

## A GRAPH APPROACH TO COMPUTING NONDETERMINACY IN SUBSTITUTIONAL DYNAMICAL SYSTEMS

TOKE M. CARLSEN<sup>1,2</sup> AND SØREN EILERS<sup>3</sup>

**Abstract.** We present an algorithm which for any aperiodic and primitive substitution outputs a finite representation of each special word in the shift space associated to that substitution, and determines when such representations are equivalent under orbit and shift tail equivalence. The algorithm has been implemented and applied in the study of certain new invariants for flow equivalence of substitutional dynamical systems.

**Mathematics Subject Classification.** 37B10, 68R15.

### 1. PRELIMINARIES

#### 1.1. INTRODUCTION

Most elements in substitutional dynamical systems, given as doubly infinite sequences, have unique pasts and futures in the sense that one one-sided infinite subsequence determines the other. The importance of those elements which do not have this property, the *special elements*, is well understood in the theory of substitutions and the dynamical systems associated to them.

---

*Keywords and phrases.* Substitution, shift spaces, special elements, orbit equivalence, shift tail equivalence.

<sup>1</sup> University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark;  
toke@math.ku.dk

<sup>2</sup> *Present address:* School of Mathematical & Physical Sciences, Mathematics Building - V123, University of Newcastle, University Drive, Callaghan NSW 2308, Australia.

<sup>3</sup> University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark;  
eilers@math.ku.dk

Determining  $K$ -groups of certain  $C^*$ -algebras we found, as described in [5], an invariant of flow equivalence (cf. [16]) — akin and related to the dimension groups considered in [8] — of substitutional systems based on combinatorial and textual properties of the special elements. For each primitive and aperiodic substitution  $\tau$  on an alphabet  $\mathbf{a}$  this invariant is an ordered group defined as a stationary inductive limit of group endomorphisms on  $\mathbb{Z}^{|\mathbf{a}|} \oplus \mathbb{Z}^{n_\tau}$  induced from a  $(|\mathbf{a}| + n_\tau) \times (|\mathbf{a}| + n_\tau)$  block matrix

$$\begin{bmatrix} \mathbf{A}_\tau & 0 \\ \mathbf{B}_\tau & \mathbf{Id} \end{bmatrix}$$

where the block  $\mathbf{A}_\tau$  is the abelianization matrix of the substitution in question. To compute the integer  $n_\tau$  and the matrix  $\mathbf{B}_\tau$  one needs a coherent finite representation of all the special words of the substitution, and to determine which among the special words are equivalent under the natural relations of *orbit* and *right shift tail* equivalence (see Def. 1.4). In a recent paper [7] we prove by example that the resulting invariant contains information not accessible by any other flow invariant known to us, such as the dimension groups in [8], the configuration graph (see p. 289 below), or the numerical index used in conjunction with the notion of weak equivalence in [1]. The example is an explicit substitution  $\tau$  on  $\{a, b, c, d\}$  such that the matrices associated to  $\tau$  and its opposite  $\tau^{-1}$ , respectively, become

$$\begin{bmatrix} 6 & 9 & 3 & 9 & 0 & 0 \\ 12 & 18 & 6 & 18 & 0 & 0 \\ 6 & 9 & 3 & 9 & 0 & 0 \\ 36 & 54 & 18 & 54 & 0 & 0 \\ 10 & 13 & 4 & 12 & 1 & 0 \\ 6 & 8 & 2 & 8 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 6 & 9 & 3 & 9 & 0 & 0 \\ 12 & 18 & 6 & 18 & 0 & 0 \\ 6 & 9 & 3 & 9 & 0 & 0 \\ 36 & 54 & 18 & 54 & 0 & 0 \\ 2 & 7 & 2 & 7 & 1 & 0 \\ 2 & 7 & 2 & 7 & 0 & 1 \end{bmatrix}$$

Needless to say, having computer based tools to compute these components of our invariant is very useful in the study of it. The project described above thus naturally lead us to concern ourselves with computability of the aforementioned words and quantities associated to the class of special elements associated to a given substitution, and failing to find algorithms meeting our needs in the literature, we developed the approach presented in the present note. Our algorithm outputs a finite representation of each special word, and determines when such representations are equivalent under *shift tail equivalence*, a naturally occurring relation of importance in our invariant.

We wish to acknowledge [2], to which our work is closely related. Although the ends and ambitions of the present note and [2] do not overlap, the means seem to do. Our method was developed independently, but we received [2] before writing up this note. Although we have not attempted to do so, the results in our Section 3 could most likely be developed using the methods in [2], and vice versa. The series of reductions based on replacing the substitutions in [2] is, however, computationally inconvenient for our purposes. Our Section 4 has no analogue there.

Our paper is organized as follows. In Section 2, after having singled out the class of elementary substitutions and explained how to reduce the problem to this case, we associate certain graphs to such substitutions and explain how they give rise to a class of adjusted left special words. We also define a class of adjusted left special words arising from  $\tau$ -periodic one-sided words. In Section 3 we then proceed to prove that each adjusted left special word is on the list generated in the previous section, and prove a separation result of importance in our paper [6]. The main technical tool is the one-sided substitute for injectivity of  $\tau$  considered as a map on its two-sided shift space, *cf.* Lemma 1.2 below, which we shall be able to derive from the work of Mossé [14]. In Section 4 we describe an algorithm for determining shift tail and orbit equivalence of the output of the algorithm described and proved in the previous section. The paper ends with a summary of the algorithm and a few remarks of relevance to related work.

The algorithm has been implemented in a Java applet [4].

### 1.2. SUBSTITUTIONS

We refer to [8, 9, 17] for a thorough introduction to this subject and shall here only lay out notation. Letting  $\mathfrak{a}$  denote a finite set or alphabet, we denote by  $\mathfrak{a}^\sharp$  the set of nonempty finite words in  $\mathfrak{a}$ . For  $w \in \mathfrak{a}^\sharp$ , we let  $|w|$  denote the number of letters and index the letters of  $w$  from 0 to  $|w| - 1$ . A *substitution* is simply a map  $\tau : \mathfrak{a} \rightarrow \mathfrak{a}^\sharp$ . We can extend  $\tau$  to  $\mathfrak{a}^\mathbb{Z}$  or to

$$\mathfrak{a}^\mathbb{Z}, \mathfrak{a}^{\mathbb{N}_0}, \mathfrak{a}^{-\mathbb{N}} \tag{1}$$

(with  $\mathbb{N}_0 = \{0, 1, \dots\}$ ,  $-\mathbb{N} = \mathbb{Z} \setminus \mathbb{N}_0$ ) in the obvious way, and define powers of  $\tau$  recursively. To define the action of  $\tau$  on  $\mathfrak{a}^\mathbb{Z}$  we need to specify that the word resulting from the substitution of the letter at index 0 of a doubly infinite sequence  $x$  will be placed starting at index 0 in  $\tau(x)$ . We thus have

$$\tau(y.x) = \tau(y).\tau(x)$$

where, as we will do in the following, we have used a dot to indicate the position separating  $-\mathbb{N}$  and  $\mathbb{N}_0$ . We denote by  $\tau^{-1}$  the *opposite substitution* defined by reversing each word  $\tau(\cdot)$ . Finally, an *abelianization matrix*  $\mathbf{A}_\tau$  is associated to  $\tau$  as the  $|\mathfrak{a}| \times |\mathfrak{a}|$ -matrix counting at row  $b$  and column  $a$  the number of occurrences of  $b$  in  $\tau(a)$ .

We equip the sequence spaces mentioned in (1) with the product topology from the discrete topology on  $\mathfrak{a}$ , and define  $\sigma : \mathfrak{a}^\mathbb{Z} \rightarrow \mathfrak{a}^\mathbb{Z}$  by  $(\sigma(x))_n = x_{n+1}$ . Maps of this type we will refer to as *shift maps*. A *two-sided shift space* is a closed subset of  $\mathfrak{a}^\mathbb{Z}$  which is mapped onto itself by  $\sigma$ . We shall refer to such spaces by “ $\underline{\mathbf{X}}$ ” with possible subscripts. Generally speaking, a *one-sided shift space* is a closed subset of  $\mathfrak{a}^{\mathbb{N}_0}$  or  $\mathfrak{a}^{-\mathbb{N}}$  which is mapped into itself by the unique shift map. We are only interested in those one-sided shift spaces which can be produced from two-sided shift spaces by projection, and denote these spaces by  $\mathbf{X}^+$  and  $\mathbf{X}^-$ , respectively. There is a rich theory of shift spaces; we refer to [12] and [11].

For  $-\infty \leq i \leq j \leq \infty$  we use interval notation  $x_{[i,j]}$  to denote the (possibly infinite) subword of  $x$  corresponding to the indices between  $i$  and  $j$ . We write  $x_{[i,j[} = x_{[i,j-1]}$  when it makes sense and is convenient. Unless specified otherwise, we index finite words by nonnegative indices starting with 0, and right or left infinite words by  $\mathbb{N}_0$  or  $-\mathbb{N}$ .

The *language* of a two-sided shift space is the subset of  $\mathfrak{a}^\# \cup \{\epsilon\}$ , where  $\epsilon$  denotes the empty word, given by

$$\mathcal{L}(\underline{X}) = \{x_{[i,j]} \mid x \in \underline{X}, i \leq j \in \mathbb{Z} \cup \{\epsilon\}\}.$$

Conversely, a subset  $\mathcal{G} \subseteq \mathfrak{a}^\# \cup \{\epsilon\}$  defines a shift space; the smallest shift space  $\underline{X}_{\mathcal{G}}$  such that  $\mathcal{G} \subseteq \mathcal{L}(\underline{X}_{\mathcal{G}})$ . With  $\mathcal{G} = \{\tau^n(a) \mid n \in \mathbb{N}, a \in \mathfrak{a}\}$  we arrive at the *substitutional dynamical system associated to  $\tau$*  denoted  $\underline{X}_\tau$  which will be our main concern in the present paper.

We single out two important properties of substitutions below. The notation “ $\mathbf{A} > 0$ ” indicates that the matrix  $\mathbf{A}$  has only positive entries.

**Definition 1.1.** A substitution  $\tau$  is *primitive* if  $|\mathfrak{a}| > 1$  and

$$\exists n \in \mathbb{N} : \mathbf{A}_\tau^n > 0.$$

A substitution  $\tau$  is *aperiodic* if  $|\underline{X}_\tau| = \infty$ .

It is decidable when a given substitution has these properties, *cf.* [15] and [19]. Primitive and aperiodic substitutions yield *minimal* shift spaces: all orbits  $\{\sigma^n(x) \mid n \in \mathbb{Z}\}$  are dense, *cf.* [17]. Consequently, there are no (ultimately)  $\sigma$ -periodic words in  $\underline{X}_\tau, \mathbf{X}_\tau^+$  or  $\mathbf{X}_\tau^-$  for such  $\tau$ : if  $x_{k+n} = x_{m+n}$  for all  $n$  in the various index sets, then  $k = m$ . Further, we have

**Lemma 1.2** ([14], *cf.* [8], Cor. 10). *The map induced by  $\tau$  on  $\underline{X}_\tau$  is injective, when  $\tau$  is primitive and aperiodic.*

**Example 1.3.** The following substitutions are all primitive and aperiodic:

- $\tau_1 : 1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 123;$
- $\tau_2 : 0 \mapsto 003210, 1 \mapsto 00, 2 \mapsto 00, 3 \mapsto 00220;$
- $\tau_3 : a \mapsto aba, b \mapsto baab;$
- $\tau_4 : 0 \mapsto 10, 1 \mapsto 0;$
- $\tau_5 : a \mapsto accdadbb, b \mapsto acdcbadb, c \mapsto aacdcdbb, d \mapsto accbdadb;$
- $\tau_6 : a \mapsto accbbadd, b \mapsto accdbabd, c \mapsto acbbccdd, d \mapsto acbcdabd.$

The following notation is convenient. When  $w_0, \dots, w_{n-1}$  is a finite list of words in  $\mathcal{L}(\underline{X}_\tau)$ , we define

$$\begin{aligned} [w_0, \dots, w_{n-1}]^+ &= w_0\tau(w_1) \cdots \tau^{n-1}(w_{n-1})\tau^n(w_0)\tau^{n+1}(w_1) \cdots \in \mathfrak{a}^{\mathbb{N}_0}, \\ [w_{n-1}, \dots, w_0]^- &= \cdots \tau^{n+1}(w_1)\tau^n(w_0)\tau^{n-1}(w_{n-1}) \cdots \tau(w_1)w_0 \in \mathfrak{a}^{-\mathbb{N}}. \end{aligned}$$

1.3. ORBIT CLASSES AND SPECIAL ELEMENTS

**Definition 1.4.** Let  $\underline{X}$  be a two-sided shift space. We define three equivalence relations on  $x, y \in \underline{X}$  in the following way:

- (i) If there exists an  $n$  such that  $x_m = y_{n+m}$  for all  $m \in \mathbb{Z}$  then we say that  $x$  and  $y$  are *orbit equivalent* and write  $x \sim_o y$ .
- (ii) If there exist an  $n$  and an  $N$  such that  $x_m = y_{n+m}$  for all  $m > N$  then we say that  $x$  and  $y$  are *right shift tail equivalent* and write  $x \sim_r y$ .
- (iii) If there exist an  $n$  and an  $N$  such that  $x_m = y_{n+m}$  for all  $m < N$  then we say that  $x$  and  $y$  are *left shift tail equivalent* and write  $x \sim_l y$ .

Notice that  $x \sim_o y$  implies that  $x \sim_r y, x \sim_l y$ , so  $\sim_r$  and  $\sim_l$  induce equivalence relations on  $\underline{X}/\sim_o$  which we also will denote by  $\sim_r$  and  $\sim_l$ . We call an orbit class  $[x]$  in  $\underline{X}/\sim_o$  *left special* ([10], Sect. 5) if there exists  $[y] \in \underline{X}/\sim_o$  such that  $[x] \neq [y]$ , but  $[x] \sim_r [y]$ . A *left special word*  $x \in \underline{X}$  is a representative of such an orbit class with the property that  $y \in \underline{X}$  exists with

$$x_{-1} \neq y_{-1} \quad x_{[0, \infty[} = y_{[0, \infty[}.$$

We say that the left special word  $x$  is *adjusted* if  $\sigma^{-n}(x)$  is not left special for any  $n \in \mathbb{N}$ .

The symmetric definition defines a class of (adjusted) right special words. Classical results ([17], p. 107 [3] Th. 3.9) give:

**Theorem 1.5.** *When  $\tau$  is aperiodic and primitive, then the number of (left or right) special orbit classes is finite, but nonzero.*

Note that as a consequence of this, there is always an adjusted special word representing each special orbit class.

A nice way of describing the structure of special words using the equivalence relations  $\sim_r$  and  $\sim_l$  on  $\underline{X}/\sim_o$ , suggested to us by an electronic exchange with Charles Holton, is by means of a bipartite graph defined as follows. The vertex set of the graph will be contained in the disjoint union of  $\underline{X}/\sim_r$  and of  $\underline{X}/\sim_l$ , and for each orbit class  $[x]_o$  with  $x$  a special element, we let an edge connect  $A \in \underline{X}/\sim_r$  with  $B \in \underline{X}/\sim_l$  if  $[x]_o \in A$  and  $[x]_o \in B$ . Removing all vertices with no edges, we arrive at a bipartite graph which we shall denote as the *configuration graph* of  $\underline{X}$ . The theorem above shows that this is a finite graph when the shift space arises from a substitution. Examples are given in 4.7 below.

**Lemma 1.6.** *The configuration graph is an invariant of conjugacy ([12], Def. 1.5.9) and flow equivalence ([12], Sect. 13.5, [16]) of the substitutional dynamical systems.*

*Proof.* Since a conjugacy is a sliding block code (cf. [12], 1.5), one easily sees that it must preserve special words and all the relevant equivalence relations. Similarly, any *expansion map* induced by

$$a_0 \mapsto a_0b \quad a_i \mapsto a_i, i > 0$$

sending biinfinite sequences on the alphabet  $\mathbf{a} = \{a_0, a_1, \dots, a_n\}$  to biinfinite sequences on  $\mathbf{a} \cup \{b\}$  will take special words to special words in a manner preserving all the relations, and since the same can be said about the inverse of this map which deletes all occurrences of  $b$ , we get that expansion maps preserve configuration graphs. This proves the second claim since flow equivalence on shift spaces is generated by conjugacy and expansion according to [13] Lemma 2.1.  $\square$

Note that the lemma implies that also the number of orbit or shift tail classes of special elements is a flow invariant.

## 2. COLLECTING SPECIAL ELEMENTS

### 2.1. ELEMENTARY AND SIMPLIFIABLE SUBSTITUTIONS

We recall from [18] p. 17 that a substitution  $\tau$  on the alphabet  $\mathbf{a}$  is *simplifiable* if it can be factored  $\tau = f \circ g$  for maps

$$f : \mathbf{b} \longrightarrow \mathbf{a}^\# \quad g : \mathbf{a} \longrightarrow \mathbf{b}^\#$$

where  $|\mathbf{b}| < |\mathbf{a}|$ . We say that the substitution  $v = g \circ f$  is a simplification of  $\tau$  in this case. In case  $\tau$  is not simplifiable, we call it *elementary*.

It is decidable whether a substitution is simplifiable or elementary, cf. [18] p. 17, and a succession of simplifications, ending with an elementary substitution, can be computed in the simplifiable case. Composing the  $2n$  maps involved in a simplification in  $n$  steps to the elementary substitution  $v$ , we get  $f, g$  with the property

$$\tau^n = f \circ g \quad v^n = g \circ f. \tag{2}$$

This was used in [15] to provide an algorithm for deciding aperiodicity by reducing to the elementary case. We shall use a similar strategy to compute the set of special elements for a given substitution, based on Proposition 2.2 below.

First, however, we need to concern ourselves with establishing our key substitution properties for simplifications. Simplifications are known to preserve aperiodicity, but a simplification of a primitive substitution may fail to be primitive. However, the following holds:

**Lemma 2.1.** *If a primitive and aperiodic substitution  $\tau$  is simplified to an elementary substitution  $v$ , then  $v$  is primitive and aperiodic.*

*Proof.* That simplifications preserve aperiodicity is a key observation in [15]. It follows easily that  $\mathbf{A}_\tau$  and  $\mathbf{A}_v$  are strongly shift equivalent, cf. [12]. Note further that  $\mathbf{A}_v$  must be essential, as otherwise a letter could be deleted from the alphabet. Applying [12], Proposition 4.5.10, we get the desired result.  $\square$

**Proposition 2.2.** *Let  $\tau$  be a primitive and aperiodic simplifiable substitution and let  $v$  be an elementary simplification with maps  $f, g$  satisfying (2) above. The map from  $\underline{X}_v$  to  $\underline{X}_\tau$  induced by  $f$  preserves orbit and shift tail equivalences, and maps the (left, right) special orbits of  $\underline{X}_v$  bijectively onto the set of (left, right) special orbits of  $\underline{X}_\tau$ .*

*Proof.* Clearly the maps induced by  $f$  and  $g$  preserve all three kinds of equivalence. Note also that they are injective because of Lemma 1.2; in the case of  $f$  because  $v$  (and  $v^n$ ) is primitive by Lemma 2.1.

Clearly, then, both maps send special elements to special elements. Let  $x_1, \dots, x_n$  be a choice of orbit inequivalent special words of  $\underline{X}_\tau$ , representing all such orbit classes. We have that  $\tau(x_1), \dots, \tau(x_n)$  are orbit inequivalent special words of  $\underline{X}_\tau$ , since  $\tau$  is injective and  $\underline{X}_\tau$  is aperiodic. Hence each orbit class of special elements is realized by a representative of the form  $f(g(x_i))$ , where  $g(x_i)$  is special.  $\square$

**Example 2.3.** The substitutions  $\tau_1, \tau_3, \tau_4, \tau_5$  and  $\tau_6$  are elementary, but  $\tau_2$  is simplifiable to

$$p \mapsto ppqp, q \mapsto pprpppppp, r \mapsto pp$$

using  $f$  given by  $p \mapsto 0, q \mapsto 321, r \mapsto 2$  and  $g$  given by  $0 \mapsto ppqp, 1 \mapsto pp, 2 \mapsto pp, 3 \mapsto pprpp$ .

## 2.2. NS-COVERS AND THEIR GRAPHS

In the following, we assume that the alphabet  $\mathfrak{a}$  is equipped with some well-ordering “ $>$ ”; in the examples, we just use alphabetical or numerical order.

Let  $\mathcal{W}$  be a finite set of nonempty words. By  $\widehat{\mathcal{W}}\mathcal{W}$  we denote the set

$$\{(v, w) \mid v, w \in \mathcal{W}, v_{|v|-1} > w_{|w|-1}\}$$

consisting of pairs of words from  $\mathcal{W}$  which end in different letters, arranged so that the word ending in the first letter according to “ $>$ ” is first among the two.

**Definition 2.4.** Let  $\tau$  be a primitive and aperiodic substitution. We say that the finite family  $\mathcal{W} \subseteq \mathcal{L}(\underline{X}_\tau)$  is an NS-cover of  $\tau$  (a nonsuffix cover) if

$$\text{Cyl}^-(w) = \{x \in \underline{X}_\tau \mid x_{[-|w|, -1]} = w\}, \quad w \in \mathcal{W}$$

forms a disjoint partition of  $\underline{X}_\tau$ , and if for every pair  $(v_0, v_1) \in \widehat{\mathcal{W}}\mathcal{W}$  one can write

$$\tau(v_0) = t_0 w_0 z \quad \tau(v_1) = t_1 w_1 z \tag{3}$$

where  $t_0, t_1, z \in \mathcal{L}(\underline{X}_\tau)$  with  $t_0, t_1 \neq \epsilon$ , and where either

$$(w_0, w_1) \in \mathcal{W} \widehat{\times} \mathcal{W} \tag{4.0}$$

or

$$(w_1, w_0) \in \mathcal{W} \widehat{\times} \mathcal{W}. \tag{4.1}$$

Not every primitive and aperiodic substitution possesses an  $NS$ -cover – our example  $\tau_2$  provides an example of this behavior as seen in Example 5.2 below. However, the following shall suffice for our purposes:

**Proposition 2.5.** *If a primitive and aperiodic substitution is elementary, it possesses an  $NS$ -cover. Indeed, there is a computable integer  $n$  such that the set*

$$\{w \in \mathcal{L}(\underline{X}_\tau) \mid |w| = n\}$$

*is an  $NS$ -cover.*

*Proof.* By [18] Theorem 1.6, p. 126, the integer

$$p = \sum_{a \in \mathfrak{a}} (|\tau(a)| - 1) + \max_{a \in \mathfrak{a}} |\tau(a)|$$

has the property that if words  $v, w \in \mathcal{L}(\underline{X}_\tau)$  end in different letters, and both  $\tau(v)$  and  $\tau(w)$  have the suffix  $z$ , then  $|z| \leq p$ .

By primitivity, one letter  $a \in \mathfrak{a}$  has the property  $|\tau(a)| \geq 2$ , and we may find  $m$  such that  $a$  occurs in  $\tau^m(b)$  for every  $b \in \mathfrak{a}$ , and with

$$n = 2(p + 1) \max\{|\tau^m(b)| \mid b \in \mathfrak{a}\}$$

we thus have that  $a$  occurs  $p + 1$  times in  $v$  if  $|v| = n$ . Thus

$$|\tau(v)| \geq p + 1 + |v| = p + n + 1$$

for each such  $v$ . In (3), this leaves  $n$  letters to read off  $v', w'$  ending in different letters, and at least one more letter to read off  $t, u$ . □

**Remark 2.6.** In practice one finds that the value of  $n$  determined above is often much larger than needed. It is hence recommendable to simply try  $n = 1, n = 2$ , etc. until one reaches a sufficiently large length. We do not know whether the bound given is reachable.

**Example 2.7.** For the substitutions considered in Example 1.3, the smallest number  $n$  such that the set of all words in the associated language is an  $NS$ -cover is

$\tau_1$	$\tau_1^{-1}$	$v_2$	$v_2^{-1}$	$\tau_3 = \tau_3^{-1}$	$\tau_4$	$\tau_4^{-1}$
2	3	3	4	1	4	2

Let now  $\tau$  be a primitive and aperiodic substitution with an  $NS$ -cover  $\mathcal{W}$ . We define a multiply labeled graph  $\mathcal{G}_{\tau, \mathcal{W}}$  of  $\tau$  and  $\mathcal{W}$  as follows. Choose as vertex set  $\mathcal{V}_{\tau, \mathcal{W}} = \mathcal{W} \widehat{\times} \mathcal{W}$  and define for each  $(v_0, v_1) \in \mathcal{W} \widehat{\times} \mathcal{W}$  a threefold labeled edge

$$(w_0, w_1) \xrightarrow{z, t_0, t_1} (v_0, v_1),$$

where  $w_0, w_1$  and  $z, t_0, t_1$  are the (obviously unique) elements satisfying (3). Let  $\mathcal{E}_{\tau}$  denote the set of all such edges with  $(v_0, v_1)$  ranging over  $\mathcal{W} \widehat{\times} \mathcal{W}$  and define labelings

$$\mathfrak{L} : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathcal{X}}_{\tau}), \mathfrak{L}_0 : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathcal{X}}_{\tau}), \mathfrak{L}_1 : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathcal{X}}_{\tau})$$

accordingly, associating  $z, t_0, t_1$ , respectively, to the edge in question. Finally, we label any edge of  $\mathcal{G}_{\tau}$  by

$$\mathfrak{L}_5 : \mathcal{E} \longrightarrow \mathbb{Z}_2,$$

according to which of the alternatives (4.0) and (4.1) in Definition 2.4 is met.

In the following definition, the *essential part* of a given graph is the subgraph defined by deleting all vertices which may not be visited by a doubly infinite path on the graph.

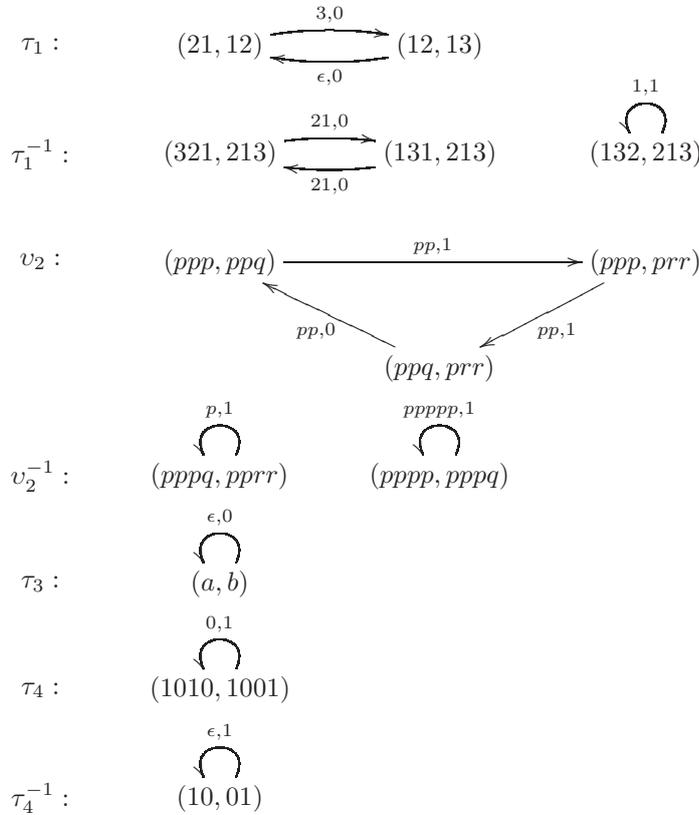
**Definition 2.8.** The graph  $\mathcal{G}_{\tau, \mathcal{W}}$  is the essential part of  $(\mathcal{V}_{\tau, \mathcal{W}}, \mathcal{E}_{\tau, \mathcal{W}})$  labeled by the restrictions of the labelings.

**Corollary 2.9.** *For every primitive and aperiodic substitution  $\tau$  with an  $NS$ -cover  $\mathcal{W}$ ,  $\mathcal{G}_{\tau, \mathcal{W}}$  is a nonempty forest of cycles.*

*Proof.* By construction, each vertex of  $\mathcal{G}_{\tau}$  has only one incoming edge. Since  $\mathcal{G}_{\tau, \mathcal{W}}$  is essential, each vertex has at least one, and thus exactly one, outgoing edge. We conclude that  $\mathcal{G}_{\tau, \mathcal{W}}$  is a forest of cycles. Since  $(\mathcal{V}_{\tau, \mathcal{W}}, \mathcal{E}_{\tau, \mathcal{W}})$  defined above has at least one cycle, so does  $\mathcal{G}_{\tau, \mathcal{W}}$ .  $\square$

**Example 2.10.** For each substitution in Example 1.3 we state  $\mathcal{G}_{\tau}$  associated to the  $NS$ -covers consisting of all words of a certain length, as found in Example 2.7.

The graphs are decorated with  $\mathfrak{L}, \mathfrak{L}_s$ .



**Lemma 2.11.** *Let  $\tau$  be a primitive and aperiodic substitution. When  $\mathcal{W}$  is an NS-cover for  $\tau$ , then it is also an NS-cover for  $\tau^N$  for  $N \geq 1$ . The underlying graph of  $\mathcal{G}_{\tau^N, \mathcal{W}}$  is identical to the higher power graph  $(\mathcal{G}_{\tau, \mathcal{W}})^N$ , with edges representing paths on  $\mathcal{G}_{\tau, \mathcal{W}}$  of length  $N$ . It is labeled by*

$$\begin{aligned}
 \mathfrak{L}^N(e_0, \dots, e_{N-1}) &= \mathfrak{L}(e_0)\tau(\mathfrak{L}(e_1)) \cdots \tau^{N-1}(\mathfrak{L}(e_{N-1})), \\
 \mathfrak{L}_s^N(e_0, \dots, e_{N-1}) &= \sum_{i=0}^{N-1} \mathfrak{L}_s(e_i),
 \end{aligned}$$

and

$$\mathfrak{L}_j^N(e_0, \dots, e_{N-1}) = \tau^{N-1}(\mathfrak{L}_{s_{N-1}}(e_{N-1})) \cdots \tau(\mathfrak{L}_{s_1}(e_1))\mathfrak{L}_{s_0}(e_0),$$

where  $j, s_i \in \mathbb{Z}_2$  and the  $s_i$  are defined recursively by  $s_0 = j$  and  $s_{i+1} = \mathfrak{L}_s(e_i) + s_i$ .

*Proof.* To see the first claim, note that when  $(v, w) \in \mathcal{W} \widehat{\times} \mathcal{W}$

$$\begin{aligned} \tau^2(v) &= \tau(tv'z) = \tau(t)\tau(v')\tau(z) = \tau(t)t'v''z'\tau(z), \\ \tau^2(w) &= \tau(tw'z) = \tau(u)\tau(w')\tau(z) = \tau(u)u'w''z'\tau(z), \end{aligned}$$

where  $(v', w')$  and  $(v'', w'')$  are elements of  $\mathcal{W} \widehat{\times} \mathcal{W}$ . This forms the basis of an induction argument proving that  $\mathcal{W}$  is an  $NS$ -cover for  $\tau^N$ . The remaining claims are straightforward.  $\square$

**Proposition 2.12.** *Let  $\tau$  be a primitive and aperiodic substitution. There is an  $N$  such that  $\mathcal{G}_{\tau^N, \mathcal{W}}$  is a forest of loops all labeled 0 by  $\mathfrak{L}_s^N$ .*

*Proof.* As seen in the proof of Corollary 2.9,  $\mathcal{G}_{\tau, \mathcal{W}}$  is a forest of cycles. The power  $m$  defined as the least common multiple of all the lengths of cycles will lead to a graph with loops only. Then  $N = 2m$  will suffice.  $\square$

Note that the non-essential part of the graph  $(\mathcal{V}_{\tau^N, \mathcal{W}}, \mathcal{E}_{\tau^N, \mathcal{W}})$  in this case consists of a collection of trees rooted at one or more of the loops. With the result above in hand we see that one to every primitive and aperiodic substitution with an  $NS$ -cover  $\mathcal{W}$  can associate a number  $N$  with the following property:

**Definition 2.13.** We shall say that  $N$  is a  $\mathcal{W}$ -basic power of the primitive and aperiodic substitution  $\tau$  (relative to the  $NS$ -cover  $\mathcal{W}$ ) if  $\mathcal{G}_{\tau^N, \mathcal{W}}$  meets the conditions of Proposition 2.12, and if every edge in  $(\mathcal{V}_{\tau^N, \mathcal{W}}, \mathcal{E}_{\tau^N, \mathcal{W}})$  begins at a vertex of  $\mathcal{G}_{\tau^N, \mathcal{W}}$ .

Suppose further that  $\tau^{-1}$  has an  $NS$ -cover  $\mathcal{W}'$ . We say that  $N$  is an  $\mathcal{W}, \mathcal{W}'$ -bibasic power if  $N$  is  $\mathcal{W}$ -basic for  $\tau$  and  $\mathcal{W}'$ -basic for  $\tau^{-1}$ .

A class of left special words on bracket form can be read of the graph thus associated to a  $\mathcal{W}$ -basic power of a substitution. Indeed, whenever  $(v, w)$  is a vertex in the graph,  $e$  is the loop at that vertex, and whenever  $\mathfrak{L}(e) \neq \epsilon$  we have that

$$\tau^{mN}(v) = \tau^{N(m-1)}(\mathfrak{L}_0^N(e)) \cdots \tau^N(\mathfrak{L}_0^N(e)) \mathfrak{L}_0^N(e) v \mathfrak{L}^N(e) \tau^N(\mathfrak{L}^N(e)) \cdots \tau^{N(m-1)}(\mathfrak{L}^N(e))$$

$$\tau^{mN}(w) = \tau^{N(m-1)}(\mathfrak{L}_1^N(e)) \cdots \tau^N(\mathfrak{L}_1^N(e)) \mathfrak{L}_1^N(e) w \mathfrak{L}^N(e) \tau^N(\mathfrak{L}^N(e)) \cdots \tau^{N(m-1)}(\mathfrak{L}^N(e))$$

are words growing to infinity in both directions, as  $\mathfrak{L}_0^N(e) \neq \epsilon, \mathfrak{L}_1^N(e) \neq \epsilon$  by definition of  $NS$ -covers. Thus both

$$[\overbrace{\epsilon, \dots, \epsilon}^{N-1}, \mathfrak{L}_0^N(e)]^- v. [\mathfrak{L}^N(e), \overbrace{\epsilon, \dots, \epsilon}^{N-1}]^+$$

and

$$[\overbrace{\epsilon, \dots, \epsilon}^{N-1}, \mathfrak{L}_1^N(e)]^- w. [\mathfrak{L}^N(e), \overbrace{\epsilon, \dots, \epsilon}^{N-1}]^+$$

are elements of  $\underline{X}_\tau$ . Since  $v$  and  $w$  end in different letters, these elements are left special. We shall denote the class of such left special words by  $\mathcal{S}_\mathcal{W}$  (lat. *sinister*). By considering opposite substitutions and reverting the output of the procedure described above, we get a set of right special elements which we denote by  $\mathcal{D}_\mathcal{W}$  (lat. *dexter*).

**Remark 2.14.** It is a theoretical convenience to work with special words read off a graph associated to basic powers, but as a consequence of our construction the words may also be read off graphs associated to smaller powers, notably  $N = 1$ .

When the graph is no longer a forest of 0-labeled loops, one proceeds as follows. For each vertex  $(v, w)$  in  $\mathcal{G}_{\tau, \mathcal{W}}$ , one follows outgoing edges

$$e_0, \dots, e_n$$

until  $e_n$  ends at  $(v, w)$ . One defines  $s_i \in \mathbb{Z}_2$  recursively by

$$s_0 = 0, \quad s_{i+1} = \mathfrak{L}_s(e_i) + s_i.$$

If  $s_n = 0$  one records

$$\begin{aligned} & [\mathfrak{L}_{s_n}(e_n), \dots, \mathfrak{L}_{s_0}(e_0)]^- \cdot v \cdot [\mathfrak{L}(e_0), \dots, \mathfrak{L}(e_n)]^+, \\ & [\mathfrak{L}_{1+s_n}(e_n), \dots, \mathfrak{L}_{1+s_0}(e_0)]^- \cdot w \cdot [\mathfrak{L}(e_0), \dots, \mathfrak{L}(e_n)]^+. \end{aligned}$$

If  $s_n = 1$  one needs to consider

$$\begin{aligned} & [\mathfrak{L}_{1+s_n}(e_n), \dots, \mathfrak{L}_{1+s_0}(e_0), \mathfrak{L}_{s_n}(e_n), \dots, \mathfrak{L}_{s_0}(e_0)]^- \cdot v \cdot [\mathfrak{L}(e_0), \dots, \mathfrak{L}(e_n)]^+, \\ & [\mathfrak{L}_{s_n}(e_n), \dots, \mathfrak{L}_{s_0}(e_0), \mathfrak{L}_{1+s_n}(e_n), \dots, \mathfrak{L}_{1+s_0}(e_0)]^- \cdot w \cdot [\mathfrak{L}(e_0), \dots, \mathfrak{L}(e_n)]^+. \end{aligned}$$

Obviously, we just get different – shorter – bracket representations of the elements of  $\mathcal{S}_\mathcal{W}$  this way.

**Example 2.15.** Reading off elements on the graphs found in Example 2.10 we get

$$\begin{aligned}
 \tau_1 : \mathcal{S}_{\mathcal{W}_2} &= \{[1, 13]^- 12.[\epsilon, 3]^+, [12, 12]^- 13.[\epsilon, 3]^+, \\
 &\quad [13, 1]^- 21.[3, \epsilon]^+, [12, 12]^- 12.[3, \epsilon]^+\}; \\
 \mathcal{D}_{\mathcal{W}_3} &= \{[12, 12]^- .123[12, 23]^+, [12, 12]^- .312[13, 13]^+, \\
 &\quad [12, 12]^- .131[23, 12]^+, [1, 1]^- .231[213, 312]^+, \\
 &\quad [1, 1]^- .312[312, 213]^+\}. \\
 v_2 : \mathcal{S}_{\mathcal{W}_3} &= \{[pqpppq, pqpppqpppppp, pqp]^- ppp.[pp, pp, pp]^+, \\
 &\quad [pqp, pqpppq, pqpppqpppppp]^- prr.[pp, pp, pp]^+, \\
 &\quad [pqpppqpppppp, pqp, pqpppq]^- ppq.[pp, pp, pp]^+\}; \\
 \mathcal{D}_{\mathcal{W}_4} &= \{[ppppp, ppppp]^- .qppp[qpp, pqpppqpppqpp]^+, \\
 &\quad [ppppp, ppppp]^- .rrpp[pqpppqpppqpp, qpp]^+, \\
 &\quad [p, p]^- .pppp[rrpppqpppqpppqpp, qpppqpppqpp]^+, \\
 &\quad [p, p]^- .qppp[qpppqpppqpp, rrpppqpppqpppqpp]^+\}. \\
 \tau_3 : \mathcal{S}_{\mathcal{W}_1} &= \emptyset; \\
 \mathcal{D}_{\mathcal{W}_1} &= \emptyset. \\
 \tau_4 : \mathcal{S}_{\mathcal{W}_4} &= \{[0, 0]^- 1010.[0, 0]^+, [0, 0]^- 1001.[0, 0]^+\}; \\
 \mathcal{D}_{\mathcal{W}_4} &= \emptyset.
 \end{aligned}$$

### 2.3. $\tau$ -PERIODIC POINTS

We call elements  $y \in X^+$ , respectively  $x \in X^-$ ,  $\tau$ -periodic when  $\tau^n(y) = y$ , respectively  $\tau^n(x) = x$ , for some  $n \geq 1$ . Let  $a$  be the last letter of a  $\tau$ -periodic  $x$  and  $b$  the first letter of a  $\tau$ -periodic  $y$ . If  $ab \in \mathcal{L}(\underline{X}_\tau)$ , then because every finite subword of  $x.y$  is contained in  $\tau^{nk}(ab)$  for some  $k$ , we have that  $x.y \in \underline{X}_\tau$ . And if another  $\tau$ -periodic word  $x' \in X^-$  ends in  $a' \neq a$  for which  $a'b \in \mathcal{L}(\underline{X}_\tau)$ , then  $x.y$  and  $x'.y$  are left special elements.

The class  $\mathcal{S}_p$  of left special elements obtained this way is computable. For a letter  $a \in \mathfrak{a}$  gives rise to a  $\tau$ -periodic word precisely when there is an  $n \in \mathbb{N}$  such that  $\tau^n(a)$  begins or ends in  $a$ , and there is a computable smallest integer  $N$  such that all possible first and last letters are attained at some power  $n \leq N$ . Furthermore, the set of two-letter words of  $\mathcal{L}(\underline{X}_\tau)$  is computable.

**Definition 2.16.** For any  $NS$ -cover  $\mathcal{W}$  we write

$$\mathcal{S}_{p\mathcal{W}} = \mathcal{S}_{\mathcal{W}} \cup \mathcal{S}_p.$$

The symmetric notation is applied to  $\mathcal{D}$  as well.

We also note that such a left special element can be written on bracket form. Indeed,

$$x.y = \overbrace{[\epsilon, \dots, \epsilon, v]}^{N-1} a.b \overbrace{[w, \epsilon, \dots, \epsilon]}^{N-1} +$$

where  $\tau^N(a) = va$  and  $\tau^N(b) = bw$ . Note that  $v, w \neq \epsilon$  by primitivity.

**Example 2.17.**  $\mathcal{S}_p$  and  $\mathcal{D}_p$  are empty for  $\tau_1, \tau_2, \tau_5, \tau_6$ , but

$$\begin{aligned} \tau_3 : \mathcal{S}_p &= \{[ab]^- a.a[ba]^+, [baa]^- b.a[ba]^+\}; \\ \mathcal{D}_p &= \{[ab]^- a.a[ba]^+, [ab]^- a.b[aaab]^+\}. \\ \tau_4 : \mathcal{S}_p &= \emptyset; \\ \mathcal{D}_p &= \{[1]^- 0.0[10, \epsilon]^+, [1]^- 0.1[0, \epsilon]^+\}. \end{aligned}$$

### 3. THE STRUCTURE OF SPECIAL WORDS

In the previous section we defined two classes of left special words which we denoted by  $\mathcal{S}_{\mathcal{W}}$  and  $\mathcal{S}_p$ , respectively, and let  $\mathcal{S}_{p\mathcal{W}}$  denote their union. In the present section we are going to prove that  $\mathcal{S}_{p\mathcal{W}}$  coincides with the set of left special words.

#### 3.1. AUXILIARY RESULTS

The following is a one-sided substitute for Lemma 1.2. It is proved using techniques from [14].

**Lemma 3.1.** *For a primitive and aperiodic substitution  $\tau$ , let  $x, y \in \underline{X}_\tau$ . If  $\tau(x) \sim_r \tau(y)$ , then  $x \sim_r y$ .*

*Proof.* We first note, as in [14], that there exists  $M \in \mathbb{N}$  such that

$$\tau^M(a) = \tau^M(b) \iff \tau^{M-1}(a) = \tau^{M-1}(b)$$

for any  $a, b \in \mathbf{a}$ . Since  $\tau$  is primitive, we may choose  $M$  so large that an element  $u \in \underline{X}_\tau$  has the property  $\tau^M(u) = u$ . As a consequence of [14], cf. [8] Corollary 12, we may choose  $x', y' \in \underline{X}_\tau$  such that

$$x \sim_o \tau^{M-1}(x') \quad y \sim_o \tau^{M-1}(y').$$

By assumption,  $\tau^M(x') \sim_r \tau^M(y')$ . We fix  $i, j \in \mathbb{Z}$  such that  $\tau^M(x')_{i+n} = \tau^M(y')_{j+n}$  for all  $n \in \mathbb{N}_0$ .

We define for each  $k \in \mathbb{N}_0$

$$a_k = |\tau^M(x'_{[0,k]})| - i \quad b_k = |\tau^M(y'_{[0,k]})| - j$$

and let  $A = \{a_0, a_1, \dots\}$ ,  $B = \{b_0, b_1, \dots\}$ . Our first goal is to prove that  $A \cap [L, \infty[ = B \cap [L, \infty[$  for an  $L \in \mathbb{N}$  chosen according to Mossé's two-sided recognizability property for  $\tau^M$ . This property states that with  $e_h^M = |\tau^M(u_{[0,h]})|$

and

$$u_{[e_h^M - L, e_h^M + L]} = u_{[l - L, l + L]}$$

then  $l \in \{e_k^M \mid k \in \mathbb{N}_0\}$ , cf. [14] Definition 1.2. Hence let  $k \in A, k \geq L$  and choose, using minimality, integers  $r, s$  such that

$$u_{[r, r+k+L+i]} = x'_{[0, k+L+i]} \quad u_{[s, s+k+L+j]} = y'_{[0, k+L+j]}.$$

Then

$$\{h + e_r^M + i \mid h \in \{a_0, \dots, a_{k+i+L}\}\} = \{e_h^M \mid r \leq h \leq r + k + i + L\}, \quad (5)$$

and

$$\{h + e_s^M + j \mid h \in \{b_0, \dots, b_{k+j+L}\}\} = \{e_h^M \mid s \leq h \leq s + k + j + L\}. \quad (6)$$

Choose by (5)  $h \in \{r, \dots, r + k + i + L\}$  such that  $e_h^M = k + e_r^M + i$ . Since

$$\begin{aligned} u_{[e_h^M - L, e_h^M + L]} &= u_{[e_r^M + k + i - L, e_r^M + k + i + L]} \\ &= \tau^M(x')_{[k+i-L, k+i+L]} \\ &= \tau^M(y')_{[k+j-L, k+j+L]} \\ &= u_{[e_s^M + k + j - L, e_s^M + k + j + L]} \end{aligned}$$

it follows by [14] (Th. 3.1 bis), that  $e_s^M + k + j = e_{h'}^M$  for some  $h'$ , where obviously  $s \leq h' \leq s + k + j$ . Using (6) we get that  $k \in B$ , as required to prove  $A \cap [L, \infty[ \subseteq B \cap [L, \infty[$ . The symmetric argument proves the other inclusion.

Let

$$n = \min\{h \in \mathbb{N}_0 \mid a_h \geq L\} \quad m = \min\{h \in \mathbb{N}_0 \mid b_h \geq L\}.$$

By what we have already proved,  $\{a_n, a_{n+1}, \dots\} = \{b_m, b_{m+1}, \dots\}$ , so

$$\begin{aligned} \tau^M(x'_{n+d}) &= \tau^M(x')_{[a_{n+d+i}, a_{n+d+1+i}[} \\ &= \tau^M(y')_{[a_{n+d+j}, a_{n+d+1+j}[} \\ &= \tau^M(y')_{[b_{m+d+j}, b_{m+d+1+j}[} \\ &= \tau^M(y'_{m+d}) \end{aligned}$$

for every  $d \in \mathbb{N}_0$ . By our initial assumption on  $M$ , also  $\tau^{M-1}(x'_{n+d}) = \tau^{M-1}(y'_{m+d})$ . Consequently,

$$x \sim_r \tau^{m-1}(x') \sim_r y,$$

as desired. □

3.2. THE MAIN THEOREM

**Lemma 3.2.** *Let  $\tau$  be a primitive and aperiodic substitution with an NS-cover  $\mathcal{W}$ , and assume that  $N$  is a  $\mathcal{W}$ -basic power for  $\tau$ . Suppose that  $x_0, x_1 \in \underline{X}_\tau$  are elements of the form*

$$x_i = \tilde{x}_i v_i \cdot \tilde{x}$$

where  $(v_0, v_1) \in \mathcal{W} \widehat{\times} \mathcal{W}$ . Then  $(v_0, v_1) \in \mathcal{G}_{\tau^N, \mathcal{W}}$ , and there exist  $y_0, y_1 \in \underline{X}_\tau$  of the form

$$y_i = \tilde{y}_i v_i \cdot \tilde{y}$$

such that  $x_i = \sigma^{-|\mathcal{L}(e)|}(\tau^N(y_i))$ , where  $e$  is the unique loop at  $(v_0, v_1)$ .

*Proof.* Let  $x_i$  be of the described form. As a consequence of [14], cf. [8] Corollary 12, there exist  $w_i \in \underline{X}_\tau$ , and integers  $m_i$  with  $0 \leq m_i < |\tau^N((w_i)_0)|$  such that

$$x_i = \sigma^{m_i}(\tau^N(w_i)).$$

We get that

$$(\tau^N(w_0))_{[m_0, \infty[} = (x_0)_{[0, \infty[} = (x_1)_{[0, \infty[} = (\tau^N(w_1))_{[m_1, \infty[}$$

and since  $\tau^N$  is also aperiodic and primitive, Lemma 3.1 applies to yield  $n_i \in \mathbb{Z}$  such that

$$(w_0)_{[n_0, \infty[} = (w_1)_{[n_1, \infty[}. \tag{7}$$

We may and shall assume that among the pairs  $(n_0, n_1) \in \mathbb{Z}^2$  satisfying (7), the sum  $n_0 + n_1$  is minimal. For if  $(n_0^j, n_1^j)$  satisfied (7) with  $n_0^j + n_1^j \rightarrow -\infty$ , we could use the fact that there are no  $\sigma$ -periodic points in  $X_\tau^+$  to prove that  $n_0^j - n_1^j$  is constant and then conclude that  $w_0 \sim_o w_1$ . This would lead to the contradiction  $x_0 \sim_o x_1$ .

Now by minimality

$$(w_0)_{n_0-1} \neq (w_1)_{n_1-1}.$$

Choose  $\ell_i \in \mathbb{Z}$  and  $u_0, u_1 \in \mathcal{W}$  such that

$$(w_i)_{[\ell_i, \infty[} = u_i \tilde{y}$$

where  $\tilde{y} = (w_0)_{[n_0, \infty[} = (w_1)_{[n_1, \infty[}$ . By definition of  $(\mathcal{V}_{\tau^N, \mathcal{W}}, \mathcal{E}_{\tau^N, \mathcal{W}})$ ,

$$\tau^N(u_j \tilde{y}) = \mathcal{L}_j^N(e') u'_j \mathcal{L}^N(e') \tau^N(\tilde{y})$$

for both  $j \in \mathbb{Z}_2$ , where  $e'$  is the unique edge in  $(\mathcal{V}_{\tau^N, \mathcal{W}}, \mathcal{E}_{\tau^N, \mathcal{W}})$  with range  $(u_0, u_1)$ , and  $(u'_0, u'_1)$  is the source of  $e'$ . Since this is a segment of  $x_i$ , we get that  $\tilde{x} = \mathcal{L}^N(e') \tau^N(\tilde{y})$  and that  $v_i = u'_i$ . Thus, since  $N$  is a  $\mathcal{W}$ -basic power for  $\tau$ , we have that  $(v_0, v_1)$  is a vertex of  $\mathcal{G}_{\tau^N, \mathcal{W}}$ .

We can now repeat the argument with

$$y_i = \tilde{y}_i u_i \cdot \tilde{y}$$

instead of

$$x_i = \tilde{x}_i v_i \tilde{x},$$

where  $\tilde{y}_i = (w_i)_{]-\infty, t_i[}$ , and we then get that  $(u_0, u_1)$  is a vertex of  $\mathcal{G}_{\tau^N, \mathcal{W}}$ , and hence that  $e'$  is a loop and  $(u_0, u_1) = (u'_0, u'_1) = (v_0, v_1)$ , and the result is established.  $\square$

**Theorem 3.3.** *Let  $\tau$  be a primitive and aperiodic substitution with an NS-cover  $\mathcal{W}$ . If  $x \in \underline{X}_\tau$  is a left special word, then  $x \in \mathcal{S}_{p\mathcal{W}}$ .*

*Proof.* Choose a  $\mathcal{W}$ -basic power  $N$  for  $\tau$ . Let  $x_0, x_1 \in \underline{X}_\tau$  be given with  $(x_0)_{[0, \infty[} = (x_1)_{[0, \infty[}$ , but  $(x_0)_{-1} \neq (x_1)_{-1}$ . By definition of NS-covers,  $x_i \in \text{Cyl}^-(v_i)$  for some unique  $v_i \in \mathcal{W}$ . Note that after interchanging  $x_0$  and  $x_1$  if necessary, we may assume that  $(v_0, v_1) \in \mathcal{W} \hat{\times} \mathcal{W}$ . Apply Lemma 3.2 to get elements  $y_i$  and an edge  $e$  with the properties stated there, and note that these elements are also left special and satisfy that  $y_i \in \text{Cyl}^-(v_i)$ .

When  $|\mathcal{L}(e)| = 0$ , Lemma 3.2 can be iterated to get that

$$x_i \in \bigcap_{i=1}^{\infty} \tau^{iN}(\underline{X}_\tau),$$

and this in turn implies that the  $x_i$  are  $\tau$ -periodic. When  $|\mathcal{L}(e)| > 0$ , we let  $z = \mathcal{L}(e)$  and note that Lemma 3.2 may be iterated to prove that

$$x_i = [\overbrace{\epsilon, \dots, \epsilon}^{N-1}, t_i]^- u_i \cdot [z, \overbrace{\epsilon, \dots, \epsilon}^{N-1}]^+$$

where  $t_i = \mathcal{L}_i^N(e)$ . We note that this word lies in  $\mathcal{S}_{\mathcal{W}}$  by construction.  $\square$

**Remark 3.4.** Note that when  $|\mathcal{L}(e)| > 0$  for all edges of  $\mathcal{G}_{\tau^N, \mathcal{W}}$  the first case does not occur and every left special word can be found in  $\mathcal{S}_{\mathcal{W}}$ . This will always be the case when  $\tau$  is *proper* in the sense that there exists  $M \in \mathbb{N}$  and letters  $r, l \in \mathbf{a}$  such that every word  $\tau^M(a), a \in \mathbf{a}$  begins in  $l$  and ends in  $r$ . This leads to the following separation result of importance in our paper [6].

**Proposition 3.5.** *Let  $\tau$  be a proper, primitive and aperiodic substitution with an NS-cover  $\mathcal{W}$ . For every left special word  $x$  of  $\mathbf{X}_\tau^+$  two words  $v_0, v_1 \in \mathcal{W}$  may be chosen with the property*

$$\forall y \in \mathbf{X}_\tau^+ : v_0 y, v_1 y \in \mathbf{X}_\tau^+ \implies x = y.$$

*Proof.* As seen in Remark 3.4,  $x \in \mathcal{S}_{\mathcal{W}}$ . Let  $(v_0, v_1)$  be a node of  $\mathcal{G}_{\tau^N, \mathcal{W}}$ , where  $N$  is a  $\mathcal{W}$ -basic power, from which we may initiate an infinite walk on  $\mathcal{G}_{\tau^N, \mathcal{W}}$  with labels yielding  $x$ . Using Lemma 3.2 as in the proof of Theorem 3.3 we get that when  $v_i y \in \mathbf{X}_\tau^+$ , then  $y$  will coincide with the labels read off this infinite walk, and hence  $y = x$ .  $\square$

#### 4. DECIDING SHIFT TAIL AND ORBIT EQUIVALENCE

We have managed to generate all special words of an aperiodic and primitive substitution. We have not yet, however, given an algorithm to decide which special words are adjusted. Similarly, since our method may output two or more elements which are orbit equivalent, or even identical, we have not yet explained how to count the number of orbit classes of special words or to compute the configuration graph. We solve these problems in the present section.

##### 4.1. A DECIDABLE RELATION

In this section we consider the following auxiliary relation, prove that it is decidable on words on the form  $[v]^-u.[w]^+$ , and that it is closely related to right shift tail equivalence.

**Definition 4.1.** Let  $\tau$  be aperiodic and primitive, and consider  $x, y \in \underline{X}_\tau$ . We write  $x \hookrightarrow y$  when there exists  $n \in \mathbb{N}_0$  such that  $x_{[0, \infty[} = y_{[n, \infty[}$ .

We first see in Lemma 4.2 that it is decidable when  $[v]^+ = ([w]^+)_{[k, \infty[}$  for  $k$  up to a certain predefined integer, and then pass to general  $k$  in Lemma 4.3.

**Lemma 4.2.** *Let  $\tau$  be aperiodic and primitive, and let  $v, w \in \mathcal{L}(\underline{X}_\tau) \setminus \{\epsilon\}$  and  $0 \leq k < |w|$  be given. We have*

$$[v]^+ = ([w]^+)_{[k, \infty[}$$

if and only if

$$v = w_{[k, |w|[} \tau(w_{[0, k[}).$$

*Proof.* If the second equality holds, we get

$$\begin{aligned} [v]^+ &= [w_{[k, |w|[} \tau(w_{[0, k[})]^+ \\ &= w_{[k, |w|[} \tau(w_{[0, k[}) \tau(w_{[k, |w|[} \tau(w_{[0, k[}) \tau^2(w_{[k, |w|[} \tau(w_{[0, k[}) \dots \\ &= w_{[k, |w|[} \tau(w) \tau^2(w) \dots \\ &= ([w]^+)_{[k, \infty[}. \end{aligned}$$

In the other direction, first note that if  $[v]^+ = [w]^+_{[k, \infty[}$  and  $0 \leq k < |w|$ ,

$$\begin{aligned} ([v]^+)_{[|v|+k, \infty[} &= \tau([v]^+)_{[k, \infty[} \\ &= \tau([w]^+_{[k, \infty[})_{[k, \infty[} \\ &= \tau([w]^+)_{[k+|\tau(w_{[0, k[})|, \infty[} \\ &= [w]^+_{[k+|\tau(w_{[0, k[})|+|w|, \infty[} \\ &= ([v]^+)_{[\tau(w_{[0, k[})|+|w|, \infty[} \end{aligned}$$

using the assumption on  $k$  in the third step. Since otherwise a subword of  $[v]^+$  would be  $\sigma$ -periodic, the two segments must agree, so the length of  $v$  must coincide with the length of  $w_{[k,|w|]}[\tau(w_{[0,k]})]$ . Reading off letters from the left we get equality of the words themselves.  $\square$

**Lemma 4.3.** *Let  $\tau$  be aperiodic and primitive, and consider  $v, w \in \mathcal{L}(\underline{X}_\tau) \setminus \{\epsilon\}$ . If  $[v]^+ \hookrightarrow [w]^+$  then*

$$v = (\tau^i(w)\tau^{i+1}(w))_{[j, j+|v|]}$$

where  $i \in \mathbb{N}_0$  satisfies  $|\tau^i(w)| \leq |v| \leq |\tau^{i+1}(w)|$  and  $0 \leq j < |\tau^i(w)|$ .

*Proof.* Assume that  $[v]^+ = ([w]^+)_{[n, \infty[}$  and write

$$n = \sum_{k=0}^{i-1} |\tau^k(w)| + j$$

with  $j \in \{0, \dots, |\tau^i(w)| - 1\}$ . Then  $[v]^+ = ([w]^+)_{[n, \infty[} = ([\tau^i(w)]^+)_{[j, \infty[}$ , and by Lemma 4.2 applied with  $\tau^i(w)$  in place of  $w$  we get that  $v$  is a subword of  $\tau^i(w)\tau^{i+1}(w)$ , and that

$$|v| = |\tau^i(w)| - j + |\tau(\tau^i(w))_{[0, j]}|.$$

Since  $|\tau(\tau^i(w))_{[0, j]}| \geq j$  we have  $|v| \geq |\tau^i(w)|$ , and since we have, for any word  $u$  and any  $\ell \leq |u|$ ,

$$\ell + |\tau(u_{[\ell, |u|]})| \geq |u|$$

we can apply this to  $u = \tau^i(w)$  and get that

$$\begin{aligned} |\tau^{i+1}(w)| &= |\tau(\tau^i(w))_{[0, j]}| + |\tau(\tau^i(w))_{[j, |\tau^i(w)|]}| \\ &= |v| - |\tau^i(w)| + j + |\tau(\tau^i(w))_{[j, |\tau^i(w)|]}| \\ &\geq |v|, \end{aligned}$$

as desired.  $\square$

Note that when  $w$  and  $v$  are given in the lemma above, there is only a finite number of  $i$  satisfying

$$|\tau^i(w)| \leq |v| \leq |\tau^{i+1}(w)|.$$

Thus “ $\hookrightarrow$ ” becomes decidable for elements of  $\underline{X}_\tau$  given on the form  $[v]^-u.[w]^+$ . To tie this in with right shift tail equivalence, we note:

**Proposition 4.4.** *Let  $\tau$  be aperiodic and primitive, and consider a finite set  $\mathcal{B} \subseteq \underline{X}_\tau$  which contains all left special words of  $\underline{X}_\tau$ . Then the equivalence relation induced by  $\hookrightarrow$  on  $\mathcal{B}$  coincides with right shift tail equivalence.*

*Proof.* Since the other implication is obvious, let  $x, x' \in \mathcal{B}$  and assume that  $x \sim_r x'$  to find a series of elements in  $\mathcal{B}$ , related by “ $\hookrightarrow$ ”, passing between  $x$  and  $x'$ . More precisely, assume that  $x_{[m, \infty[} = x'_{[m', \infty[}$ , where we may and shall assume that the pair  $(m, m')$  is chosen such that among pairs of nonnegative integers with this

property,  $m + m'$  is least possible. If  $m = 0$  or  $m' = 0$  we have  $x \hookrightarrow x'$  or  $x' \hookrightarrow x$ . If  $m, m' > 0$ , we get by the minimality assumption that  $x_{[m-1, \infty[} \neq x'_{[m'-1, \infty[}$  holds, whence  $\sigma^m(x)$  is left special, and thus  $x_{[m, \infty[} = x'_{[m', \infty[} = x''_{[0, \infty[}$  with  $x'' \in \mathcal{B}$ . Consequently,  $x'' \hookrightarrow x$  and  $x'' \hookrightarrow x'$ , wherefrom it follows that  $x$  and  $x'$  are related by any symmetric and transitive relation extending " $\hookrightarrow$ ".  $\square$

**Example 4.5.** For  $\tau_4$  our algorithm has produced a set

$$\{[0, 0]^-1010.[0, 0]^+, [0, 0]^-1001.[0, 0]^+, [1]^-0.0[10, \epsilon]^+, [1]^-0.1[0, \epsilon]^+\}$$

of special words which we enumerate  $x_1, \dots, x_4$ . Applying the results of the present section to  $\tau_4^2$  we get that  $x_1 \hookrightarrow x_2, x_2 \hookrightarrow x_1, x_3 \hookrightarrow x_1, x_4 \hookrightarrow x_1$ , but  $x_3 \not\hookrightarrow x_4$  and  $x_4 \not\hookrightarrow x_3$ . This demonstrates that the symmetrized relation induced by " $\hookrightarrow$ " is not an equivalence relation, and hence not the same as right shift tail equivalence.

#### 4.2. DECIDING EQUIVALENCES

**Theorem 4.6.** *Let  $\tau$  be a primitive and aperiodic substitution with an NS-cover  $\mathcal{W}$ , and with an NS-cover  $\mathcal{W}'$  given for  $\tau^{-1}$ . On a finite set  $\mathcal{B}$  with*

$$\mathcal{S}_{p\mathcal{W}} \cup \mathcal{D}_{p\mathcal{W}'} \subseteq \mathcal{B} \subseteq \underline{\mathcal{X}}_\tau$$

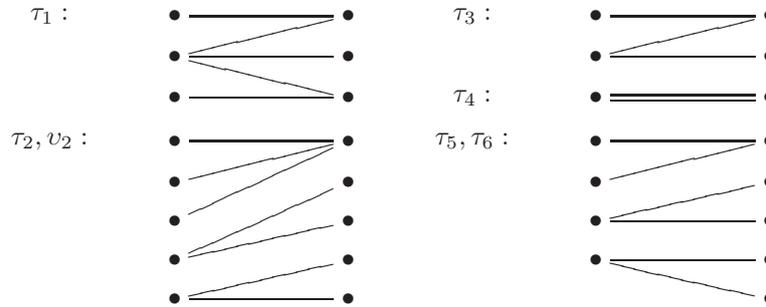
*of elements finitely presented on the form  $[v]^-u.[w]^+$ , right and left shift tail equivalence, as well as orbit equivalence, is decidable. Furthermore, it is decidable which special words are adjusted.*

*Proof.* Let  $x = [v]^-u.[w]^+$  and  $x' = [v']^-u'.[w']^+$  denote generic elements of  $\mathcal{B}$ . We have seen in Proposition 4.4 that right shift tail equivalence is generated by a relation which is decidable by Lemma 4.3. By symmetry, the same is true for left shift tail equivalence. To decide orbit equivalence, we first decide whether the shift tail equivalences hold, noting that the algorithms described above provide us with integers satisfying

$$([v]^-)_{]-\infty, l]} = ([w']^-)_{]-\infty, l']} \quad ([w]^+)_{[m, \infty[} = ([w']^+)_{[m', \infty[}$$

Checking orbit equivalence is hence reduced to comparing finite segments containing  $u$  and  $u'$ . The adjusted left special words are then those left special words  $x$  with the property that  $y \hookrightarrow x$  for each other left special  $y \in \mathcal{B}$  in the same orbit class.  $\square$

**Example 4.7.** Our algorithm leads to the following configuration data graphs:



### 5. CONCLUSION

In conclusion, our algorithm is laid out as follows. We have implemented it as a Java applet, see [4].

- (a) Check that  $\tau$  is aperiodic and primitive [15, 19].
- (b) Decide whether  $\tau$  is simplifiable or elementary [18].
- (c) If  $\tau$  is elementary, let  $v = \tau$ . If  $\tau$  is simplifiable, compute a simplification  $v$  of  $\tau$  [18].
- (d) Compute an integer such that  $\{w \in \mathcal{L}(\underline{X}_v) \mid |w| = n\}$  is an  $NS$ -cover for  $v$  and  $\{w \in \mathcal{L}(\underline{X}_{v^{-1}}) \mid |w| = n\}$  is an  $NS$ -cover for  $v^{-1}$  (2.5).
- (e) Compute graphs  $\mathcal{G}_{v, \mathcal{W}_n}$  and  $\mathcal{G}_{v^{-1}, \mathcal{W}'_n}$  and read off sets  $\mathcal{S}_{\mathcal{W}}$  and  $\mathcal{D}_{\mathcal{W}'}$  (2.8, 2.14).
- (f) Compute sets  $\mathcal{S}_p$  and  $\mathcal{D}_p$  (2.16).
- (g) Determine shift tail and orbit equivalence among elements in  $\mathcal{S}_{p\mathcal{W}} \cup \mathcal{D}_{p\mathcal{W}'}$  (4.6).
- (h) If  $\tau \neq v$ , transfer special elements back to the alphabet of  $\tau$  (2.2).

**Remark 5.1.** Obviously steps (c) and (h) are redundant when  $\tau$  is already elementary. And as noted in Remark 3.4, we may skip step (f) and work directly with  $\mathcal{S}_{\mathcal{W}}$  and  $\mathcal{D}_{\mathcal{W}'}$  when  $\mathfrak{L}(e) \neq \epsilon$  for all edges of  $\mathcal{G}_{\tau^N, \mathcal{W}}$  and  $\mathcal{G}_{(\tau^{-1})^N, \mathcal{W}'}$ , when  $N$  is an  $\mathcal{W}, \mathcal{W}'$ -bibasic power. This will always be the case when  $\tau$  is *proper* in the sense defined there.

**Example 5.2.** Applying step (h) of the algorithm one gets that two of the right special elements of  $\tau_2$  are  $[00000]^- .123000[12300, 012300012300012300]^+$  and  $[00000]^- .2200[012300012300012300, 12300]^+$ . This proves that  $\tau_2$  has no  $NS$ -cover, for since  $[00000]^{-1}$  and  $[00000]^{-2}$  are mapped to the same sequence under  $\tau_2$ , there are words of any length with this problematic behavior.

**Remark 5.3.** Our paper [5] shows how finer data associated to special words may be used to distinguish the flow classes of  $\tau_5$  and  $\tau_6$  even though their configuration graphs coincide.

## REFERENCES

- [1] M. Barge and B. Diamond, A complete invariant for the topology of one-dimensional substitution tiling spaces. *Ergodic Theory Dynam. Systems* **21** (2001) 1333–1358.
- [2] M. Barge, B. Diamond and C. Holton, Asymptotic orbits of primitive substitutions. *Theoret. Comput. Sci.* **301** (2003) 439–450.
- [3] M. Boyle and D. Lind, Exponential subdynamics. *Trans. Amer. Math. Soc.* **349** (1997) 55–102.
- [4] T.M. Carlsen and S. Eilers, *Java applet*, [www.math.ku.dk/~eilers/papers/cei](http://www.math.ku.dk/~eilers/papers/cei) (2002).
- [5] T.M. Carlsen and S. Eilers, Augmenting dimension group invariants for substitution dynamics. *Ergodic Theory Dynam. Systems* **24** (2004) 1015–1039.
- [6] T.M. Carlsen and S. Eilers, Matsumoto K-groups associated to certain shift spaces. *Doc. Math.* **9** (2004) 639–671.
- [7] T.M. Carlsen and S. Eilers, Ordered K-groups associated to substitutional dynamics. *J. Funct. Anal.* **238** (2006) 99–117.
- [8] F. Durand, B. Host, and C. Skau, Substitutional dynamical systems, Bratteli diagrams and dimension groups. *Ergodic Theory Dynam. Systems* **19** (1999) 953–993.
- [9] N.P. Fogg, Substitutions in dynamics, arithmetics and combinatorics. *Lect. Notes Math.* **1794** Springer-Verlag, Heidelberg (2002).
- [10] C. Holton and L.Q. Zamboni, Directed graphs and substitutions. *Theory Comput. Syst.* **34** (2001) 545–564.
- [11] B.P. Kitchens, *Symbolic dynamics; One-sided, two-sided and countable state Markov shifts*. Springer-Verlag, Berlin (1998).
- [12] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge (1995).
- [13] K. Matsumoto, Bowen-Franks groups as an invariant for flow equivalence of subshifts. *Ergodic Theory Dynam. Systems* **21** (2001) 1831–1842.
- [14] B. Mossé, Puissances de mots et reconnaissabilité des points fixes d’une substitution. *Theoret. Comput. Sci.* **99** (1992) 327–334.
- [15] J.-J. Pansiot, Decidability of periodicity for infinite words. *RAIRO-Theor. Inf. Appl.* **20** (1986) 43–46.
- [16] B. Parry and D. Sullivan, A topological invariant of flows on 1-dimensional spaces. *Topology* **14** (1975) 297–299.
- [17] M. Queffélec, *Substitution dynamical systems—spectral analysis*. Springer-Verlag, Berlin (1987).
- [18] G. Rozenberg and A. Salomaa, *The mathematical theory of L systems*. Academic Press Inc. Harcourt Brace Jovanovich Publishers, New York (1980).
- [19] H. Wielandt, Unzerlegbare, nicht negative Matrizen. *Math. Z.* **52** (1950) 642–648.

Communicated by J. Karhumäki.

Received July 5, 2004. Accepted August 1, 2006.