A graph approach to computing nondeterminacy in substitutional dynamical systems

Toke M. Carlsen and Søren Eilers

December 2002

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.

Determining K-groups of certain C^*-algebras we found (see [3]) an invariant of flow equivalence — akin and related to the dimension groups considered in [5] — of substitutional systems based on combinatorial and textual properties of the special elements. To investigate the invariant, we were naturally lead to concern ourselves with computability of certain 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 *tail equivalence*, a naturally occurring relation of importance in our invariant.

We wish to acknowledge [1], to which our work is closely related. Although the ends and ambitions of the present note and [1] do not overlap, the means seem to do. Our method was developed independently, but we recieved [1] 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 [1], and vice versa. The series of reductions based on replacing the substitutions in [1] is, however, computationally inconvenient for our purposes. Our Section 4 has no analogue there.

1.2 Substitutions

We refer to [6], [5], and [13] 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 $\mathfrak{a}$. A *substitution* is simply a map $\tau : \mathfrak{a} \longrightarrow \mathfrak{a}^\sharp$. We can extend τ to $\mathfrak{a}^\sharp$ or to

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

(with $\mathbb{N}_0 = \{0, 1, \dots\}, -\mathbb{N} = \mathbb{Z}\backslash\mathbb{N}_0$) in the obvious way, and define powers of τ recursively. To define the action of τ 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 the *opposite substitution* defined by reversing each word $\tau(\cdot)$ by τ^{-1}. Finally, an $|\mathfrak{a}|\times|\mathfrak{a}|$-matrix A_τ is associated to τ, 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 $\sigma : \mathfrak{a}^{\mathbb{Z}} \longrightarrow \mathfrak{a}^{\mathbb{Z}}$ 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 σ. We shall refer to such spaces by "$\underline{\mathsf{X}}$" 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 X^+ and X^-, respectively. There is a rich theory of shift spaces; we refer to [8] and [7].

For $-\infty \leq i < j \leq \infty$ 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 left or right infinite words by $\mathbb{N}_0$ or $-\mathbb{N}$.

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

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

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

We single out two important properties of substitutions below. The notation "$A > 0$" indicates that the matrix A has only positive entries.

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

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

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

It is decidable when a given substitution has these properties, cf. [11] and [15]. Primitive and aperiodic substitutions yield *minimal* shift spaces: all orbits $\{\sigma^n(x) \mid n \in \mathbb{Z}\}$ are dense in the product topology induded by the discrete topology on $\mathfrak{a}$, cf. [13]. Consequently, there are no (ultimately) *σ-periodic words* in $\underline{\mathsf{X}}_\tau, \mathsf{X}_\tau^+$ or X_τ^- for such τ: if $x_{k+n} = x_{m+n}$ for all n in the various index sets, then $k = m$. Further, we have

Lemma 1.2 *[[10], cf. [5, Corollary 10]] The map induced by τ on $\underline{\mathsf{X}}_\tau$ is injective, when τ is primitive and aperiodic.*

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

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

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

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

1.3 Orbit classes and special elements

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

(i) If there exists a 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 a n and a N such that $x_m = y_{n+m}$ for all $m > N$ then we say that x and y are *right tail equivalent* and write $x \sim_r y$.

(iii) If there exist a n and a N such that $x_m = y_{n+m}$ for all $m < N$ then we say that x and y are *left 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 $\underline{\mathsf{X}}/\sim_o$ which we also will denote by $\sim_r$ and $\sim_l$. We call an orbit class $[x]$ in $\underline{\mathsf{X}}/\sim_o$ left or right *special*, respectively, if there exists $[y] \in \underline{\mathsf{X}}/\sim_o$ such that $[x] \neq [y]$, but $[x] \sim_r [y]$ or $[x] \sim_l [y]$. A *left special word* $x \in \underline{\mathsf{X}}$ is a representative of such an orbit class with the property that $y \in \underline{\mathsf{X}}$ exists with

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

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

The symmetric definition defines a class of (adjusted) right special words. Classical results ([13, p. 107], [2, Theorem 3.9]) give:

Theorem 1.5 *When τ 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. Clearly this word is unique.

A nice way of describing the structure of special words using the equivalence relations $\sim_r$ and $\sim_l$ on $\underline{\mathsf{X}}/\sim_o$, 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 $\underline{\mathsf{X}}/\sim_r$ and of $\underline{\mathsf{X}}/\sim_l$, and for each orbit class $[x]_o$ with x a special element, we let an edge connect $A \in \underline{\mathsf{X}}/\sim_r$ with $B \in \underline{\mathsf{X}}/\sim_l$ if $[x]_o \in A$ and $[x]_o \in B$, and we label that edge $[x]_o$. Removing all vertices with no edges, we arrive at a bipartite graph which we shall denote as the *configuration graph* of $\underline{\mathsf{X}}$. 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.

Using the fact that flow equivalence ([12]) on shift spaces is generated by conjugacy and expansion (see [9, Lemma 2.1]) one easily sees that the configuration graph, and hence the number of orbit or tail classes of special elements, is a flow invariant.

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 τ-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. The main technical tool is the one-sided substitute for Lemma 1.2 which we shall be able to derive from the work of Mossé ([10]). In Section 4 we describe an algorithm for determining 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.

2 Collecting special elements

2.1 Elementary and simplifiable substitutions

We recall from [14, p. 17] that a substitution τ on the alphabet $\mathfrak{a}$ is *simplifiable* if it can be factored $\tau = f \circ g$ for maps

$$f : \mathfrak{b} \longrightarrow \mathfrak{a}^\sharp \qquad g : \mathfrak{a} \longrightarrow \mathfrak{b}^\sharp$$

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

It is decidable whether a substitution is simplifiable or elementary, cf. [14, 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 υ, we get f, g with the property

$$\tau^n = f \circ g \qquad \upsilon^n = g \circ f. \tag{2}$$

This was used in [11] to provide an algorithm for deciding aperodicity 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 preserve aperiodicity – this is a key observation in [11] – 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 τ is simplified to an elementary substitution υ, then υ is primitive and aperiodic.*

Proof: It follows easily that A_τ and A_υ are strongly shift equivalent, cf. [8]. Note further that A_υ must be essential, as otherwise a letter could be deleted from the alphabet. Applying [8, Proposition 4.5.10], we get the desired result. $\square$

Proposition 2.2 *Let τ be a simplifiable substitution and let υ be an elementary simplification with maps f, g satisfying* (2) *above. The map induced by f preserves orbit and tail equivalence, and maps the (left, right) special orbits of $\underline{\mathsf{X}}_\upsilon$ bijectively onto the set of (left, right) special orbits of $\underline{\mathsf{X}}_\tau$.*

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 υ (and υ^n) is primitive by Lemma 2.1.

Clearly, then, both maps send special elements to special elements. Let $x_1, \ldots, x_n$ be a choice of orbit inequivalent special words of $\underline{\mathsf{X}}_\tau$, representing all such orbit classes. We have that $\tau(x_1), \ldots, \tau(x_n)$ are orbit inequivalent special words of $\underline{\mathsf{X}}_\tau$, since τ is injective and $\underline{\mathsf{X}}_\tau$ 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 τ_6 are elementary, but τ_2 is simplifiable to*

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

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

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

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

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

$$\tau(v) = tv'z \qquad \tau(w) = uw'z \tag{3}$$

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

$$(v', w') \in \mathcal{W}\widehat{\times}\mathcal{W} \tag{+}$$

or

$$(w', v') \in \mathcal{W}\widehat{\times}\mathcal{W}. \tag{−}$$

Not every primitive and aperiodic substitution posseses an NS-cover – our example τ_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 posseses an NS-cover. Indeed, there is a computable integer n such that the set*

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

is an NS-cover.

Proof: By [14, 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 $v, w \in \mathcal{L}(\underline{\mathsf{X}}_\tau)$ 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 $a \dashv \tau^m(b)$ 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. $\square$

Implementation 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.*

Example 2.7 *For the substitutions considered in Example 2.8, the smallest number n such that the set of all words in the associated language is an NS-cover is*

τ_1	τ_1^{-1}	υ_2	υ_2^{-1}	$\tau_3 = \tau_3^{-1}$	τ_4	τ_4^{-1}
2	3	3	4	1	4	2

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

$$(v', w') \xrightarrow{z,t,u} (v, w),$$

where v', w' and z, t, u are the (obviously unique) elements satisfying (3). Let $\mathcal{E}_\tau$ denote the set of all such edges with (v, w) ranging over $\mathcal{W}\widehat{\times}\mathcal{W}$ and define labelings

$$\mathfrak{L} : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathsf{X}}_\tau), \mathfrak{L}_+ : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathsf{X}}_\tau), \mathfrak{L}_- : \mathcal{E} \longrightarrow \mathcal{L}(\underline{\mathsf{X}}_\tau)$$

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

$$\mathfrak{L}' : \mathcal{E} \longrightarrow \mathfrak{s} = \{+, -\},$$

according to which of the alternatives in Definition 2.4 is met. We will need to consider $\{+, -\}$ as the group $\mathbb{Z}_2$, but find this notation more suggestive.

In the following definition, the *essential part* of a given graph is the subgraph defined by deleting all vertices which do not have both incoming and outgoing edges.

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

Example 2.10 *For each substitution in Example 2.8 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 $\mathfrak{L}, \mathfrak{L}'$.

$$\tau_1 : \quad \begin{tikzcd} (21,12) \arrow[r, bend left, "{3,+}"] & (12,13) \arrow[l, bend left, "{\epsilon,+}"] \end{tikzcd}$$

$$\tau_1^{-1} : \quad \begin{tikzcd} (321,213) \arrow[r, bend left, "{21,+}"] & (131,213) \arrow[l, bend left, "{21,+}"] & (132,213) \arrow[loop above, "{1,-}"] \end{tikzcd}$$

$$\upsilon_2 : \quad \begin{tikzcd} (ppp,ppq) \arrow[rr, "{pp,-}"] & & (ppp,prr) \arrow[dl, "{pp,-}"] \\ & (ppq,prr) \arrow[ul, "{pp,+}"] & \end{tikzcd}$$

$$\upsilon_2^{-1} : \quad \begin{tikzcd} (pppq,pprr) \arrow[loop above, "{p,-}"] & (pppp,pppq) \arrow[loop above, "{ppppp,-}"] \end{tikzcd}$$

$$\tau_3 : \quad \begin{tikzcd} (a,b) \arrow[loop above, "{\epsilon,+}"] \end{tikzcd}$$

$$\tau_4 : \quad \begin{tikzcd} (1010,1001) \arrow[loop above, "{0,-}"] \end{tikzcd}$$

$$\tau_4^{-1} : \quad \begin{tikzcd} (10,01) \arrow[loop above, "{\epsilon,-}"] \end{tikzcd}$$

Lemma 2.11 *When $\mathcal{W}$ is an NS-cover for τ, then it is also an NS-cover for τ^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*

$$\mathfrak{L}(e_0, \ldots, e_{N-1}) = \mathfrak{L}(e_0)\tau(\mathfrak{L}(e_1)) \cdots \tau^{N-1}(\mathfrak{L}(e_{n-1})),$$
$$\mathfrak{L}'(e_0, \ldots, e_{N-1}) = \prod_{i=0}^{N-1} \mathfrak{L}'(e_i),$$

and

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

where $s_i \in \{+,-\}$ is defined recursively by $s_0 = \pm$ and $s_{i+1} = \mathfrak{L}'(e_i)s_i$.

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

$$\tau^2(v) = \tau(tv'z) = \tau(t)\tau(v')\tau(z) = \tau(t)t'v''z'\tau(z),$$
$$\tau^2(w) = \tau(tw'z) = \tau(u)\tau(w')\tau(z) = \tau(u)u'w''z'\tau(z),$$

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 τ^N. The remaining claims are straightforward. □

Proposition 2.12 *There is an N such that $\mathcal{G}_{\tau^N,\mathcal{W}}$ is a forest of loops all labeled $+$ by $\mathfrak{L}'$.*

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. □

We shall say that N is a $\mathcal{W}$*-basic power* of the primitive and aperiodic substitution τ (relative to the NS-cover $\mathcal{W}$) if $\mathcal{G}_{\tau^N,\mathcal{W}}$ meets the conditions of Proposition 2.12. Our result above proves that every primitive and aperiodic substitution with an NS-cover $\mathcal{W}$ has an $\mathcal{W}$-basic power. Suppose further that τ^{-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 τ and $\mathcal{W}'$-basic for τ^{-1}.

A class of left special words on bracket form can be read of the graph thus associated to a $\mathcal{W}$-basic 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

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

are words growing to infinity in both directions, as $\mathfrak{L}_+(e) \neq \emptyset, \mathfrak{L}_-(e) \neq \emptyset$ by definition of NS-covers. Thus both

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

and

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

are elements of $\underline{\mathsf{X}}_\tau$. 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*).

Implementation remark 2.13 *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 $+$-labeled loops, one proceeds as follows. For each vertex (v, w) in $\mathcal{G}_{\tau,\mathcal{W}}$, one follows outgoing edges

$$e_0, \ldots, e_n$$

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

$$s_0 = +, \qquad s_{i+1} = \mathfrak{L}'(e_i)s_i.$$

If $s_n = +$ one records

$$[\mathfrak{L}_{s_n}(e_n), \ldots, \mathfrak{L}_{s_0}(e_0)]^- v.[\mathfrak{L}(e_0), \ldots, \mathfrak{L}(e_n)]^+,$$
$$[\mathfrak{L}_{-s_n}(e_n), \ldots, \mathfrak{L}_{-s_0}(e_0)]^- w.[\mathfrak{L}(e_0), \ldots, \mathfrak{L}(e_n)]^+.$$

If $s_n = -$ one needs to consider

$$[\mathfrak{L}_{-s_n}(e_n), \ldots, \mathfrak{L}_{-s_0}(e_0), \mathfrak{L}_{s_n}(e_n), \ldots, \mathfrak{L}_{s_0}(e_0)]^- v.[\mathfrak{L}(e_0), \ldots, \mathfrak{L}(e_n)]^+,$$
$$[\mathfrak{L}_{s_n}(e_n), \ldots, \mathfrak{L}_{s_0}(e_0), \mathfrak{L}_{-s_n}(e_n), \ldots, \mathfrak{L}_{-s_0}(e_0)]^- w.[\mathfrak{L}(e_0), \ldots, \mathfrak{L}(e_n)]^+.$$

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

Example 2.14 *Reading off elements on the graphs found in Example 2.10* *we get*

$$
\begin{aligned}
\tau_1 :& \mathcal{S}_{\mathcal{W}_2} = \{[1,13]^{-}12.[\epsilon,3]^{+}, [12,12]^{-}13.[\epsilon,3]^{+}, [13,1]^{-}21.[3,\epsilon]^{+}, [12,12]^{-}12.[3,\epsilon]^{+}\};\\
& \mathcal{D}_{\mathcal{W}_3} = \{[12,12]^{-}.123[12,23]^{+}, [12,12]^{-}.312[13,13]^{+},\\
&\qquad [12,12]^{-}.131[23,12]^{+}, [1,1]^{-}.231[213,312]^{+},\\
&\qquad [1,1]^{-}.312[312,213]^{+}\};\\
\upsilon_2 :& \mathcal{S}_{\mathcal{W}_3} = \{[pqpppqp, pqpppqpppppp, pqp]^{-}ppp.[pp,pp,pp]^{+},\\
&\qquad [pqp, pqpppqp, pqpppqpppppp]^{-}prr.[pp,pp,pp]^{+},\\
&\qquad [pqpppqpppppp, pqp, pqpppqp]^{-}ppq.[pp,pp,pp]^{+}\};\\
& \mathcal{D}_{\mathcal{W}_4} = \{[ppppp, ppppp]^{-}.qppp[qpp, pqpppqpppqpp]^{+},\\
&\qquad [ppppp, ppppp]^{-}.rrpp[pqpppqpppqpp, qpp]^{+},\\
&\qquad [p,p]^{-}.pppp[rrpppqpppqpppqpp, qpppqpppqpp]^{+},\\
&\qquad [p,p]^{-}.qppp[qpppqpppqpp, rrpppqpppqpppqpp]^{+}\};\\
\tau_3 :& \mathcal{S}_{\mathcal{W}_1} = \emptyset;\\
& \mathcal{D}_{\mathcal{W}_1} = \emptyset.\\
\tau_4 :& \mathcal{S}_{\mathcal{W}_4} = \{[0,0]^{-}1010.[0,0]^{+}, [0,0]^{-}1001.[0,0]^{+}\}\\
& \mathcal{D}_{\mathcal{W}_4} = \emptyset.
\end{aligned}
$$

2.3 τ-periodic points

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

Definition 2.15 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,

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

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

Example 2.16 *$\mathcal{S}_p$ is empty for $\tau_1, \tau_2, \tau_5, \tau_6$, but*

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

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 [10].

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

Proof: We first note, as in [10], that there exists $M \in \mathbb{N}$ such that

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

for any $a, b \in \mathfrak{a}$. Since τ is primitive, we may choose M so large that an element $u \in \underline{\mathsf{X}}_\tau$ has the property $\tau^M(u) = u$. As a consequence of [10], cf. [5, Corollary 12], we may choose $x', y' \in \underline{\mathsf{X}}_\tau$ such that

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

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 $k \in \mathbb{N}_0$

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

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 reckognizability property for τ^M. This property states that with $e_h^M = |\tau^M(u_{[0,h[})|$ and

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

then $l \in \{e_k^M \mid k \in \mathbb{N}_0\}$, cf. [10, Définition 1.2]. Hence let $k \in A, k \geq L$ and choose, using minimality, integers r, s such that

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

Then

$$\{h + e_r^M + i \mid h \in \{a_0, \dots, a_{k+i+L}\}\} = \{e_h^M \mid r \leq h \leq r + k + i + L\}, \quad (4)$$

and

$$\{h + e_s^M + j \mid h \in \{b_0, \dots, b_{k+j+L}\}\} = \{e_h^M \mid s \leq h \leq s + k + j + L\}. \quad (5)$$

Choose by (4) $h \in \{r, \dots, r + k + i + L\}$ 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 [10, Théoréme 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 (5) 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

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

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, $x \sim_r y$, as desired. $\square$

3.2 The main theorem

Lemma 3.2 *Let τ be a primitive and aperiodic substitution with an NS-cover $\mathcal{W}$, and assume that N is a $\mathcal{W}$-basic power for τ. Suppose that $x_1, x_2 \in \underline{\mathsf{X}}_\tau$ are elements of the form*

$$x_i = \widetilde{x}_i v_i . \tilde{x}$$

where $(v_1, v_2) \in \mathcal{W} \widehat{\times} \mathcal{W}$. Then there exist $y_1, y_2 \in \underline{\mathsf{X}}_\tau$ of the form

$$y_i = \widetilde{y}_i v_i . \tilde{y}$$

such that $x_i = \sigma^{-|\mathfrak{L}(e)|}(\tau^N(y_i))$, where e is the unique loop at $(v_1, v_2) \in \mathcal{G}_{\tau,\mathcal{W}}$.

Proof: As a consequence of [10], cf. [5, Corollary 12], there exist $w_i \in \underline{\mathsf{X}}_\tau$, and integers m_i with $0 \le m_i < |\tau^N((w_i)_{[0]})|$ such that

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

We get that

$$(\tau^N(w_1))_{[m_1,\infty[} = (x_1)_{[0,\infty[} = (x_2)_{[0,\infty[} = (\tau^N(w_2))_{[m_2,\infty[}$$

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

$$(w_1)_{[n_1,\infty[} = (w_2)_{[n_2,\infty[}. \tag{6}$$

We may and shall assume that among the pairs $(n_1, n_2) \in \mathbb{Z}^2$ satisfying (6), the sum $n_1 + n_2$ is minimal. For if (n_1^j, n_2^j) satisfied (6) with $n_1^j + n_2^j \longrightarrow -\infty$, we could use the fact that there are no σ-periodic points in X_τ^+ to prove that

$n_1^j - n_2^j$ is constant and then conclude that $w_1 \sim_o w_2$. This would lead to the contradiction $x_1 \sim_o x_2$.

Now by minimality

$$(w_1)_{[n_1-1]} \neq (w_2)_{[n_2-1]}.$$

Choose $\ell_i \in \mathbb{Z}$ and $v_1, v_2 \in \mathcal{W}$ such that

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

where $\tilde{y} = (w_1)_{[n_1,\infty[} = (w_2)_{[n_2,\infty[}$. Since we are working with a $\mathcal{W}$-basic power τ^N, we have

$$\tau^N(v_i\tilde{y}) = t_i v_i z \tau^N(\tilde{y}),$$

and since this is a segment of x_i, we get that z begins at index 0. Consequently, $(u_1, u_2) = (v_1, v_2)$. $\square$

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

Proof: Choose a $\mathcal{W}$-basic power τ^N. Let $x_1, x_2 \in \underline{\mathsf{X}}_\tau$ be given with $(x_1)_{[0,\infty[} = (x_2)_{[0,\infty[}$, but $(x_1)_{[-1]} \neq (x_2)_{[-1]}$. By definition of NS-covers, $(x_i) \in \mathsf{Cyl}^-(u_i)$ for some unique $v_i \in \mathcal{W}$. Note that after interchanging x_1 and x_2 if necessary, we may assume that $(v_1, v_2) \in \mathcal{W}\widehat{\times}\mathcal{W}$. Apply Lemma 3.2 to get y_i with the property stated there, and note that these elements are also left special and satisfy that $(y_i) \in \mathsf{Cyl}^-(v_i)$.

When $z = \epsilon$, Lemma 3.2 can be iterated to get that

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

and this in turn implies that the x_i are τ-periodic. When $z \neq \epsilon$, Lemma 3.2 may be iterated to prove that

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

We note that this word lies in $\mathcal{S}_\mathcal{W}$ by construction. $\square$

We note in passing that when $\mathfrak{L}(e) \neq \epsilon$ for all edges of $\mathcal{G}_{\tau^N,\mathcal{W}}$ the case $z = \epsilon$ above does not occur and every left special word can be found in $\mathcal{S}_\mathcal{W}$.

4 Deciding 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 tail equivalence.

Definition 4.1 Let τ be aperiodic and primitive, and consider $x, y \in \underline{\mathsf{X}}_{\tau}$. 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 τ be aperiodic and primitive, and let $v, w \in \mathcal{L}(\underline{\mathsf{X}}_{\tau})\backslash\{\epsilon\}$ and $0 \le k \le |w|$ be given. We have*

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

if and only if

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

Proof: If the equality of words holds, we get

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

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

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

using the assumption on k in the third step. Since otherwise a subword of v would be σ-periodic, the two powers 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. □

Lemma 4.3 *Let τ be aperiodic and primitive, and consider $v, w \in \mathcal{L}(\underline{X}_\tau)\backslash\{\epsilon\}$. If $[v]^+ = ([w]^+)_{[n,\infty[}$ for some $n \in \mathbb{N}_0$ then*

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

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

Proof: Write

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

with $j \in \{0, \ldots, |\tau^i(w)| - 1\}$. Then $[w]^+ = ([w]^+)_{[n,\infty[} = ([\tau^i(w)]^+)_{[n+j,\infty[}$, 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 $|\tau(\tau^i(w)_{[0,j[})| \ge j$ we have $|v| \ge |\tau^i(w)|$, and since we have, for any word u and any $\ell \le |u|$,

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

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

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

as desired. $\square$

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

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

In fact, unless a letter $a \in \mathfrak{a}$ exists with $|\tau(a)| = 1$, i is uniquely determined. Thus "$\hookrightarrow$" becomes decidable for elements of $\underline{\mathsf{X}}_\tau$ given on the form $[v]^-u.[w]^+$. To tie this in with right tail equivalence, we note:

Proposition 4.4 *Let τ be aperiodic and primitive, and consider a finite set $\mathcal{B} \subseteq \underline{\mathsf{X}}_\tau$ which contains all left special words of $\underline{\mathsf{X}}_\tau$. Then the equivalence relation induced by $\hookrightarrow$ on $\mathcal{B}$ coincides with left 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 "$\hookrightarrow$", leading from x to 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 $x'' \hookrightarrow x'$, proving the claim. $\square$

Example 4.5 *For τ_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 $x_1, \ldots, x_4$. Applying the results of the present section to τ_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 tail equivalence.

4.2 Deciding equivalences

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

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

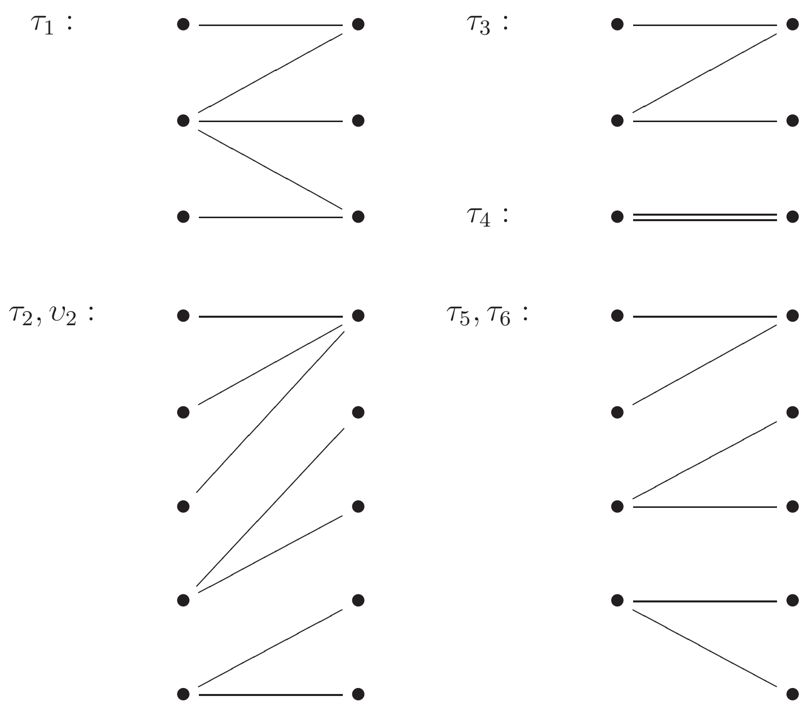
of elements finitely presented on the form $[v]^- u.[w]^+$, right and left tail equivalence, as well as orbit equivalence, is decidable. Furthermore, it is decidable which special words are adjusted.

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 tail equivalence is generated by a relation which is decidable by Lemma 4.3. By symmetry, the same is true for left tail equivalence. To decide orbit equivalence, we first decide whether the tail equivalences hold, noting that the algorithms described above provide us with integers satisfying

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

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 τ is aperiodic and primitive ([11],[15]).

(b) Decide whether τ is simplifiable or elementary ([14]).

(c) If τ is elementary, let $\upsilon = \tau$. If τ is simplifiable, compute a simplification υ of τ ([14]).

(d) Compute an integer such that $\mathcal{W}_n$ is an NS-cover for υ (2.5).

(e) Compute graphs $\mathcal{G}_{\upsilon,\mathcal{W}_n}$ and $\mathcal{G}_{\upsilon^{-1},\mathcal{W}_n}$ and read off sets $\mathcal{S}_{\mathcal{W}}$ and $\mathcal{D}_{\mathcal{W}'}$ (2.8, 2.13).

(f) Compute sets $\mathcal{S}_p$ and $\mathcal{D}_p$ (2.15).

(g) Determine tail and orbit equivalence among elements in $\mathcal{S}_{p\mathcal{W}} \cup \mathcal{D}_{p\mathcal{W}'}$ (4.6).

(h) If $\tau \neq \upsilon$, transfer special elements back to the alphabet of τ (2.2).

Remark 5.1 Obviously steps (c) and (h) are redundant when τ is already elementary. And as noted after the proof of Theorem 3.3, 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 τ is *proper* in the sense that there exists $M \in \mathbb{N}$ and letters $r, l \in \mathfrak{a}$ such that every word $\tau^M(a), a \in \mathfrak{a}$ begins in l and ends in r.

Example 5.2 *Applying step (h) of the algorithm one gets that two of the right special elements of* τ_4 *are* $[00000]^-.123000[12300, 012300012300012300]^+$ *and* $[00000]^-.2200[012300012300012300, 12300]^+$. *This proves that* τ_4 *has no* NS*-cover, for since* $[00000]^-1$ *and* $[00000]^-2$ *are mapped to the same sequence under* τ_4*, there are words of any length with this problematic behavior.*

Remark 5.3 Our paper [3] shows how finer data associated to special words may be used to distinguish the flow classes of τ_5 and τ_6 even though their configuration graphs coincide.

References

[1] M. Barge, B. Diamond, and C. Holton, *Asymptotic orbits of primitive substitutions*, preprint, 2002.

[2] M. Boyle and D. Lind, *Exponential subdynamics*, Trans. Amer. Math. Soc. **349** (1997), 55–102.

[3] T.M. Carlsen and S. Eilers, *Augmenting dimension group invariants for substitution dynamics*, preprint, `www.math.ku.dk/~eilers/papers/ceii`, 2002.

[4] ———, *Java applet*, `www.math.ku.dk/~eilers/papers/cei`, 2002.

[5] F. Durand, B. Host, and C. Skau, *Substitutional dynamical systems, Bratteli diagrams and dimension groups*, Ergodic Theory Dynam. Systems **19** (1999), no. 4, 953–993. MR 2000i:46062

[6] N.P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Mathematics, vol. 1794, Springer-Verlag, Heidelberg, 2002.

[7] B.P. Kitchens, *Symbolic dynamics*, Springer-Verlag, Berlin, 1998, One-sided, two-sided and countable state Markov shifts. MR 98k:58079

[8] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995. MR 97a:58050

[9] K. Matsumoto, *Bowen-Franks groups as an invariant for flow equivalence of subshifts*, Ergodic Theory Dynam. Systems **21** (2001), no. 6, 1831–1842. MR 2002i:37015

[10] B. Mossé, *Puissances de mots et reconnaissabilité des points fixes d'une substitution*, Theoret. Comput. Sci. **99** (1992), no. 2, 327–334. MR 93f:68076

[11] J.-J. Pansiot, *Decidability of periodicity for infinite words*, RAIRO Inform. Théor. Appl. **20** (1986), no. 1, 43–46. MR 87m:68048

[12] B. Parry and D. Sullivan, *A topological invariant of flows on* 1-*dimensional spaces*, Topology **14** (1975), no. 4, 297–299. MR 53 #9179

[13] M. Queffélec, *Substitution dynamical systems—spectral analysis*, Springer-Verlag, Berlin, 1987. MR 89g:54094

[14] G. Rozenberg and A. Salomaa, *The mathematical theory of L systems*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. MR 82g:68053

[15] H. Wielandt, *Unzerlegbare, nicht negative Matrizen*, Math. Z. **52** (1950), 642–648. MR 11,710g

Department of Mathematics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
`toke@math.ku.dk` `eilers@math.ku.dk`