

# Bispecial Factors in the Brun $S$ -adic System

Sébastien Labbé\* and Julien Leroy\*\*

Université de Liège, Institut de mathématique,  
Allée de la découverte 12 (B37), 4000 Liège, Belgium  
slabbe@ulg.ac.be, j.leroy@ulg.ac.be

**Abstract.** We study the bispecial factors in the  $S$ -adic system associated with the Brun Multidimensional Continued Fraction algorithm. More precisely, by describing how strong and weak bispecial words can appear, we get a sub-language of the Brun language for which all bispecial words are neutral.

**Keywords:** Substitutions, Brun, factor complexity, bispecial.

## 1 Introduction

Sturmian sequences [14] are infinite sequences on a binary alphabet in which appear exactly  $n + 1$  distinct finite subsequences of consecutive  $n$  letters for each  $n \in \mathbb{N}$ . It is known that the symbolic dynamical system associated to a sturmian sequence (with the shift transformation) is minimal and is measure-theoretically isomorphic to an irrational rotation on the unit circle  $\mathbb{T}_1$ . The result was extended to higher dimensions when Rauzy proved in [15] that the symbolic dynamical system associated to the fixed point of the Tribonacci substitution  $\sigma : 1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 1$ , which has  $p(n) = 2n + 1$  factors of length  $n$ , is measure-theoretically isomorphic to an irrational translation on the torus  $\mathbb{T}_2$ . Proving that Rauzy's result hold in a more general setting is still an open question known as the Pisot conjecture [1] in the case of all Pisot unimodular substitutions. However substitutive dynamical systems obtained from the iteration of *one* substitution is quite limited (frequencies of letters must be algebraic) and do not form a satisfactory generalization to larger alphabets of sturmian systems (achieving all irrational frequencies of letters).

A generalization of the Pisot conjecture was proposed in [7] in the case of  $S$ -adic symbolic dynamical systems. These shift spaces are obtained by iterating substitutions from a set  $S$ , generalizing the substitutive case where  $\text{Card } S = 1$ . Like it is the case for sturmian words, the sequence of substitutions is obtained from the continued fractions algorithm or some multidimensional version of it [8,17]. They proved using results from [4,3,10] that almost all  $S$ -adic shifts based on Brun's Multidimensional Continued Fraction Algorithm [9] are measurably conjugate to a translation on the torus  $\mathbb{T}_2$ . They also proved that these shifts

---

\* Postdoctoral Marie Curie fellowship (BeIPD-COFUND).

\*\* Postdoctoral FNRS fellowship.

provide a natural coding of almost all rotations on  $\mathbb{T}_2$  providing a reverent generalization of sturmian systems to a three-letter alphabet.

One statement about Brun  $S$ -adic systems has remained unproved: the factor complexity. As mentioned in [7], it is believed that any Brun  $S$ -adic shift has a linear factor complexity and this is the subject of this contribution. In this work, we initiate a study of bispecial factors in the Brun  $S$ -adic system pushing further methods already used in [6,12] and also in [5] where it was proved that  $p(n) \leq \frac{5}{2}n + 1$  for Arnoux-Rauzy-Poincaré  $S$ -adic system. In the Brun system, it appears that left extensions of length 1 are not enough to study the evolution of bispecial factors. Also, some neutral bispecial factors can split into a pair of strong and weak bispecial factors which can later on merge again into a neutral bispecial factor. These phenomenon are not possible in the case of the Arnoux-Rauzy-Poincaré algorithm and these are reasons why the factor complexity for the Brun  $S$ -adic system has shown to be harder to prove.

## 2 Brun's Algorithm

Brun's algorithm [9] is a Multidimensional Continued Fraction Algorithm [8,17] which subtracts the second largest entry to the largest entry of a nonnegative vector in  $\mathbb{R}_+^d$ . In the most often used version of Brun's algorithm, the entries are sorted after each iteration. Keeping the entries sorted has the advantage of reducing the number of branches of the algorithm at each step but the disadvantage of losing the symmetry between them. In this work, we prefer to keep the symmetry and present below the unsorted version of Brun's algorithm which has 6 branches when  $d = 3$ . On  $\Lambda = \mathbb{R}_+^3$ , the *unsorted Brun's algorithm* is the map  $F(x_1, x_2, x_3) = (x'_1, x'_2, x'_3)$  defined by

$$x'_{\pi_1} = x_{\pi_1}, \quad x'_{\pi_2} = x_{\pi_2}, \quad x'_{\pi_3} = x_{\pi_3} - x_{\pi_2}$$

where  $\pi \in \mathcal{S}_3$  is the permutation of  $\{1, 2, 3\}$  such that  $x_{\pi_1} < x_{\pi_2} < x_{\pi_3}$ . Equivalently, the map  $F$  on  $\Lambda$  can be defined as a linear application  $F\mathbf{x} = M(\mathbf{x})^{-1}\mathbf{x}$  with  $M(\mathbf{x}) = M_\pi$  if and only if  $\mathbf{x} \in \Lambda_\pi$  where  $\Lambda_\pi = \{(x_1, x_2, x_3) \in \Lambda \mid x_{\pi_1} < x_{\pi_2} < x_{\pi_3}\}$  defines a partition of the positive cone  $\Lambda = \cup_{\pi \in \mathcal{S}_3} \Lambda_\pi$  up to a set of Lebesgue measure zero and  $M_\pi$  are the following elementary matrices:

$$M_{123} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad M_{132} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_{213} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

$$M_{231} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_{312} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_{321} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The algorithm  $F$  defines a *cocycle*  $M_n : \Lambda \rightarrow SL(d, \mathbb{Z})$

$$M_0(\mathbf{x}) = I \quad \text{and} \quad M_n(\mathbf{x}) = M(\mathbf{x})M(F\mathbf{x}) \cdots M(F^{n-1}\mathbf{x})$$

with the cocycle property  $M_{m+n}(\mathbf{x}) = M_m(\mathbf{x})M_n(F^m\mathbf{x})$ . Since Brun's algorithm is strongly convergent almost everywhere when  $d = 3$  [13] (also when  $d = 4$  [16]),

the columns of  $M_n(\mathbf{x})$  are good rational approximations of  $\mathbf{x}$ . Indeed, an MCF algorithm is *strongly convergent* at  $\mathbf{x} \in \Lambda$  with  $\|\mathbf{x}\| = 1$  if for all  $i$  with  $1 \leq i \leq d$ , we have

$$\lim_{n \rightarrow \infty} M_n(\mathbf{x})\mathbf{e}_i - \|M_n(\mathbf{x})\mathbf{e}_i\|\mathbf{x} = 0. \quad (1)$$

### 3 Brun $S$ -adic System

#### 3.1 $S$ -adic words

Let  $S$  be a set of substitutions. A word  $\mathbf{w} \in A^{\mathbb{N}}$  is said to be  $S$ -adic if there is a sequence  $(\sigma_n : A_{n+1}^* \rightarrow A_n^*)_{n \in \mathbb{N}} \in S^{\mathbb{N}}$  and a sequence of letters  $(a_n \in A_n)_{n \in \mathbb{N}}$  such that  $A_0 = A$  and

$$\mathbf{w} = \lim_{n \rightarrow +\infty} \sigma_0 \sigma_1 \cdots \sigma_n(a_{n+1}).$$

For all  $r \in \mathbb{N}$  we define the  $S$ -adic word

$$\mathbf{w}^{(r)} = \lim_{n \rightarrow +\infty} \sigma_r \sigma_{r+1} \cdots \sigma_{r+n}(a_{k+n+1}).$$

In our setting we will consider the alphabet  $A = \{1, 2, 3\}$  and we usually use the set  $\{i, j, k\}$  to represent  $A$ .

A directive sequence of substitutions  $(\sigma_n : A_{n+1}^* \rightarrow A_n^*)_{n \in \mathbb{N}}$  is *primitive* if for all  $r \in \mathbb{N}$ , there exists  $s \geq r$  such that for all  $a \in A_r$  and all  $b \in A_{s+1}$ , the letter  $a$  occurs in  $\sigma_r \cdots \sigma_s(b)$ . Primitiveness of a directive sequence of substitutions implies the uniform recurrence of the associated  $S$ -adic word [11].

#### 3.2 Brun substitutions and Brun words

For every totally irrational  $\mathbf{x} \in \Lambda$ , the Brun's algorithm  $F$  defines a sequence of substitutions  $(\sigma(F^n \mathbf{x}))_{n \in \mathbb{N}}$ , where  $\sigma(\mathbf{x}) = \beta_{jk}$  if and only if  $\mathbf{x} \in \Lambda_{ijk}$  and  $\beta_{jk} : i \mapsto i, j \mapsto jk, k \mapsto k$  is a substitution called *Brun substitution*. Note that the incidence matrix of  $\beta_{jk}$  is  $M_{ijk}$  for all  $ijk \in \mathcal{S}_3$ .

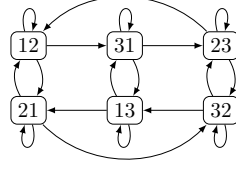
One can see that the allowed product of two consecutive Brun substitutions is restricted among the possibilities. One can show that after each  $\beta_{ij}$  only three of the six substitutions are allowed:

$$\{\sigma(\mathbf{x})\sigma(F\mathbf{x}) \mid \mathbf{x} \in \Lambda\} = \{\beta_{ij}\beta_{ij}, \beta_{ij}\beta_{ji}, \beta_{ij}\beta_{ki} \mid ijk \in \mathcal{S}_3\}.$$

Writing  $\mathcal{S}_{\mathcal{B}} = \{\beta_{ij} \mid ijk \in \mathcal{S}_3\}$ , the *Brun language* is:

$$\begin{aligned} \mathcal{L}_{\mathcal{B}} &= \{\sigma(\mathbf{x})\sigma(F\mathbf{x}) \cdots \sigma(F^{n-1}\mathbf{x}) \mid \mathbf{x} \in \Lambda, n \in \mathbb{N}\} \\ &= \mathcal{S}_{\mathcal{B}}^* \setminus \mathcal{S}_{\mathcal{B}}^* \{\beta_{ij}\beta_{ik}, \beta_{ij}\beta_{jk}, \beta_{ij}\beta_{kj} \mid ijk \in \mathcal{S}_3\} \mathcal{S}_{\mathcal{B}}^*. \end{aligned}$$

It is a regular language accepted by the automaton represented in Figure 1 where the label of an edge is  $\beta_{ij}$  whenever the edge goes to the state  $ij$ .



**Fig. 1.** The Brun language  $\mathcal{L}_{\mathcal{B}}$  is regular.

If  $\mathbf{x} \in A$  is totally irrational, then there are infinitely many  $n \in \mathbb{N}$  such that  $\sigma(F^n \mathbf{x})\sigma(F^{n+1} \mathbf{x}) \in \{\beta_{ij}\beta_{kj} \mid ijk \in \mathcal{S}_3\}$ . This is equivalent to say that for all  $ijk \in \mathcal{S}_3$ ,  $\text{Card}(\{n \in \mathbb{N} \mid \sigma(F^n \mathbf{x}) \in \{\beta_{ij}, \beta_{ik}\}\}) = +\infty$ . This implies that  $\lim_{n \rightarrow +\infty} \min_{i \in A} |\sigma(\mathbf{x})\sigma(F\mathbf{x}) \cdots \sigma(F^n \mathbf{x})(i)| = +\infty$ . As all Brun substitutions are prolongable on every letter, this allows us to define the  $S$ -adic infinite word

$$\lim_{n \rightarrow +\infty} \sigma(\mathbf{x})\sigma(F\mathbf{x}) \cdots \sigma(F^{n-1} \mathbf{x})(1)$$

whose letter frequencies exist and are proportional to  $\mathbf{x}$  by (1).

**Definition 1 (Brun word).** A word  $\mathbf{w} \in A^{\mathbb{N}}$  is a Brun word if it is an  $\mathcal{S}_{\mathcal{B}}$ -adic word whose directive sequence  $(\sigma_n)_{n \in \mathbb{N}} \in \mathcal{S}_{\mathcal{B}}^{\mathbb{N}}$  is such that for all  $n \in \mathbb{N}$ ,  $\sigma_0 \sigma_1 \cdots \sigma_n \in \mathcal{L}_{\mathcal{B}}$  and for all  $ijk \in \mathcal{S}_3$ ,  $\text{Card}(\{n \in \mathbb{N} \mid \sigma_n \in \{\beta_{ij}, \beta_{ik}\}\}) = +\infty$ .

**Proposition 2.** If  $\mathbf{s} = (\sigma_n)_{n \in \mathbb{N}} \in \mathcal{S}_{\mathcal{B}}^{\mathbb{N}}$  is the directive sequence of a Brun word  $\mathbf{w}$ , then  $(\sigma_n)_{n \in \mathbb{N}}$  is primitive. In particular,  $\mathbf{w}$  is uniformly recurrent.

### 3.3 Relations with Arnoux-Rauzy and Poincaré substitutions

The Brun substitutions share some relations with other well-known substitutions. For all  $\{i, j, k\} = A$ , we let  $\alpha_i$  denote the *Arnoux-Rauzy substitution* [2] and  $\pi_{ij}$  denote the *Poincaré substitution* [5]:

$$\alpha_i : i \mapsto i, j \mapsto ji, k \mapsto ki, \quad \pi_{ij} : i \mapsto ij, j \mapsto j, k \mapsto kij. \quad (2)$$

These are products of Brun substitutions. More precisely, for all  $ijk \in \mathcal{S}_3$ , we have  $\pi_{ij} = \beta_{ij}\beta_{ki}$  and  $\alpha_i = \beta_{ji}\beta_{ki} = \beta_{ki}\beta_{ji}$ .

Note that a Poincaré substitution can appear as a product of two consecutive Brun substitutions in the Brun  $S$ -adic system, but not an Arnoux-Rauzy one. We let  $\mathcal{S}_{\mathcal{A}}$  and  $\mathcal{S}_{\mathcal{P}}$  respectively denote the set of Arnoux-Rauzy substitutions and the set of Poincaré substitutions:  $\mathcal{S}_{\mathcal{A}} = \{\alpha_1, \alpha_2, \alpha_3\}$  and  $\mathcal{S}_{\mathcal{P}} = \{\pi_{ij} \mid ijk \in \mathcal{S}_3\}$ .

Now we show that Poincaré substitutions appear infinitely often as products of two consecutive Brun substitutions. This will be useful to study the extension type (defined in Section 4.1) of the empty word in Brun words.

**Lemma 3.** Let  $\mathbf{w}$  be a Brun word with directive sequence  $(\sigma_n)_{n \in \mathbb{N}} \in \mathcal{S}_{\mathcal{B}}^{\mathbb{N}}$ . There exist infinitely many integers  $n \in \mathbb{N}$  such that  $\sigma_n \sigma_{n+1} \in \mathcal{S}_{\mathcal{P}}$ . Moreover, if  $\sigma_r = \beta_{ij}$ , then the smallest integer  $\ell \geq r$  such that  $\sigma_\ell \sigma_{\ell+1} \in \mathcal{S}_{\mathcal{P}}$  satisfies  $\sigma_\ell \sigma_{\ell+1} \in \{\pi_{ij}, \pi_{ji}\}$ . Finally, if  $\sigma_\ell \sigma_{\ell+1} = \pi_{xy}$  with  $xy \in \{ij, ji\}$ , then  $(\sigma_n)_{n \geq \ell} \in L_{xy}$  where

$$L_{xy} = \pi_{xy}\beta_{yk}\mathcal{S}_{\mathcal{B}}^{\mathbb{N}} \cup \pi_{xy}\{\beta_{xk}, \beta_{kx}\}^* \{\pi_{xk}, \pi_{kx}\}\mathcal{S}_{\mathcal{B}}^{\mathbb{N}}. \quad (3)$$

### 3.4 Other substitutions used for Brun's algorithm in the literature

In Jolivet's thesis or also in [4], they proposed the following substitutions for Brun's algorithm in its sorted version. Note that it was in the purpose of generating discrete planes. Therefore, their incidence matrix is dual to the matrix associated with the execution of the sorted Brun algorithm:

$$1 \mapsto 1, 2 \mapsto 2, 3 \mapsto 32, \quad 1 \mapsto 1, 2 \mapsto 3, 3 \mapsto 23, \quad 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 13.$$

In [10], they use the reversal of the above three substitutions for the sorted algorithm. For the unsorted one, they propose the six Brun substitutions:

$$\gamma_{ij} : j \mapsto ij, i \mapsto i, k \mapsto k, \quad \text{for each } ijk \in \mathcal{S}_3,$$

with a language of allowed words of length two:  $\{\gamma_{ij}\gamma_{ij}, \gamma_{ij}\gamma_{ji}, \gamma_{ij}\gamma_{jk} \mid ijk \in \mathcal{S}_3\}$ . More recently, in [7], they present the Brun's algorithm in its sorted version using the following substitutions:

$$1 \mapsto 1, 2 \mapsto 23, 3 \mapsto 3, \quad 1 \mapsto 1, 2 \mapsto 3, 3 \mapsto 23, \quad 1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 23.$$

One observes that any  $S$ -adic word obtained by the above substitutions of sorted Brun algorithm can be obtained as a Brun word with the unsorted algorithm.

## 4 Bispecial Factors under Brun Substitutions

In this section we define the extension type of a word. We also describe the extensions of a word under the application of a Brun substitution.

### 4.1 Special factors and extension type

Let  $\mathbf{w}$  be a (infinite) word over  $A$ . We let  $\text{Fac}(\mathbf{w})$  denote the set of factors of  $\mathbf{w}$ :

$$\text{Fac}(\mathbf{w}) = \{u \in A^* \mid \exists i \in \mathbb{N} : \mathbf{w}_i \cdots \mathbf{w}_{i+|u|-1} = u\}.$$

Let  $u \in \text{Fac}(\mathbf{w})$  and  $\ell \in \mathbb{N}$ . The  $\ell$ -extension set of  $u$  is the set  $E_\ell(u, \mathbf{w}) = \{(a, b) \in A^\ell \times A^1 \mid aub \in \text{Fac}(\mathbf{w})\}$ . We represent it by a tabular of the form

$$E_\ell(u, \mathbf{w}) = \begin{array}{c|ccc} u & 1 & \cdots & k \\ \hline v_1 & & \times & \times \\ \cdots & & & \\ v_n & & & \times \end{array},$$

where a symbol  $\times$  in position  $(v_i, j)$  means that  $(v_i, j)$  belongs to  $E_\ell(u, \mathbf{w})$ . When the context is clear we omit the information on  $\mathbf{w}$  and simply write  $E_\ell(u)$ . In this paper we will only work with  $\ell \in \{1, 2\}$ .

Two extension sets  $E_\ell(u)$  and  $E_\ell(v)$  are said to be *equivalent* if one can be obtained from the other by a permutation of the alphabet. The equivalent class of an extension set is called an *extension type*.

Given an extension set  $E_\ell(u, \mathbf{w})$ , we consider the corresponding set of *left extensions*  $E_\ell^-(u, \mathbf{w}) = \pi_1(E_\ell(u, \mathbf{w}))$  (resp. of *right extensions*  $E_\ell^+(u, \mathbf{w}) = \pi_2(E_\ell(u, \mathbf{w}))$ ), where  $\pi_1$  (resp.  $\pi_2$ ) represents the projection on the first (resp. second) component. We assume that the reader is familiar with the notion of left, right and bispecial words. For definitions, see [6, Chap. 4].

#### 4.2 Antecedents, extended images and their extension types

The next lemma allows to define the antecedents of a word under  $\beta_{ij}$ . It directly follows from the fact that the set  $\{ij, j, k\}$  forms a prefix code.

**Lemma 4 (Synchronization lemma).** *Let  $i, j, k$  such that  $\{i, j, k\} = A$ . Consider a word  $u \in A^*$  and let  $w$  be a factor of  $\beta_{ij}(u)$ .*

- (i) *If  $w$  is empty or belongs to  $\{i, k\}A^*$ , there exists a unique word  $v \in A^*$  and a unique  $s \in \{\varepsilon, i\}$  such that  $w = \beta_{ij}(v) \cdot s$ . We say that  $v$  is the antecedent of  $w$  under  $\beta_{ij}$ .*
- (ii) *If  $w \in jA^*$ , there is a unique word  $v \in A^*$  and a unique  $s \in \{\varepsilon, i\}$  such that  $w = j \cdot \beta_{ij}(v) \cdot s = \beta_{ij}(jv) \cdot s$ . We say that  $v$  and  $jv$  are the two antecedents of  $w$  under  $\beta_{ij}$ .*

**Definition 5.** *Suppose that  $v$  is an antecedent of  $w$  under  $\sigma$  as in Lemma 4. In this case, we say that  $w$  is an extended image of  $v$ . In particular, if  $w$  is a left special (resp. right special, bispecial) factor in  $\sigma(u)$ , then we say that it is a left special (resp. right special, bispecial) extended image of  $v$  under  $\sigma$ .*

The next lemma provides the link between the extensions of a word and those of its extended images.

**Lemma 6 (Extensions).** *Let  $i, j, k$  such that  $\{i, j, k\} = A$ . Let  $u \in A^*$  and  $v$  be a factor of  $u$ . We assume that for all  $(a, b) \in E_1(v)$ , there exists a letter  $e$  such that  $eavb$  is also a factor of  $u$ . The extensions of  $v$  in  $u$  are related to the extensions of  $\beta_{ij}(v)$  and  $j\beta_{ij}(v)$  considered as factors of  $\beta_{ij}(u)$  as follows:*

$$\begin{aligned}
(i, b) \in E_1(v) & \iff (j, b) \in E_1(\beta_{ij}(v)) \quad \text{and} \quad (i, b) \in E_1(j\beta_{ij}(v)), \\
(ij, b) \text{ or } (jj, b) \in E_2(v) & \iff (j, b) \in E_1(\beta_{ij}(v)) \quad \text{and} \quad (j, b) \in E_1(j\beta_{ij}(v)), \\
(kj, b) \in E_2(v) & \iff (j, b) \in E_1(\beta_{ij}(v)) \quad \text{and} \quad (k, b) \in E_1(j\beta_{ij}(v)), \\
(k, b) \in E_1(v) & \iff (k, b) \in E_1(\beta_{ij}(v)), \\
v = jv' \text{ with } \{(j, b), (i, b)\} \subseteq E_1(v') & \iff (i, b) \in E_1(\beta_{ij}(v)).
\end{aligned}$$

**Lemma 7 (Extended images).** *Consider the same hypothesis as in Lemma 6.*

1. *If  $v$  is right special in  $u$ , then  $\beta_{ij}(v)$  is a right special factor of  $\beta_{ij}(u)$ .*
2. *If  $v$  is left special in  $u$ , then  $v$  has at least one left special extended image in  $\beta_{ij}(u)$ .*
3. *If  $v$  is bispecial factor of  $u$  such that  $\beta_{ij}(v)$  is not a left special factor of  $\beta_{ij}(u)$ , then  $j\beta_{ij}(v)$  is a right special factor of  $\beta_{ij}(u)$ .*

Therefore, if  $v$  is a bispecial factor of  $u$ , then it has one or two bispecial extended images under  $\beta_{ij}$  in  $\beta_{ij}(u)$ ; they are  $\beta_{ij}(v)$  or  $j \cdot \beta_{ij}(v)$ .

**Lemma 8 (Antecedents).** *Let  $i, j, k$  such that  $\{i, j, k\} = A$ . Consider a word  $u \in A^*$  and  $w$  a bispecial factor of  $\beta_{ij}(u)$ . Then at least one antecedent of  $w$  under  $\beta_{ij}$  is a bispecial factor of  $u$ . We call it a bispecial antecedent of  $w$ .*

## 5 Bispecial Words in the Brun System

In this section we study the set of bispecial words in a Brun word. We first show that any bispecial factor can be canonically desubstituted to the empty word. We then define the descendants of a bispecial word and describe those of the empty word. At the end of the section, we give an example that illustrates some results that we give (see Figure 2).

**General assumption** *In all what follows, we assume that  $\mathbf{w}$  is a Brun word with directive sequence  $(\sigma_n)_{n \in \mathbb{N}}$ .*

### 5.1 Desubstitution of bispecial words

**Definition 9 (nth-antecedent).** *Let  $u$  be a bispecial factor of  $\mathbf{w}$ . Let  $u^{(0)} = u$  and  $u^{(i+1)}$  be the shortest bispecial antecedent of  $u^{(i)}$  under  $\sigma_i$  for  $i \geq 0$ . We say that  $u^{(n)}$  is the  $n$ -th antecedent of  $u$ . Observe that  $u^{(n)}$  is a factor of  $\mathbf{w}^{(n)}$ .*

With Brun substitutions, as opposed to the Arnoux-Rauzy-Poincaré substitutions (2), we are unable to prove  $|v| < |w|$  for any antecedent  $v$  of a bispecial word  $w \neq \varepsilon$ . This is not a problem since it holds for the  $n$ -th antecedent of  $w$ , for some  $n \geq 1$ , under the hypothesis that  $w$  is a factor of an  $S$ -adic Brun word.

**Lemma 10.** *If  $u \neq \varepsilon$  is a bispecial word of  $\mathbf{w}$ , there is  $s \geq 1$  such that  $u^{(s)} = \varepsilon$ .*

**Definition 11 (Descendants).** *Let  $u$  be a bispecial factor of  $\mathbf{w}^{(s)}$  for some  $s \in \mathbb{N}$ . A bispecial factor  $v$  of  $\mathbf{w}^{(r)}$ ,  $r < s$ , is called a descendant of  $u$  if there exists a sequence  $(u_r, u_{r+1}, \dots, u_s)$  such that  $u_r = v$ ,  $u_s = u$  and each  $u_\ell$ ,  $r \leq \ell < s$  is a bispecial extended image of  $u_{\ell+1}$ . We let  $\text{desc}(u)$  denote the set of descendants of  $u$  and, for  $r < s$ , we let  $\text{desc}_r(u)$  denote the set of bispecial factors of  $\mathbf{w}^{(r)}$  that are descendants of  $u$ .*

### 5.2 Extension type of the empty word

The aim of this paper is to study bispecial words in  $\mathbf{w}$ . Since any such bispecial is a descendant of the empty word in some  $\mathbf{w}^{(s)}$ , the first step is to study the possible extension types of the empty word. The next result in particular ensures that the empty word is always a neutral bispecial factor. In the next section we will show that the extension type of a bispecial word essentially governs the extension type of any of its bispecial extended image. However, as seen in Lemma 6, we sometimes need to consider left extensions of length 2 to be able to describe those of a bispecial extended image. In the next result, we thus describe the 2-extension types of the empty word.

**Theorem 12.** *For all  $s \in \mathbb{N}$ , the empty word is a neutral bispecial factor of  $\mathbf{w}^{(s)}$ . More precisely, if  $\sigma_s = \beta_{ij}$  and  $\mathbf{s} = (\sigma_r)_{r \geq s}$  for some  $\{i, j, k\} = A$ , then the 2-extension type of the empty word is*

$$\begin{aligned} E_1 \text{ if } \mathbf{s} \in M_{ij}, & & E_4 \text{ if } \mathbf{s} \in (\beta_{ij}\beta_{ji}\{\beta_{ij}\beta_{ji}\}^*L_{ji}) \cup \beta_{ij}L_{ji}, \\ E_2 \text{ if } \mathbf{s} \in (\beta_{ij}\beta_{ji}\{\beta_{ij}, \beta_{ji}\}^*L_{ij}) \cup N_{ij}, & & E_5 \text{ if } \mathbf{s} \in \beta_{ij}\beta_{ij}\{\beta_{ij}, \beta_{ji}\}^*L_{ji}, \\ E_3 \text{ if } \mathbf{s} \in (\beta_{ij}\beta_{ij}\{\beta_{ij}, \beta_{ji}\}^*L_{ij}) \cup \beta_{ij}L_{ij}, & & \end{aligned}$$

where  $L_{ij}$  as defined in Equation (3) and  $L_{ij} = M_{ij} \cup N_{ij}$  with

$$\begin{aligned} M_{ij} &= \pi_{ij}\{\beta_{ik}, \beta_{ki}\}^* \pi_{ki} \mathcal{S}_{\mathcal{B}}^{\mathbb{N}} \cup \pi_{ij}\beta_{jk} \mathcal{S}_{\mathcal{B}}^{\mathbb{N}}, \\ N_{ij} &= \pi_{ij}\{\beta_{ik}, \beta_{ki}\}^* \pi_{ik} \mathcal{S}_{\mathcal{B}}^{\mathbb{N}}, \end{aligned}$$

and  $E_1, E_2, E_3, E_4$  and  $E_5$  are as follows:

$E_1$	$i$	$j$	$k$	$E_2$	$i$	$j$	$k$	$E_3$	$i$	$j$	$k$	$E_4$	$i$	$j$	$k$	$E_5$	$i$	$j$	$k$
$ji$		$\times$		$ji$		$\times$		$ji$		$\times$		$ji$		$\times$		$ji$		$\times$	
$ki$		$\times$		$ki$		$\times$		$ki$		$\times$		$ij$	$\times$	$\times$	$\times$	$ij$		$\times$	
$ij$	$\times$	$\times$	$\times$	$ij$	$\times$	$\times$	$\times$	$ij$	$\times$	$\times$	$\times$	$jj$	$\times$	$\times$	$\times$	$jj$	$\times$	$\times$	$\times$
$jj$		$\times$		$jj$		$\times$		$jj$	$\times$	$\times$	$\times$	$kj$	$\times$			$kj$		$\times$	
$jk$	$\times$			$jk$	$\times$			$jk$	$\times$			$jk$	$\times$			$jk$		$\times$	

### 5.3 Left extensions of length 2 are sufficient

As already stated and as seen in Lemma 6, we sometimes need to consider left extensions of length 2 to be able to determine the left extensions of the longer extended images. In this section, we show that considering 2-extensions is sufficient to recover 2-extensions of any descendant. For a word  $u$  and an integer  $x \geq 1$ , we let  $u_{[-x:]}$  denote the suffix of length  $x$  of  $u$ .

**Definition 13.** *Assume that  $i, j, k$  are such that  $A = \{i, j, k\}$ . We define the function  $\varphi_{ij} : A^2 \rightarrow A^2$  and the partial function  $\psi_{ij} : A^2 \rightarrow A^2$  by*

$$\varphi_{ij}(x) = (\beta_{ij}(x))_{[-2:]} \quad \text{and} \quad \psi_{ij}(x) = \begin{cases} (j\beta_x(x)j^{-1})_{[-2:]} & \text{if } \beta_{ij}(x) \in A^*j, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

**Proposition 14.** *Let  $s \in \mathbb{N}$  and assume that  $\sigma_s = \beta_{ij}$  and that  $u$  is a factor of  $\mathbf{w}^{(s+1)}$ .*

1.  $E_2(j\beta_{ij}(u), \mathbf{w}^{(s)}) = \{(\psi_{ij}(a), b) \mid (a, b) \in E_2(u, \mathbf{w}^{(s+1)}), \beta_{ij}(a) \in A^*j\}$ .
2.  $E_2(\beta_{ij}(u), \mathbf{w}^{(s)}) = \{(\varphi_{ij}(a), b) \mid (a, b) \in E_2(u, \mathbf{w}^{(s+1)})\}$  if  $u \in \{i, k\}A^*$  or  $u = ju'$  for some  $u' \in A^*$  which is not left special.

Note that if  $u = ju'$  for some left special  $u' \in A^*$ , then the equation in item 2. above does not hold but the equation  $\beta_{ij}(u) = j\beta_{ij}(u')$  allows to use item 1.



#### 5.4 First descendants of the empty word

In this section we show that the first descendants of the empty word are always neutral bispecial words. We also show that if the empty word of  $\mathbf{w}^{(s)}$  has no descendant which has 3 left extensions of length 1 and 3 left extensions of length 2, then all its descendants are neutral bispecial. Below, we denote the left valence of a factor  $v$  by  $d_\ell^-(v) = \text{Card}(E_\ell^-(v))$  for  $\ell \in \{1, 2\}$ . Given a bispecial word  $u$  of  $\mathbf{w}^{(s)}$ , for all  $\ell < s$  we consider the multiset  $D_\ell(u) = \{(d_1^-(v), d_2^-(v)) \mid v \in \text{desc}_\ell(u)\}$ .

**Theorem 15.** *Let  $s \geq 1$  and consider  $\varepsilon$  as a bispecial factor of  $\mathbf{w}^{(s)}$ . One of the following occurs.*

1. For all  $r < s$ ,  $(3, 3) \notin D_r(\varepsilon)$  and all bispecial words in  $\text{desc}(\varepsilon)$  are neutral.
2. There exists  $r < s$  such that  $(3, 3) \in D_r(\varepsilon)$  and one of the following occurs:
  - (a)  $D_r(\varepsilon) = \{(3, 3)\}$ . Furthermore, the bispecial word  $v$  such that  $\text{desc}_r(\varepsilon) = \{v\}$  is neutral.
  - (b)  $D_r(\varepsilon) = \{(2, 2), (3, 3)\}$ . Furthermore, if  $v_1$  and  $v_2$  are the bispecial words such that  $\text{desc}_r(\varepsilon) = \{v_1, v_2\}$ , with  $d_1^-(v_1) = 3$  and  $d_1^-(v_2) = 3$ , then  $v_1$  is neutral,  $v_2$  is ordinary and the longest proper suffix of  $v_2$  is not left special.

Finally, if  $r$  is the greatest such integer, then all bispecial words in the set  $\bigcup_{r < \ell < s} \text{desc}_\ell(\varepsilon)$  are neutral.

**Proposition 16.** *Let  $s \geq 1$  and assume that  $u$  is a non-empty bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_2^-(u) = 2$  and whose longest proper suffix is not left special. Then for all  $r < s$ ,  $\text{desc}_r(u)$  contains a unique bispecial word and this word has the same bispecial multiplicity as  $u$ .*

#### 5.5 Descendance of bispecial factors $u$ with $d_2^-(u) = d_1^-(u) = 3$

By Theorem 15 and Proposition 16, strong and weak bispecial words can only occur as descendant of a neutral bispecial words  $u$  with  $d_2^-(u) = d_1^-(u) = 3$ . In this section we show that such a word  $u$  can have a descendant with the same property and we describe the sub-language of  $\mathcal{L}_B$  that makes this happen.

**Definition 17.** *For  $ijk \in \mathcal{S}_3$ , we define the regular language  $\Gamma_{ijk} = \beta_{kj}\beta_{jk}^+\beta_{ij}$ .*

By definition of  $\mathcal{L}_B$ , we have  $\Gamma_{xyz}\beta_{ij} \subset \mathcal{L}_B$  if and only if  $xyz \in \{ijk, jik, jki\}$ . Furthermore, if  $\sigma_0 \cdots \sigma_s \beta_{ij} \in \mathcal{L}_B$ , then there exists  $r < s$  such that  $\sigma_{[r,s]} = \sigma_r \cdots \sigma_s$  is a suffix of a word in  $\Gamma_{ijk} \cup \Gamma_{jik} \cup \Gamma_{jki}$ . The next result concerns the descendance of a bispecial word with left valence 3 under the application of a product of substitution in some  $\Gamma_{xyz}$ .

**Theorem 18.** *We assume that  $\sigma_s = \beta_{ij}$  for some  $s \geq 1$  and that  $u$  is a neutral bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_1^-(u) = d_2^-(u) = 3$ . We also suppose that there exists  $r < s$  such that  $\sigma_{[r,s]} \in \Gamma_{ijk} \cup \Gamma_{jik} \cup \Gamma_{jki}$ . We have the following.*

1.  $\text{desc}_r(u) = \{v\}$ , where  $v$  is a neutral bispecial factor of  $\mathbf{w}^{(r)}$  such that  $d_1^-(v) = d_2^-(v) = 3$ . In particular,
  - (a) if  $u$  is ordinary in  $\mathbf{w}^{(s)}$ , then  $v$  is ordinary in  $\mathbf{w}^{(r)}$ ;
  - (b) if  $u$  is not ordinary in  $\mathbf{w}^{(s)}$ , then  $v$  is not ordinary in  $\mathbf{w}^{(r)}$ .
2. for all  $\ell$  such that  $r < \ell < s$ ,  $\text{Card}(\text{desc}_\ell(u)) \in \{1, 2\}$  and all bispecial words in  $\text{desc}_\ell(u)$  have left valence 2. Furthermore, if  $\text{Card}(\text{desc}_\ell(u)) = 2$ , then one is the longest proper suffix of the other.

The previous result describes what happens for the descendants of a bispecial factor  $u$  of  $\mathbf{w}^{(s)}$  with  $d_1^-(u) = d_2^-(u) = 3$  and  $\sigma_s = \beta_{ij}$  when some product of substitution  $\sigma_{[r,s]}$  belongs to  $\Gamma_{ijk} \cup \Gamma_{jik} \cup \Gamma_{jki}$ . To describe what happens when this is not the case, we need the following notation: given a language  $L$ ,  $\text{Suff}(L)$  is the set of suffixes of words in  $L$ . For a bispecial word  $v$ , we also let  $m(v)$  denote its bispecial multiplicity.

**Proposition 19.** *We assume that  $\sigma_s = \beta_{ij}$  for some  $s \in \mathbb{N}$  and that  $u$  is a neutral bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_1^-(u) = d_2^-(u) = 3$ . We also assume that there exists  $r < s - 1$  such that  $\sigma_{[r+1,s]} \in \text{Suff}(\Gamma_{ijk} \cup \Gamma_{jik} \cup \Gamma_{jki})$  and  $\sigma_{[r,s]} \notin \text{Suff}(\Gamma_{ijk} \cup \Gamma_{jik} \cup \Gamma_{jki})$ . Then for all  $\ell < s$ , we have the equality of multisets*

$$\{m(v) \mid v \in \text{desc}_\ell(u)\} = \{m(v) \mid v \in \text{desc}_{s-1}(u)\}.$$

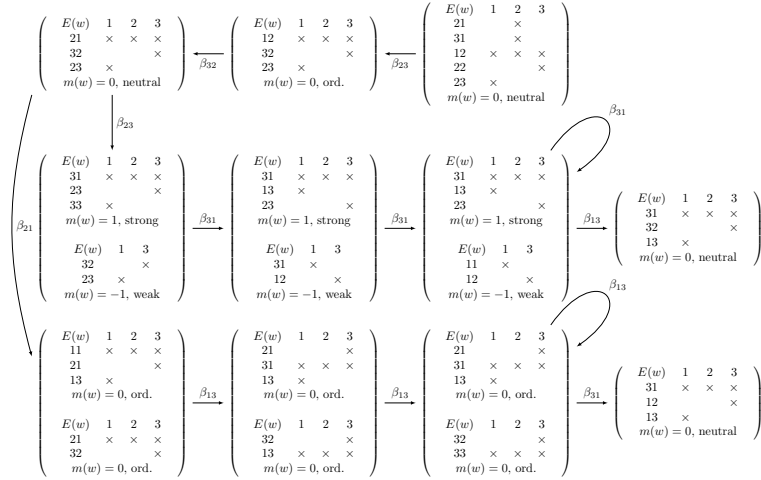
## 5.6 Occurrences of strong and weak bispecial factors

Let  $u$  be a neutral bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_1^-(u) = d_2^-(u) = 3$ . By Theorem 18 and Proposition 19, the bispecial multiplicity of the descendants of  $u$  is completely determined by what happens between two occurrences of a bispecial word  $v$  with  $d_1^-(v) = d_2^-(v) = 3$  in the sequence  $(\text{desc}_r)_{0 \leq r \leq s}$ . The first result of this section shows that strong and weak bispecial words can only appear when  $u$  is not ordinary.

**Proposition 20.** *Assume that  $u$  is an ordinary bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_1^-(u) = d_2^-(u) = 3$ . All bispecial words in  $\text{desc}(u)$  are ordinary.*

If  $u$  is a neutral bispecial, then strong and weak bispecial words can appear in  $\text{desc}(u)$  depending on which letter  $a \in A$  is such that  $d^+(au) = 3$  and which  $\text{Suff}(\Gamma_{xyz})$  the product  $\sigma_{[r+1,s]}$  of Proposition 19 belongs to. This can be explained using Proposition 14 as follows. When we apply a Brun morphism  $\beta_{xy}$  on  $\mathbf{w}^{(s)}$ , the lines  $L_x = E^+(xu)$  and  $L_y = E^+(yu)$  are merged to one line in the extension set of  $\beta_{xy}(u)$ . For the other bispecial extended image  $y\beta_{xy}(u)$ , its extension set has two lines that are copies of  $L_x$  and  $L_y$ . Depending on whether  $a = z$  or  $a \in \{x, y\}$ , we get a pair of strong and weak bispecial words or we get ordinary bispecial words.

**Proposition 21.** *Assume that  $\sigma_s = \beta_{ij}$  for some  $s \geq 1$  and that  $u$  is a neutral non-ordinary bispecial factor of  $\mathbf{w}^{(s)}$  such that  $d_1^-(u) = d_2^-(u) = 3$ . Let also  $a \in A$  such that  $d^+(au) = 3$ . Let finally  $r < s$  such that  $\sigma_{[r,s]} \in \text{Suff}(\Gamma_{xyz})$  with  $xyz \in \{ijk, jik, jki\}$ . One of the following occurs:*



**Fig. 2.** Neutral bispecial words with left valence 3 can split either into two neutral bispecial words or, into a pair of a strong one and a weak one. Above  $\beta_{32}\beta_{23}$  is applied on  $\varepsilon$ . Then we can apply morphisms in  $\Gamma_{321} \cup \Gamma_{231} \cup \Gamma_{213}$ . The figure illustrates  $\Gamma_{213} = \beta_{31}\beta_{13}^+\beta_{21}$  and  $\Gamma_{231} = \beta_{13}\beta_{31}^+\beta_{23}$  where strong and weak bispecial factors are created.

1.  $a \in \{x, y\}$  and all bispecial words in  $\text{desc}_{s-1}(u)$  are ordinary;
2.  $a = z$  and  $\text{desc}_{s-1}(u) = \{v_1, v_2\}$ , where  $m(v_1) = +1$  and  $m(v_2) = -1$ . In particular, we have  $v_1 = \beta_{xy}(u)$  and  $v_2 = y\beta_{xy}(u)$ .

We now give an exemple that illustrates all results that we obtained (see Figure 2). In that example, we describe the first elements of the sequence

$$(\text{desc}_s(\varepsilon), \text{desc}_{s-1}(\varepsilon), \text{desc}_{s-2}(\varepsilon), \dots),$$

where  $\varepsilon$  is considered as a bispecial of  $\mathbf{w}^{(s)}$  whose 2-extension set corresponds to  $E_1$  in Theorem 12. In this example, the extension set on the top left is the one of a bispecial word  $u$  with  $d_2^-(u) = d_1^-(u) = 3$ . We illustrate the fact that the multiplicity of its descendants depends on which  $\Gamma_{xyz}$  the product  $\sigma_{[r,s]}$  belongs to.

## 6 Further Work

The results of the paper allow to understand how can appear strong and weak bispecial factors in a Brun word. These are preliminary results to perfectly understand the factor complexity of a Brun word. The missing information to complete this knowledge concerns the length of bispecial words. To ensure a linear complexity, we need to prove that strong and weak bispecial factors are “well distributed” in the sequence of bispecial factors ordered by length. Experimentally, strong and weak bispecial factors come by pairs and alternate. This is

supported by Proposition 21 where we show that when they appear, the strong one is a suffix of the weak one. However this property is not preserved under the application of Brun morphisms so more work needs to be done. If strong and weak bispecial words indeed alternate, then the factor complexity of any Brun word is always between  $2n + 1$  and  $3n + 1$ . Like in the Arnoux-Rauzy-Poincaré system, the upper bound should be improvable to  $\frac{5}{2}n + 1$ .

## References

1. Akiyama, S., Barge, M., Berthé, V., Lee, J.Y., Siegel, A.: On the Pisot substitution conjecture. Preprint (2014)
2. Arnoux, P., Rauzy, G.: Représentation géométrique de suites de complexité  $2n+1$ . *Bull. Soc. Math. France* 119(2), 199–215 (1991)
3. Avila, A., Delecroix, V.: Some monoids of Pisot matrices. preprint (Jun 2015), <http://arxiv.org/abs/1506.03692>
4. Berthé, V., Bourdon, J., Jolivet, T., Siegel, A.: A combinatorial approach to products of Pisot substitutions. *Ergodic Theory Dyn. Syst. FirstView*, 1–38 (Mar 2015), [http://journals.cambridge.org/article\\_S0143385714001412](http://journals.cambridge.org/article_S0143385714001412)
5. Berthé, V., Labbé, S.: Factor complexity of  $S$ -adic words generated by the Arnoux-Rauzy-Poincaré algorithm. *Adv. in Appl. Math.* 63, 90–130 (2015)
6. Berthé, V., Rigo, M. (eds.): *Combinatorics, Automata and Number Theory, Encyclopedia of Mathematics and its Applications*, vol. 135. Cambridge University Press, Cambridge (2010)
7. Berthé, V., Steiner, W., Thuswaldner, J.M.: Geometry, dynamics, and arithmetic of  $S$ -adic shifts. preprint (2014), <http://arxiv.org/abs/1410.0331>
8. Brentjes, A.J.: *Multidimensional continued fraction algorithms*. Mathematisch Centrum, Amsterdam (1981)
9. Brun, V.: Algorithmes euclidiens pour trois et quatre nombres. In: *Treizième congrès des mathématiciens scandinaves, tenu à Helsinki 18-23 août 1957*, pp. 45–64. Mercators Tryckeri, Helsinki (1958)
10. Delecroix, V., Hejda, T., Steiner, W.: Balancedness of Arnoux-Rauzy and Brun words. In: *WORDS. LNCS*, vol. 8079, pp. 119–131. Springer (2013)
11. Durand, F.: Corrigendum and addendum to: “Linearly recurrent subshifts have a finite number of non-periodic subshift factors”. *Ergodic Theory Dyn. Syst.* 23, 663–669 (2003)
12. Klouda, K.: Bispecial factors in circular non-pushy D0L languages. *Theoret. Comput. Sci.* 445, 63–74 (2012), <http://dx.doi.org/10.1016/j.tcs.2012.05.007>
13. Lagarias, J.C.: The quality of the Diophantine approximations found by the Jacobi-Perron algorithm and related algorithms. *Monatsh. Math.* 115(4), 299–328 (1993), <http://dx.doi.org/10.1007/BF01667310>
14. Morse, M., Hedlund, G.A.: Symbolic Dynamics II. Sturmian Trajectories. *American Journal of Mathematics* 62(1), 1–42 (Jan 1940), <http://www.jstor.org/stable/2371431>
15. Rauzy, G.: Nombres algébriques et substitutions. *Bulletin de la Société Mathématique de France* 110, 147–178 (1982)
16. Schratzberger, B.R.: The quality of approximation of Brun’s algorithm in three dimensions. *Monatsh. Math.* 134(2), 143–157 (2001), <http://dx.doi.org/10.1007/s006050170004>
17. Schweiger, F.: *Multidimensional Continued Fraction*. Oxford Univ. Press, New York (2000)