+1}E\gamma^{k+2}E\cdots\gamma^{k+n-1}E\gamma^{k+n}(0^\omega).$$

We then have $\mathbf{w}_{\gamma,E} = \mathbf{w}^{(0)} = \Gamma_k(\mathbf{w}^{(k)})$, for all $k \geq 1$. For all $i \geq 1$ we also consider the word $u_i = \gamma^i(10) = 1\gamma^i(0)$. As 100 and 101 are factors of $E(\mathbf{w}^{(k+1)})$, and as 100 and 001 are factors of $\gamma^{k-j}\gamma(E(\mathbf{w}^{(k+1)}))$ for all j with $1 \leq j \leq k$ (observe also that $\gamma(100) = \gamma(10)001$ and $\gamma(001) = 00\gamma(10)1$), we can deduce that u_i is a right special factor of $\mathbf{w}^{(k)}$ for all i with $1 \leq i \leq k+1$. As words $\Gamma_k(0)$ and $\Gamma_k(1)$ start with different letters, for all integers i such that $1 \leq i \leq k+1$, the word $\Gamma_k(u_i)$ is a right special factor of $\mathbf{w}_{\gamma,E}$, and so are all of its suffixes.

For $1 \leq i \leq k+1$, the word u_i end with 1001^i, so that the longest common suffix between u_i and u_{i+1} is 1^i. It follows that the longest common suffix between $\Gamma_k(u_i)$ and $\Gamma_k(u_{i+1})$ is the word $p_k\Gamma_k(1^i)$ where $p_0 = \varepsilon$ and for $k \geq 1$,

$$p_k = 1\Gamma_1(1^2)\Gamma_2(1^3)\cdots\Gamma_{k-1}(1^k).$$

This is indeed a consequence of the fact (we let the reader verify it) that for any word u containing at least one occurrence of 0 and 1, the word $\Gamma^k(u)$ ends with ap_k where a is the last letter of u when k is even, and where a is the opposite letter to the last letter of u when k is odd.

From what precedes $\mathbf{w}_{\gamma,E}$ has at least $n(k)$ right special words of length $|p_k\Gamma_k(1^k)| + 1$, where $n(k)$ denotes the number of integers i between 1 and $k+1$ such that $|\Gamma(u_i)| > |p_k\Gamma_k(1^k)|$. Next fact allows to estimate $n(k)$.

Fact: With $f(k) = \frac{k^2+k+2}{2}$, for all $k \geq 2$,

1. for all $1 \leq i \leq k+1$, $|\Gamma_k(u_i)| = 2^i 2^{f(k)}$,

2. $|p_k\Gamma_k(1^k)| \leq k^2 2^{f(k)}$.

Proof of the fact. 1. Let i be between 1 and $k+1$. By induction, one can verify that $|\gamma^i(0)|_0 = 2^i$ and $|\gamma^i(0)|_1 = 2^i - 1$. As $u_i = 1\gamma^i(0)$, we have $|u_i|_0 = |u_i|_1 = 2^i$. Observe that for any word v such that $|v|_0 = |v|_1$, we have $|\gamma(v)| = 2|v|$ and $|\gamma(v)|_0 = |\gamma(v)|_1 = |v|$ and $|E(v)|_0 = |E(v)|_1 = |v|_0$.

Thus, as $\Gamma_k = \gamma E\gamma^2 E\cdots\gamma^{k-1}E\gamma^k E$, $|\Gamma_k(u_i)| = 2^{i+1}2^{\sum_{j=1}^k j} = 2^{i+1}2^{\frac{k(k+1)}{2}} = 2^i 2^{f(k)}$.

2. Before estimating $|p_k\Gamma_k(1^k)|$, we need an estimate of $|\Gamma_k(1)|$.

$$\begin{aligned}
|\Gamma_k(1)| &= \left|\gamma E\gamma^2 E\cdots\gamma^k E(1)\right| \\
&= \left|\gamma^k(0)\right|_0 \left|\gamma E\gamma^2 E\cdots\gamma^{k-1}E(0)\right| + \left|\gamma^k(0)\right|_1 \left|\gamma E\gamma^2 E\cdots\gamma^{k-1}E(1)\right| \\
&= 2^k\left(|\Gamma_{k-1}(01)| - |\Gamma_{k-1}(1)|\right) + \left(2^k - 1\right)|\Gamma_{k-1}(1)| \\
&= 2^{\frac{k^2+k+2}{2}} - |\Gamma_{k-1}(1)| < 2^{f(k)}.
\end{aligned}$$

Observe that $|\Gamma_k(1^k)| = k|\Gamma_k(1)|$. One can verify that, for all j with $1 \le j \le k$, $\Gamma_{j-1}(1^j) \le \Gamma_k(1^k)$. Moreover as $k \ge 2$, $|1\Gamma_1(1^2)| \le |\Gamma_k(1^k)|$.

Hence $|p_k\Gamma_k(1^k)| = |1\Gamma_1(1^2)| + \sum_{j=3}^{k} |\Gamma_{j-1}(1^j)| + |\Gamma_k(1^k)| < k|\Gamma_k(1^k)| < k^2 f(k)$. □

To end the proof of proposition 3.9, it suffices to notice that for all i such that $\log_2 k^2 < i \le k+1$, $|\Gamma_k(u_i)| > |p_k\Gamma_k(1^k)|$, so that $n(k) \ge k + 1 - \lceil \log_2 k^2 \rceil$. In other words, the number of right special words of $\mathbf{w}_{\gamma,E}$ of length $|p_k\Gamma_k(1^k)| + 1$ is unbounded, which shows by Theorem 3.7 that $\mathbf{w}_{\gamma,E}$ does not have linear complexity. □

Remark 3.10. The previous result is even stronger than just considering sets S of morphisms with fixed points of linear complexity. Indeed, the sequence also has *bounded partial quotients*, i.e., all morphisms occur with bounded gaps in the directive word (over $\{\gamma E, E\gamma\}$).

The opposite question to the one previously answered is to ask whether S-adic sequences can have linear complexity when S contains a morphism that admits a fixed point which does not have linear complexity. Example 3.5 positively answers that question, and we can prove even more. Indeed, the following example provides a uniformly recurrent S-adic sequence with linear complexity such that a "bad" morphism occurs with arbitrary high powers in its directive word.

Example 3.11. Let us consider the morphisms

$$\beta : \begin{cases} 0 \mapsto 010 \\ 1 \mapsto 1112 \\ 2 \mapsto 2 \end{cases} \qquad \text{and} \qquad M : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 1 \\ 2 \mapsto 1 \end{cases}$$

and the sequence

$$\mathbf{w}_{\beta,M} = \lim_{n\to+\infty} M\beta M\beta^2 M\beta^3 M \cdots \beta^{n-1} M\beta^n(0^\omega).$$

Note that β is not everywhere-growing; its set of bounded letters is $A_{\mathfrak{B},\beta} = \{2\}$ and all words in $A^*_{\mathfrak{B},\beta}$ are factors of $\beta^\omega(0)$. Hence by Theorem 2.2(2), $p_{\beta^\omega(0)}$ is in $\theta(n^2)$.

Proposition 3.12. *The sequence $\mathbf{w}_{\beta,M}$ defined just above has linear complexity. More precisely, for all n we have $p(n+1) - p(n) \in \{1, 2\}$.*

The proof of this proposition uses the link, proved by Cassaigne [Cas97], between bispecial factors and the first difference s of the complexity function. Given an infinite word $\mathbf{w}$, we recall that the *bilateral order* of a word u is the value

$$m(u) = \mathrm{Card}(\mathrm{Fac}(\mathbf{w}_{\beta,M}) \cap AuA) - \mathrm{Card}(\mathrm{Fac}(\mathbf{w}_{\beta,M}) \cap Au) - \mathrm{Card}(\mathrm{Fac}(\mathbf{w}_{\beta,M}) \cap uA) + 1.$$

When $m(u) > 0$, the word u is a bispecial factor of $\mathbf{w}$ called a *strong bispecial* factor, and when $m(u) < 0$, the word u is also a bispecial factor of $\mathbf{w}$ called a *weak bispecial* factor.

Let $s_{\mathbf{w}}$ (or simply s) be the function defined for $n \geq 0$ by $s(n) = p(n+1) - p(n)$. We have $s(0) = 1$, and Cassaigne proved (see [Cas97]) that

$$s(n+1) - s(n) = \sum_{u \in \mathrm{Fac}_n(\mathbf{w})} m(u) \tag{5}$$

In the next useful result, $\mathcal{B}_0$ is the identity morphism and, for $k > 0$, we have $\mathcal{B}_k = M\beta M\beta^2 \cdots M\beta^k$.

Lemma 3.13. *A word u is a strong bispecial factor of $\mathbf{w}_{\beta,M}$ if and only if $u = \mathcal{B}_k M\beta^i(1)$ for some $k \geq 0$ and some i in $\{0, \ldots, k\}$.*

A word u is a weak bispecial factor of $\mathbf{w}_{\beta,M}$ if and only if $u = \mathcal{B}_k M\beta^i(101)$ for some $k \geq 0$ and some i in $\{0, \ldots, k\}$.

Proof. For all integers $k \geq 0$, let $\mathbf{w}^{(k)}$ be the sequence $\mathbf{w}^{(k)}$ whose directive word is (M, β^{k+1}, M, $\beta^{k+2}, \ldots$). Note that, for $k \geq 0$, $\mathbf{w}_{\beta,M} = \mathcal{B}_k(\mathbf{w}^{(k)})$.

All words 010, 011, 110, 111 occur in $\mathbf{w}^{(k+1)}$ and all words 010, 011, 120, 111 occur in $\beta^j(\mathbf{w}^{(k+1)})$ for $j \geq 1$. Thus $M\beta^i(1)$ for i in $\{0, 1, \ldots, k+1\}$ are strong bispecial factors of $\mathbf{w}^{(k+1)}$ (observe that $M\beta^i(120) = M\beta^i(210)$ contains the factor $1M\beta^i(1)0$). As $\mathcal{B}_k(0)$ starts and ends with 0, and as $\mathcal{B}_k(1)$ starts and ends with 1, $\mathcal{B}_k M\beta^i(1)$ for $i \in \{0, 1, \ldots, k\}$ are strong bispecial factors of $\mathbf{w}_{\beta,M}$.

Now let $i \in \{0, 1, \ldots, k\}$. The words $0M\beta^i(101)1$ and $1M\beta^i(101)0$ are factors of $\mathbf{w}^{(k)}$ (respectively factors of $M\beta^{i+1}(01)1$ and $1M\beta^{i+1}(10)$). So $0\mathcal{B}_k(M\beta^i(101))1$ and $1\mathcal{B}_k(M\beta^i(101))0$ are factors of $\mathbf{w}_{\beta,M}$. One can verify that the existence of the factor $0\mathcal{B}_k(M\beta^i(101))0$ in $\mathbf{w}_{\beta,M}$ would imply the existence of factors $M(0\beta^j(101)0)$ in $\beta^{(k+1-j)}(\mathbf{w}^{(k+1)})$ for $0 \leq j \leq i$, and the existence of the factor 01010 in $\beta^{(k+1-i)}(\mathbf{w}^{(k+1)})$, which is impossible. Similarly the existence of the factor $1\mathcal{B}_k(M\beta^i(101))1$ in $\mathbf{w}_{\beta,M}$ would imply the existence of 11011 in $\beta^{(k+1-i)}(\mathbf{w}^{(k+1)})$, which is also impossible.

We now prove that there are no other strong and weak bispecial factors. Let u be a bispecial factor of $\mathbf{w}_{\beta,M}$. If u contains no occurrence of the letter 0, then $u = 1^n$ for some $n \geq 0$. As $\mathbf{w}_{\beta,M}$ contains arbitrarily large powers of 1 and infinitely many occurrences of 0, the words 11^n1, 01^n1, 11^n0 are all factors of $\mathbf{w}_{\beta,M}$. When 01^n0 is a factor of $\mathbf{w}_{\beta,M}$, one can verify that $1^n = \mathcal{B}_k M\beta^i(1)$ for some $k \geq 0$ and some integer i in $\{0, \ldots, k\}$. Thus 1^n is a strong bispecial factor of $\mathbf{w}_{\beta,M}$ if and only if $1^n = \mathcal{B}_k M\beta^i(1)$ for some $k \geq 0$ and some integer i in $\{0, \ldots, k\}$.

Note that any factor of $\mathbf{w}_{\beta,M}$ containing only one occurrence of 0 is of the form 01^n, 101^n, 1^n0 or 1^n01 for some $n \geq 0$, and is not bispecial.

Thus assume u contains at least two occurrences of 0. From the form of the morphisms $\mathcal{B}_k$, there is a unique integer k and a unique sequence of words

$$v, p_0, s_0, p_1, s_1, \ldots, p_{k-1}, s_{k-1}$$

over $\{0, 1\}$ such that

$$u = p_0\mathcal{B}_1(p_1)\mathcal{B}_2(p_2)\cdots\mathcal{B}_{k-1}(p_{k-1})\mathcal{B}_k(v)\mathcal{B}_{k-1}(s_{k-1})\cdots\mathcal{B}_2(s_2)\mathcal{B}_1(s_1)s_0,$$

where v is a proper factor of $M\beta^{k+1}(010)$ and for all i, $0 \le i \le k-1$, the word p_i (resp., s_i) is a suffix (resp., prefix) of $\mathcal{B}_{i+1}(0)$ or of $\mathcal{B}_{i+1}(1)$ but is neither equal to $\mathcal{B}_{i+1}(0)$ nor equal to $\mathcal{B}_{i+1}(1)$.

Since for all i, $\mathcal{B}_i(0)$ and $\mathcal{B}_i(1)$ do not have any common prefix or suffix, the factor u is bispecial if and only if all words $p_0, s_0, p_1, s_1, \ldots, p_{k-1}, s_{k-1}$ are empty words and v is bispecial in $\mathbf{w}^{(k)}$. Moreover, v contains an occurrence of the letter 0. We also have that u is strongly (resp., weakly) bispecial if and only if v is also. Then it can be easily verified that the only bispecial factors in $\mathbf{w}^{(k)}$ that contain an occurrence of 0 and that are factors of $M\beta^{k+1}(010)$ are the words $M\beta^i(101)$ for $i \in \{0, 1, \ldots, k\}$, which are weak bispecial factors. ☐

Proof of Proposition 3.12. Let $(u_n)_{n\ge 0}$ be the sequence of strong and weak bispecial factors of $\mathbf{w}_{\beta,M}$. Observe that, for all $k \ge 0$,

- $\mathcal{B}_k M\beta^{k+1}(1) = \mathcal{B}_{k+1}(1) = \mathcal{B}_{k+1} M\beta^0(1)$.
- $|\mathcal{B}_k(M\beta^i(1))| < |\mathcal{B}_k(M\beta^i(101))| < |\mathcal{B}_k(M\beta^{i+1}(1))|$, for all i in $\{0, 1, \ldots, k\}$.

Therefore, by Lemma 3.13, the word u_n is a strong bispecial factor of $\mathbf{w}_{\beta,M}$ if n is even, and it is a weak bispecial factor of $\mathbf{w}_{\beta,M}$ if n is odd. As for any factor u we have $m(u) = 1$ if it is strong bispecial, $m(u) = -1$ if it is weak bispecial, and $m(u) = 0$ otherwise, by Formula 5, for all $n \ge 0$, we have

- $s(|u_n| + 1) - s(|u_n|) = 1$ if n is even,
- $s(|u_n| + 1) - s(|u_n|) = -1$ if n is odd,
- $s(n + 1) - s(n) = 0$ if $n \notin \{|u_m| \mid m \ge 0\}$.

Consequently, since $s(0) = p(1) - p(0) = 1$, for all $n \ge 0$, we have $p(n + 1) - p(n) = s(n) \in \{1, 2\}$. ☐

3.3 Another (easier?) version of the S-adic conjecture

A natural idea to try to understand the S-adic conjecture is to consider examples composed of well-known morphisms.

Example 3.14. The Sturmian word which is the fixed point of the *Fibonacci morphism* $\varphi = [01, 0]$ and the Thue-Morse word which is the fixed point of the *Thue-Morse morphism* $\mu = [01, 10]$ both have linear complexity.

Proposition 3.15. *If $S = \{\varphi, \mu\}$, where φ and μ are defined above, any S-adic sequence is linearly recurrent.*

Proof. Let $S = \{\mu, \varphi\}$. Let $\mathbf{w}$ be an S-adic sequence directed by $(\sigma_n)_{n\in\mathbb{N}}$ and, for all $k \in \mathbb{N}$, let $\mathbf{w}^{(k)}$ be the S-adic sequence directed by $(\sigma_n)_{n\ge k}$.

By the definition of φ and μ, as $\mathbf{w}^{(k)} \in \{\varphi(\mathbf{w}^{(k+1)}), \mu(\mathbf{w}^{(k+1)})\}$, the word 111 does not occur in $\mathbf{w}^{(k)}$ for all $k \ge 1$. Note also that 000 does not occur in $\mu(\{0,1\}^*)$, and 0000 does not occur in $\varphi(\mathbf{w}^{(k+1)})$, since otherwise 111 would occur in $\mathbf{w}^{(k+1)}$. Thus, for all $k \ge 0$, neither 0000 nor 111 occurs in $\mathbf{w}^{(k)}$. This implies that the gap between two occurrences of

01 and 10 is at most 5 (7 is the maximal length of a shortest word of the form $10u10$ or of the form $01u01$).

Now observe that 11 occurs in $\mathbf{w}^{(k)}$ only in the image of $\sigma_k(10)$ (and only if $\sigma_k = \mu$). This implies that the gap between two consecutive occurrences of 11 is at most 12.

Finally observe that 00 occurs in $\mathbf{w}^{(k)}$ only in the image of $\sigma_1(10)$ or in the image of $\sigma_k(11)$ (when $\sigma_k = \varphi$). As any factor of $\mathbf{w}^{(k)}$ that starts with 10 or 11, ends with 10 or 11, and does not contain any other occurrences of 10 or 11 has length at most 7, the gap between two consecutive occurrences of 00 is at most 12.

We have just proved that the length of the largest gap between two occurrences of a word of length 2 in all words $\mathbf{w}^{(k)}$ is bounded. By the choice of S, we see that $\mathbf{w}$ is primitive S-adic. Hence by a result of Durand [Dur03], we know that $\mathbf{w}$ is linearly recurrent. □

Let us say that a set of morphisms S is *good-adic* if all S-adic sequences have linear complexity. Proposition 3.15 provides an example of such a set. Many other examples are known. For example:

- any singleton $\{f\}$ with f a morphism with fixed points of linear complexity;
- the set of Sturmian morphisms $\{[01, 1], [10, 1], [0, 01], [0, 10]\}$;
- the set of episturmian morphisms over an alphabet A: $S = \{L_a, R_a \mid a \in A\}$ with for all $a \in A$, $L_a(a) = R_a(a) = a$, $L_a(b) = ab$ and $R_a(b) = ab$ for $b \neq a$;
- any finite set S that contains only uniform morphisms (see Corollary 3.2);
- any finite set S that contains only strongly primitive morphisms (see Corollary 3.4).

Note that if S is good-adic then, for any morphism f in S admitting an infinite fixed point, this fixed point must have linear complexity. But this necessary condition is certainly not the only one.

Question 3.16. *What are good-adic sets of morphisms?*

3.4 About return words

If u is a factor of a sequence $\mathbf{w}$, a *(left) return word* to u is a word r such that ru is a factor of $\mathbf{w}$ having u as a prefix and containing only two occurrences of u. Return words to u in $\mathbf{w}$ actually correspond to the gaps between two consecutive occurrences of u in $\mathbf{w}$.

Durand [Dur98] proved that primitive morphic sequences, i.e., sequences defined as the image, under a morphism, of a fixed point of a primitive morphism, can be characterised using return words. Hence, it is quite natural to ask whether such a result exists for S-adic sequences with linear complexity.

There exist many examples of S-adic sequences $\mathbf{w}$ with linear complexity for which there exists an integer K such that all factors of $\mathbf{w}$ have at most K return words. For instance, Vuillon [Vui01] proved that Sturmian sequences are exactly the sequences whose factors have exactly two return words. Justin and Vuillon [JV00] extended this result to the family of Arnoux-Rauzy sequences (also called *strict episturmian sequences*) whose factors have exactly m return words, with m the cardinality of the alphabet. Balková, Pelantová and Steiner [BPS08] characterized sequences that do not contain any weak bispecial factors and whose factors have the same number of return words by showing they all have linear complexity. Improving a result by Ferenczi [Fer96], Leroy [Ler12a] proved that sequences whose complexity satisfies $1 \leq p(n+1)-p(n) \leq 2$ for all sufficiently large n are S-adic and infinitely many of their factors admit two or three return words.

Despite the previous example, the next proposition shows that there exist S-adic sequences with linear complexity such that any long factor has many return words ("many" depending on "long").

Example 3.17. Let us consider the three morphisms p, σ and s defined by

$$p : \begin{cases} a_1 \mapsto 0 \\ a_2 \mapsto 1 \\ b_1 \mapsto 1 \\ b_2 \mapsto 0 \end{cases} \qquad \sigma : \begin{cases} a_1 \mapsto a_1 a_2 a_1 \\ a_2 \mapsto a_2 a_2 a_2 \\ b_1 \mapsto b_1 b_2 b_1 \\ b_2 \mapsto b_2 b_2 b_2 \end{cases} \qquad s : \begin{cases} 0 \mapsto a_1 \\ 1 \mapsto b_1 \end{cases}$$

For all $n \geq 1$ we let π_n denote the morphism $p\sigma^n s$. Then we have

$$\pi_n : \begin{cases} 0 \mapsto 01^{\mathbf{e}_0}01^{\mathbf{e}_1}01^{\mathbf{e}_2} \cdots = \left(\prod_{i=0}^{2^{n-1}-1} 01^{\mathbf{e}_i}\right) 0 \\ 1 \mapsto 10^{\mathbf{e}_0}10^{\mathbf{e}_1}10^{\mathbf{e}_2} \cdots = \left(\prod_{i=0}^{2^{n-1}-1} 10^{\mathbf{e}_i}\right) 1 \end{cases},$$

where $\mathbf{e}$ is the fixed point of the morphism Exp defined over the infinite alphabet $\{3^n \mid n \in \mathbb{N}\}$ by

$$\forall n \in \mathbb{N}, \quad \mathrm{Exp}(3^n) = 13^{n+1},$$

i.e.,

$$\mathbf{e} = 1313^2 1313^3 1313^2 1313^4 1313^2 1313^3 1313^2 13 \cdots .$$

Now let us consider the sequence

$$\mathbf{w}_\pi = \lim_{n \to +\infty} \pi_1 \pi_2 \cdots \pi_n(0^\omega).$$

Proposition 3.18. *The sequence $\mathbf{w}_\pi$ defined above is uniformly recurrent, has linear complexity and for all integers k, there is a length ℓ_k such that all factors of $\mathbf{w}_\pi$ of length at least ℓ_k have at least k return words in $\mathbf{w}_\pi$.*

Proof. Uniform recurrence is a direct consequence of Theorem 2.12. Since $\mathbf{w}$ is S-adic with $S = \{p, \sigma, s\}$, it is a consequence of Corollary 3.2 that $\mathbf{w}_\pi$ has linear complexity.

Let us prove that all sufficiently long factors of $\mathbf{w}_\pi$ have many return words. For all $k \geq 1$, we let $\mathbf{w}^{(k)}$ denote the sequence

$$\mathbf{w}^{(k)} = \lim_{n \to \infty} \pi_k \pi_{k+1} \cdots \pi_n(0^\omega).$$

We obviously have $\mathbf{w}^{(1)} = \mathbf{w}_\pi$ and for all $k \geq 1$, $\mathbf{w}_\pi = \pi_1 \cdots \pi_k(\mathbf{w}^{(k+1)})$. We also have $|\pi_k(0)| = |\pi_k(1)| = 3^k$ for all k and we let ℓ_k denote the length

$$|\pi_1 \pi_2 \cdots \pi_k(0)| = |\pi_1 \pi_2 \cdots \pi_k(1)| = 3^{\frac{k(k+1)}{2}}.$$

Due to the form of the morphisms π_1, any factor u of $\mathbf{w}_\pi$ of length at least equal to 24 can be uniquely decomposed into images $\pi_1(0)$ and $\pi_1(1)$, i.e., there is a unique word v in $\mathbf{w}^{(2)}$ such that $u \in \mathrm{Fac}(\pi_1(v))$ and such that if $u \in \mathrm{Fac}(\pi_1(v'))$, then $v \in \mathrm{Fac}(v')$. Indeed, given such a factor u, either it contains an occurrence of 00 or of 11 (which uniquely determine how to decompose the factor), or it is equal to $\pi_1(10101010)$ or to $\pi_1(01010101)$ (neither 101010101, nor 010101010 occur in $\mathbf{w}^{(2)}$).

Similarly, we see that, for all $k \geq 2$, any factor of $\mathbf{w}^{(k)}$ of length at least equal to 7 can be uniquely decomposed into $\pi_k(0)$ and $\pi_k(1)$.

Let n be a positive integer; suppose it is large. The sequence $(l_k)_{k\geq 1}$ is increasing, so there is a unique positive integer k such that $l_{k-1} \leq n < l_k$. Consequently, all factors of length n of $\mathbf{w}_\pi$ belong to $\mathrm{Fac}\left(\pi_1 \cdots \pi_k\left(\{0,1\}^2\right)\right)$. From what precedes, there is a unique word v in $\mathrm{Fac}\left(\pi_k\left(\{0,1\}^2\right)\right) \subset \mathrm{Fac}\left(\mathbf{w}^{(k)}\right)$ such that $u \in \mathrm{Fac}(\pi_1 \cdots \pi_{k-1}(v))$ and any word v' over $\{0,1\}$ such that $u \in \mathrm{Fac}\left(\pi_1 \cdots \pi_{k-1}(v')\right)$ contains v as a factor.

By the uniqueness of v, the number of return words to u in $\mathbf{w}_\pi$ is at least equal to the number of return words to v in $\mathbf{w}^{(k)}$ (some return words might also occur in $\pi_1 \cdots \pi_{k-1}(v)$), so we only have to show that the number of return words to v in $\mathbf{w}_k$ is at least linear in k. This will prove the result since k increases with n.

First, if $|v| \geq 7$, we have seen that there is a unique word $x \in \{0,1,00,01,10,11\}$ such that $v \in \mathrm{Fac}(\pi_k(x))$ and such that if $v \in \mathrm{Fac}(\pi_k(y))$, then $x \in \mathrm{Fac}(y)$. The number of return words to v in $\mathbf{w}^{(k)}$ is at least equal to the number of return words to x in $\mathbf{w}^{(k+1)}$, and the reader can verify that it is at least linear in k.

Let us suppose that $|v| < 7$ and that the factor 000 occurs in v (the case $111 \in \mathrm{Fac}(v)$ is similar). Then, from the form of π_k, the word v only occurs in $\mathbf{w}^{(k)}$ as factor of $\pi_k(1)$. The number of return words to v in $\mathbf{w}^{(k)}$ is then at least equal to the number of return words to 1 in $\mathbf{w}^{(k+1)}$, and the reader can verify that it is at least linear in k.

Now suppose that neither 000 nor 111 occur in v (still with $|v| < 7$). Since the number of return words to a factor is the same as the number of return words to the smallest bispecial factor containing it, we only have to count the number of return words to bispecial factors of $\mathbf{w}^{(k)}$ that are smaller than 7 and that do not contain either 000 or 111 as factors. The reader can now verify that these are given by

$$\{0,1,00,01,10,11,010,101,0101,1010,01010,01011,11010,10101,010101,101010\}.$$

The bispecial factors 01010, 10101, 010101 and 101010 only occur in $\mathbf{w}^{(k)}$ as factors of $\pi_k(01)$ and $\pi_k(01)$ but are neither factors of $\pi_k(0)$ nor of $\pi_k(1)$. Since the number of words filling the gaps between occurrences of 01 and 10 in $\mathbf{w}^{(k+1)}$ is linear in k, these bispecial factors have a number of return words that is linear in k, too.

The reader can now verify that the other bispecial factors have a number of return words in $\mathbf{w}^{(k)}$ which is linear in k (each return word containing a different highest power of 0 or of 1 depending on the bispecial factor). This completes the proof. $\square$

To end with return words, let us observe that it is possible to build sequences whose complexity is not linear and in which infinitely many factors have a bounded number of return words. Indeed, the sequence $\gamma^\omega(0)$ (see Example 3.5) has quadratic complexity and any factor $\gamma^n(1)$ admits two return words. After the last observation, it seems difficult to characterize sequences with linear complexity using return words. Nevertheless, the next interesting question is open.

Question 3.19. *Let* $\mathbf{w}$ *be a sequence such that, for some integer* K*, all factors of* $\mathbf{w}$ *have at most* K *return words. Is it true that* $\mathbf{w}$ *is* S*-adic for some suitable set* S *and has linear complexity?*

3.5 Number of distinct high powers

As already mentioned, it is known that the factor complexity of k-power-free pure morphic sequences grows at most like $n \log n$. In this section, we explore links between bounding the number of distinct exponents of factors instead of the maximal exponent of factors and the factor complexity.

Let us introduce the following notation: given a language L and a word $u = u_1 \cdots u_{|u|} \in L$, $\mathrm{Pow}(u, L)$ is the set of all non-negative integers i such that there exist words p and s such that $pu^i s$ belongs to L, $|p| \leq |u|$, $|s| \leq |u|$, p is not a suffix of u, and s is not a prefix of u.

It is certainly well known that for any Sturmian sequence $\mathbf{w}$ and any factor u of $\mathbf{w}$, we have $\mathrm{Card}\,(\mathrm{Pow}(u, \mathrm{Fac}(\mathbf{w}_\pi))) = \{1, 2\}$. Indeed if $\mathrm{Card}\,(\mathrm{Pow}(u, \mathrm{Fac}(\mathbf{w}_\pi))) \geq 3$ for a factor u of a Sturmian sequence $\mathbf{w}$, any factor of length $|u|$ not in $\mathrm{Fac}(u^\omega)$ (such a factor exists since $p_{\mathbf{w}}(|u|) = |u| + 1$)) would have at least 3 return words, contradicting the result of Vuillon [Vui01] that all factors of a Sturmian sequence have two return words.

Example 3.20 (Example 3.17 continued)**.** With the notation of previous section, we have $\mathrm{Pow}(\pi_1 \cdots \pi_n(0), \mathrm{Fac}(\mathbf{w}_\pi)) = \{3^0, 3^1, \ldots, 3^n\}$, which proves the existence of an S-adic sequence with at most linear complexity that has factors with an unbounded number of distinct exponents.

Example 3.21 (Example 3.5 continued)**.** The previous phenomenon also holds for S-adic sequences that do not have linear complexity. Indeed, this is the case of the S-adic sequence considered in Proposition 3.6 when the sequence $(k_n)_{n \in \mathbb{N}}$ is unbounded, as one can observe that

$$\mathrm{Pow}(\gamma^{k_0}\mu\gamma^{k_1}\mu\gamma^{k_2}\mu \cdots \gamma^{k_n}\mu(1), \mathrm{Fac}(\mathbf{w}_{\gamma,\mu})) = \{1, 2, \ldots, k_{n+1} + 2\}.$$

It seems difficult from the previous discussion to get an S-adic characterization of sequences with linear complexity using the number of distinct powers of their factors. Nevertheless, inspired by the word $\mathbf{w}_{\gamma,\mu}$, Corollary 3.24 provides a new criterion for proving that a sequence does not have linear complexity. Recall that a nonempty word u is *primitive* if it is not a power of a smaller word v, i.e., there is no integer $k \geq 2$ such that $u = v^k$.

Lemma 3.22. *Let* $\mathbf{w}$ *be a recurrent sequence over* A*. If* $u \in \mathrm{Fac}(\mathbf{w})$ *is a primitive word such that there exist positive integers* k *and* $i < k/2$ *with* $k - i \in \mathrm{Pow}(u, \mathrm{Fac}(\mathbf{w}))$*, then there is a subset* F_i *of factors of length* $k|u|$ *of* $\mathbf{w}$ *such that* $\mathrm{Card}(F_i) = i|u|$*. Moreover, for* $i \neq i'$ *we have* $F_i \cap F_{i'} = \emptyset$*.*

Before proving the result, let us recall the following classical result (see, for instance, [Lot97, Prop. 1.3.2]).

Proposition 3.23. *If x and y are two nonempty words such that $xy = yx$, then there is a word v shorter than or the same length as x and y such that $x = v^n$ and $y = v^m$ for some positive integers n and m.*

Proof. (of Lemma 3.22) By hypothesis the word u^{k-i} is a factor of $\mathbf{w}$ and occurs infinitely many times in $\mathbf{w}$. Let us study the return words to it in $\mathbf{w}$.

Since u is a primitive word, a return word r to it is either u or has length $|r| > (k-i)|u|$. Indeed, if not, one of the following holds:

1. $|r| = j|u|$ with $2 \leq j \leq k-i$; in this case, we have $r = u^j$ which is not a return word to u because it contains more than two occurrences of u;

2. $|r| < (k-i)|u|$ and $|r|$ is not a multiple of $|u|$; in this case, there is a nonempty word u' smaller than u such that $u'u = uu'$ and, by Proposition 3.23, u is not a primitive word.

Now, since $k-i$ belongs to $\mathrm{Pow}(u, \mathrm{Fac}(\mathbf{w}))$ and $\mathbf{w}$ is recurrent, there are two words p_i and s_i such that $p_i u^{k-i} s_i$ belongs to $\mathrm{Fac}(\mathbf{w})$, $|p_i| \leq |u|$, $|s_i| \leq |u|$ and u does not admit p_i as a suffix or s_i as a prefix. Thus, from what precedes, p_i is a suffix of a return word r_i to u^{k-i} with $|r_i| > (k-i)|u|$ and such that $r_i u^{k-i} s_i$ is a factor of the recurrent word $\mathbf{w}$.

By hypothesis, we have $k - i > k/2$, so we get

$$|r_i u^{k-i}| > k|u|.$$

Let v_i denote the suffix of length $k|u|$ of $r_i u^{k-i}$ and let x_i denote a factor of $\mathbf{w}$ of length $(k+i)|u|$ that admits $v_i s_i$ as a prefix. For a given word $w = w_1 \cdots w_{|w|}$, we let $w[r:s]$ denote its factor $w_r \cdots w_{s-1}$. By definition of return words, the word u^{k-i} occurs only once in v_i. Thus all the words

$$x_i[\ell : \ell + k|u|], \quad \ell \in \{1, 2, \ldots, i|u|\},$$

are different. We take $F_i = \{x_i[\ell : \ell + k|u|] \mid \ell \in \{1, 2, \ldots, i|u|\}\}$.

Let us show that if i and j are such that $1 \leq i < j < k/2$, then $F_i \cap F_j = \emptyset$, i.e., all words

$$\begin{aligned} x_i[\ell : \ell + k|u|], &\quad \ell \in \{1, 2, \ldots, i|u|\} \\ x_j[m : m + k|u|], &\quad m \in \{1, 2, \ldots, j|u|\} \end{aligned}$$

are distinct. Suppose that on the contrary there are some integers $\ell \in \{1, 2, \ldots, i|u|\}$ and $m \in \{1, 2, \ldots, j|u|\}$ such that

$$x_i[\ell : \ell + k|u|] = x_j[m : m + k|u|] = X.$$

Since $\ell \leq i|u|$ and $m \leq j|u|$, there are some words α, β, γ and δ such that

$$X = \alpha u^{k-i} \beta = \gamma u^{k-j} \delta.$$

We have $|\alpha| > |\gamma|$; otherwise u^{k-j} would occur twice in v_j. Let us show that we also have $|\alpha| < |\gamma| + (k-j)|u| - |u|$. As $k - i > k - j > k/2$, we have

$$(k-i)|u| \geq (k-j+1)|u| > k|u|/2 + |u|.$$

Thus, if $|\alpha| \geq |\gamma| + (k-j)|u| - |u|$, we have

$$\begin{aligned} |\alpha| + (k-i)|u| &\geq |\gamma| + (k-j)|u| - |u| + (k-j+1)|u| \\ &> |\gamma| + k|u|, \end{aligned}$$

which is a contradiction because $|\alpha u^{k-i}\beta| = k|u|$.

We have just proved that

$$|\gamma| < |\alpha| < |\gamma| + (k-j)|u| - |u| = |\gamma u^{k-j-1}|.$$

This implies that, if p is the prefix of length $(|\alpha| - |\gamma|) \mod |u|$ of u, we have $up = pu$. Since $|\alpha| - |\gamma| \neq 0 \mod |u|$ (otherwise s_j would be a prefix of u), Proposition 3.23 implies that u is not primitive, a contradiction. Thus we have $F_i \cap F_j = \emptyset$. $\square$

Corollary 3.24. *Let* $\mathbf{w}$ *be a recurrent sequence over* A. *If*

$$\sup_{\substack{u \in \mathrm{Fac}(\mathbf{w}),\\ u \text{ primitive word},\\ k \geq 1}} \sum_{\substack{0<i<k/2,\\ k-i \in \mathrm{Pow}(u,\mathrm{Fac}(\mathbf{w}))}} \frac{i}{k} = +\infty, \tag{6}$$

then $\mathbf{w}$ *does not have linear complexity.*

Proof. Let u be a primitive word in $\mathrm{Fac}(\mathbf{w})$ and let i and k be integers as in Lemma 3.22. This lemma permits us to obtain a lower bound on the number of factors of length $k|u|$. Indeed, we immediately deduce that

$$p_{\mathbf{w}}(k|u|) \geq \sum_{\substack{0<i<k/2,\\ k-i \in \mathrm{Pow}(u,\mathrm{Fac}(\mathbf{w}))}} i|u| = |u| \sum_{\substack{0<i<k/2,\\ k-i \in \mathrm{Pow}(u,\mathrm{Fac}(\mathbf{w}))}} i.$$

But, as we have

$$\sup_{\substack{u \in \mathrm{Fac}(\mathbf{w}),\\ u \text{ primitive word},\\ k \geq 1}} \frac{1}{k} \sum_{\substack{0<i<k/2,\\ k-i \in \mathrm{Pow}(u,\mathrm{Fac}(\mathbf{w}))}} i = +\infty,$$

the sequence $\mathbf{w}$ does not have linear complexity. $\square$

Remark 3.25. The hypothesis of primitivity of the words u in previous results cannot be avoided. Indeed, consider the fixed point $\mathbf{w} = \sigma^\omega(0)$ with $\sigma(0) = 010$ and $\sigma(1) = 111$. The sequence $(1^{3n})_{n\in\mathbb{N}}$ satisfies (6) (without primitivity) but $p_{\mathbf{w}}(n) = 2n - 1$ for all $n \geq 2$.

The next example shows that the converse of Corollary 3.24 does not hold.

Example 3.26. It is well known that the Thue-Morse sequence $\mathbf{t} \in \{0,1\}^{\mathbb{N}}$ is cubefree [Thu12], i.e., it does not contain any factor of the form xxx. It is also strongly recurrent [Sal10], which means that the direct product of it with any uniformly recurrent sequence is uniformly recurrent. Let $\mathbf{w}$ be a uniformly recurrent sequence over $\{0,1\}$ with exponential complexity (such a word exists from Theorem 2.11). The direct product $\mathbf{t}\times\mathbf{w}$ is then a uniformly recurrent sequence over the four-letter alphabet $\{(0,0),(0,1),(1,0),(1,1)\}$ which has exponential complexity and which is cubefree (because so is $\mathbf{t}$). Thus we have

$$\sup_{\substack{u\in\mathrm{Fac}(\mathbf{t}\times\mathbf{w}),\\ u \text{ primitive word},\\ k\geq 1}} \sum_{\substack{0<i<k/2,\\ k-i\in\mathrm{Pow}(u,\mathrm{Fac}(\mathbf{t}\times\mathbf{w}))}} \frac{i}{k} = 0.$$

Despite the fact that the converse of Corollary 3.24 does not hold, Corollary 3.28 below provides a family of sequences satisfying (6) and whose complexities are very close to linear. First, we need the following result, which is obtained by a simple adaptation of Durand's result [Dur03, Prop. 2.1] (which is equivalent to Proposition 3.1 on page 9).

Proposition 3.27. *Let $\mathbf{w}$ be an everywhere-growing S-adic sequence whose directive word is $(\sigma_n)_{n\in\mathbb{N}}$ with $\sigma_n : A_{n+1}^* \to A_n^*$ and A_0 the alphabet of $\mathbf{w}$. If there is a constant K such that for all n we have $\mathrm{Card}(A_n)\leq K$, and if there is a function $D:\mathbb{N}\to\mathbb{R}$ such that for all n and all $a\in A_{n+2}$, $b\in A_{n+1}$ we have*

$$|\sigma_0\cdots\sigma_{n+1}(a)| \leq D(n)|\sigma_0\cdots\sigma_n(b)|, \tag{7}$$

then for all m, $p_{\mathbf{w}}(m)\leq K^2 D(n(m))m$, where $n(m)$ is the greatest integer such that

$$\min_{a\in A_{n(m)+1}} |\sigma_0\cdots\sigma_{n(m)}(a)| \leq m.$$

Proof. Let $m\geq 1$. Since the sequence $\left(\inf_{a\in A_{k+1}}|\sigma_0\cdots\sigma_k(a)|\right)_{k\in\mathbb{N}}$ is non-decreasing and tends to $+\infty$, there is a greatest integer $n(m)$ such that

$$\inf_{a\in A_{n(m)+1}} |\sigma_0\cdots\sigma_{n(m)}(a)| \leq m \leq \inf_{a\in A_{n(m)+2}} |\sigma_0\cdots\sigma_{n(m)+1}(a)|.$$

Consequently, for any word u in $\mathrm{Fac}_m(\mathbf{w})$, there exist letters b and c in $A_{n(m)+2}$ and a positive integer $i\leq|\sigma_0\cdots\sigma_{n(m)+1}(b)|$ such that $u=\sigma_0\cdots\sigma_{n(m)+1}(bc)[i:i+m]$. It comes

$$\begin{aligned} p_{\mathbf{w}}(m) &\leq K^2 \sup_{a\in A_{n(m)+2}} |\sigma_0\cdots\sigma_{n(m)+1}(a)| \\ &\leq K^2 D(n(m)) \inf_{a\in A_{n(m)+1}} |\sigma_0\cdots\sigma_{n(m)}(a)| \\ &\leq K^2 D(n(m))m. \end{aligned}$$

□

The idea of the next result is to build a directive word $(\sigma_n)_{n\in\mathbb{N}}$ satisfying (7) with slowly increasing functions $D(n)$ and $n(m)$ while producing more and more powers with morphisms σ_n.

Corollary 3.28. *Let $(\phi(n))_{n\geq 0}$ be a non-decreasing sequence of integers greater or equal to 2 such that $\lim_n \phi(n)=+\infty$. There exists a non-periodic sequence $\mathbf{w}\in\{0,1\}^{\mathbb{N}}$ such that:*

i. $\mathbf{w}$ *satisfies* (6) *in Corollary 3.24;*

ii. $p_{\mathbf{w}}(n) \leq 4\phi(\log_2(n))n$ *for all* n*;*

iii. $p_{\mathbf{w}}(n) \leq 8n$ *for infinitely many* n*.*

Proof. Let μ be the Thue-Morse morphism (see Example 3.5 on Page 10). For all integers $n \geq 2$, let τ_n be the uniform morphism over $\{0,1\}$ defined by:

$$\tau_n : \begin{cases} 0 \mapsto 0u0 \\ 1 \mapsto 1v1 \end{cases},$$

where u and v are the prefixes of length $n-2$ of $\prod_{i\geq 1} 1^i 0$ and $\prod_{i\geq 1} 0^i 1$ respectively. For instance, we have

$$\tau_9 : \begin{cases} 0 \mapsto 010110110 \\ 1 \mapsto 101001001 \end{cases} \qquad \tau_{10} : \begin{cases} 0 \mapsto 0101101110 \\ 1 \mapsto 1010010001 \end{cases}.$$

Let $(\sigma_n)_{n\in\mathbb{N}}$ be the sequence of morphisms defined, for all $n \geq 0$, by

$$\sigma_{2n} = \mu \qquad \text{and} \qquad \sigma_{2n+1} = \tau_{\phi(n)}.$$

By construction, the sequence $\mathbf{w} = \lim_{n\to+\infty} \sigma_0 \cdots \sigma_n(0)$ satisfies the hypothesis of Corollary 3.24, so it does not have linear complexity.

Let us give a general upper bound for $p_{\mathbf{w}}(m)$, using Proposition 3.27 and its terminology $D(n)$ and $n(m)$. For all n, we have $|\sigma_{n+1}(0)| = |\sigma_{n+1}(1)| \leq \phi(\lfloor n/2 \rfloor) \leq \phi(n)$. Thus we can choose $D(n) \leq \phi(n)$. For all m we have $|\sigma_0 \cdots \sigma_m(0)| = |\sigma_0 \cdots \sigma_m(1)| \geq 2^m$. Thus $n(m) \leq \log_2(m)$ and we obtain

$$p_{\mathbf{w}}(m) \leq 4\phi(\log_2(m))m.$$

Now let us prove that $p_{\mathbf{w}}(m) \leq 8m$ for infinitely many m. Consider $m = |\sigma_0\sigma_1 \cdots \sigma_{2k+1}(0)|$. We have $n(m) = 2k+1$ and we can set $D(2k) = |\sigma_{2k}(0)| = |\mu(0)| = 2$; hence

$$p_{\mathbf{w}}(m) \leq 8m.$$

□

4 Beyond linearity

Until now, we have provided several examples showing that various natural approaches to characterize S-adic sequences that have linear complexity do not appear to be promising. To conclude this paper, we raise a new problem related to S-adicity and, more precisely, to everywhere-growing S-adic sequences.

For pure morphic sequences, the complexity function can have only 5 asymptotic behaviors and only depends on the growth rate of images (see Theorem 2.2). For S-adic sequences, we have seen in Section 2.2 that things are more complicated. However, Ferenczi [Fer96] (see also [Ler12b, LR]) proved that any uniformly recurrent sequence with linear complexity is everywhere-growing S-adic with S finite. This is a kind of generalization of the third point of Theorem 2.2. Moreover, one can check that all examples considered in previous sections (and more generally all S-adic representations of well-known families of sequences such as codings of rotations, codings of interval exchanges, etc.) are everywhere-growing with S finite. It is also interesting to note that for pure morphic sequences, the class of highest complexity $\Theta(n^2)$ can be reached only by morphisms with bounded letters (Theorem 2.2). A similar phenomenon occurs for S-adic sequences. Indeed, Cassaigne's constructions (Proposition 2.9) allow building S-adic sequences with arbitrarily high complexity and they admit several bounded letters. Consequently, the fact that the length of all images tends to infinity with n seems to be important to get a *reasonably low complexity*. This is confirmed by the following two lemmas. The first one can be found in a paper written by Boyle and Handelman [BH94] and certainly elsewhere. We let denote $h(\mathbf{w})$ the entropy of the sequence $\mathbf{w}$ (see Section 2.2).

Lemma 4.1. *Let L be a set of m finite words of length at least l and at most k. Let $\mathbf{w}$ be an infinite concatenation of elements of L. Then*

$$p_{\mathbf{w}}(n) \leq km^{2+\frac{n}{l}} \text{ and } h(\mathbf{w}) \leq \frac{\log m}{l}.$$

As a corollary we obtain the following lemma whose conclusion was certainly known in the 1970's in ergodic theory, for instance to prove that measure-theoretical dynamical systems of bounded rank have entropy zero (see [BKMS]).

Lemma 4.2. *Let $\mathbf{w}$ be an S-adic sequence whose directive word is $(\sigma_n)_{n\in\mathbb{N}}$ with $\sigma_n : A_{n+1}^* \to A_n^*$ and A_0 the alphabet of $\mathbf{w}$. Then, for all n such that $\min_{a\in A_{n+1}} |\sigma_0 \cdots \sigma_n(a)| \neq 0$,*

$$h(\mathbf{w}) \leq \frac{\log\left(\mathrm{Card}(A_{n+1})\right)}{\min_{a\in A_{n+1}} |\sigma_0 \cdots \sigma_n(a)|}.$$

Thus, if $(\mathrm{Card}(A_n))_{n\in\mathbb{N}}$ is bounded and $(\min_{a\in A_{n+1}} |\sigma_0 \cdots \sigma_n(a)|)_{n\in\mathbb{N}}$ tends to infinity, then $h(\mathbf{w}) = 0$.

Remark 4.3. The previous lemma shows that to build an S-adic sequence with a very high complexity (i.e., with positive entropy), either we need bounded letters, or we need the set S to be infinite and the alphabets A_{n+1} must grow exponentially in the minimal length of the images $\sigma_0 \cdots \sigma_n(a)$.

Moreover, the converse of previous lemma is not true: there exist sequences of zero entropy without S-adic representation satisfying the hypothesis. Indeed, as shown by Williams [Wil84] (also see [GJ00]) there exist (Toeplitz) sequences x with zero entropy whose generated subshift (X, S) has infinitely many ergodic measures. It is part of the folklore of ergodic theory that, in our context and for uniformly recurrent sequences, when $(|A_n|)_n$ is bounded by some constant k and $(\min_{a\in A_n} |\sigma_0 \cdots \sigma_n(a)|)_n$ tends to infinity, then the subshift generated by the S-adic sequence has at most k ergodic measures. Thus x has zero entropy but does not have an S-adic representation with $(|A_n|)_n$ bounded and $(\min_{a\in A_n} |\sigma_0 \cdots \sigma_n(a)|)_n$ going to infinity.

Observe that, given an S-adic sequence $\mathbf{w}$, the everywhere-growing property is not a necessary condition for $\mathbf{w}$ to have low complexity. Indeed, Cassaigne's constructions also hold for sequences with low complexity. One can also recall the *Chacon substitution* ϱ defined by $\varrho(0) = 0010$ and $\varrho(1) = 1$ whose fixed point $\varrho^\omega(0)$ has complexity $p(n) = 2n - 1$ for all n (see [Fer95]). However, *the existence* of an everywhere-growing S'-adic representation of $\mathbf{w}$ might be necessary (this is the case for the Chacon substitution and for any uniformly recurrent sequence with linear complexity).

That property is neither a sufficient condition since the sequence $\mathbf{w}_{\gamma,E}$ of Example 3.5 satisfies it and does not always have linear complexity. However, it seems natural to look for Pansiot-like results.

Question 4.4. *Let* $\mathbf{w}$ *be an everywhere-growing* S*-adic sequence directed by* $(\sigma_n)_{n\in\mathbb{N}}$ *such that the alphabets* A_n *have bounded cardinality. How can the ratio*[5] $\max_{a,b\in A_{n+1}} \frac{|\sigma_0\cdots\sigma_n(a)|}{|\sigma_0\cdots\sigma_n(b)|}$ *influence* $p_{\mathbf{w}}$*?*

This question seems to be a new and non-trivial problem. Proposition 4.5 below provides a partial answer to this question. Indeed, it deals with *expansive* S-adic sequences, i.e., with S-adic sequences such that for all morphisms σ in S and all letters a, we have $|\sigma(a)| \geq 2$. Leroy's proof [Ler12b] involves techniques similar to those used by Ehrenfeucht, Lee, and Rozenberg [ELR75] for D0L systems.

Proposition 4.5. *If* $\mathbf{w}$ *is an expansive* S*-adic sequence with* $\mathrm{Card}(S) < +\infty$*, then* $p_{\mathbf{w}}(n) \in \mathcal{O}(n \log n)$*.*

The next example shows that the bound of Proposition 4.5 is the best one we can obtain.

Example 4.6. Let ϑ be the morphism

$$\vartheta : \begin{cases} 0 \mapsto 0120 \\ 1 \mapsto 11 \\ 2 \mapsto 222 \end{cases}$$

and consider its fixed point $\mathbf{w}_\vartheta = \vartheta^\omega(0)$. It can be seen as an expansive $\{\vartheta\}$-adic sequence and, by Theorem 2.2, we have that $p_{\mathbf{w}_\vartheta}(n) \in \Theta(n \log n)$.

5 Acknowledgments

Many thanks to the two referees for their valuable comments.

References

[AB07] B. Adamczewski and Y. Bugeaud, On the complexity of algebraic numbers I: Expansions in integer bases, *Ann. Math.* **165** (2007), 547–565.

[5] Considering only the growth rate of $\min_{a\in A_{n+1}} |\sigma_0 \cdots \sigma_n(a)|$ is not enough since it would suffices to add occurrences of the identity morphism in $(\sigma_n)_{n\in\mathbb{N}}$ to change this growth rate.

[AB11] B. Adamczewski and Y. Bugeaud, Nombres réels de complexité sous-linéaire: mesures d'irrationalité et de transcendance, *J. Reine Angew. Math.* **658** (2011), 65–98.

[Abe03] A. Aberkane, Words whose complexity satisfies $\lim \frac{p(n)}{n} = 1$, *Theoret. Comput. Sci.* **307** (2003), 31–46.

[All94] J.-P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc. Simon Stevin* **1** (1994), 133–143.

[AR91] P. Arnoux and G. Rauzy, Représentation géométrique de suites de complexité $2n+1$, *Bull. Soc. Math. France* **119** (1991), 199–215.

[AZ98] J. P. Allouche and L. Q. Zamboni, Algebraic irrational binary numbers cannot be fixed points of non-trivial constant length or primitive morphisms, *J. Number Theory* **69** (1998), 119–124.

[Ber80] J. Berstel, Mots sans carré et morphismes itérés, *Discrete Math.* **29** (1980), 235–244.

[BH94] M. Boyle and D. Handelman, Entropy versus orbit equivalence for minimal homeomorphisms, *Pacific J. Math.* **164** (1994), 1–13.

[BHZ06] V. Berthé, C. Holton, L. Q. Zamboni, Initial powers of Sturmian words, *Acta Arith.* **122** (2006), 315–347.

[BKMS] S. Bezuglyi, J. Kwiatkowski, K. Medynets, and B. Solomyak, Finite rank Bratteli diagrams: structure of invariant measures, Preprint. Available at http://arxiv.org/abs/1003.2816.

[Bos84] M. Boshernitzan, A unique ergodicity of minimal symbolic flows with linear block growth, *J. Analyse Math.*, **44** (1984/85), 77–96.

[Bos92] M. Boshernitzan, A condition for unique ergodicity of minimal symbolic flows, *Ergodic Theory Dynam. Systems*, **12** (1992), 425–428.

[BPS08] L. Balková, E. Pelantová, and W. Steiner, Sequences with constant number of return words, *Monatsh. Math.* **155** (2008), 251–263.

[BR10] V. Berthé and M. Rigo, eds., *Combinatorics, Automata and Number Theory*, Vol. 135 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, 2010.

[Cas96] J. Cassaigne, Special factors of sequences with linear subword complexity, In *Developments in Language Theory, II*, World Sci. Publ., 1996, pp. 25–34.

[Cas97] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc. Simon Stevin* **4** (1997), 67–88.

[Cas03] J. Cassaigne, Constructing infinite words of intermediate complexity, in *Developments in Language Theory*, *Lect. Notes in Comput. Sci.*, Vol. 2450, Springer, 2003, pp. 173–184.

[CN03] J. Cassaigne and F. Nicolas, Quelques propriétés des mots substitutifs, *Bull. Belg. Math. Soc. Simon Stevin* **10** (2003), 661–676.

[Cob68] A. Cobham, On the Hartmanis-Stearns problem for a class of tag machines, in *Proc. of 9th Annual Symposium on Switching and Automata Theory*, IEEE Computer Society, 1968, pp. 51–60.

[CS01] V. Canterini and A. Siegel, Geometric representation of substitutions of Pisot type, *Trans. Amer. Math. Soc.* **353** (2001), 5121–5144.

[Dev08] R. Deviatov, On subword complexity of morphic sequences, in *Computer Science – Theory and Applications*, *Lect. Notes in Comput. Sci.*, Vol. 5010, Springer, 2008, pp. 146–157.

[DL06] D. Damanik and D. Lenz, Substitution dynamical systems: characterization of linear repetitivity and applications, *J. Math. Anal. Appl.* **321** (2006), 766–780.

[Dur] F. Durand, Decidability of uniform recurrence of morphic sequences, *Internat. J. Found. Comput. Sci.*, to appear.

[Dur98] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math.* **179** (1998), 89–101.

[Dur00] F. Durand, Linearly recurrent subshifts have a finite number of non-periodic subshift factors, *Ergodic Theory Dynam. Systems* **20** (2000), 1061–1078.

[Dur03] F. Durand, Corrigendum and addendum to: "Linearly recurrent subshifts have a finite number of non-periodic subshift factors", *Ergodic Theory Dynam. Systems* **23** (2003), 663–669.

[ELR75] A. Ehrenfeucht, K. P. Lee, and G. Rozenberg, Subword complexities of various classes of deterministic developmental languages without interactions, *Theoret. Comput. Sci.* **1** (1975), 59–75.

[ER83] A. Ehrenfeucht and G. Rozenberg, On the subword complexity of m-free D0L languages, *Inform. Process. Lett.* **17** (1983), 121–124.

[Fer95] S. Ferenczi, Les transformations de Chacon: combinatoire, structure géométrique, lien avec les systèmes de complexité $2n+1$, *Bull. Soc. Math. France* **123** (1995), 271–292.

[Fer96] S. Ferenczi, Rank and symbolic complexity, *Ergodic Theory Dynam. Systems* **16** (1996), 663–682.

[Fer99] S. Ferenczi, Complexity of sequences and dynamical systems, *Discrete Math.* **206** (1999), 145–154.

[FM97] S. Ferenczi and C. Mauduit, Transcendence of numbers with a low complexity expansion, *J. Number Theory* **67** (1997), 146–161.

[Fog02] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, *Lecture Notes in Mathematics*, Vol. 1794, Springer-Verlag, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit, and A. Siegel.

[Fog11] N. Pytheas Fogg, Terminologie S-adique et propriétés, Preprint available at http://tinyurl.com/8opdb8s, 2011.

[GJ00] R. Gjerde and Ø. Johansen, Bratteli-Vershik models for Cantor minimal systems: applications to Toeplitz flows, *Ergodic Theory Dynam. Systems*, **20** (2000), 1687–1710.

[GJ09] A. Glen and J. Justin, Episturmian words: a survey, *Theor. Inform. Appl.* **43** (2009), 403–442.

[Gri73] C. Grillenberger, Constructions of strictly ergodic systems. I. Given entropy, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **25** (1972/73), 323–334.

[HZ98] C. Holton and L.Q. Zamboni, Geometric realizations of substitutions, *Bull. Soc. Math. France* **126** (1998), 149–179.

[JV00] J. Justin and L. Vuillon, Return words in Sturmian and episturmian words, *RAIRO Theor. Inform. Appl.* **34** (2000), 343–356.

[Kos98] M. Koskas, Complexités de suites de Toeplitz, *Discrete Math.*, **183** (1998), 161–183.

[Ler12a] J. Leroy, *Contribution to the resolution of the S-adic conjecture*, PhD thesis, Université de Picardie Jules Verne, 2012.

[Ler12b] J. Leroy, Some improvements of the S-adic conjecture, *Adv. in Appl. Math.* **48** (2012), 79–98.

[LR] J. Leroy and G. Richomme, A combinatorial proof of S-adicity for sequences with sub-affine complexity, *Integers*, to appear.

[LR07] F. Levé, G. Richomme, Quasiperiodic Sturmian words and morphisms, *Theoret. Comput. Sci.* **372** (2007), 15–25.

[Lot97] M. Lothaire, *Combinatorics on Words*, Cambridge Mathematical Library, Cambridge University Press, 1997. (Corrected reprint of the 1983 original.)

[Lot02] M. Lothaire, *Algebraic Combinatorics on Words*, Vol. 90 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, 2002.

[MH38] M. Morse and G. A. Hedlund, Symbolic dynamics, *Amer. J. Math.* **60** (1938), 815–866.

[MH40] M. Morse and G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, *Amer. J. Math.* **62** (1940), 1–42.

[MM10] C. Mauduit and C. G. Moreira, Complexity of infinite sequences with zero entropy, *Acta Arith.* **142** (2010), 331–346.

[NP09] F. Nicolas and Y. Pritykin, On uniformly recurrent morphic sequences, *Internat. J. Found. Comput. Sci.* **20** (2009), 919–940.

[Pan83] J.-J. Pansiot, Hiérarchie et fermeture de certaines classes de tag-systèmes, *Acta Inform.* **20** (1983), 179–196.

[Pan84] J.-J. Pansiot, Complexité des facteurs des mots infinis engendrés par morphismes itérés, in *Automata, Languages and Programming*, *Lect. Notes in Comput. Sci.*, Vol. 172, Springer, 1984, pp. 380–389.

[Pan85] J.-J. Pansiot, Subword complexities and iteration, *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, No. 26 (1985), 55–62.

[Rau82] G. Rauzy, Nombres algébriques et substitutions, *Bull. Soc. Math. France* **110** (1982), 147–178.

[Rot94] G. Rote, Sequences with subword complexity $2n$, *J. Number Theory* **46** (1994), 196–213.

[RS80] G. Rozenberg and A. Salomaa, *The Mathematical Theory of L systems*, Vol. 90 of *Pure and Applied Mathematics*, Academic Press, 1980.

[Sal10] P. V. Salimov, On uniform recurrence of a direct product, *Discrete Math. Theor. Comput. Sci.* **12** (2010), 1–8.

[Séé85] P. Séébold, Sequences generated by infinitely iterated morphisms, *Discrete Appl. Math.* **11** (1985), 255–264.

[Thu12] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Kra. Vidensk. Selsk. Skrifter. I. Math. Nat. Kl.* **1** (1912), 1–67.

[Vui01] L. Vuillon, A characterization of Sturmian words by return words, *European J. Combin.* **22** (2001), 263–275.

[Wil84] S. Williams Toeplitz minimal flows which are not uniquely ergodic, *Z. Wahrsch. Verw. Gebiete*, **67** (1984), 95–107.

2010 *Mathematics Subject Classification*: Primary 68R15; Secondary 37B10.

Keywords: factor complexity, S-adicity, S-adic conjecture, morphism, subword complexity, linear complexity, uniform recurrence.

(Concerned with sequences A003849, A010060, and A049320.)

Received June 23 2012; revised version received October 9 2012. Published in *Journal of Integer Sequences*, March 2 2013.

Return to Journal of Integer Sequences home page.