Convergence and Factor Complexity for the Arnoux-Rauzy-Poincaré Algorithm

Valérie Berthé and Sebastien Labbé

LIAFA, Université Paris Diderot, Paris 7 - Case 7014, F-75205 Paris Cedex 13
{berthe,sebastien.labbe}@liafa.univ-paris-diderot.fr

Abstract. We introduce a multidimensional continued fraction algorithm based on Arnoux-Rauzy and Poincaré algorithms, and we study its associated S-adic system. An S-adic system is made of infinite words generated by the composition of infinite sequences of substitutions with values in a given finite set of substitutions, together with some restrictions concerning the allowed sequences of substitutions, expressed in terms of a regular language. We prove that these words have a factor complexity $p(n)$ with $\limsup p(n)/n < 3$, which provides a proof for the convergence of the associated algorithm by unique ergodicity.

1 Introduction

Given a vector of frequencies $(f_1, f_2, \cdots, f_d) \in \mathbb{R}^d_+$ (with $\sum f_i = 1$), our goal here is to propose a construction of an infinite word $\mathbf{w}$ over the alphabet $\mathcal{A} = \{1, 2, \cdots, d\}$ such that the frequency of each letter $i \in \mathcal{A}$ exists and is equal to f_i. We also would like the word $\mathbf{w}$ to have particular combinatorial properties, namely a linear factor complexity and a bounded balance. In dimension two, the question is completely answered. The Sturmian words form a well-known family of infinite balanced words having a linear factor complexity ($p(n) = n + 1$). But the situation is more contrasted in higher dimensions.

In [BL11, Lab12], we considered this question under the approach of multidimensional continued fraction algorithms and S-adic systems. Experimentations suggested that Brun multidimensional continued fraction algorithm as well as a fusion of Arnoux-Rauzy and Poincaré algorithms were the two best choices to investigate for such an approach. In this article, we focus on the Arnoux-Rauzy-Poincaré algorithm (a bit better than Brun experimentally), and construct an infinite word for Lebesgue almost each frequency vector $(f_1, f_2, f_3) \in \mathbb{R}^3_+$. We show that such words have a linear factor complexity, namely $p(n+1) - p(n) \in \{2, 3\}$ for all $n \geq 0$, by describing extensively the life of every bispecial factor, including strong and weak ones which come in pairs (as proved in Lemma 10 below).

More precisely, we introduce an S-adic system associated with a set of 9 substitutions. Three of them are substitutions known under the name of Arnoux-Rauzy substitutions [AR91], and the other six are named Poincaré substitutions after Poincaré algorithm [Nog95]. The execution of the Arnoux-Rauzy-Poincaré algorithm yields restrictions to the allowed infinite sequences of substitutions,

J. Karhumäki, A. Lepistö, and L. Zamboni (Eds.): WORDS 2013, LNCS 8079, pp. 71–82, 2013.

expressed in terms of a regular language. We show that we have a bijection (up to a set of zero measure) between infinite words and $\mathbb{R}^3_+$. We show that these words have a factor complexity $p(n)$ that satisfies with $\limsup p(n)/n < 3$. The proof relies on the fact that weak and strong bispecial factors are alternating in the sequence (ordered by increasing length) of non neutral bispecial factors. Then, by using a result of Boshernitzan [Bos85], we deduce the existence of (uniform) frequency of any factor, and thus of the letters. This provides a combinatorial proof of convergence for this multidimensional continued fraction algorithm.

The article is structured as follows. In Section 2, we introduce the Arnoux-Rauzy-Poincaré multidimensional continued fraction algorithm, with its nine associated substitutions, as well as our main result on the factor complexity and on the convergence. In Section 3, we study bispecial factors under Arnoux-Rauzy and Poincaré substitutions with no restriction on the application of substitutions. In Section 4, we prove the result on factor complexity of the associated S-adic system where the language of substitutions is restricted to a regular language defined by a finite automaton.

2 The Arnoux-Rauzy-Poincaré Algorithm

The Arnoux-Rauzy-Poincaré multidimensional continued fraction algorithm belongs to the family of multidimensional continued fraction algorithms defined in terms of triangle maps such as introduced in [Gar01]. It combines the two classical algorithms that are Poincaré algorithm and Arnoux-Rauzy algorithm, which are respectively defined in dimension 3 as follows: Poincaré algorithm acts on a triple of non-negative entries by subtracting the smallest entry to the median and the median to the largest, whereas Arnoux-Rauzy algorithm acts by subtracting the sum of the two smallest entries to the largest, when possible. Our fusion algorithm privilegiates an Arnoux-Rauzy step if possible, otherwise it perfoms a Poincaré step.

We follow here the formalism described in Section 2.1 of [DFG$^+$12]. The Arnoux-Rauzy-Poincaré multidimensional continued fraction algorithm is a fusion algorithm such as introduced in [BL11, Lab12]. It is defined on the 2-simplex

$$\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3_+ : x_1 + x_2 + x_3 = 1\}$$

whose vertices are the vectors $\mathbf{e}_1 = (1,0,0)^\top$, $\mathbf{e}_2 = (0,1,0)^\top$ and $\mathbf{e}_3 = (0,0,1)^\top$. In order to partition Δ, we consider the following fifteen matrices:

$$A_1 = \begin{pmatrix} 1&1&1\\0&1&0\\0&0&1 \end{pmatrix}, P_{21} = \begin{pmatrix} 1&1&1\\0&1&1\\0&0&1 \end{pmatrix}, P_{31} = \begin{pmatrix} 1&1&1\\0&1&0\\0&1&1 \end{pmatrix}, H_{21} = \begin{pmatrix} 1&0&0\\0&1&0\\1&0&1 \end{pmatrix}, H_{31} = \begin{pmatrix} 1&0&0\\1&1&0\\0&0&1 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 1&0&0\\1&1&1\\0&0&1 \end{pmatrix}, P_{12} = \begin{pmatrix} 1&0&1\\1&1&1\\0&0&1 \end{pmatrix}, P_{32} = \begin{pmatrix} 1&0&0\\1&1&1\\1&0&1 \end{pmatrix}, H_{12} = \begin{pmatrix} 1&0&0\\0&1&0\\0&1&1 \end{pmatrix}, H_{32} = \begin{pmatrix} 1&1&0\\0&1&0\\0&0&1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 1&0&0\\0&1&0\\1&1&1 \end{pmatrix}, P_{13} = \begin{pmatrix} 1&1&0\\0&1&0\\1&1&1 \end{pmatrix}, P_{23} = \begin{pmatrix} 1&0&0\\1&1&0\\1&1&1 \end{pmatrix}, H_{13} = \begin{pmatrix} 1&0&0\\0&1&1\\0&0&1 \end{pmatrix}, H_{23} = \begin{pmatrix} 1&0&1\\0&1&0\\0&0&1 \end{pmatrix},$$

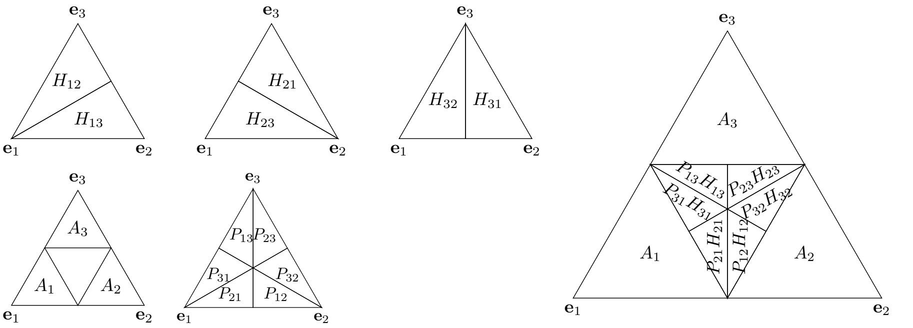

Fig. 1. Left: the three Arnoux-Rauzy matrices, the six Poincaré matrices and the six half triangles. Right: the partition of Arnoux-Rauzy-Poincaré algorithm.

whose column vectors are represented at Figure 1. Then, the column vectors of A_1, A_2, A_3, $P_{31}H_{31}$, $P_{13}H_{13}$, $P_{23}H_{23}$, $P_{32}H_{32}$, $P_{12}H_{12}$ and $P_{21}H_{21}$ describe a disjoint partition of Δ depicted in Figure 1. This partition then allows the definition of the following map:

$$T : \Delta \to \mathbb{R}^3_+$$
$$\mathbf{x} \mapsto \begin{cases} A_k^{-1}\mathbf{x}, & \text{if } \mathbf{x} \in A_k\Delta \text{ (subtract the sum of the two smallest entries} \\ & \text{to the largest),} \\ P_{jk}^{-1}\mathbf{x}, & \text{if } \mathbf{x} \in P_{jk}H_{jk}\Delta \text{ (subtract the smallest entry to the median} \\ & \text{and the median to the largest).} \end{cases}$$

The Arnoux-Rauzy-Poincaré multidimensional continued fractions algorithm is defined as the iteration of the function $\overline{T} : \Delta \to \Delta$, $\mathbf{x} \mapsto \dfrac{T(\mathbf{x})}{||T(\mathbf{x})||}$ with $||\mathbf{x}|| = x_1 + x_2 + x_3$. For each $\mathbf{x}$, it generates a sequence of matrices $(M_n)_n$ with values in the set $\{A_k,\ P_{jk} \mid j, k \in \{1,2,3\},\ j \neq k\}$.

2.1 The Arnoux-Rauzy-Poincaré S-Adic System

We recall below the definition of an S-adic system. For more on S-adic words see [BD13, CN10, DLR13, Ler12]. We say that the infinite word $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ admits an *S-adic representation* if there exist a finite set $\mathcal{S}$ of substitutions defined on the alphabet $\mathcal{A}$, a sequence $s = (\sigma_n)_{n\in\mathbb{N}} \in \mathcal{S}^{\mathbb{N}}$ of substitutions that all belong to $\mathcal{S}$, and $(a_n)_{n\in\mathbb{N}}$ a sequence of letters in $\mathcal{A}$ such that $\mathbf{w} = \lim_{n\to\infty} \sigma_0\sigma_1 \cdots \sigma_n(a_n)$. The word $\mathbf{w}$ is said to be S-adic, and the sequence s is called the *directive sequence*. An *S-adic system* is obtained by adding restrictions on the set of allowed directive sequences $s \in \mathcal{S}^{\mathbb{N}}$: an S-adic system is given by a finite directed strongly connected graph $\mathcal{G}$ labeled by the substitutions in S, with each infinite path giving rise to a directive sequence.

Here sequences of letters $(a_n)_n$ are constant sequences. Let i, j, k be such that $\{i, j, k\} = \{1, 2, 3\}$. A *Poincaré substitution* is a substitution of the form $\pi_{jk} : i \mapsto ijk, j \mapsto jk, k \mapsto k$. An *Arnoux-Rauzy substitution* is given by $\alpha_k : i \mapsto ik, j \mapsto jk, k \mapsto k$. Let

$$\mathcal{S} := \{\pi_{jk}, \alpha_k \mid j, k \in \{1, 2, 3\},\ j \neq k\}.$$

For each $\{i, j, k\} = \{1, 2, 3\}$, P_{jk} is the incidence matrix of the substitution π_{jk} and A_k is the incidence matrix of α_k.

The automaton $\mathcal{G} = (Q, \mathcal{S}, T, I)$ is defined by the states $Q = \{\Delta, H_{12}, H_{13}, H_{21}, H_{23}, H_{31}, H_{32}\}$, the transitions $T = \{(\Delta, \alpha_k, \Delta), (\Delta, \pi_{jk}, H_{jk}), (H_{jk}, \alpha_j, H_{jk}), (H_{jk}, \alpha_i, \Delta), (H_{jk}, \pi_{ij}, H_{ij}), (H_{jk}, \pi_{ki}, H_{ki}), (H_{jk}, \pi_{ji}, H_{ji}) :$ for each $\{i, j, k\} = \{1, 2, 3\}\} \subset Q \times \mathcal{S} \times Q$ and the initial state $I = \{\Delta\}$ (see Figure 2). We consider the S-adic system associated with the regular language $\mathcal{L}(\mathcal{G})$. This language corresponds to directive sequences $(s_n)_n$ for which the sequence of incidence matrices $(M_{s_n})_n$ is generated by the execution of the Arnoux-Rauzy-Poincaré algorithm.

Proposition 1 (ARP regular language). *The set of directive sequences produced by the Arnoux-Rauzy-Poincaré algorithm is exactly the set of labeled infinite paths starting in Δ in the graph $\mathcal{G}$ illustrated in Figure 2.*

We now state the main theorem. Its proof is given in Section 4. Let us say that $\mathbf{x} \in \Delta$ is *totally irrational* if x_1, x_2, x_3 are linearly independent over $\mathbb{Q}$.

Theorem 1 (Factor Complexity). *Let $\mathbf{w}$ be an S-adic word generated by the Arnoux-Rauzy-Poincaré algorithm applied to a totally irrational vector $\mathbf{x} \in \Delta$. Then the factor complexity of $\mathbf{w}$ is such that $p(n) \leq 3n+1$, $p(n+1) - p(n) \in \{2, 3\}$ for all $n \geq 0$, and $\limsup_{n\to\infty} \frac{p(n)}{n} < 3$.*

Theorem 2 (Frequencies and Convergence). *Let $\mathbf{w}$ be an S-adic word generated by the Arnoux-Rauzy-Poincaré algorithm applied to a totally irrational*

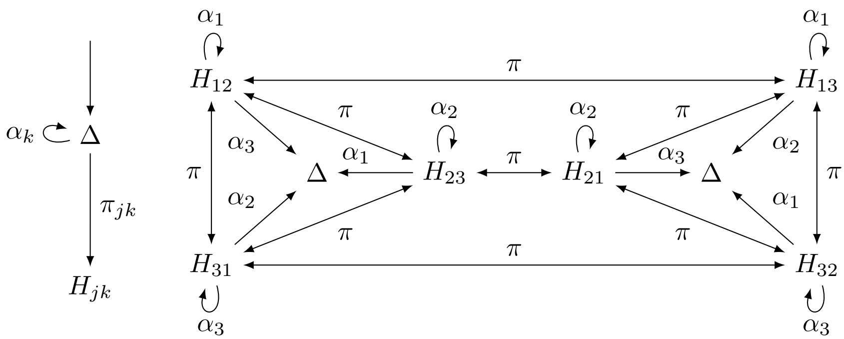

Fig. 2. The deterministic automaton $\mathcal{G}$. To avoid crossing arrows, the initial state Δ is drawn at three places. Also, the indices of π transitions are not written because they are determined by the indices of the arrival state: $\xrightarrow{\pi} H_{jk}$ means $\xrightarrow{\pi_{jk}} H_{jk}$.

vector $\mathbf{x} \in \Delta$. *Then the symbolic dynamical system generated by* $\mathbf{w}$ *is uniquely ergodic, and the frequencies of letters are proved to exist in* $\mathbf{w}$ *and to be equal to the coordinates of* $\mathbf{x}$.

Furthermore, the Arnoux-Rauzy-Poincaré algorithm is a weakly convergent algorithm, that is, for Lebesgue almost every $\mathbf{x} \in \Delta$, *if* $(M_n)_n$ *stands for the sequence of matrices produced by the Arnoux-Rauzy-Poincaré algorithm, then one has* $\cap_n M_0 \cdots M_n(\mathbb{R}^3_+) = \mathbb{R}_+\mathbf{x}$.

Theorem 2 is a direct consequence of Theorem 1 together with Theorem 1.5 of [Bos85] for the unique ergodicity statement (see also [FM10]). The weak convergence comes from the unique ergodicity. Usual proofs of convergence rely on linear algebra and on the use of the Hilbert projective metric (see e.g. [Sch00]). Let us stress the fact that we provide here a purely combinatorial proof of convergence for a multidimensional continued fraction algorithm.

3 Bispecial Factors under Arnoux-Rauzy and Poincaré Substitutions

3.1 Bispecial Factors and Extension Types

The proof of Theorem 4 requires some preparation. In this section, we follow the notation of [CN10]. Let w be a factor of a recurrent infinite word $\mathbf{u}$. We let $E^+(w) = \{x \in \mathcal{A} \mid wx \in \mathcal{L}(\mathbf{u})\}$ denote the set of right extensions of w in $\mathbf{u}$. The *right valence* $d^+(w) = \operatorname{Card} E^+(w)$ of w (in $\mathbf{u}$) is the number of distinct right extensions of w. *Left extensions* $E^-(w)$ and *left valence* $d^-(w)$ are defined similarly. A factor whose right valence is at least 2 is called *right special*. A factor whose left valence is at least 2 is called *left special*. A factor which is both left and right special is called *bispecial*. The set of bispecial factors of length n are identified by $BS_n(\mathbf{u})$. The *extension type* $E_{\mathbf{u}}(w)$ of a factor w of $\mathbf{u}$ is the set of pairs (a, b) of $\mathcal{A} \times \mathcal{A}$ such that w can be extended in both directions as awb :

$$E_{\mathbf{u}}(w) = \{(a, b) \in \mathcal{A} \times \mathcal{A} \mid awb \in \mathcal{L}(\mathbf{u})\}.$$

We let denote $E_{\mathbf{u}}(w)$ by $E(w)$ when the context is clear. The *bilateral multiplicity* of a factor w is the number

$$m(w) = \operatorname{Card} E(w) - d^-(w) - d^+(w) + 1.$$

A bispecial factor is said *strong* if $m(w) > 0$, *weak* if $m(w) < 0$ and *neutral* if $m(w) = 0$. A bispecial factor whose extension type satisfies

$$E(w) \subseteq (\{a\} \times \mathcal{A}) \cup (\mathcal{A} \times \{b\}) \qquad \text{for a pair of letters} \qquad (a, b) \in E(w) \quad (1)$$

is said *ordinary*. An ordinary bispecial factor is neutral, but the converse is not true for $|\mathcal{A}| > 2$. It is convenient to represent extension type $E(w)$ of a bispecial factor w graphically. Often represented as a bipartite graph, we choose a table representation: a cross $(\times)$ is drawn at the intersection of row a and column b if and only if $(a, b) \in E(w)$ (see Figure 3).

	1	2	3
1		×	
2		×	
3	×	×	×

$m(w) = 0$
neutral and ordinary

	1	2	3
1		×	
2			×
3	×	×	×

$m(w) = 0$
neutral but not ordinary

	1	2	3
1		×	
2			
3			×

$m(w) = -1$
weak

	1	2	3
1			
2		×	×
3	×	×	×

$m(w) = 1$
strong

Fig. 3. We represent the extension type $E(w)$ of a bispecial factor w by a table. A cross ($\times$) is at the intersection of row a and of column b if and only if $(a, b) \in E(w)$.

Definition 1 (Left equivalence). *Let w and w' be two bispecial factors defined on the alphabet $\mathcal{A}$. We say that their extension types are* left equivalent *if there exists a permutation τ acting on $\mathcal{A}$ such that $E(w') = \{(\tau(a), b) \mid (a, b) \in E(w)\}$.*

Right equivalence is defined similarly. Left equivalence can be interpreted on the table representation of the extension type as follows: one representation can be obtained from the other by a permutation of the rows:

$E(w) =$

	1	2	3
1			×
2			
3	×	×	×

and

$E(w') =$

	1	2	3
1	×	×	×
2			×
3			

Substitutions considered in this article preserve the first letter and thus preserve the right extensions. Then, the notion of left-equivalence is sufficient for our need. When the extension type of two words are equivalent, they share common properties. In particular, being ordinary, strong or weak is preserved under equivalence.

3.2 Factor Complexity

Let $p(n)$ be the factor complexity function of $\mathbf{w}$. The sequences of *finite differences of order* 1 *and* 2 respectively of $p(n)$, that is, $s(n) = p(n+1) - p(n)$ and $b(n) = s(n+1) - s(n)$, are used to show upper bounds for $p(n)$.

Lemma 1. *Suppose $|\mathcal{A}| = 3$. Then, $p(n+1) - p(n) \in \{2, 3\}$ if and only if $\sum_{\ell=0}^{n-1} b(\ell) \in \{0, 1\}$. Also, if the sequence of finite differences of order* 2 *is such that $(b(\ell))_\ell = 0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0, 1, 0, \ldots$ then $\sum_{\ell=0}^{n-1} b(\ell) \in \{0, 1\}$.*

Proof. Since $|\mathcal{A}| = 3$, then $p(1) = 3$ and $s(0) = p(1) - p(0) = 3 - 1 = 2$. We have $p(n+1) - p(n) = s(n) = s(0) + \sum_{\ell=0}^{n-1} b(\ell) = 2 + \sum_{\ell=0}^{n-1} b(\ell)$.

Function $b(n)$ is related to the multiplicity of bispecial factors.

Theorem 3. [CN10, Theorem 4.5.4] *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ be an infinite recurrent word. Then, for all $n \in \mathbb{N}$: $b(n) = \sum_{w \in BS_n(\mathbf{u})} m(w)$.*

3.3 Synchronization Lemmas

The goal of the next sections is to describe strong and weak bispecial factors. From now on, the alphabet is set to $\mathcal{A} = \{1, 2, 3\}$. The next lemma describes the preimage of a factor under Arnoux-Rauzy (**AR**) and Poincaré (**P**) substitutions.

Lemma 2 (Synchronization). *Let $u \in \mathcal{A}^*$ and w be a factor of $\alpha_k(u)$ for some $\{i, j, k\} = \{1, 2, 3\}$.*

(i) *If w is empty or if the first letter of w is i or j, then there exists a unique $v \in \mathcal{A}^*$ and a unique $s \in \{\varepsilon, i, j\}$ such that $w = \alpha_k(v) \cdot s$.*
(ii) *If the first letter of w is k, then there exists a unique $v \in \mathcal{A}^*$ and a unique $s \in \{\varepsilon, i, j\}$ such that $w = k \cdot \alpha_k(v) \cdot s$.*

Let $u \in \mathcal{A}^*$ and w be a factor of $\pi_{jk}(u)$ for some $\{i, j, k\} = \{1, 2, 3\}$.

(i) *If w is empty or if the first letter of w is i, then there exists a unique $v \in \mathcal{A}^*$ and a unique $s \in \{\varepsilon, i, j, ij\}$ such that $w = \pi_{jk}(v) \cdot s$.*
(ii) *If $w = j$, then there exists a unique $v(= \varepsilon)$ such that $w = j \cdot \pi_{jk}(v)$.*
(iii) *If the first letter of w is j and $|w| > 1$, then there exists a unique $v \in \mathcal{A}^*$ and a unique $s \in \{\varepsilon, i, j, ij\}$ such that $w = jk \cdot \pi_{jk}(v) \cdot s$.*
(iv) *If the first letter of w is k, then there exists a unique $v \in \mathcal{A}^*$ and a unique $s \in \{\varepsilon, i, j, ij\}$ such that $w = k \cdot \pi_{jk}(v) \cdot s$.*

Proof. The sets $\{ik, jk, k\}$ and $\{ijk, jk, k\}$ form a prefix code.

Definition 2 (Antecedent, extended image). *Let $\sigma = \alpha_k$ or $\sigma = \pi_{jk}$, $u \in \mathcal{A}^*$ and w be a factor of $\sigma(u)$. We say that the* antecedent of w under σ *is the unique word v as defined by Lemma 2. If v is the antecedent of a word w, then we say that the word w is an* extended image *of v.*

While the antecedent is unique, a word v may have more than one extended image. For example, $w_1 = 23\pi_{23}(11)1 = 231231231$ and $w_2 = 3\pi_{23}(11)2 = 31231232$ are two distinct extended images of $v = 11$. This is why the situation becomes here quite intricate especiallly for bispecial factors (it happens that strong and weak bispecial words appear in pairs, see Lemma 10 below).

Definition 3 (Bispecial extended image). *We shall say that a* bispecial extended image *w of v under σ is a bispecial word which is an extended image of v under σ.*

3.4 Antecedents and Images of Bispecial Words

Lemma 3 (AR - Bispecial extended image). *Let v be a bispecial factor. There is a unique bispecial extended image $w = k\alpha_k(v)$ of v under α_k.*

Lemma 4 (AR - Antecedent of a bispecial). *Let $u \in \mathcal{A}^*$ and $w \neq \varepsilon$ be a bispecial factor of $\alpha_k(u)$. Let v be the unique antecedent of w under α_k such that $w = k\alpha_k(v)$. Then, v is bispecial and it has the same extension type $E_{\alpha_k(u)}(w) = E_u(v)$ and same multiplicity $m(w) = m(v)$ as w.*

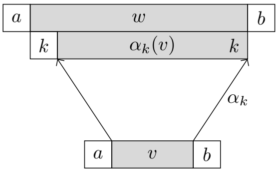

Fig. 4. The preimage of the bispecial word w under α_k

Proof. One checks that $(a, b) \in E(v)$ if and only if $(a, b) \in E(k\alpha_k(v))$ (see Figure 4). Then $E(k\alpha_k(v)) = E(v)$. We deduce that $E^+(k\alpha_k(v)) = E^+(v)$ and $E^-(k\alpha_k(v)) = E^-(v)$. From this we conclude that $m(k\alpha_k(v)) = m(v)$.

Lemma 5 (P - Bispecial extended images). *Let v be a bispecial factor. There are at most two distinct bispecial extended images of v under π_{jk}. They are either $k\pi_{jk}(v)$ or $jk\pi_{jk}(v)$.*

Proof. Let w be a bispecial extended images of v under π_{jk}. Since it is a bispecial factor, it must start with letter j or k and end with letter k. From Lemma 2, $w \in \{jk\pi_{jk}(v), k\pi_{jk}(v)\}$.

Lemma 6 (P - Antecedent of a bispecial). *Let $u \in \mathcal{A}^*$ and $w \neq \varepsilon$ be a bispecial factor of $\pi_{jk}(u)$. Let v be the unique antecedent of w under π_{jk} such that $w = k\pi_{jk}(v)$ or $w = jk\pi_{jk}(v)$. Then, v is bispecial.*

Now we want to describe more precisely under which conditions a bispecial word v has a unique bispecial extended image under Poincaré substitutions and give its extension type. In general, this depends on its left extensions $E^-(v)$. However, if the value of the left valence $d^-(v) = 2$, we deduce the unicity of the bispecial extended image as well as important information on the extension type of the extended image.

Lemma 7 (P - Bispecial extended images in details). *Let v be a bispecial factor.*

(i) *If $d^-(v) = 2$, v admits a unique bispecial extended image $w \in \{k\pi_{jk}(v), jk\pi_{jk}(v)\}$ under π_{jk} and $d^-(w) = 2$. Moreover, the extension types $E(v)$ and $E(w)$ are left equivalent.*

(ii) *If $d^-(v) = 3$, then v admits either one, or two bispecial extended images $w \in \{k\pi_{jk}(v), jk\pi_{jk}(v)\}$ under π_{jk}. In any case, $d^-(w) = 2$ and the two non empty rows of $E(w)$ are obtained by projection of rows of $E(v)$.*

3.5 Life of a Bispecial Factor under Arnoux-Rauzy-Poincaré Substitutions

In this section, the life of a bispecial factor is analyzed more precisely under the application of Arnoux-Rauzy and Poincaré substitutions in the spirit of Section

4.2.2 of [Cas97] where bispecial factors are described under the image of circular morphisms. To achieve this, we need to understand exactly the left extensions which will give information about the multiplicity of the bispecial factors. We denote by $\mathcal{S}_\alpha$, $\mathcal{S}_\pi$, respectively the following sets of substitutions:

$$\mathcal{S}_\alpha = \{\alpha_1, \alpha_2, \alpha_3\},\ \mathcal{S}_\pi = \{\pi_{12}, \pi_{13}, \pi_{23}, \pi_{21}, \pi_{31}, \pi_{32}\},\ \text{with } \mathcal{S} = \mathcal{S}_\alpha \cup \mathcal{S}_\pi.$$

Let w be a factor of $\lim_{k\to\infty} \sigma_0\sigma_1\cdots\sigma_k(a_k)$, $a_k \in \mathcal{A}$, where $\sigma_i \in \mathcal{S}$. Let $w_0 = w$ and w_{i+1} be the unique antecedent of w_i under σ_i for $i \geq 0$. If $|w_i| > 0$, then $|w_{i+1}| < |w_i|$, then there exists n such that $w_n = \varepsilon$.

Definition 4 (Age). *The smallest of those integers n is called the* age *of w and is noted* age(w)*.*

Thus, w_1 is the antecedent of w_0 under σ_0 and w_2 is the antecedent of w_1 under σ_1. If $n = \text{age}(w)$, w_n is the antecedent of w_{n-1} under σ_{n-1} and the extension type $E(w_n)$ of $w_n = \varepsilon$ depends on σ_n.

Definition 5 (History, life). *We say that the finite sequence $\sigma_0\sigma_1\cdots\sigma_n$ is the* history *and the sequence $(w_i)_{0\leq i\leq n}$ is the* life *of the bispecial word w.*

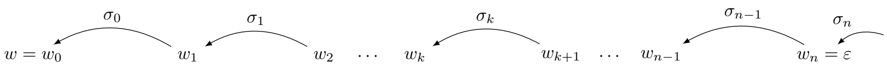

Fig. 5. Life and history of a factor w

Lemma 8. *Let $n \geq 0$ be an integer. Let B_n be the set of all bispecial factors of age n of $\lim_{n\to\infty} \sigma_0\sigma_1\cdots\sigma_n(a_n)$, $a_n \in \mathcal{A}$, where $\sigma_i \in \mathcal{S}$. Then* $\text{Card}\, B_n \leq 2$.

The life $(w_i)_{0\leq i\leq n}$ of bispecial factors starts as an empty word at $i = n$. The word w_i for $i < n$ is the concatenation of one or two letters and $\sigma_i(w_{i+1})$. These letters depend on the extension type $E(w_{i+1})$ and recursively on the extension type $E(w_n)$ of $w_n = \varepsilon$. Thus, it is important to understand properly what are the possible extension types of the empty word under the application of Arnoux-Rauzy and Poincaré substitutions. Below, the extension type $E(\varepsilon)$ of the empty word considered as a bispecial factor in the language of $\sigma(u)$ is denoted by $E_{\sigma(u)}(\varepsilon)$.

Lemma 9. *Let $\mathbf{u} \in \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$ be such that all letters of $\mathcal{A}$ appear as proper factors of $\mathbf{u}$. Considered as a bispecial factor of the language of the word $\alpha_k(\mathbf{u})$, the empty word ε is ordinary. Considered as a bispecial factor of the language of the word $\pi_{jk}(\mathbf{u})$, the empty word ε is neutral but not ordinary:*

$$E_{\alpha_k(\mathbf{u})}(\varepsilon) = \begin{array}{c|ccc} & i & j & k \\ \hline i & & & \times \\ j & & & \times \\ k & \times & \times & \times \end{array} \quad \textit{and} \quad E_{\pi_{jk}(\mathbf{u})}(\varepsilon) = \begin{array}{c|ccc} & i & j & k \\ \hline i & & \times & \\ j & & & \times \\ k & \times & \times & \times \end{array}.$$

In the next lemma, we describe exactly what are the bispecial factors associated with each possible history.

Lemma 10. *Let* $\mathbf{u} = \lim_{n\to\infty} \sigma_0\sigma_1\cdots\sigma_n(a_n)$. *Let* w *be a bispecial factor of* $\mathbf{u}$ *such that* $n = \mathrm{age}(w)$ *and* $\lim_{m\to\infty} \sigma_{n+1}\sigma_{n+2}\cdots\sigma_m(a_m)$ *contains all letters of* $\mathcal{A}$ *as proper factors. Let* z *be the other bispecial factor of the same age as* w *if it exists. Then the history* $\sigma_0\sigma_1\cdots\sigma_n$ *of* w *determine the left valence, multiplicity and extension type of* w *and* z *according to the following table.*

$\sigma_0\sigma_1\cdots\sigma_n \in$	$d^-(w)$	$m(w)$	*ordinary*	$d^-(z)$	$m(z)$	*ordinary*
$\mathcal{S}_\alpha^*\mathcal{S}_\alpha$	3	0	yes			
$\mathcal{S}_\alpha^*\mathcal{S}_\pi$	3	0	no			
$\mathcal{S}^*\,\pi_{jk}\,\mathcal{S}_\alpha^*\{\alpha_k\}$	2	0	yes			
$\mathcal{S}^*\,\pi_{jk}\,\mathcal{S}_\alpha^*\{\alpha_i,\alpha_j\}$	2	0	yes	2	0	yes
$\mathcal{S}^*\,\pi_{jk}\,\mathcal{S}_\alpha^*\{\pi_{ji},\pi_{ki},\pi_{ij},\pi_{kj}\}$	2	0	yes	2	0	yes
$\mathcal{S}^*\,\pi_{jk}\,\mathcal{S}_\alpha^*\{\pi_{ik},\pi_{jk}\}$	2	$+1$	no	2	-1	no

Strong and weak bispecial words thus appear in pairs under the application of Poincaré substitutions each time π_{jk} is followed by π_{jk} or π_{ik} for $\{i,j,k\} = \{1,2,3\}$ with possibly some Arnoux-Rauzy substitutions α_k, $k \in \{1,2,3\}$, in between.

4 Proof of Theorem 1

Restricted to the language of the automaton $\mathcal{G}$, illustrated in Figure 2, the history of a strong or weak bispecial factor necessarily contains Arnoux-Rauzy substitutions.

Lemma 11. *Let* $\mathbf{u} = \lim_{n\to\infty} \sigma_0\sigma_1\cdots\sigma_n(a_n)$. *Let* w *be a bispecial factor of* $\mathbf{u}$ *such that* $n = \mathrm{age}(w)$ *and* $\lim_{m\to\infty} \sigma_{n+1}\sigma_{n+2}\cdots\sigma_m(a_m)$ *contains all letters of* $\mathcal{A}$ *as proper factors. If* w *is weak or strong and the history of* w *is in the regular language* $\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{L}(\mathcal{G})$, *then*

$$\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{S}^*\,\pi_{jk}\{\alpha_j\}^*\,\alpha_i\,\mathcal{S}_\alpha^*\ \{\pi_{ik},\pi_{jk}\}$$

for some $\{i,j,k\} = \{1,2,3\}$.

Lemma 12. *Let* z^+ *and* z^- *be two bispecial factors of a word* u *of the same age* $\mathrm{age}(z^+) = \mathrm{age}(z^-)$. *Suppose that* z^- *is weak and* z^+ *is strong. Then* $|z^+| < |z^-|$.

Lemma 13. *Let* z^- *and* w^+ *be two bispecial factors of a word* u *such that* z^- *is weak and* w^+ *is strong. If* $\mathrm{age}(z^-) < \mathrm{age}(w^+)$, *then* $|z^-| < |w^+|$.

Lemma 14. *Let* z^-, w^+ *and* w^- *be bispecial factors of a word* $\mathbf{u}$ *such that* z^- *is weak,* w^+ *is strong and* w^- *is weak. If* $\mathrm{age}(z^-) < \mathrm{age}(w^+) = \mathrm{age}(w^-)$, *then* $|w^+| - |z^-| > |w^-| - |w^+|$.

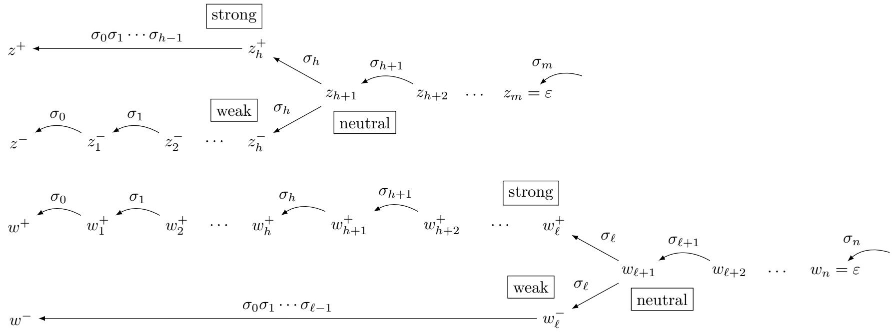

Fig. 6. Lifes of two pairs of strong and weak bispecial factors: z^+, z^- and w^+, w^-

We now have gathered all the elements for giving a proof of Theorem 1. We show that strong and weak bispecial words alternate when the length increases and make use of Lemma 1 (see Figure 6). Note that the notion of alternance was used to prove Theorem 4.11.2 in [CN10, p. 238].

Proof (of Theorem 1). Note first that the assumption on $\mathbf{x}$, i.e., $\mathbf{x}$ is totally irrational, is required for applying Lemma 11 for bispecial factors of all age. The set of bispecial factors of length n contains at most one weak or strong bispecial factor. Indeed, suppose on the contrary that it contains two of them: w and z. They cannot have the same age according to Lemma 12 since this would otherwise imply $|w| \neq |z|$. Also, if one is older, e.g. $\text{age}(w) > \text{age}(z)$, then $|w| > |z|$ from Lemma 13. Then $b(n) \in \{-1, 0, +1\}$ according to Theorem 3. Finally, it remains to prove that the assumptions of Lemma 1 are satisfied. The first non-zero value of $b(n)$ is $+1$ because strong and weak bispecial factors come in pairs and the strong one is smaller than the weak one from Lemma 12. Moreover, non-zero values are alternating. Indeed, let z^+ and w^+ be two strong bispecial factors such that $\text{age}(w^+) > \text{age}(z^+)$. Let z^- be the weak bispecial factor such that $\text{age}(z^-) = \text{age}(z^+)$. From Lemma 12 and Lemma 13, $|z^+| < |z^-| < |w^+|$. Hence, there is always a -1 between two $+1$ in the sequence $(b(n))_{n\geq 0}$. This shows that $p(n+1) - p(n) \in \{2, 3\}$ (Lemma 1), so that $p(n) \leq 3n + 1$ for $n \geq 0$. Moreover, $p(n) < 3n$ for each $n > 0$ since $p(1) = 3$ and $p(2) = 5$. We can show even more. From Lemma 14, the range of consecutive values of 2 for $p(n+1) - p(n)$ is larger than the range of consecutive values of 3 which follows immediately. From this we conclude that $\limsup_{n\to\infty} \frac{p(n)}{n} \leq \frac{5}{2}$.

5 Concluding Remarks

The restriction to the regular language $\mathcal{L}(\mathcal{G})$ is clearly important; there exist examples of S-adic words constructed with the alphabet of substitutions $\mathcal{S}$ for which the upper bound of $3n$ is false otherwise. Moreover, a quadratic complexity is even also achievable (fixed point of $\pi_{23}\pi_{13}$). Hence, this gives some more insight on a statement of the S-adic conjecture which is to find conditions for which S-adic sequences have a linear complexity (see e.g. [DLR13, Ler12]).

Factor complexity of Poincaré and Arnoux-Rauzy substitutions can be described exactly by considering left and right extensions of length one. It is not always the case, and Brun substitutions seems to be an example for which extensions of length longer than 1 are necessary to describe bispecial factors. Recently, Klouda [Klo12] described bispecial factors in fixed point of morphisms where extensions of length longer than one were considered. Extending this work to S-adic words deserves further research.

Balance of the Poincaré and Arnoux-Rauzy S-adic system also has nice properties and its study will be part of a extended version of this article.

References

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

[BD13] Berthé, V., Delecroix, V.: Beyond substitutive dynamical systems: s-adic expansions (preprint, 2013)

[BL11] Berthé, V., Labbé, S.: Uniformly balanced words with linear complexity and prescribed letter frequencies. In: Proc. 8th Int. Conf. on Words. EPTCS, vol. 63, pp. 44–52. Open Publishing Association (2011)

[Bos85] Boshernitzan, M.: A unique ergodicity of minimal symbolic flows with linear block growth. J. Analyse Math. 44, 77–96 (1984/1985)

[Cas97] Cassaigne, J.: Complexité et facteurs spéciaux. Bull. Belg. Math. Soc. Simon Stevin 4(1), 67–88 (1997); Journées Montoises (Mons, 1994)

[CN10] Cassaigne, J., Nicolas, F.: Factor complexity. In: Combinatorics, Automata and Number Theory. Encyclopedia Math. Appl., vol. 135, pp. 163–247. Cambridge Univ. Press, Cambridge (2010)

[DFG+12] Dasaratha, K., Flapan, L., Garrity, T., Lee, C., Mihaila, C., Neumann-Chun, N., Peluse, S., Stroffregen, M.: Cubic irrationals and periodicity via a family of multi-dimensional continued fraction algorithms. arXiv:1208.4244 (2012)

[DLR13] Durand, F., Leroy, J., Richomme, G.: Do the Properties of an S-adic Representation Determine Factor Complexity? J. of Int. Seq. 16 (2013)

[FM10] Ferenczi, S., Monteil, T.: Infinite words with uniform frequencies, and invariant measures. In: Combinatorics, Automata and Number Theory. Encycl. Math. Appl., vol. 135, pp. 373–409. Cambridge Univ. Press (2010)

[Gar01] Garrity, T.: On periodic sequences for algebraic numbers. J. Number Theory 88(1), 86–103 (2001)

[Klo12] Klouda, K.: Bispecial factors in circular non-pushy D0L languages. Theoret. Comput. Sci. 445, 63–74 (2012)

[Lab12] Labbé, S.: Structure des pavages, droites discèrtes 3D et combinatoire des mots. PhD thesis, Université du Québec à Montréal (May 2012)

[Ler12] Leroy, J.: Some improvements of the S-adic conjecture. Adv. in Appl. Math. 48(1), 79–98 (2012)

[Nog95] Nogueira, A.: The three-dimensional Poincaré continued fraction algorithm. Israel J. Math. 90(1-3), 373–401 (1995)

[Sch00] Schweiger, F.: Multidimensional continued fractions. Oxford Science Publications. Oxford University Press, Oxford (2000)