**Theoretical Informatics and Applications** Will be set by the publisher
Informatique Théorique et Applications

# STUDY OF IRREDUCIBLE BALANCED PAIRS FOR SUBSTITUTIVE LANGUAGES

Bernat Julien

**Abstract.** Let $\mathcal{L}$ be a language. A balanced pair $(u, v)$ consists of two words $u$ and $v$ in $\mathcal{L}$ which have the same number of occurences of each letter. It is irreducible if the pairs of strict prefixes of $u$ and $v$ of the same length do not form balanced pairs.

In this article, we are interested in computing the set of irreducible balanced pairs on several examples. We make connections with the balanced pairs algorithm and discrete geometrical constructions related to substitutive languages. We characterize substitutive languages which admit infinitely many irreducible balanced pairs of a given form.

**AMS Subject Classification.** — Give AMS classification codes —.

## Keywords

## Introduction

This article deals with substitutions, that is, free morphisms on the monoid $\mathcal{A}^*$ generated by a finite set $\mathcal{A}$ with the concatenation. A substitution $\sigma$ naturally defines a symbolic dynamical system $(\mathcal{X}_\sigma, S)$, which may be splitted in topological factors. In the following, we only consider primitive substitutions; in this case, $(\mathcal{X}_\sigma, S)$ admits a unique topological factor.

When $\sigma$ is a $d$-letter Pisot type substitution, it is possible to construct a geometrical representation of $(\mathcal{X}_\sigma, S)$ known as the Rauzy fractal of the substitution. The Rauzy fractal $\mathcal{T}$ is a compact subset of $\mathbb{R}^{d-1}$ which is equal to the adherence of its inner points, and which has positive Lebesgue measure ( [35]). Furthermore, $(\mathcal{X}_\sigma, S)$ admits a minimal translation on the torus $\mathbb{R}^d$ / $\mathcal{T}$ as a topological factor ( [13]).

When the substitution $\sigma$ is unimodular, several combinatorial conditions known as coincidence conditions provide additional knowledge on $(\mathcal{X}_\sigma, S)$. Under the strong coincidence property, $(\mathcal{X}_\sigma, S)$ is measure-theoretically isomorphic to the exchange of $d$ domains defined almost everywhere on the associated Rauzy fractal ( [5]). When the super-coincidence property holds (see [22]), the Rauzy fractal generates a periodic tiling of the space, that is, there exists a lattice $\Lambda$ such that $(\mathcal{X}_\sigma, S)$ admits a toral representation as a fundamental domain for $\mathbb{R}^{d-1}$ / $\Lambda$, see [5]. Among unimodular Pisot type substitutions, the super-coincidence property is also equivalent to $(\mathcal{X}_\sigma, S)$ having a discrete spectrum, or being metrically isomorphic to a translation on a compact abelian group (see [21, 29, 30]). Note that the super coincidence condition implies the strong coincidence condition. At the moment, we do not know any Pisot type unimodular substitution which does not satisfy either the super-coincidence or the strong coincidence property.

We are interested in the connection between the study of coincidence conditions and combinatorial properties of $\mathcal{L}_\sigma$, the language generated by $\sigma$. For a language $\mathcal{L}$ and $u, v \in \mathcal{L}$, we say that the pair $(u, v)$ is *balanced* if:

$$\text{for any } l \in \mathcal{A}_\sigma, |u|_l = |v|_l.$$

Balanced pairs $(l, l)$, where $l \in \mathcal{A}_\sigma$, are called *trivial pairs*. The notion of balanced pairs has been introduced by Queffélec in [30], following an idea of Host. Note that, by definition, $(u, v)$ cannot be a balanced pair if $|u| \neq |v|$. Hence we define the *length* of the balanced pair $(u, v)$ as the quantity $|u| = |v|$.

**Remark 0.1.** *Trivial pairs are introduced in [34] as* coincidences. *We prefer to avoid this term, since it has an other meaning in the framework of substitutive dynamical systems.*

When $(u, v)$ is a balanced pair with $|u| \geqslant 2$, we say that $(u, v)$ is an *irreducible balanced pair* if the pairs of strict prefixes of $u$ and $v$ are not balanced. For any balanced pair $(u, v)$, we call *reduction* of $(u, v)$ the set of irreducible balanced pairs $(u^{(i)}, v^{(i)})_{i \in [\![1,m]\!]}$ such that $u = u^{(1)} \ldots u^{(m)}$ and $v = v^{(1)} \ldots v^{(m)}$. Obviously, $(u, v)$ is irreducible if and only if $m = 1$.

**Remark 0.2.** *We consider that trivial pairs are irreducible balanced pairs as well. However, we choose the convention that $(\varepsilon, \varepsilon)$ is not an irreducible balanced pair.*

**Example 0.3.** *Let $\mathcal{A} = \{a, b\}$. The pair $(aab, baa)$ is irreducible. The pair $(abaaba, babaaa)$ is reducible. The reduction of $(abaaba, babaaa)$ is $\{(ab, ba), (aab, baa), (a, a)\}$.*

Any substitution $\sigma$ defines an action on irreducible balanced pairs. This action may be represented by the balanced pairs algorithm, detailed in Section 3. Queffélec noticed in [30] that, if the balanced pairs algorithm associated with a substitutive language terminates, and if the dominant eigenvalue of the incidence matrix of the algorithm is less than the dominant eigenvalue of the incidence matrix of $\sigma$, then $(\mathcal{X}_\sigma, S)$ has a (purely) discrete spectrum. More recently, it is proven in the forthcoming study [10] that, for any Pisot type substitution, the super-coincidence holds if and only if the strong coincidence holds and the balanced pairs algorithm terminates. Hence the study of irreducible balanced pairs associated with substitutive languages, and the action of $\sigma$ on these pairs, is closely related to the determination of combinatorial and ergodic properties for $(\mathcal{X}_\sigma, S)$; see also [7, 8, 26, 34].

This article is structured in the following form. In Section 1, we introduce the definitions and notation. In Section 2, we study the set of irreducible balanced pairs for Sturmian languages, and for languages associated with Arnoux-Rauzy words. We prove that the set of irreducible balanced pairs for the Fibonacci case may be explicitly computed (Corollary 2.3). Although we are not able to compute the whole set of irreducible balanced pairs for Arnoux-Rauzy words, we compute a particular class of irreducible balanced pairs with Proposition 2.5.

In Section 3, we introduce the balanced pair algorithm. We study with Proposition 3.4 the action of the balanced pair algorithm on the Fibonacci example. Finally, we establish a connection with the discrete geometrical representation of Rauzy fractals in Section 4. Notably, we characterize with Theorem 4.9 substitutive languages that admit infinitely many irreducible balanced pairs. We end our study by a non-exhaustive list of questions in Section 5.

## 1. Definitions

For convenience, we denote by $[\![i,j]\!]$ the set of integers $k$ such that $i \leqslant k \leqslant j$.

### 1.1. Words

Let $\mathcal{A}$ be an *alphabet*, that is, a finite set of elements called *letters*. Endowed with the concatenation map, $\mathcal{A}$ generates a free monoid that is denoted by $\mathcal{A}^*$. We denote by $\varepsilon$ the empty word. A *language* $\mathcal{L}$ is a subset of $\mathcal{A}^*$; its elements are the *finite words*. Note that any word $u$, finite or infinite, naturally defines a language if we consider the set of *factors* of $u$, that is, the blocs of consecutive letters occuring in $u$.

Let $u = u_1 \dots u_n \in \mathcal{A}^*$; $n$ is the *length* of $u$. For any $i \in [\![0,n]\!]$, $\mathrm{pref}_i(u) = u_1 \dots u_i$ is the *prefix* of length $i$ of $u$. We set the *center* of $u$ as $\varepsilon$ if $n$ is even, or as the letter $u_{\frac{n+1}{2}}$ if $n$ is odd. The *mirror image* is the map: $\mathcal{A}^* \to \mathcal{A}^*, u_1 \dots u_n \longmapsto u_n \dots u_1$. Fixed points for the mirror image map are called *palindromes*.

For any $l \in \mathcal{A}$, we denote by $|u|_l$ the number of occurences of the letter $l$ in $u$. The language $\mathcal{L}$ is said to be $k$-balanced, with $k \in \mathbb{N}^*$, if for all words $u, v \in \mathcal{L}$, one has $\max_{l \in \mathcal{A}} \{|u|_l - |v|_l\} \leqslant k$. See also [1, 37] for a study of balanced languages.

**Remark 1.1.** *A language that is* 1*-balanced is often said* balanced. *We do not use this terminology, which stands in our study for a property on pairs of words.*

For any integer $n \in \mathbb{N}^*$ and any language $\mathcal{L}$, we set $i_{\mathcal{L}}(n)$ as the number of irreducible balanced pairs in $\mathcal{L}$ of length $n$. The *complexity map* of the language $\mathcal{L}$ is the map $p_{\mathcal{L}} : \mathbb{N} \to \mathbb{N}$, $p_{\mathcal{L}}(n)$ being the number of distinct words of length $n$ in $\mathcal{L}$. The language $\mathcal{L}$ is said to be *Sturmian* if $\mathcal{L}$ is the set of factors of a one-sided sequence, such that $p_{\mathcal{L}}(n) = n + 1$ for any $n \in \mathbb{N}$.

Let $u \in \mathcal{L}$. If there exist $a, b \in \mathcal{A}$ such that $au$ and $bu \in \mathcal{L}$ (respectively $ua$ and $ub \in \mathcal{L}$), $u$ is said to be a *left special factor* (resp. *right special factor*). A left (right) special factor $u$ is *total* if for any $a \in \mathcal{A}$, $au \in \mathcal{L}$ ($ua \in \mathcal{L}$). A *(total) bispecial factor* is a word which is both (total) left special and (total) right special.

When $u$ is an infinite word, we say that $u$ is a *left special factor* if any prefix of $u$ is a left special factor.

An infinite word $u$ is said to be *uniformly recurrent* if, for any factor $w$, there exists $C \in \mathbb{N}$ such that any word of length $C$ admits $w$ as a factor. Let $\mathcal{L}$ be the language which consists of the factors of a uniformly recurrent word $w$. We say that $w$ is an *Arnoux-Rauzy* word ( [6]) if, for any $n \in \mathbb{N}^*$, there exist a unique left-special factor and a unique right-special factor of length $n$, which are both total. In the following, we call *Arnoux-Rauzy language* the language which consists of the set of factors of an Arnoux-Rauzy word. An Arnoux-Rauzy word $u$ is *characteristic* if the set of left special factors coincides with the set of prefixes of $u$.

It is proven in [31] that, for any Arnoux-Rauzy language and any $n \in \mathbb{N}$, the right-special factor of length $n$ is the mirror image of the left-special factor of the same length. As a consequence, any Arnoux-Rauzy language is stable under mirror image.

As a generalization of the study of Rauzy graphs associated with Arnoux-Rauzy words on a three-letter alphabet [6], one gets the following result:

**Proposition 1.2.** *Any Arnoux-Rauzy language admits infinitely many bispecial words.*

### 1.2. Substitutions

Let $\mathcal{A}$ be the $d$-letter alphabet $\{a_i\}_{i\in[\![1,d]\!]}$. A *substitution* is a morphism of monoid on $\mathcal{A}^*$. The substitution $\sigma$ is said *non-erasing* if for every $a \in \mathcal{A}$, $\sigma(a) \neq \varepsilon$. The *incidence matrix* of $\sigma$ is defined as the square matrix $M_\sigma \in \mathcal{M}_d(\mathbb{N})$ such that, for every $(i,j) \in [\![1,d]\!]^2$, $M_\sigma[i,j] = |\sigma(a_j)|_{a_i}$.

The substitution $\sigma$ is *primitive* if there exists $n \in \mathbb{N}$ such that for every $(i,j) \in [\![1,d]\!]^2$, $M_\sigma^n[i,j] \geqslant 1$. We denote by $\mathcal{L}_\sigma$ the set of factors of the words $\{\sigma^k(l), l \in \mathcal{A}, k \in \mathbb{N}\}$. When $\sigma$ is primitive, $\mathcal{L}_\sigma$ is the set of factors of $\{\sigma^k(l), k \in \mathbb{N}\}$ for any letter $l \in \mathcal{A}$. The substitution $\sigma$ is said to be *unimodular* if $|\det M_\sigma| = 1$. See [30] for a study of the properties of dynamical systems generated by primitive substitutions.

The substitution $\sigma$ is said to be *of Pisot type* if the eigenvalues of the incidence matrix $M_\sigma$ satisfy the following: there exists a dominant eigenvalue $\beta > 1$ such that, for every other eigenvalue $\alpha$, one has $0 < |\alpha| < 1$. Note that the characteristic polynomial of $M_\sigma$ is irreducible when $\sigma$ is of Pisot type. Let us remind that a *Pisot number* is an algebraic integer such that any Galois conjugate $\alpha \neq \beta$ satisfies $|\alpha| < 1$. In particular, the dominant eigenvalue of $M_\sigma$ is a Pisot number when $\sigma$ is of Pisot type. It is proven in [3] that, for any Pisot type substitution $\sigma$, there exists $k \in \mathbb{N}$ such that $\mathcal{L}_\sigma$ is $k$-balanced.

For any pair of words $(u,v)$, let us denote $(\sigma(u), \sigma(v))$ by $\sigma(u,v)$ for convenience. Let $\sigma$ be a substitution and $(u,v)$ be a balanced pair. Then $\sigma^k(u,v)$ is a balanced pair as well for any $k \in \mathbb{N}^*$. If there exists $k \in \mathbb{N}^*$ such that the reduction of $\sigma^k(u,v)$ contains a trivial pair, one says that $(u,v)$ *leads to a coincidence*.

## 2. Irreducible balanced pairs on several examples

First, we study the case of the Fibonacci substitution; we explicitly compute the set of irreducible balanced pairs for this example. Then, we compute a subset of irreducible balanced pairs for a family of substitutions which generalize the Fibonacci substitution.

### 2.1. Sturmian languages

We consider here the case of Sturmian languages. First, let us remind several well-known results concerning Sturmian sequences and their corresponding languages (see for instance Chapter 6 in [29]):

(1) any Sturmian language is closed under the mirror image map,
(2) a two-sided sequence is Sturmian if and only if it is balanced and not eventually periodic.

**Lemma 2.1.** *Let $\mathcal{L}$ be a 1-balanced language on $\{a,b\}$. Then the set of irreducible balanced pairs of length $\geqslant 2$ is the set of pairs $\{(aub, bua); u$ bispecial$\}$.*

**Proof** Let $(u,v)$ be an irreducible balanced pair of length $\geqslant 2$. By definition, $u$ and $v$ admit distincts strict prefixes and suffixes; we may additionally assume without loss of generality that $u = au'$ and $v = bv'$. Let $p$ be the longest common prefix of $u'$ and $v'$. If $u'$ admits $pb$ as a strict prefix and $v'$ admits $pa$ as a strict prefix, then $(u,v)$ is not irreducible since it admits $(apb, bpa)$ as a balanced pair in its reduction. On the other hand, if $u'$ admits $pa$ as a prefix and $v'$ admits $pb$ as a prefix, then $apa$ and $bpb$ both belong to $\mathcal{L}_\sigma$, which contradicts the fact that $\mathcal{L}_\sigma$ is 1-balanced. Hence $u' = pb$ and $v' = pa$, that is, $(u, v) = (apb, bpa)$, and $p$ is bispecial.

Conversely, let $u$ be a bispecial word. Since $\mathcal{L}$ is 1-balanced, $aua$ and $bub$ cannot both belong to $\mathcal{L}$. Since $\mathcal{L}$ is extendable, $aub$ and $bua \in \mathcal{L}$. Hence $(aub, bua)$ is an irreducible balanced pair, which ends the proof. □

Note that, since there exists a unique bispecial word of a given length, we obtain the following corollary, where $i(n)$ denotes the number of irreducible balanced pairs of length $n$.

**Corollary 2.2.** *Let $\mathcal{L}$ be a Sturmian language. Then one has $i(n) \in \{0, 1\}$ for any $n \geqslant 2$.*

Sturmian languages may be generated by substitutions. This is for instance the case of the *Fibonacci substitution* $\sigma$, defined on the two-letter alphabet $\{a, b\}$ as $\sigma(a) = ab$ and $\sigma(b) = a$. The Fibonacci substitution admits a unique right-sided fixed point, that we denote by $\omega$, known as the *Fibonacci word*. In particular, $\omega$ is a characteristic Arnoux-Rauzy word. See also [11, 15, 16, 25, 29] for further details concerning the Fibonacci substitution and Sturmian sequences.

**Corollary 2.3.** *The irreducible balanced pairs of the Fibonacci languages are $(a\omega^{(n)}b, b\omega^{(n)}a)$, where $(\omega^{(n)})_{n\in\mathbb{N}}$ is the sequence of palindromic prefixes of $\omega$.*

**Proof** In a Sturmian language, bispecial words are exactly left special factors that are palindromes. Since the Fibonacci word $\omega$ is characteristic, bispecial words are exactly the palindromes that are prefixes of $\omega$. Hence, due to Lemma 2.1, any irreducible balanced pair is of the form $(a\omega^{(n)}b, b\omega^{(n)}a)$. □

### 2.2. Arnoux-Rauzy words and confluent Parry unit substitutions

As Arnoux-Rauzy words generalize in some sense Sturmian sequences, we are interested in determining which properties satisfied by Sturmian languages still hold for Arnoux-Rauzy languages. Note that, due to [14], there exist Arnoux-Rauzy words that are not $k$-balanced for any $k \in \mathbb{N}^*$, which points out that Arnoux-Rauzy languages and Sturmian languages may admit significant differences.

*Parry numbers*, introduced in [28, 33], are numbers $\beta > 1$ for which the $T_\beta$-orbit of 1 is finite, where $T_\beta$ is the map:

$$T_\beta : [0, 1] \to [0, 1], x \longmapsto \beta x \mod 1.$$

When the $T_\beta$-orbit of 1 ends in 0, $\beta$ is said to be a *simple Parry number*, a *non-simple Parry number* otherwise.

A particular class of Parry numbers is the set of *confluent Parry numbers*. They are the positive roots of polynomials $X^d - \sum_{i=1}^{d-1} kX^i - l$, where $k, l \in \mathbb{N}^*$ with $k \geqslant l$ and $d \geqslant 2$. Confluent Parry numbers are introduced in [20] and mainly studied in [2, 9, 27].

It is possible to define for any Parry number an associated substitution ( [19, 36]), which admits a unique right-sided fixed point $\omega$. This property implies that the strong coincidence condition is satisfied. In the particular case of confluent Parry numbers, we obtain the class of *confluent Parry substitutions*, which are substitutions $\sigma$ defined on the $d$-letter alphabet

$\{a_i\}_{i\in[\![1,d]\!]}$ by:

$$\sigma(a_i) = a_1^k a_{i+1} \text{ if } i < d \text{ and } \sigma(a_d) = a_1^l. \tag{1}$$

Among the class of Parry substitutions, confluent Parry substitutions are exactly those for which any of the following equivalent properties hold ( [2]):

$$\omega \text{ admits infinitely many prefixes that are palindromes,} \tag{2}$$

$$\mathcal{L}_\sigma \text{ is stable under mirror image.} \tag{3}$$

Moreover, due to [2], confluent Parry unit substitutions (for which $l = 1$) define characteristic Arnoux-Rauzy sequences.

**Example 2.4.** *The 3-letter substitution $\sigma$ defined as $\sigma(a) = ab$, $\sigma(b) = ac$, $\sigma(c) = a$, known as the* Tribonacci substitution*, is a confluent Parry unit substitution, with $k = 1$ and $d = 3$.*

The following proposition provides a partial generalization of Corollary 2.3.

**Proposition 2.5.** *Let $\mathcal{L}$ be an Arnoux-Rauzy language. There are infinitely many irreducible balanced pairs of the form $(lul', l'ul)$, where $u$ is a total bispecial factor and $l, l'$ are distincts letters.*

**Proof** Let $\mathcal{L}$ be an Arnoux-Rauzy language. Due to Proposition 1.2, there exist infinitely many bispecial words $(u^{(n)})_{n\in\mathbb{N}}$; moreover these bispecial words are palindromes. Let $n \in \mathbb{N}$. Since there exists a unique left special factor $w$ of length $k$ for any $k \in \mathbb{N}$, the left special factor $w$ of length $|u^{(n)}| + 1$ admits $u^{(n)}$ as a prefix. Hence there exists $l \in \mathcal{A}$ such that $w = u^{(n)}l$; additionally, there exists $l' \neq l$ such that $l'u^{(n)}l \in \mathcal{L}$. Since the language generated by an Arnoux-Rauzy word is stable under mirror image, one has $lu^{(n)}l' \in \mathcal{L}$. As a consequence, $(lu^{(n)}l', l'u^{(n)}l)$ is an irreducible balanced pair. □

**Remark 2.6.** *The irreducible balanced pairs computed in Corollary 2.3 and Proposition 2.5 are of the form $(aub, bua)$, and of the form $(u, \widetilde{u})$ as well. However, there may exist irreducible balanced pairs that are not of these forms in the case of Arnoux-Rauzy languages. For instance, $(cababa, baabac)$ is an irreducible balanced pair for the language generated by the Tribonacci substitution.*

*Note that there does not exist an irreducible balanced pair of any length. For instance, the Tribonacci language does not contain any irreducible balanced pair of length 8.*

## 3. Balanced pairs algorithm

First, we introduce in Section 3.1 the balanced pairs algorithm. Then, we study in Section 3.2 the action of the balanced pairs algorithm for the Fibonacci language.

### 3.1. Definition of the balanced pairs algorithm

Let us recall the notion of return word. For any recurrent infinite word $u$ and any factor $w$, a *return word* $v$ is a finite word such that $vw$ is a factor of $u$, $w$ is a prefix of $v$ and $w$ occurs exactly twice in $vw$. In particular, if there exists $C > 0$ such that, for any factor $w$ of $u$, any return word $v$ satisfies $|v| < C|w|$, the language is said *linearly recurrent*. Durand has proven in [17] that primitive substitutions generate linearly recurrent languages.

Let $\sigma$ be a primitive substitution. Let $\omega$ be a right-sided $\sigma$-periodic point, and $w$ be a non-empty prefix of $\omega$. Since $\sigma$ is primitive, there exist $N \in \mathbb{N}$ and a finite set of return words $(X_i)_{i\leqslant N}$ for $w$ (see [17]). Each pair $(wX_i, X_iw)$ splits in finitely many irreducible balanced pairs. Hence $I_1(w)$, the set of irreducible balanced pairs obtained when reducing $(S^{|w|}\omega, \omega)$, is finite. Then, one defines by recurrence the set $I_{n+1}(w)$ as the set of irreducible balanced pairs occuring in the reduction of $(\sigma(u), \sigma(v))$ for all irreducible balanced pairs $(u, v)$ of $I_n(w)$. Let $I(w) = \bigcup_{n\in\mathbb{N}^*} I_n(w)$. When $I(w)$ is finite, it is said that the balanced pair algorithm associated with the prefix $w$, denoted by $A(w)$, terminates.

**Example 3.1.** *Let $\sigma$ be the Fibonacci substitution and $w = a$. Return words for $a$ are $a$ and $ab$. One has $I_1(a) = \{(a, a), (ab, ba)\}$, $I_2(a) = \{(a, a), (ab, ba), (b, b)\}$ and $I_3(a) = I_2(a)$. Hence $I(a) = I_2(a)$ and $A(a)$ terminates.*

Let us study more closely the following example, introduced by Martensen as Example 3.2 in [26]. Let $\sigma$ be defined on the two-letter alphabet $\{a, b\}$ by $\sigma(a) = aab$, $\sigma(b) = abb$. The words $\sigma(a)$ and $\sigma(b)$ both admit $a$ as a prefix and $b$ as a suffix. Let $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ be the sequences defined by $u_0 = ab$, $v_0 = ba$, such that for all $n \in \mathbb{N}$, $au_{n+1}b = \sigma(u_n)$ and $av_{n+1}b = \sigma(v_n)$. Note that, by definition of $\sigma$, for any $w = w_1 \dots w_n$, one has $\sigma(w) = w'$, where $|w'| = 3n$ with $w'_i = a$ if $i = 1 \mod 3$, $w'_i = b$ if $i = b \mod 3$ and $w'_i = w_{\frac{i+1}{3}}$ if $i = 2 \mod 3$.

Let $(u, v)$ be an irreducible balanced pair. Assume that $|\mathrm{pref}_k(u)|_a > |\mathrm{pref}_k(v)|_a$ for all $k \in [\![1, |u| - 1]\!]$. Then, by definition of $\sigma$, one has $|\mathrm{pref}_k(\sigma(u))|_a > |\mathrm{pref}_k(\sigma(v))|_a$ for any $k \in [\![2, |\sigma(u)| - 2]\!]$. As a consequence, for any $n \in \mathbb{N}$, $(u_n, v_n)$ is an irreducible balanced pair, and the reduction of $(\sigma(u_n), \sigma(v_n))$ is exactly $\{(a, a), (u_{n+1}, v_{n+1}), (b, b)\}$. We deduce that, if $i, k \in \mathbb{N}^*$ are such that $(u_i, v_i)$ occurs in $I_k(w)$, then, for any $n \in \mathbb{N}$, $(u_{i+n}, v_{i+n})$ occurs in $I_{k+n}(w)$. Hence, for any prefix $w \neq \varepsilon$ of $\omega$, $A(w)$ does not terminate.

**Remark 3.2.** *We do not know whether there exists a primitive substitution for which $I(w)$ has finitely many elements for any prefix $w$, but $I(w)$ is not uniformly bounded.*

The balanced pairs algorithm enables an algorithmic criterion which determines whether the Rauzy fractal of a given Pisot type substitution generates a periodic tiling of the space. More precisely, one has the following result initially obtained by Livshits:

**Theorem 3.3.** *[24, 34] Let $\sigma$ be a primitive substitution such that $\mathrm{pref}_1(\sigma(a)) = a$. Let $\omega = \sigma^\infty(a)$.*

1. *If there exists $i \in \mathbb{N}^*$ such that $A(\mathrm{pref}_i(\omega))$ terminates, and such that any balanced pair in $I(\mathrm{pref}_i(\omega))$ leads to a coincidence, then the spectrum of $(\mathcal{X}_\sigma, S)$ is discrete.*
2. *If there exists $i \in \mathbb{N}^*$ such that $\mathrm{pref}_{i+1}(\omega) = \mathrm{pref}_i(\omega)a$, and if the spectrum of $(\mathcal{X}_\sigma, S)$ is discrete, then any balanced pair in $I(\mathrm{pref}_i(\omega))$ leads to a coincidence.*

### 3.2. Action of the balanced pairs algorithm on the irreducible balanced pairs of Fibonacci

We prove here with Proposition 3.4 that, for the Fibonacci case, for any $i \in \mathbb{N}^*$, $I(\text{pref}_i(\omega))$ may contain at most 4 elements, which are $(a,a),(b,b),(ab,ba)$ and $(aab,baa)$. As a consequence, the balanced pairs algorithm $A(w)$ terminates in the Fibonacci case for any non-empty prefix $w$ of $\omega$.

As we have seen in Corollary 2.3, the irreducible balanced pairs for the Fibonacci language are of the form $(a\omega^{(n)}b, b\omega^{(n)}a)$. The following Proposition describes the action of the Fibonacci substitution $\sigma$ on the set of irreducible balanced pairs, where $(F_n)_{n\in\mathbb{Z}}$ denotes the Fibonacci sequence, that is, the sequence defined by $F_{n+2} = F_{n+1} + F_n$ for any $n \in \mathbb{Z}$, with $F_0 = F_1 = 1$.

**Proposition 3.4.** *Let $n \in \mathbb{N}$. The reduction of $\sigma(a\omega^{(n)}b, b\omega^{(n)}a)$ splits in:*

1. *one occurence of the trivial pair $(a,a)$,*
2. *$F_n + 1$ occurences of the irreducible balanced pairs $(ab,ba)$,*
3. *$F_{n+1} - 1$ occurences of the irreducible balanced pairs $(aab,baa)$.*

The more general case of the computation of the reduction of a balanced pair $(u,v)$ is depicted in Figure 1. Let us introduce the following definition, which is useful for the proof of the Proposition.

**Definition 3.5.** *Let $u$ and $v$ be two finite words. The pair $(vu,u)$ is a* potential balanced pair *if there exists a finite word $w$ such that $(vuw,uwv)$ is a balanced pair.*

**Proof** The image of $(u,v) = (a\omega^{(n)}b, b\omega^{(n)}a)$ under $\sigma$ is $(ab\sigma(\omega^{(n)})a, a\sigma(\omega^{(n)})ab)$. Note that, since $\omega^{(n)}$ is both bispecial and a palindrome, it admits $a$ as a prefix and as a suffix.

Since $u_1 = a$ and $v_1 = b$, $\sigma(u_1,v_1)$ splits in $(a,a)$ and $(b,\varepsilon)$. For any $i \in [\![1, |\omega^{(n)}|]\!]$, let $l_i$ denote the $i$-th letter of $\omega^{(n)}$. First, suppose that $i \neq |\omega^{(n)}|$.

If $l_i = a$, and if the reduction of $\sigma(al_1\dots l_{i-1}, bl_1\dots l_{i-1})$ provides the potential balanced pair $(b,\varepsilon)$, then the reduction of $\sigma(al_1\dots l_i, bl_1\dots l_i)$ provides the potential balanced pair $(bab,ab)$. This latter potential balanced pair splits in $(ba,ab)$ and $(b,\varepsilon)$. If $l_i = b$, and if the reduction of $\sigma(al_1\dots l_{i-1}, bl_1\dots l_{i-1})$ provides the potential balanced pair $(b,\varepsilon)$, then the reduction of $\sigma(al_1\dots l_i, bl_1\dots l_i)$ provides the potential balanced pair $(ba,a)$. Since $bb$ is not a word in the Fibonacci language, one has $l_{i+1} = a$, and the reduction of $\sigma(al_1\dots l_{i+1}, bl_1\dots l_{i+1})$ provides the potential balanced pair $(baab,aab)$, which splits in $(baa,aab)$ and the potential balanced pair $(b,\varepsilon)$.

If $i = |\omega^{(n)}|$, then $l_i = a$ and the potential balanced pair $(b,\varepsilon)$ obtained when reducing $\sigma(al_1\dots l_{i-1}, bl_1\dots l_{i-1})$ is completed as $(ba,ab)$. Hence the reduction of $\sigma(a\omega^{(n)}b, b\omega^{(n)}a)$ splits in one occurence of the trivial pair $(a,a)$, $1 + |\omega^{(n)}|_a - |\omega^{(n)}|_b$ irreducible balanced pairs of the form $(ab,ba)$ and $|\omega^{(n)}|_b$ irreducible balanced pairs of the form $(aab,baa)$.

Finally, since for any $n \in \mathbb{N}$, $\omega^{(n+1)} = \sigma(\omega^(n))a$, we have $|\omega^{(n+1)}|_a = |\omega^{(n)}|_a + |\omega^{(n+1)}|_b + 1$ and $|\omega^{(n+1)}|_b = |\omega^{(n)}|_a$. Since $\omega^{(0)} = \varepsilon$ and $\omega^{(1)} = a$, we obtain that, for any $n \in \mathbb{N}$, one has $|\omega^{(n)}|_a = F_{n+2} - 1$ and $|\omega^{(n)}|_b = F_{n+1} - 1$. Hence $|\omega^{(n)}|_a - |\omega^{(n)}|_b + 1 = F_{n+2} - F_{n+1} + 1 = F_n + 1$, which ends the proof. □

As a direct consequence of Proposition 3.4, it is possible to compute all the irreducible balanced pairs produced by the algorithm $A(w)$, independently from $w$.

**Corollary 3.6.** *For any prefix $w$ of $\omega$, the algorithm $A(w)$ terminates, and one has $I(w) \subset \{(a,a),(b,b),(ab,ba),(aab,baa)\}$.*

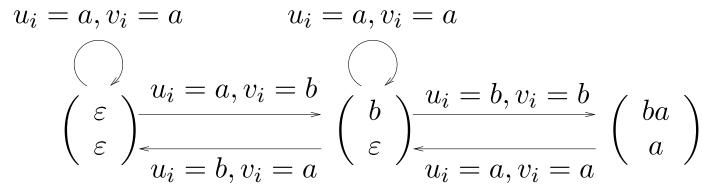


FIGURE 1. Potential balanced pairs produced by the action of $\sigma$ on $(u, v)$ for the Fibonacci case

## 4. Irreducible balanced pairs and discrete geometry

Up to now, we have only characterized irreducible balanced pairs of the form $(lul', l'ul)$ or $(u, \tilde{u})$, and we have seen that, except for the Sturmian case, there may exist other kind of irreducible balanced pairs. In this section, we are interested in irreducible balanced pairs of the form $(u, S_c^k(u))$, where $S_c$ denotes the *circular shift map*, defined as the map $\mathcal{A}^* \to \mathcal{A}^*$ such that $S_c(u_1 \dots u_n) = u_2 \dots u_n u_1$.

We provide in Section 4.1 a discrete geometrical construction inspired by [12, 32], and we explain how this construction is related to the study of irreducible balanced pairs. Then, we characterize in Section 4.2 with Theorem 4.9 substitutive languages that admit infinitely many irreducible balanced pairs of the form $(u, S_c^k(u))$.

### 4.1. Definition

Let $\{\vec{v_i}\}_{i\in[\![1,d]\!]}$ be $d$ vectors in $\mathbb{Z}^{d-1}$. We say that $\{\vec{v_i}\}_{i\in[\![1,d]\!]}$ is $(d-1)$-*independent* if any collection of $d-1$ vectors is linearly independent. In this case, $\Lambda = \sum_{i=2}^{d} \mathbb{Z}(\vec{v_i} - \vec{v_1})$ is a lattice of rank $d-1$, and there exists a fundamental domain for $\mathbb{Z}^{d-1}/\Lambda$ which may be seen as the union of unit hypercubes, not necessarily connected, in $\mathbb{R}^{d-1}$. Let us see how codings are naturally defined on fundamental domains.

Let $u$ be a word on $\{a_i\}_{i\in[\![1,d]\!]}$. Let $\mathcal{C}$ be a collection of unit cubes in $\mathbb{R}^{d-1}$. We say that $u$ is a *roundwalk* of $\mathcal{C}$, with vectors $\{\vec{v_i}\}_{i\in[\![1,d]\!]}$, if the following relation holds:

$$\sum_{i=1}^{d} \vec{v_i} |u|_{a_i} = 0. \tag{4}$$

For any $k \in [\![0, |u|-1]\!]$, we set $\overrightarrow{T}_k^{(u)} = \sum_{i=1}^{d} |\mathrm{pref}_k(u)|_{a_i} \vec{v_i}$. If $\{\overrightarrow{T}_k^{(u)}\}_{k\in[\![0,|u|-1]\!]}$ consists of $|u|$ distinct elements, $u$ is said to be an *irreducible roundwalk* of $\mathcal{C}$. Say differently, $u$ is an irreducible roundwalk if, for all prefixes $p$, $p'$ of $u$ such that $0 \leqslant |p| < |p'| < |u|$, one has $\sum_{i=1}^{d} \vec{v_i} |p|_{a_i} \neq \sum_{i=1}^{d} \vec{v_i} |p'|_{a_i}$.

**Remark 4.1.** *The irreducible roundwalk $u$ of $\mathcal{C}$ enables a labelling of the elements of $\mathcal{C} = \bigcup_{k\in[\![0,|u|-1]\!]} \{c_k\}$, with $c_k = \overrightarrow{T}_k^{(u)}(0)$.*

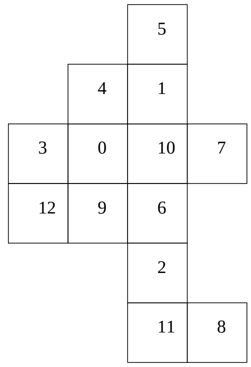


FIGURE 2. *abcaababcabca* is a roundwalk of $\mathcal{C}$

**Example 4.2.** *There exists $\mathcal{C} \subset \mathbb{R}^2$ such that abcaababcabca is an irreducible roundwalk of $\mathcal{C}$, with vectors $\vec{v_1} = \vec{e_1} + \vec{e_2}$, $\vec{v_2} = -3\vec{e_2}$ and $\vec{v_3} = -2\vec{e_1} + 2\vec{e_2}$. This example is depicted in Figure 2, with the corresponding labelling.*

**Lemma 4.3.** *Let $\mathcal{L}$ be a language, and $u \in \mathcal{L}$ such that the integers $\{|u|_{a_i}\}_{i \in [\![1,d]\!]}$ are positive and relatively primes. There exists $d$ vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$, which are $(d-1)$-independent, and $\mathcal{C} \subset \mathbb{R}^{d-1}$ such that $u$ is an irreducible roundwalk of $\mathcal{C}$ with vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$.*

**Proof** Let $u \in \mathcal{L}$ such that $|u|_{a_1}, \dots, |u|_{a_d}$ are positive and relatively primes. Consider the set of vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$ defined by $\vec{v_1} = \sum_{i=2}^{d} |u|_{a_i} \vec{e_{i-1}}$ and $\vec{v_i} = -|u|_{a_i} \vec{e_1}$ for all $i \geqslant 2$, where $\{\vec{e_i}\}_{i \in [\![1,d-1]\!]}$ is the canonical basis of $\mathbb{Z}^{d-1}$. Clearly, (4) holds, and $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$ is $(d-1)$-independent. Moreover, since the integers $\{|u|_{a_i}\}_{i \in [\![1,d]\!]}$ are relatively primes, the elements of $\{\overrightarrow{T}_k^{(u)}\}_{k \in [\![0,|u|-1]\!]}$ are distinct, that is, $u$ is an irreducible roundwalk of $\mathcal{C} = \bigcup_{k \in [\![0,|u|-1]\!]} \{\overrightarrow{T}_k^{(u)}(0)\}$. $\square$

We remind that $S_c$ denotes the circular shift map, defined by $S_c(u_1 \dots u_n) = u_2 \dots u_n u_1$.

**Lemma 4.4.** *Let $u$ be an irreducible roundwalk of $\mathcal{C}$ with vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$. Then, for any $i \in [\![1, |u|-1]\!]$, $S_c^i(u)$ is an irreducible roundwalk of $\mathcal{C} - \overrightarrow{T}_i^{(u)}$.*

**Proof** Without loss of generality, one may assume that $i = 1$. Clearly, if $u$ is an irreducible roundwalk, then $v = S_c(u)$ is an irreducible roundwalk as well, with the same corresponding vectors. Since any strict prefix of $u$ is obtained as the concatenation of $u_1$ and a prefix of $v$, one has $\overrightarrow{T}_k^{(u)} = \overrightarrow{T}_{k-1}^{(v)} + \overrightarrow{T}_1^{(u)}$. As a consequence, $v$ is an irreducible roundwalk of $\mathcal{C} - \overrightarrow{T}_1^{(u)}$. $\square$

The computation of the set of irreducible balanced pairs of a given language seems to be a tough problem. We are interested here in irreducible balanced pairs of the form $(u, S_c^k(u))$. Note that $S_c^k(u)$ may not belong to $\mathcal{L}$; a sufficient condition for having balanced pairs of the form $(u, S_c^k(u))$ is that $uu \in \mathcal{L}$.

**Proposition 4.5.** *Let $u$ be an irreducible roundwalk of $\mathcal{C}$. Let $\overrightarrow{T}$ be such that $\mathcal{C} \cap (\mathcal{C} + \overrightarrow{T})$ contains $n \geqslant 1$ elements. Let $I$ be the set of indices $i \in [\![0, |u|-1]\!]$ such that $\mathcal{C} \cap (\mathcal{C} + \overrightarrow{T}) = \{c_i\}_{i \in I}$. Let $i_0 = \min I$ and $j_0$ be such that $c_{j_0} + \overrightarrow{T} = c_{i_0}$. Then $(S^{i_0}(u), S^{j_0}(u))$ is a balanced pair which reduces in $n$ irreducible balanced pairs.*

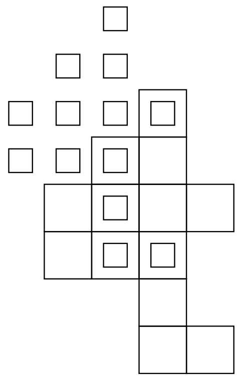

FIGURE 3. $\mathcal{C} \cap \vec{T_2}(\mathcal{C})$ contains 5 elements

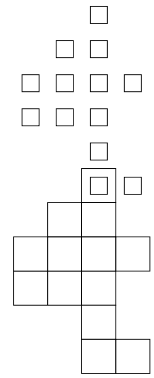

FIGURE 4. $\vec{T_5}(\mathcal{C}) \cap \vec{T_{11}}(\mathcal{C})$ is a unit square

**Proof** Let $u$ be an irreducible roundwalk of $\mathcal{C}$. Let $\overrightarrow{T}$ be such that $\mathcal{C} \cap (\mathcal{C} + \overrightarrow{T})$ contains $n \geqslant 1$ elements. Let $I$ be the set of indices $i \in [\![0, |u|-1]\!]$ such that $\mathcal{C} \cap (\mathcal{C} + \overrightarrow{T}) = \{c_i\}_{i \in I}$. Let $i_0 = \min I$ and $j_0$ be such that $c_{j_0} + \overrightarrow{T} = c_{i_0}$. Let $v = S_c^{i_0}(u)$ and $w = S_c^{j_0}(u)$. Clearly, $(v, w)$ is a balanced pair, since for any letter $l$, one has $|u|_l = |v|_l = |w|_l$. Moreover, for any $i \in [\![1, |u|]\!]$, let $p$ and $q$ be respectively the prefix of $v$ and $w$ of length $i$. Then $(p, q)$ is a balanced pair if and only if $c_{i_0} + \overrightarrow{T}_i^{(v)} = c_{j_0} + \overrightarrow{T} + \overrightarrow{T}_i^{(w)}$. Due to Lemma 4.4, this equality is equivalent to $c_{i_0+i} = \overrightarrow{T} + c_{j'}$, with $j' = j_0 + i \mod |u|$. Hence the number of reductions in $(v, w)$ is equal to the number of elements in $I$, that is, to the number of unit cubes in $\mathcal{C} \cap \overrightarrow{T}_k^{(u)}(\mathcal{C})$. □

**Example 4.6.** *Let us carry on Example 4.2 with $u = abcaababcabca$. One has $\overrightarrow{T_2}(\mathcal{C}) = \vec{e_1} - 2\vec{e_2}$. One checks that $\mathcal{C} \cap (\overrightarrow{T_2}(\mathcal{C})) = \{c_0, c_4, c_5, c_6, c_9\}$, which means thanks to Proposition 4.5 that $(abcaababcabca, caababcabcaab) = (abca, caab)(a, a)(b, b)(abc, cab)(abca, caab)$, see Figure 3.*

*Let $p_5$ and $p_{11}$ be respectively the prefixes of length 5 and 11 of $u$. One has $\overrightarrow{T_5} = \vec{e_1} + 2\vec{e_2}$ and $\overrightarrow{T_{11}} = \vec{e_1} - 5\vec{e_2}$. One checks that $\overrightarrow{T_5}(\mathcal{C}) \cap \overrightarrow{T_{11}}(\mathcal{C})$ contains exactly one unit square (see Figure 4), that is, $(S_c^5(u), S_c^{11}(u)) = (babcabcaabcaa, aabcaababcabc)$ is an irreducible balanced pair.*

### 4.2. IRREDUCIBLE BALANCED PAIRS FOR CERTAIN SUBSTITUTIVE LANGUAGES

Let $\sigma$ be a substitution on the alphabet $\{a_i\}_{i \in [\![1,d]\!]}$. Assume that $\sigma$ satisfies the following property:

$$\text{there exists } u \in \mathcal{L} \text{ such that } |u|_{a_1}, \dots, |u|_{a_d} \text{ are relatively primes, and } uu \in \mathcal{L}_\sigma. \quad (5)$$

Then, for any $k \in \mathbb{N}$, $\sigma^k(uu) \in \mathcal{L}_\sigma$. Hence, for any $i \in [\![0, |\sigma^k(l)|]\!]$, $S_c^i \circ \sigma^k(u) \in \mathcal{L}_\sigma$.

**Proposition 4.7.** *Let $\sigma$ be a unimodular primitive $d$-letter substitution. Let $N \in \mathbb{N}$ such that, for any $l, l' \in \mathcal{A}$, $|\sigma^N(l)|_{l'} \geqslant 1$. Then for any $n \geqslant N$, there exist $(d-1)$-independent vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$ and $\mathcal{C}_n \subset \mathbb{R}^{d-1}$ such that $\sigma^n(l)$ is an irreducible roundwalk of $\mathcal{C}_n$ with vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$.*

**Proof** For any $n \in \mathbb{N}$ and any $i, j \in [\![1, d]\!]$, the number of occurences of the letter $a_i$ in $\sigma^n(a_j)$ is given by the relation:

$$|\sigma(a_j)|_i = \vec{e}_i M_\sigma^n \vec{e}_j. \tag{6}$$

Since $\sigma$ is unimodular, $M_\sigma$ is invertible. As a consequence, for any $n \in \mathbb{N}$, the coordinates of $M_\sigma^n \vec{e_1}$, which are $(|\sigma^n(a_j)|_i)_{i \in [\![1,d]\!]}$, are relatively primes, and they are positive starting from some rank $N$ since $\sigma$ is primitive. Hence, due to Lemma 4.3, for any $n \geqslant N$, $\sigma^n(l)$ is an irreducible roundwalk for the set $\mathcal{C}_n = \bigcup_{i \in [\![0, |\sigma^n(l)|-1]\!]} \{\overrightarrow{T}_i^{(\sigma^n(l))}(0)\}$. □

**Remark 4.8.** *Suppose additionally that $\sigma$ is of Pisot type; remind that, in this case, $\mathcal{H}$ is the contracting hyperspace associated with $M_\sigma$, and $\pi_\mathcal{H}$ denotes the projection on $\mathcal{H}$ along $\mathcal{D}$. Then one may define a sequence of sets $(\mathcal{C}_n)_{n \geqslant N}$ such that $(\pi_\mathcal{H}(M_\sigma^{-n}\mathcal{C}_k))_{n \geqslant N}$ consists of uniformly bounded compact sets which tends for the Hausdorff metric to the Rauzy fractal of $\sigma$. See for instance the construction introduced in [3, 4] in the framework of generalized substitutions.*

Note that Proposition 4.7 may not hold if $\sigma$ is not unimodular. For instance, let $\sigma$ be the two-letter substitution defined by $\sigma(a) = aab$ and $\sigma(b) = aa$. Let $u = \sigma^2(a) = aabaabaa$. For any vectors $\{\vec{v_i}\}_{i \in [\![0,k]\!]}$ such that (4) holds for $u$, the prefix $aaba$ of $u$ satisfies (4) as well. Hence $aabaabaa$ cannot be an irreducible roundwalk of any set.

**Theorem 4.9.** *Let $\sigma$ is a unimodular primitive substitution such that (5) holds for a finite word $u$. Then $\mathcal{L}_\sigma$ admits infinitely many irreducible balanced pairs.*

**Proof** Let $\sigma$ be a unimodular primitive substitution such that (5) holds. As seen in Proposition 4.7, there exist $u \in \mathcal{L}$ and $N \in \mathbb{N}$ such that, for any $n \geqslant N$, $\sigma^n(u)$ is an irreducible roundwalk of $\mathcal{C}_n$ with vectors $\{\vec{v_i}\}_{i \in [\![1,d]\!]}$.

Consider the labelling of $\mathcal{C}_n = \{c_i\}_{i \in [\![0, |\sigma^n(u)|-1]\!]}$ defined by the irreducible roundwalk $\sigma^n(u)$. Let $\overrightarrow{T}$ be a translation vector such that $\overrightarrow{T}(\mathcal{C}_n) \cap \mathcal{C}_n$ contains exactly one unit cube. Set $i, j$ such that $\overrightarrow{T}(\mathcal{C}_n) \cap \mathcal{C}_n = \{c_i\} = \{\overrightarrow{T}(c_j)\}$. Then, due to Proposition 4.5, $(S_c^i(\sigma^n(u)), S_c^j(\sigma^n(u)))$ is an irreducible balanced pair. Since $(|\sigma^n(u)|)_{n \in \mathbb{N}}$ is an increasing sequence, we obtain infinitely many irreducible balanced pairs of distinct lengths. □

**Remark 4.10.** *There exist Pisot type unimodular substitutions such that (5) does not hold for any word $u$. For instance, the following Pisot type unimodular substitution, which has been computed by Pascal Ochem, generates a square-free language: $\sigma(a) = acbacabcbacbc$, $\sigma(b) = acbacabacbabc$, $\sigma(c) = abacabcbacabacbc$.*

## 5. Open questions

The following non-exhaustive list of questions naturally arises from the construction of geometric representations associated with a substitutive dynamical system.

The language generated by the substitution $\sigma$ defined as $\sigma(a) = \sigma(b) = ab$ clearly contains only finitely many irreducible balanced pairs. Is it possible to characterize substitutive languages that admit finitely many irreducible balanced pairs? Also, is it possible to characterize other classes of irreducible balanced pairs than those of the form $(u, S_c^k(u))$?

As already noticed in Section 3.2, we do not know whether the balanced pair algorithm $A(w)$ may terminate for any prefix $w$ after a number of steps that does not depend on the starting irreducible balanced pairs. Say differently, does there exists $N \in \mathbb{N}$ such that $I(\text{pref}_i(\omega)) = I_N(\text{pref}_i(\omega))$ for any $i \in \mathbb{N}^*$? This question seems to be strongly related to the question of determining whether there exist finitely many irreducible balanced pairs which may appear in the reduction of the image under $\sigma$ of any balanced pair.

The number of balanced pairs of length $n$ is $\frac{p(n)(p(n)-1)}{2}$. Since substitutive languages have a sublinear complexity map ( [18]), $i(n)$ grows at most in $n^2$ up to a multiplicative constant term. However, numerical experiments suggest that this order of growth is not sharp. Is it possible to determine the exact order of growth of $i(n)$, the number of irreducible balanced pairs of length $n$? For which languages is it possible to obtain the number of irreducible balanced pairs of a given length as the solution of a closed formula? Is this formula related to a rational or an algebraic function?

## Acknowledgements

I am indebted to Pierre Arnoux and Valérie Berthé for their valuable comments and corrections. I also thank Julien Cassaigne and Pascal Ochem for their helpful remarks and examples.

## References

[1] B. Adamczewski. Balances for fixed points of primitive substitutions. *Theoret. Comput. Sci.*, 307(1):47–75, 2003. Words.

[2] P. Ambrož, C. Frougny, Z. Másáková, and E. Pelantová. Palindromic complexity of infinite words associated with simple Parry numbers. *Annales Institut Fourier, to appear*, 2006.

[3] P. Arnoux, V. Berthé, H. Ei, and S. Ito. Tilings, quasicrystals, discrete planes, generalized substitutions, and multidimensional continued fractions. In *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*, Discrete Math. Theor. Comput. Sci. Proc., AA, pages 059–078 (electronic). Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.

[4] P. Arnoux, V. Berthé, and S. Ito. Discrete planes, $\mathbb{Z}^2$-actions, Jacobi-Perron algorithm and substitutions. *Ann. Inst. Fourier (Grenoble)*, 52(2):305–349, 2002.

[5] P. Arnoux and S. Ito. Pisot substitutions and Rauzy fractals. *Bull. Belg. Math. Soc. Simon Stevin*, 8(2):181–207, 2001.

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

[7] M. Barge and B. Diamond. Proximality in Pisot Tiling Spaces. *Preprint*, 2006.

[8] M. Barge and J. Kwapisz. Geometric theory of unimodular Pisot substitutions. *Preprint*, 2005.

[9] J. Bernat. Symmetrized $\beta$-integers. *Preprint*, 2006.

[10] J. Bernat, V. Berthé, and H. Rao. On super-coincidence condition of substitutions. *Preprint*, 2006.

[11] J. Berstel. Fibonacci words - a survey. In *The Book of L*, pages 13–27. Springer-Verlag, 1986.

[12] V. Berthé and R. Tijdeman. Lattices and multi-dimensional words. *Theoret. Comput. Sci.*, 319(1-3):177–202, 2004.

[13] V. Canterini and A. Siegel. Geometric representation of substitutions of Pisot type. *Trans. Amer. Math. Soc.*, 353(12):5121–5144 (electronic), 2001.

[14] J. Cassaigne, S. Ferenczi, and L. Q. Zamboni. Imbalances in Arnoux-Rauzy sequences. *Ann. Inst. Fourier (Grenoble)*, 50(4):1265–1276, 2000.

[15] A. de Luca. A combinatorial property of the Fibonacci words. *Inform. Process. Lett.*, 12(4):193–195, 1981.

[16] X. Droubay. Palindromes in the Fibonacci word. *Inform. Process. Lett.*, 55(4):217–221, 1995.

[17] F. Durand. A characterization of substitutive sequences using return words. *Discrete Math.*, 179(1-3):89–101, 1998.

[18] F. Durand, B. Host, and C. Skau. Substitutional dynamical systems, Bratteli diagrams and dimension groups. *Ergodic Theory Dynam. Systems*, 19(4):953–993, 1999.

[19] S. Fabre. Dépendance de systèmes de numération associés à des puissances d'un nombre de Pisot. *Theoret. Comput. Sci.*, 158(1-2):65–79, 1996.

[20] C. Frougny. Confluent linear numeration systems. *Theoret. Comput. Sci.*, 106(2):183–219, 1992.

[21] B. Host. Valeurs propres des systèmes dynamiques définis par des substitutions de longueur variable. *Ergodic Theory Dynam. Systems*, 6(4):529–540, 1986.

[22] S. Ito and H. Rao. Atomic surfaces, tilings and coincidences I. Irreducible case. *Israel J. Math., to appear*, 2006.

[23] A. N. Livshits. On the spectra of adic transformations of Markov compact sets. *Uspekhi Mat. Nauk*, 42(3(255)):189–190, 1987.

[24] A. N. Livshits. Some examples of adic transformations and automorphisms of substitutions. *Selecta Math. Soviet.*, 11(1):83–104, 1992. Selected translations.

[25] M. Lothaire. *Algebraic Combinatorics On Words.* Cambridge University Press, 2002.

[26] B. F. Martensen. Generalized balanced pair algorithm. *Topology Proc.*, 28(1):163–178, 2004. Spring Topology and Dynamical Systems Conference.

[27] Z. Masáková. Subword and palindromic complexity of infinite words associated with non-standard numeration systems. *Preprint*, 2005.

[28] W. Parry. On the $\beta$-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, 11:401–416, 1960.

[29] N. Pytheas Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics (Berthé V., Ferenczi S., Mauduit C. and Siegel A.).* Springer Lecture Notes in Mathematics, 1794, Berlin, 2002.

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

[31] R. N. Risley and L. Q. Zamboni. A generalization of Sturmian sequences: combinatorial structure and transcendence. *Acta Arith.*, 95(2):167–184, 2000.

[32] S. Rosema and R. Tijdeman. The Tribonacci substitution. *INTEGERS (Electronic Journal of Combinatorial Number Theory), to appear*, 3(5):1–21, 2005. pp. A13, (Proceedings of the 2004 Number Theoretic Algorithms and Related Topics Workshop).

[33] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8:477–493, 1957.

[34] V. F. Sirvent and B. Solomyak. Pure discrete spectrum for one-dimensional substitution systems of Pisot type. *Canad. Math. Bull.*, 45(4):697–710, 2002. Dedicated to Robert V. Moody.

[35] V. F. Sirvent and Y. Wang. Self-affine tiling via substitution dynamical systems and Rauzy fractals. *Pacific J. Math.*, 206(2):465–485, 2002.

[36] W. P. Thurston. Groups, tilings and finite state automata. *Summer 1989 AMS Colloquium lectures*, 1989.

[37] L. Vuillon. Balanced words. *Bull. Belg. Math. Soc. Simon Stevin*, 10(suppl.):787–805, 2003.


Communicated by (The editor will be set by the publisher).
(The dates will be set by the publisher).