

Coding of substitution dynamical systems as shifts of finite type

PAUL SURER

*Departamento de Matemática, IBILCE - UNESP, São José do Rio Preto, Brazil
(e-mail: me@palovsky.com)*

(Received 26 November 2013 and accepted in revised form 7 July 2014)

Abstract. We develop a theory that allows us to code dynamical systems induced by primitive substitutions continuously as shifts of finite type in many different ways. The well-known prefix–suffix coding turns out to correspond to one special case. We precisely analyse the basic properties of these codings (injectivity, coding of the periodic points, properties of the presentation graph, interaction with the shift map). A lot of examples illustrate the theory and show that, depending on the particular coding, several amazing effects may occur. The results give new insights into the theory of substitution dynamical systems and might serve as a powerful tool for further researches.

1. Introduction

In 2001 Canterini and Siegel [CS01a] and, independently, Holton and Zamboni [HZ01] showed how a substitution dynamical system can be coded as a shift of finite type. This representation is frequently used for studying Rauzy fractals (see, for example, [CS01b]) and exposes the strong relation between substitution dynamical systems and numeration systems (see, for example, [BS05]).

The results in [CS01a] and [HZ01] are more or less equivalent; in particular, the respective shifts of finite types are isomorphic. The main point is that the coding map is continuous, surjective and measure-theoretically injective. Total injectivity depends on the respective substitution dynamical system. The shift map can be described by an adic transformation on the set of edges of the presentation graph. Comparable results can be found in [DHS99, For97, Ver81, VL92] from the point of view of Bratteli–Vershik systems.

The aim of the actual research is to present a crucial generalization of these results. In particular, we will see that there are a lot of different ways to code a substitution dynamical system continuously and measure-theoretically bijective as a shift of finite type. The codings from Canterini and Siegel and from Holton and Zamboni will turn out to

correspond to one special case. However, we will be able to transfer and generalize the majority of the results obtained in [CS01a, HZ01].

For a given substitution the actual coding of the induced dynamical system will depend on a combinatorial object, a so-called *coding prescription*. We summarize the main properties of the coding.

Presentation graph. The set of vertices of the presentation graph (referred to as the *coding graph*) always consists of the induced words of length two. The exact positions and labels of the edges depend on the particular coding prescription. Even the number of edges may differ.

Surjectivity and injectivity. The coding is always surjective. The injectivity depends on the actual coding prescription as well as on the structure of the periodic points. However, measure-theoretically the coding is always injective. We will also define a stronger form of injectivity, the so-called *traceability*. It ensures that the elements of the substitution dynamical system can be recovered by a consistent representation in terms of their code.

We will give necessary and sufficient conditions for injectivity and traceability. Contrary to [CS01a, HZ01], we will see that each substitution dynamical system allows a traceable and, hence, completely injective coding (possibly by considering higher powers of the substitution).

Interaction with the shift map. Similar to [CS01a, HZ01] we will define a partial ordering on the edges of the presentation graph that extends to a partial ordering on the induced shift of finite type. The induced successor map conjugates with the shift map, provided that the respective coding prescription satisfies a certain smoothness condition.

As our main innovative method we consequently exploit the inverse shift, which is uniquely defined since we will exclusively concentrate on bi-infinite substitution dynamical systems. Furthermore, we depart from the usual ‘letter-centred’ representation of the elements of substitution dynamical systems. Indeed, by the classical definition, the words are considered as bi-infinite sequences over an alphabet \mathcal{A} where the letter with the index 0 acts as point of reference. We will elaborate an alternative approach that considers the elements as a composite of a left-infinite and a right-infinite sequence that are equally treated. This will be realized by using sequences over the alphabet \mathcal{A}^2 whose ‘letters’ overlap, that is, the right part of one term equals the left part of the term on the right. Therefore, the centre is formed by a duple that depends on both the left-infinite and the right-infinite subword.

The paper is organized as follows. In §2 we state important notation, give the classical definition of substitution dynamical systems, recall the concept of recognizability given in [Mos92] and the subsequent idea of desubstitution, and summarize the results from [CS01a, HZ01]. In §3 we state the definition of substitution dynamical systems in a slightly modified way, translate several basic properties into our new notation and make some observations concerning the periodic points. In §4 we introduce coding prescriptions and develop the main tools in order to finally define our coding. We define the coding graph in §5. We will see that the coding maps a substitution dynamical system continuously and

surjectively onto the induced shift of finite type. Furthermore, we will introduce the notion of traceability and study the behaviour of the periodic points. The problem of injectivity is solved completely in §6. We describe exactly the coding prescriptions that provide an injective coding. For coding prescriptions of a special shape we may define an ordering on the edges of the coding graph that extends to an ordering of the infinite paths. We precisely elaborate this in §7. It will turn out that the induced successor map is conjugate to the shift map of the substitution dynamical system. The results of the paper are demonstrated by several examples in §8.

2. Preliminaries: classical definitions and results

In the present section we first state some formalisms concerning sequences over a finite set and introduce the notion of substitution. Then we state the classical definition of a substitution dynamical system, which is the principal mathematical object of the present paper. This definition is based on the majority of research papers concerning this topic. Note that in the next section we restate the definition using a slightly different approach. We also recall the notion of a shift of finite type. Finally, we summarize the results obtained in [CS01a, HZ01].

2.1. Words and substitutions. The main objects that we will deal with are finite and infinite words over some finite set (alphabet) \mathcal{A} . Note that in the context of substitutions it has become customary to use alphabets that consist of consecutive integers starting with 1. For the most part we will follow this convention. But note that we will meet words over other alphabets, too; in particular, for a given alphabet \mathcal{A} we will frequently consider sequences over \mathcal{A}^2 .

We denote the set of finite words over \mathcal{A} by \mathcal{A}^* . Observe that \mathcal{A}^* , endowed with the concatenation of words, forms a free monoid. The empty word $\varepsilon \in \mathcal{A}^*$ acts as the neutral element. For finite words over \mathcal{A} we use upper-case letters.

For a finite word A we denote by $|A|$ the length of A , by A_i the i th letter of A and by $A_{[i,j]}$ the subword formed by the letters from position i to j (provided that $1 \leq i \leq j \leq |A|$). For convenience we also allow negative values as indices. For $-1 \geq j \geq i \geq -|A|$ let $A_i := A_{|A|+i+1}$ be the i th letter from the right and $A_{[i,j]} = A_{[|A|+i+1, |A|+j+1]}$. However, we always set $A_{[i,j]} := \varepsilon$ for $i > j$.

Concerning infinite words over \mathcal{A} , we have to distinguish between one-sided infinite words and bi-infinite words. They will always be denoted by lower-case letters.

For a left-infinite word u the rightmost letter will be u_{-1} , the letter to its left will be u_{-2} , and so on. Therefore, we write $\mathcal{A}^{\mathbb{Z}^-}$ for the set of left-infinite words over \mathcal{A} . A right-infinite word v starts with the letters v_1, v_2 , and so on. Consequently, the set of right-infinite words is denoted by $\mathcal{A}^{\mathbb{Z}^+}$. Bi-infinite words will be indexed by integers and, thus, identified with $\mathcal{A}^{\mathbb{Z}}$.

For addressing subwords of a (bi)-infinite word w we proceed analogously as for finite words: we denote by $w_{[i,j]}$ the finite word starting at the index i and ending at the index j . Note that we must have $i \leq j \leq -1$ for left-infinite sequences, $1 \leq i \leq j$ for right-infinite sequences and $i \leq j$ for bi-infinite sequences, otherwise the expression yields the empty word ε .

For right-infinite and bi-infinite words w we define the left shift $S : \mathcal{A}^{\mathbb{Z}^+} \longrightarrow \mathcal{A}^{\mathbb{Z}^+}$ such that $(Sw)_n = w_{n+1}$ for all n . Note that the shift is not reversible for right-infinite words. For bi-infinite words S is a bijective map and the inverse shift S^{-1} is well defined.

Finally, we define the language $\mathfrak{L}(w) \subset \mathcal{A}^*$ of an infinite word w to be the set of finite words (factors) that appear in w .

A substitution is a morphism defined over \mathcal{A}^* that satisfies some additional requirements.

Definition 2.1. (Substitution) We call a map $\sigma : \mathcal{A} \rightarrow \mathcal{A}^*$ and its extension to \mathcal{A}^* by concatenation a *substitution* over the alphabet \mathcal{A} if:

- (i) for all $a \in \mathcal{A} : \sigma(a) \neq \varepsilon$;
- (ii) there exists $a \in \mathcal{A} : \lim_{n \rightarrow \infty} |\sigma^n(a)| = \infty$.

A substitution is called *primitive* if there exists an integer $n \in \mathbb{N}$ such that a appears in $\sigma^n(b)$ for every pair of letters $a, b \in \mathcal{A}$.

A substitution σ can also be extended to one-sided infinite words by concatenation. For applying σ on bi-infinite words we need a point of reference which usually is the letter indexed with 0. In particular, we have

$$\sigma(\cdots w_{-2}w_{-1}w_0w_1w_2\cdots) = \cdots w'_{-2}w'_{-1}w'_0w'_1w'_2\cdots$$

with $w'_0w'_1w'_2\cdots = \sigma(w_0)\sigma(w_1)\sigma(w_2)\cdots$ and $\cdots w'_{-3}w'_{-2}w'_{-1} = \cdots \sigma(w_{-2})\sigma(w_{-1})$.

A substitution σ induces the set $L \subset \mathcal{A}^{\mathbb{Z}^-}$ of left-periodic words and the set $R \subset \mathcal{A}^{\mathbb{Z}^+}$ of right-periodic words that are invariant with respect to a finite number of applications of σ :

$$L := \{u \in \mathcal{A}^{\mathbb{Z}^-} \mid \exists n \geq 1 : \sigma^n(u) = u\},$$

$$R := \{v \in \mathcal{A}^{\mathbb{Z}^+} \mid \exists n \geq 1 : \sigma^n(v) = v\}.$$

By [Que10] the sets L and R are both non-empty. Note that each element of L (R) is uniquely identified by its rightmost (leftmost) letter. Hence, $|L| \leq |\mathcal{A}|$ and $|R| \leq |\mathcal{A}|$.

Analogously we may define the set

$$P' := \{z \in \mathcal{A}^{\mathbb{Z}^+} \mid \exists n \geq 1 : \sigma^n(z) = z\}$$

of bi-infinite periodic words. Obviously each element of $z \in P'$ is a unique composite of an element of L and an element of R , that is, $P' \cong L \times R$. We will see that in the context of substitution dynamical systems this set P' has to be handled with care.

2.2. Different types of symbolic dynamical systems. Symbolic dynamical systems are shift-invariant subsets of the infinite words over a finite alphabet \mathcal{A} . We differentiate between one-sided and two-sided ones. Symbolic dynamical systems can be equipped with the product topology of the discrete topology on each copy of \mathcal{A} .

We will deal with two different types of symbolic dynamical systems. The first are substitution dynamical systems. These consist of bi-infinite sequences over an alphabet \mathcal{A} that are defined via the language induced by a primitive substitution σ . We will summarize several important facts concerning this topic. For a more detailed survey on substitution dynamical systems we refer to [Fog02, Que10].

Definition 2.2. (Substitution dynamical system: classical) Let σ be a primitive substitution over an alphabet \mathcal{A} and

$$\mathcal{L}(\sigma) := \{A \in \mathcal{A}^* \mid \exists a \in \mathcal{A}, k \in \mathbb{N} \text{ such that } A \text{ is a factor of } \sigma^k(a)\}.$$

The symbolic dynamical system (Ω_σ, S) , where

$$\Omega_\sigma = \{w \in \mathcal{A}^{\mathbb{Z}} \mid \mathcal{L}(w) \subseteq \mathcal{L}(\sigma)\},$$

is the *substitution dynamical system induced by σ* .

Note that a great deal of research, especially less recent studies, considers one-sided substitution dynamical systems. This approach is justified by the fact that substitution dynamical systems have topological entropy zero. However, the use of bi-infinite systems becomes more and more relevant. The main reason seems to be the important results of Mossé (see Theorem 2.3) that do not hold for one-sided systems in general. Our theory will emphasize the importance of bi-infiniteness.

Define the set of *periodic points*

$$P := \{z \in \Omega_\sigma \mid \exists n \geq 1 : \sigma^n(z) = z\} = P' \cap \Omega_\sigma$$

and observe that for a $z \in P$ we have

$$\Omega_\sigma = \{w \in \mathcal{A}^{\mathbb{Z}} \mid \mathcal{L}(w) \subseteq \mathcal{L}(z)\} = \overline{\{S^m z \mid m \in \mathbb{Z}\}}$$

(the closure with respect to the above mentioned product topology). By the primitivity of the substitution, the choice of the particular periodic point is irrelevant. Actually, in several research papers substitution dynamical systems are defined in this way. Observe that we have to be quite careful here since $P = P'$ does not hold in general (see Example 8.2). For a word $z \in P' \setminus P$ (these words are frequently called *not admissible*), the above equation obviously does not hold. However, note that for each $z \in P'$ we have $z \in P$ if and only if $z_{[-1,0]} \in \mathcal{L}(\sigma)$. In particular, a substitution induces not admissible words if and only if $|P| < |L| \cdot |R|$.

Different substitutions may induce the same substitution dynamical system. In particular, for a given substitution σ we have $\Omega_\sigma = \Omega_{\sigma^2} = \Omega_{\sigma^3} = \dots$. For certain primitive substitutions it may happen that the induced substitution dynamical system is a finite set. Such substitutions are called *periodic*. Obviously, this case is of no great interest. Thus, throughout the paper, we will implicitly restrict to non-periodic substitutions. Note that it can be decided algorithmically whether a substitution is periodic or not (see [HL86, Pan86]).

As already indicated, symbolic dynamical systems are topological dynamical systems. In particular, a substitution dynamical system forms a compact space. A basis of the topology is given by the so-called cylinder sets. These are clopen sets defined for $A, B \in \mathcal{A}^*$ by[†]

$$[A, B] := \{w \in \Omega_\sigma \mid w_{[-|A|, |B|-1]} = AB\}.$$

[†] Observe that usually a point instead of the comma is used to separate A and B . Since the point will play an important role in our further discourse we use the comma here to avoid confusion.

The corresponding metric is given by

$$d : \Omega_\sigma \times \Omega_\sigma \rightarrow \mathbb{R}, (w, w') \mapsto \begin{cases} 0 & \text{if } w = w', \\ \exp(-\min\{|j| \mid j \in \mathbb{Z}, w_j \neq w'_j\}) & \text{otherwise.} \end{cases} \quad (1)$$

Note that substitution dynamical systems are minimal and uniquely ergodic, hence, there exists a uniquely defined probability measure that is invariant with respect to the shift map S (cf. [Que10]).

The second type of symbolic dynamical systems that we deal with are shifts of finite type. We will state some important facts in order to obtain a consistent notation. For a detailed discourse on this topic consult [LM95], which is more or less a standard reference.

Consider a finite directed graph G (presentation graph) whose labels are distinct, and denote by G^∞ the set of infinite paths, more precisely the sequences of the particular labels. Then (G^∞, S) is a shift of finite type. Hence, shifts of finite type are one-sided symbolic dynamical systems.

Analogously to the two-sided case the cylinder sets

$$[A] := \{g \in G^\infty \mid g_{[1, |A|]} = A\},$$

with $A \in \mathcal{A}^*$ and \mathcal{A} the set of labels of G , form a basis of the topology on G^∞ .

2.3. Recognizability, desubstitution and further results. One result, already cited, that seems to be of extraordinary importance, especially in the present context, is due to Mossé. For completeness and in order to emphasize their significance we will quickly summarize the most important facts. Then we explain the codings presented in [CS01a, HZ01] and outline the major properties.

As already mentioned, for a primitive substitution σ the closure of the shift orbit of an arbitrary $z \in P$ yields the set Ω_σ . Consider a finite subword $A = z_{[i, i+|A|]}$. Since $\sigma(z) = z$ there exist a finite subword B of z and indices $1 \leq p \leq |\sigma(B_1)|$, $0 \leq q < |\sigma(B_{-1})|$ such that

$$A = \sigma(B)_{[p, |\sigma(B)|-q]}.$$

The triple (p, B, q) is the so-called *1-cutting at the rank i of A* and the subword B is usually called the *ancestor word of A* .

The uniqueness of the 1-cutting is a priori not clear and for a long time it had to be ensured by additional recognizability requirements. The problem was finally solved in [Mos92] by showing that all primitive substitutions satisfy the so-called bilateral recognizability.

THEOREM 2.3. (Cf. [Fog02, Mos92, Mos96]) *Let σ be a primitive substitution and $z \in P$. There exists $L > 0$ such that for all $i, j \in \mathbb{Z}$, $l \in \mathbb{N}$ the words A with $A = z_{[i, i+l]}$ and B with $B = z_{[j, j+l]}$ have the same ancestor word and the same 1-cutting at the ranks i and j whenever $z_{[i-L, i+l+L]} = z_{[j-L, j+l+L]}$.*

Note that the application of the theorem without restrictions requires consideration of bi-infinite words. Counterexamples show that a one-sided version (unilateral recognizability) does not hold in general. For a more detailed summary (in English) concerning this topic we refer to [Fog02].

The following theorem is a consequence of Theorem 2.3. It describes the relation between Ω_σ and $\sigma(\Omega_\sigma)$ and forms the basis for various research papers (for example, [CS01a, DHS99, HZ01]). Our further results will be based on it, too.

THEOREM 2.4. (Cf. [DHS99]) *Let σ be a primitive substitution over an alphabet \mathcal{A} :*

- (i) $\sigma : \Omega_\sigma \longrightarrow \Omega_\sigma$ *is injective and open;*
- (ii) *for each $w \in \Omega_\sigma$ there is a unique pair $(w', k) \in \Omega_\sigma \times \mathbb{N}$ with $0 \leq k < |\sigma(w_0)|$ such that*

$$w = S^k \sigma(w').$$

We can illustrate the statement of Theorem 2.4 for a word w in the following way:

$$w = \cdots \mid \underbrace{\cdots \cdots \cdots}_{\sigma(w'_{-2})} \mid \underbrace{\cdots w_{-k-1}}_{\sigma(w'_{-1})} \mid \underbrace{w_{-k} \cdots w_{-1} w_0 \cdots w_{k'-1}}_{\sigma(w'_0)} \mid \underbrace{w_{k'} \cdots}_{\sigma(w'_1)} \mid \underbrace{\cdots \cdots \cdots}_{\sigma(w'_2)} \mid \cdots .$$

It is easy to see that $w = S^k \sigma(w')$. On the other hand, the illustration immediately suggests the use of S^{-1} instead of S . Provided that $k \neq 0$ we can set $w'' := S w'$ in order to obtain $w = S^{-k'} \sigma(w'')$. This idea will be essential for our coding and we will discuss it in §4.

The possibility of a unique desubstitution provided by Theorem 2.4 was exploited in [CS01a, HZ01] in order to represent substitution dynamical systems as shifts of finite type. As the notation we develop is more in the spirit of [HZ01], we will explain the results from this point of view (and call the respective coding the HZ coding). However, we will also quickly outline the connection with the results from [CS01a].

The main idea is to successively apply Theorem 2.4 on the elements of a substitution dynamical system. In particular, given a $w = w^{(0)} \in \Omega_\sigma$, we obtain a unique sequence $(w^{(n)}, k_n)_{n \geq 0} \in (\Omega_\sigma \times \mathbb{N})^\mathbb{N}$ such that $w^{(n-1)} = S^{k_n} \sigma(w^{(n)})$ for all $n \geq 1$. The HZ code of w is given by

$$\Gamma_{\text{HZ}}(w) := (w_0^{(n)}, k_n)_{n \geq 1} \in (\mathcal{A} \times \mathbb{N})^\mathbb{N}.$$

The following statements summarize the main properties of the HZ coding.

Retrieving the pre-image. If $\Gamma_{\text{HZ}}(w) = (a_n, k_n)_{n \geq 1}$ is the HZ code of some element $w \in \Omega_\sigma$ then we can uniquely recover w by

$$w = \lim_{n \rightarrow \infty} (S^{k_1} \circ \sigma \circ S^{k_2} \circ \sigma \circ \cdots \circ S^{k_n} \circ \sigma)[\varepsilon, a_n], \quad (2)$$

provided that $k_n > 0$ as well as $k_n < |\sigma(a_n)| - 1$ hold for infinitely many indices n . The only elements of Ω_σ that do not satisfy this property are those that are contained in the shift orbits of the periodic points. In particular, we have:

- $k_n = 0$ for all $n \geq 1$ if and only if $w \in P$;
- $k_n = 0$ for all $n \geq n_0$ if and only if $S^m w \in P$ for some $m \leq 0$;
- $k_n = |\sigma(a_n)| - 1$ for all $n \geq 1$ if and only if $S w \in P$;
- $k_n = |\sigma(a_n)| - 1$ for all $n \geq n_0$ if and only if $S^m w \in P$ for some $m \geq 1$.

Coding as a shift of finite type. The terms of the sequences obtained by the HZ coding are elements of the finite set $\{(a, k) \mid a \in \mathcal{A}, 0 \leq k < |\sigma(a)|\}$. Define the finite, directed graph G_{HZ} with set of vertices \mathcal{A} and an edge from b to a labelled by (a, k) whenever $\sigma(a)_{k+1} = b$. Each infinite walk on G_{HZ} corresponds to the HZ code of an element

of Ω_σ . On the other hand, for each element $w \in \Omega_\sigma$ $\Gamma_{\text{HZ}}(w)$ is given by an infinite walk on G_{HZ} starting at w_0 . Since the edges of G_{HZ} are all distinct, the set of infinite walks G_{HZ}^∞ forms a shift of finite type. Γ_{HZ} is a continuous and surjective map from Ω_σ onto G_{HZ}^∞ .

Injectivity. The HZ coding is possibly not injective for the points that are contained in the shift orbit of the periodic points. More precisely, for $w, w' \in \Omega_\sigma$ with $w \neq w'$ we have $\Gamma_{\text{HZ}}(w) = \Gamma_{\text{HZ}}(w')$ if and only if one of following statements holds:

- there exists an integer $m \geq 0$ such that $S^{-m}w, S^{-m}w' \in P$ and $w_n = w'_n$ for all $n \geq -m$;
- there exists an integer $m \geq 1$ such that $S^mw, S^mw' \in P$ and $w_n = w'_n$ for all $n < m$.

Thus, the HZ coding is finite-to-one for the elements contained in the shift orbits of the periodic points. Since these shift orbits form a set of measure zero, the HZ coding is measure-theoretically injective. The HZ coding is completely injective if and only if $|L| = |R| = |P|$.

Interaction with the shift map. One can equip the edges of the graph G_{HZ} with a partial ordering $<$ by defining

$$(a, k) < (a', k') \iff a = a', k < k'.$$

Edges of the shape $(a, 0)$ are minimal and edges of the form $(a, |\sigma(a)| - 1)$ are maximal with respect to the ordering. We call them minimal edges and maximal edges, respectively. This ordering extends to a partial ordering on G_{HZ}^∞ :

$$(a_n, k_n)_{n \geq 1} < (a'_n, k'_n)_{n \geq 1} \\ \iff \text{there exists } n_0 : (a_{n_0}, k_{n_0}) < (a'_{n_0}, k'_{n_0}) \wedge (\forall n > n_0 : (a_n, k_n) = (a'_n, k'_n)).$$

The minimal elements of G_{HZ}^∞ consist of minimal edges only and are the Γ_{HZ} -images of the periodic points. The maximal elements are the Γ_{HZ} -images of the periodic points under the inverse shift S^{-1} and consist of maximal edges only. For the non-maximal elements of G_{HZ}^∞ one can define the successor map T on G_{HZ}^∞ with respect to the partial ordering (adic transformation in the sense of [Ver81]). Now, Γ_{HZ} conjugates the shift map with T making (Ω_σ, S) and $(G_{\text{HZ}}^\infty, T)$ topologically conjugated dynamical systems.

The coding presented in [CS01a] is frequently called prefix-suffix coding. We will denote it by Γ_{CS} , and it is isomorphic to the HZ coding. In particular, if $\Gamma_{\text{HZ}}(w) = (a_n, k_n)_{n \geq 1}$ then

$$\Gamma_{\text{CS}}(w) = (\sigma(a_n)_{[1, k_n]}, \sigma(a_n)_{k_n+1}, \sigma(a_n)_{[k_n+2, |\sigma(a_n)|]})_{n \geq 1} \in (\mathcal{A}^* \times \mathcal{A} \times \mathcal{A}^*)^{\mathbb{N}}.$$

Thus, the corresponding graph G_{CS} (frequently called a prefix-suffix graph) can be obtained by replacing each label (a, k) of G_{HZ} by $(\sigma(a)_{[1, k]}, \sigma(a)_{k+1}, \sigma(a)_{[k+2, |\sigma(a)|]})$. Note that the edges of G_{CS} do not necessarily have distinct labels. However, G_{CS}^∞ is always a shift of finite type since, for each $w \in \Omega_\sigma$ with $\Gamma_{\text{CS}}(w) = (A^{(n)}, a_n, B^{(n)})_{n \geq 1}$, we can retrieve $\Gamma_{\text{HZ}}(w)$ by a two-block code:

$$\Gamma_{\text{HZ}}(w) = (a_{n+1}, |A^{(n)}|)_{n \geq 1}.$$

Therefore, the results concerning the HZ coding apply straightforwardly to the prefix–suffix coding. The advantage of the prefix–suffix coding is a symbolic way to realize the shift map. In particular, suppose that w is not contained in the shift orbit of a periodic point and $\Gamma_{\text{CS}}(w) = (A^{(n)}, a_n, B^{(n)})_{n \geq 1}$. Then the analogue to (2) is given by

$$w = \lim_{n \rightarrow \infty} [\sigma^{n-1}(A^{(n)}) \cdots \sigma^2(A^{(3)})\sigma(A^{(2)})A^{(1)}, a_1 B^{(1)}\sigma(B^{(2)}) \cdots \sigma^{n-1}(B^{(n)})].$$

We will adopt some ideas from prefix–suffix coding to show several results in a more convenient way.

3. Substitution dynamical systems from an alternative point of view

3.1. *Words over \mathcal{A}^2 .* As an important aspect of our investigation we will consequently work with (finite and infinite) words over \mathcal{A}^2 instead of \mathcal{A} whose terms somehow overlap. As a general rule we will mark these mathematical objects with a tilde. In particular, for an alphabet \mathcal{A} we define the following subsets of $(\mathcal{A}^2)^*$ and $(\mathcal{A}^2)^{\mathbb{Z}}$, respectively:

$$\begin{aligned}\widetilde{\mathcal{A}}^* &:= \{\tilde{A} \in (\mathcal{A}^2)^* \mid \forall i \in \{1, \dots, |\tilde{A}| - 1\} : (\tilde{A}_i)_{-1} = (\tilde{A}_{i+1})_1\}, \\ \widetilde{\mathcal{A}}^{\mathbb{Z}} &:= \{\tilde{w} \in (\mathcal{A}^2)^{\mathbb{Z}} \mid \forall i \in \mathbb{Z} : (\tilde{w}_i)_{-1} = (\tilde{w}_{i+1})_1\}\end{aligned}$$

(note that each term is considered as word over \mathcal{A} of length 2).

Observe that each element $\tilde{A} \in \widetilde{\mathcal{A}}^*$ corresponds exactly to one element $A \in \mathcal{A}^*$. We have $|A| = |\tilde{A}| + 1$ and $A_{[i, i+1]} = \tilde{A}_i$ ($A_{[-i-1, -i]} = \tilde{A}_{-i}$) for each $i \in \{1, \dots, |\tilde{A}|\}$. Note that a reverse correspondence also holds except for the words of \mathcal{A}^* whose length equals 1.

Without additional conventions there is no unique correspondence between $\widetilde{\mathcal{A}}^{\mathbb{Z}}$ and $\mathcal{A}^{\mathbb{Z}}$. Indeed, for $\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}}$ we may choose $w \in \mathcal{A}^{\mathbb{Z}}$ such that $w_0 w_1 = \tilde{w}_0$ as well as $w_{-1} w_0 = \tilde{w}_0$. In fact, a bi-infinite word $\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}}$ is the unique composite of a left-infinite sequence $u \in \mathcal{A}^{\mathbb{Z}^-}$ and a right-infinite sequence $v \in \mathcal{A}^{\mathbb{Z}^+}$. We define the binary composition map \cdot by

$$\cdot : \mathcal{A}^{\mathbb{Z}^-} \times \mathcal{A}^{\mathbb{Z}^+} \longrightarrow \widetilde{\mathcal{A}}^{\mathbb{Z}}, u \cdot v \longmapsto \tilde{w} \quad \text{with } \tilde{w}_n = \begin{cases} u_{n-1} u_n & \text{for } n < 0, \\ u_{-1} v_1 & \text{for } n = 0, \\ v_n v_{n+1} & \text{for } n > 0. \end{cases}$$

In contrast, for each bi-infinite word $\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}}$ we denote by $\pi^{(l)}(\tilde{w}) \in \mathcal{A}^{\mathbb{Z}^-}$ (left projection) and $\pi^{(r)}(\tilde{w}) \in \mathcal{A}^{\mathbb{Z}^+}$ (right projection) the unique words such that $\tilde{w} = \pi^{(l)}(\tilde{w}) \cdot \pi^{(r)}(\tilde{w})$.

3.2. *Alternative definition of substitution dynamical systems.* Our aim is to consider substitution dynamical systems as a subset of $\widetilde{\mathcal{A}}^{\mathbb{Z}}$. Of course, in this way we will carry a lot of redundant information and need some additional formalism. However, we will see that several ambiguities vanish and results become more consistent.

We do not have to change the definition of a primitive substitution but we must describe how it acts on the elements of $\widetilde{\mathcal{A}}^*$ and $\widetilde{\mathcal{A}}^{\mathbb{Z}}$. In particular, for $\tilde{A} \in \widetilde{\mathcal{A}}^*$ and $\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}}$ set

$$\begin{aligned}\sigma(\tilde{A}) &:= \widetilde{\sigma(A)}, \\ \sigma(\tilde{w}) &:= \sigma(\pi^{(l)}(\tilde{w})) \cdot \sigma(\pi^{(r)}(\tilde{w})).\end{aligned}$$

The actual definition of substitution dynamical systems is quite analogous to Definition 2.2. One main difference is that the atomic elements are not given by the entire set \mathcal{A}^2 . In fact, we need the set

$$\mathfrak{L}_2(\sigma) := \{ab \in \mathcal{A}^2 \mid \exists a' \in \mathcal{A}, k \in \mathbb{N} \text{ such that } ab \text{ appears in } \sigma^k(a')\}.$$

The set $\mathfrak{L}_2(\sigma)$ is formed by the words of $\mathfrak{L}(\sigma)$ that have length 2.

Definition 3.1. (Substitution dynamical system: alternative) Let σ be a primitive substitution over an alphabet \mathcal{A} and

$$\widetilde{\mathfrak{L}(\sigma)} := \{\tilde{A} \in \widetilde{\mathcal{A}}^* \mid \exists ab \in \mathfrak{L}_2(\sigma), k \in \mathbb{N} \text{ such that } \tilde{A} \text{ is a factor of } \sigma^k(ab)\}.$$

The shift space $(\tilde{\Omega}_\sigma, S)$, where

$$\tilde{\Omega}_\sigma = \{\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}} \mid \mathfrak{L}(\tilde{w}) \subseteq \widetilde{\mathfrak{L}(\sigma)}\},$$

is the *substitution dynamical system induced by σ* .

Observe that for an element $\tilde{w} = u \cdot v \in \tilde{\Omega}_\sigma$ (where $u = \pi^{(l)}(\tilde{w})$ and $v = \pi^{(r)}(\tilde{w})$) we have

$$\begin{aligned} S\tilde{w} &= \cdots u_{-3}u_{-2}u_{-1}v_1 \cdot v_2v_3v_4 \cdots, \\ S^{-1}\tilde{w} &= \cdots u_{-4}u_{-3}u_{-2} \cdot u_{-1}v_1v_2 \cdots. \end{aligned}$$

Furthermore, for all $m \in \mathbb{N}$ we have

$$\begin{aligned} \sigma(S^m \tilde{w}) &= S^{|\sigma(v_{[1,m]})|} \sigma(\tilde{w}), \\ \sigma(S^{-m} \tilde{w}) &= S^{-|\sigma(u_{[-m,-1]})|} \sigma(\tilde{w}). \end{aligned} \tag{3}$$

Keep these relations in mind since we will need them frequently. We will often represent elements of $\tilde{\Omega}_\sigma$ as \cdot -composites of a left- and a right-infinite sequence.

The topology of $\tilde{\Omega}_\sigma$ is given by the product topology of the discrete topology on each copy of $\mathfrak{L}_2(\sigma)$. Again, the basis of this topology is given by cylinder sets. For $A, B \in \mathcal{A}^*$ define

$$[A \cdot B] := \{\tilde{w} \in \tilde{\Omega}_\sigma \mid \pi^{(l)}(\tilde{w})_{[-|A|,-1]} = A \wedge \pi^{(r)}(\tilde{w})_{[1,|B|]} = B\}.$$

The metric is defined analogously as in (1) and we denote it by d , too.

Let us compare the two ways of defining substitution dynamical systems. Consider the maps

$$\begin{aligned} f_1 : \widetilde{\mathcal{A}}^{\mathbb{Z}} &\longrightarrow \mathcal{A}^{\mathbb{Z}}, u \cdot v \longmapsto w \quad \text{with} \quad \begin{cases} w_i = v_{i-1} & \text{for } i \geq 0, \\ w_i = u_i & \text{for } i < 0, \end{cases} \\ f_2 : \widetilde{\mathcal{A}}^{\mathbb{Z}} &\longrightarrow \mathcal{A}^{\mathbb{Z}}, u \cdot v \longmapsto w \quad \text{with} \quad \begin{cases} w_i = u_{-i-1} & \text{for } i \geq 0, \\ w_i = u_{-i} & \text{for } i < 0. \end{cases} \end{aligned} \tag{4}$$

For each $\tilde{w} \in \widetilde{\mathcal{A}}^{\mathbb{Z}}$ the word $f_1(\tilde{w})$ is a bi-infinite sequence over \mathcal{A} that consists of the right elements of the terms of \tilde{w} . Furthermore, we have $\sigma(f_1(\tilde{w})) = f_1(\sigma(\tilde{w}))$. f_1 is a continuous bijection between $\tilde{\Omega}_\sigma$ and Ω_σ that conjugates the shift map.

The word $f_2(\tilde{w})$ is a bi-infinite sequence over \mathcal{A} formed by the left elements of the terms of \tilde{w} with a subsequent reversion of the order. Observe that $Sf_2(\tilde{w}) = f_2(S^{-1}(\tilde{w}))$

and $\sigma'(f_2(\tilde{w})) = f_2(\sigma(\tilde{w}))$ where σ' is the reversed substitution defined by

$$\sigma'(a) = a_n a_{n-1} \cdots a_1 \quad \text{for } \sigma(a) = a_1 \cdots a_n.$$

f_2 is a continuous bijection between $\tilde{\Omega}_\sigma$ and $\Omega_{\sigma'}$.

By our definition the dualism between the substitution dynamical systems induced by a substitution σ and the reversed substitution σ' becomes obvious. Since our central element, that is, the term indexed with 0, actually consists of two letters of \mathcal{A} , we ensure that both the left-infinite sequence and the right-infinite sequence, are treated equally. This cannot be realized by considering letter-centred systems (as in Definition 2.2).

The difference between $\tilde{\Omega}_\sigma$ and Ω_σ is also reflected in the metric. In particular, for $\tilde{w}, \tilde{w}' \in \tilde{\Omega}_\sigma$ we have

$$d(\tilde{w}, \tilde{w}') = \max\{d(f_1(\tilde{w}), f_1(\tilde{w}')), d(f_2(\tilde{w}), f_2(\tilde{w}'))\}.$$

Apart from these nice symmetric aspects we will profit from our notation since the coding strongly depends on the two central letters of the words. Conveniently, these two letters are united in our representation.

3.3. Periodic points. The periodic points will play an important role in our further analysis. For a given substitution σ we have already defined the sets L and R . We adapt the definition of periodic points by setting

$$\begin{aligned} \tilde{P} &:= \{\tilde{w} \in \tilde{\Omega}_\sigma \mid \exists n \geq 1 : \sigma^n(\tilde{w}) = \tilde{w}\} \\ &= \{\tilde{w} \in \tilde{\mathcal{A}}^{\mathbb{Z}} \mid \exists n \geq 1 : \sigma^n(\tilde{w}) = \tilde{w}, \tilde{w}_0 \in \mathfrak{L}_2(\sigma)\}. \end{aligned}$$

Note that

$$\begin{aligned} \tilde{P} &= \{u \cdot v \mid u \in L, v \in R, u_{-1}v_1 \in \mathfrak{L}_2(\sigma)\}, \\ L &= \{\pi^{(l)}(\tilde{z}) \mid \tilde{z} \in \tilde{P}\}, \\ R &= \{\pi^{(r)}(\tilde{z}) \mid \tilde{z} \in \tilde{P}\}. \end{aligned}$$

We have already mentioned that each element $u \in L$ is uniquely determined by the letter u_{-1} and each element $v \in R$ is uniquely determined by the letter v_1 . Each element of \tilde{P} is a unique composition of an element of L and R . Thus, each $\tilde{z} \in \tilde{P}$ is uniquely determined by $\tilde{z}_0 \in \mathfrak{L}_2(\sigma)$ and, hence, $|\tilde{P}| \leq |\mathfrak{L}_2(\sigma)|$.

σ acts as a permutation on the set \tilde{P} . More precisely, we can represent $\sigma|_{\tilde{P}}$, the restriction of σ to \tilde{P} , as a permutation in cyclic form

$$\sigma|_{\tilde{P}} : \langle \tilde{z}^{(1,0)}, \dots, \tilde{z}^{(1,p_1-1)} \rangle \langle \tilde{z}^{(2,0)}, \dots, \tilde{z}^{(2,p_2-1)} \rangle \dots \langle \tilde{z}^{(q,0)}, \dots, \tilde{z}^{(q,p_q-1)} \rangle,$$

where for all $i, i' \in \{1, \dots, q\}$, $j \in \{0, \dots, p_i - 1\}$, $j' \in \{0, \dots, p_{i'} - 1\}$ we have $\tilde{z}^{(i,j)} \in \tilde{P}$, $\tilde{z}^{(i,j)} \neq \tilde{z}^{(i',j')}$ for $(i, j) \neq (i', j')$, and $\sigma(\tilde{z}^{(i,j)}) = \tilde{z}^{(i,j+1)}$ (where the second upper index is considered modulo p_i). Furthermore, $\sum_{i=1}^q p_i = |\tilde{P}|$. The angle brackets underline the cyclic structure, thus we may rotate the entries arbitrarily.

Similarly to the periodic points \tilde{P} , the restrictions of σ to L and R , $\sigma|_L$ and $\sigma|_R$ respectively, are also permutations. Again, we represent these permutations by cycles. Note that $\pi^{(l)}$ and $\pi^{(r)}$ respect the cyclic structure of $\sigma|_{\tilde{P}}$, $\sigma|_L$ and $\sigma|_R$: cycles are mapped

onto cycles. Observe that the period of a cycle of \tilde{P} is always the multiple of the period of its image under $\pi^{(l)}$ or $\pi^{(r)}$ but the periods do not necessarily have to coincide; in particular, $\pi^{(l)}$ and $\pi^{(r)}$ are not injective in general.

Higher powers of σ may change the order inside the cycles or a cycle can split into smaller ones (of equal period). In particular, when we consider σ^n , where n is the least common multiple of the periods of $\sigma|_{\tilde{P}}$, all the cycles of $\sigma^n|_{\tilde{P}}$, $\sigma^n|_L$ and $\sigma^n|_R$ have period 1.

We analyse the cyclic structure of $\sigma|_{\tilde{P}}$, $\sigma|_L$ and $\sigma|_R$ for concrete substitutions in Examples 8.1 and 8.2.

4. The coding map

As a first step we will ‘translate’ Theorem 2.4 in order to fit it into our alternative definition of substitution dynamical systems. In fact, we are able to add a complementary result from the point of view of the inverse shift S^{-1} as already indicated.

THEOREM 4.1. *Let σ be a primitive substitution over an alphabet \mathcal{A} .*

- (i) $\sigma : \tilde{\Omega}_\sigma \longrightarrow \tilde{\Omega}_\sigma$ is injective and open.
- (ii) For each $\tilde{w} \in \tilde{\Omega}_\sigma$ there is a unique pair $(u \cdot v, k) \in \tilde{\Omega}_\sigma \times \mathbb{N}$ with $0 \leq k < |\sigma(v_1)|$ such that

$$\tilde{w} = S^k \sigma(u \cdot v).$$

- (iii) For each $\tilde{w} \in \tilde{\Omega}_\sigma$ there is a unique pair $(u' \cdot v', k') \in \tilde{\Omega}_\sigma \times \mathbb{N}$ with $0 \leq k' < |\sigma(u'_{-1})|$ such that

$$\tilde{w} = S^{-k'} \sigma(u' \cdot v').$$

Proof. Consider the maps f_1 and f_2 from (4).

f_1 maps $\tilde{\Omega}_\sigma$ bijectively and continuously onto Ω_σ and conjugates σ as well as the shift map. Additionally, for $u \cdot v \in \tilde{\Omega}_\sigma$, we have $f_1(u \cdot v)_0 = v_1$. Thus, (i) and (ii) follow directly from Theorem 2.4.

On the other hand, consider Theorem 2.4 for the system $\Omega_{\sigma'}$ induced by the reversed substitution σ' . f_2 maps $\tilde{\Omega}_\sigma$ bijectively onto $\Omega_{\sigma'}$. For $u \cdot v \in \tilde{\Omega}_\sigma$ we have $Sf_2(u \cdot v) = f_2(S^{-1}(u \cdot v))$, $\sigma'(f_2(u \cdot v)) = f_2(\sigma(u \cdot v))$ and $f_2(u \cdot v)_0 = u_{-1}$. This yields (iii). \square

Our main goal is to combine (ii) and (iii) of Theorem 4.1. To keep the desubstitution unique we have to define when we desubstitute according to (ii) (i.e. consider the left shift) and when we desubstitute with respect to (iii) (i.e. consider the right shift). This can be realized combinatorially by introducing so-called coding prescriptions.

Definition 4.2. (Coding prescription) Let σ be a primitive substitution over an alphabet \mathcal{A} and c a map with source \mathcal{A} that assigns to each letter $a \in \mathcal{A}$ a complete set of integer representatives modulo $|\sigma(a)|$ whose absolute values are smaller than $|\sigma(a)|$. Extend c to \mathcal{A}^2 by

$$c(ab) := \{k \in c(a) \mid k \leq 0\} \cup \{k \in c(b) \mid k \geq 0\}.$$

We call such a function c a *coding prescription* (with respect to σ).

By definition we have for each $a \in \mathcal{A}$ that $0 \in c(a)$, $|c(a)| = |\sigma(a)|$ and, for each $k \in c(a)$, $1 - |\sigma(a)| \leq k \leq |\sigma(a)| - 1$. $c(ab)$ consists of the non-positive elements of $c(a)$ and of the non-negative elements of $c(b)$ and, therefore, $c(ab)$ contains 0 for all $ab \in \mathcal{A}^2$.

One can easily calculate that for a given substitution σ there are exactly

$$\prod_{a \in \mathcal{A}} 2^{|\sigma(a)|-1} > 1$$

different coding prescriptions. The inequality follows directly from (i) and (ii) of Definition 2.1. In Example 8.3 we explicitly determine all coding prescriptions for a given substitution.

With the above notation we state the following theorem which forms the core of our theory.

THEOREM 4.3. *Let σ be a primitive substitution over an alphabet \mathcal{A} , c a coding prescription with respect to σ , and $\tilde{w} \in \tilde{\Omega}_\sigma$. There exists a unique pair $(\Theta(\tilde{w}), \varphi(\tilde{w})) \in \tilde{\Omega}_\sigma \times \mathbb{Z}$ such that $\varphi(\tilde{w}) \in c(\Theta(\tilde{w})_0)$ and*

$$\tilde{w} = S^{\varphi(\tilde{w})} \sigma(\Theta(\tilde{w})).$$

Proof. First we prove that there exists a pair $(\Theta(\tilde{w}), \varphi(\tilde{w}))$ that has the desired properties. By (ii) of Theorem 4.1 we can find a unique pair $(u \cdot v, k)$ with $0 \leq k < |\sigma(v_1)|$ and $\tilde{w} = S^k \sigma(u \cdot v)$. If $k \in c(v_1)$ we set $\Theta(\tilde{w}) := u \cdot v$ and $\varphi(\tilde{w}) := k$. Trivially, $\tilde{w} = S^{\varphi(\tilde{w})} \sigma(\Theta(\tilde{w}))$ and $\varphi(\tilde{w}) \in c(\Theta(\tilde{w})_0) = c(u_{-1} v_1)$.

If $k \notin c(v_1)$ then set $\varphi(\tilde{w}) := k - |\sigma(v_1)| < 0$ and $\Theta(\tilde{w}) := S(u \cdot v)$. Note that $c(v_1)$ necessarily contains $\varphi(\tilde{w})$ by Definition 4.2. Now (3) yields

$$\tilde{w} = S^k \sigma(u \cdot v) = S^k \circ S^{-|\sigma(v_1)|} \sigma(S(u \cdot v)) = S^{\varphi(\tilde{w})} \sigma(\Theta(\tilde{w})).$$

Furthermore, we have that $\Theta(\tilde{w})_0 = v_1 v_2$ and, hence, $\varphi(\tilde{w}) \in c(\Theta(\tilde{w})_0)$.

We now show the uniqueness of the pair $(\Theta(\tilde{w}), \varphi(\tilde{w}))$. Suppose that there were another pair $(u \cdot v, k) \neq (\Theta(\tilde{w}), \varphi(\tilde{w}))$ with the same properties. $\varphi(\tilde{w})$ and k cannot be both non-negative or non-positive since this would contradict Theorem 4.1. Without loss of generality we may assume that $\varphi(\tilde{w}) < 0$ and $k > 0$. This implies that $k \in c(v_1)$. Consider the word $\tilde{w}' := S(u \cdot v)$ and the integer $k' := k - |\sigma(v_1)|$. Note that $\tilde{w}'_0 = v_1 v_2$, $-|\sigma(v_1)| < k' < 0$ and $\tilde{w} = S^{k'} \sigma(\tilde{w}')$. By the uniqueness in (iii) of Theorem 4.1 we deduce that $(\Theta(\tilde{w}), \varphi(\tilde{w})) = (\tilde{w}', k')$. Therefore, $\Theta(\tilde{w})_0 = \tilde{w}'_0 = v_1 v_2$ and, hence, $k' = \varphi(\tilde{w}) \in c(v_1)$. We thus have $k, k' \in c(v_1)$ which contradicts Definition 4.2 since $k - k' = |\sigma(v_1)|$. \square

Observe that Theorem 4.3 is a direct consequence of Theorem 2.4. We do not need other results or additional restrictions in the proof. Furthermore, note that if c is chosen such that $c(a)$ consists of non-negative (non-positive) integers only for all $a \in \mathcal{A}$, then Theorem 4.3 corresponds exactly to (ii) of Theorem 4.1 ((iii) of Theorem 4.1).

For a coding prescription c with respect to a given primitive substitution σ consider the finite set

$$\mathcal{R} := \{(ab, k) \mid ab \in \mathcal{L}_2(\sigma), k \in c(ab)\}.$$

Analogously to [CS01a, HZ01] we use Theorem 4.3 to define a map Γ_c that assigns to every element of a substitution dynamical system $\tilde{\Omega}_\sigma$ an infinite word over \mathcal{R} by

$$\Gamma_c : \tilde{\Omega}_\sigma \longrightarrow \mathcal{R}^{\mathbb{Z}^+}, \tilde{w} \longmapsto (\Theta^n(\tilde{w})_0, \varphi(\Theta^{n-1}(\tilde{w})))_{n \geq 1}.$$

We summarize several basic properties of Γ_c in the following lemma.

LEMMA 4.4. *Let $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$. Then the following assertions hold:*

- (i) $\Theta^{n-1}(\tilde{w}) = S^{k_n} \sigma(\Theta^n(\tilde{w}))$ for all $n \geq 1$;
- (ii) $\Gamma_c(\sigma(\tilde{w})) = (\tilde{w}_0, 0), (a_1 b_1, k_1), (a_2 b_2, k_2) \dots$;
- (iii) $\Gamma_c(\Theta(\tilde{w})) = (a_2 b_2, k_2), (a_3 b_3, k_3) \dots$.

Proof. (i) follows directly from the definition of k_n ;

(ii) follows from the fact that $0 \in c(ab)$ for all $ab \in \mathfrak{L}_2(\sigma)$ and, hence, $0 \in c(\tilde{w}_0)$. Thus, $\Theta(\sigma(\tilde{w})) = \tilde{w}$, $\varphi(\sigma(\tilde{w})) = 0$ and, by definition, $\Gamma_c(\sigma(\tilde{w}))$ starts with $(\tilde{w}_0, 0)$;

(iii) is a consequence of the definition of Γ_c . \square

Now suppose that $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$. We obtain the following ways of representing \tilde{w} by means of $\Theta(\tilde{w})$, $\varphi(\tilde{w})$ and $\Gamma_c(\tilde{w})$. For all $n \geq 1$,

$$\begin{aligned} \tilde{w} &= S^{\varphi(\tilde{w})} \circ \sigma \circ \dots \circ S^{\varphi(\Theta^{n-2}(\tilde{w}))} \circ \sigma \circ S^{\varphi(\Theta^{n-1}(\tilde{w}))} \circ \sigma(\Theta^n(\tilde{w})) \\ &= S^{k_1} \circ \sigma \circ \dots \circ S^{k_{n-1}} \circ \sigma \circ S^{k_n} \circ \sigma(\Theta^n(\tilde{w})) \\ &= S^{r_n} \sigma^n(\Theta^n(\tilde{w})) \\ &\quad \text{with } r_n = \sum_{j=1}^n |\sigma^{j-1}(\sigma(b_j)_{[1, k_j]})| - \sum_{j=1}^n |\sigma^{j-1}(\sigma(a_j)_{[k_j, -1]})|, \end{aligned} \quad (5)$$

where the last representation follows from (3). Note that $k_j \leq 0$ gives no contribution to the first sum while $k_j \geq 0$ gives no contribution to the second one for each $j \in \{1, \dots, n\}$.

We immediately see that

$$\tilde{w} \in S^{r_n} \sigma^n[a_n \cdot b_n] = \{S^{r_n} \sigma^n(\tilde{z}) \mid \tilde{z} \in [a_n \cdot b_n]\}$$

for each $n \geq 1$ and $S^{r_n} \sigma^n[a_n \cdot b_n] \subseteq C_n$ with the cylinder set

$$C_n := S^{r_n}[\sigma^n(a_n) \cdot \sigma^n(b_n)]. \quad (6)$$

Using the ideas of the prefix–suffix coding, we can state the last considerations in a slightly different and, perhaps, clearer way. Still supposing that $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ set $a_0 b_0 := \tilde{w}_0$ and define, for each $n \geq 0$,

$$\begin{aligned} A^{(n)} &:= \sigma(a_{n+1} b_{n+1})_{[1, |\sigma(a_{n+1})| + k_{n+1} - 1]}, \\ B^{(n)} &:= \sigma(a_{n+1} b_{n+1})_{[-|\sigma(b_{n+1})| + k_{n+1} + 1, -1]}. \end{aligned} \quad (7)$$

We obviously have $\sigma(a_{n+1} b_{n+1}) = A^{(n)} a_n b_n B^{(n)}$ and, hence,

$$\sigma^n(a_n b_n) = \sigma^{n-1}(A^{(n-1)}) \dots \sigma(A^{(1)}) A^{(0)} a_0 b_0 B^{(0)} \sigma(B^{(1)}) \dots \sigma^{n-1}(B^{(n-1)}).$$

Thus, for all $n \geq 1$, we see that

$$r_n = -|\sigma^n(a_n)| + 1 + \sum_{j=0}^{n-1} |\sigma^j(A^{(j)})| = |\sigma^n(b_n)| - 1 - \sum_{j=0}^{n-1} |\sigma^j(B^{(j)})|.$$

Furthermore, we have

$$C_n = [\sigma^{n-1}(A^{(n-1)}) \dots A^{(0)} a_0 \cdot b_0 B^{(0)} \dots \sigma^{n-1}(B^{(n-1)})].$$

Remark 4.5. Using the applications f_1 and f_2 , as defined in (4), we see the connection between Γ_c and Γ_{HZ} . In particular, consider a substitution σ over \mathcal{A} and choose the coding prescription c such that $k \geq 0$ for all $a \in \mathcal{A}$, $k \in c(a)$. Let $\tilde{w} \in \tilde{\Omega}_\sigma$ and suppose that $\Gamma_c(\tilde{w}) = (a'_n b'_n, k'_n)_{n \geq 1}$. Then $\Gamma_{\text{HZ}}(f_1(\tilde{w})) = (b_n, k_n)_{n \geq 1}$.

On the other hand, choose the coding prescription c' with respect to σ such that for each $a \in \mathcal{A}$ the set $c'(a)$ consists of non-positive integers only, and suppose that $\Gamma_c(\tilde{w}) = (a'_n b'_n, k'_n)_{n \geq 1}$. Then $(a'_n, -k'_n)_{n \geq 1}$ is the HZ code of $f_2(\tilde{w})$ related to the reversed substitution σ' .

5. Basic properties of Γ_c

In this section we show several basic properties of the coding Γ_c . We will see that Γ_c maps a substitution dynamical system surjectively and continuously onto a shift of finite type. We will meet a sufficient condition for the coding Γ_c to be injective which we will call traceability. Finally, we will study the periodic points and their Γ_c -images.

Definition 5.1. (Coding graph) Let σ be a primitive substitution over an alphabet \mathcal{A} and c a coding prescription with respect to σ . The *coding graph associated with c* , denoted by $G(c)$, is the graph with the following properties. The vertices of $G(c)$ are the elements of $\mathcal{L}_2(\sigma)$. For every $k \in c(a'b')$ we draw an edge from ab to $a'b'$ with

$$ab = \begin{cases} \sigma(a')_{[k-1, k]} & \text{for } k < 0, \\ \sigma(a')_{-1} \sigma(b')_1 & \text{for } k = 0, \\ \sigma(b')_{[k, k+1]} & \text{for } k > 0, \end{cases}$$

and label it with $(a'b', k) \in \mathcal{R}$.

The coding graph $G(c)$ can easily be seen to have distinct labels. Thus, the set of infinite walks $G(c)^\infty$ forms a shift of finite type. By the construction we see that Γ_c maps $\tilde{\Omega}_\sigma$ onto $G(c)^\infty$. Observe that for $\tilde{w} \in \tilde{\Omega}_\sigma$ the infinite path $\Gamma_c(\tilde{w}) \in G(c)^\infty$ starts in \tilde{w}_0 .

In §8 coding graphs for particular examples are depicted. In particular, in Example 8.1 we will outline a practical construction method. Several structural details of the coding graph depend on the substitution only. Indeed, the set of vertices is independent of the particular coding prescription. Also the edges of the shape $(ab, 0)$ are completely determined by the substitution. Each vertex has exactly one incoming edge of this type. They form more or less the backbone of each coding graph.

The other edges strongly depend on the choice of the coding prescription. Even the number of edges is in general not constant. For a coding prescription c with respect to a given substitution σ one can easily verify the formula

$$|\mathcal{R}| = \sum_{ab \in \mathcal{L}_2(\sigma)} |c(ab)|.$$

Hence, if there exist letters $a \in \mathcal{A}$ such that

$$|\{ab \in \mathcal{L}_2(\sigma) \mid b \in \mathcal{A}\}| \neq |\{ba \in \mathcal{L}_2(\sigma) \mid b \in \mathcal{A}\}|$$

then we can minimize or maximize the number of edges by a suitable choice of c (see Example 8.2). Another interesting detail is that coding graphs are not necessarily strongly connected. We demonstrate this in Example 8.4.

The main result of the present section is that $\Gamma_c : \tilde{\Omega}_\sigma \longrightarrow G(c)^\infty$ is always a continuous and surjective map.

THEOREM 5.2. *Let σ be a primitive substitution and c a coding prescription. Γ_c is continuous and maps $\tilde{\Omega}_\sigma$ surjectively onto $G(\sigma)^\infty$. The pre-image of the cylinder*

$$U := [(a_1b_1, k_1), (a_2b_2, k_2), \dots, (a_{n_0}b_{n_0}, k_{n_0})] \subset G(c)^\infty$$

is exactly the open set

$$O := S^{r_{n_0}}\sigma^{n_0}[a_{n_0} \cdot b_{n_0}] \subset \tilde{\Omega}_\sigma$$

with r_n as in (5).

Proof. We first show that Γ_c is surjective. Let $(a_nb_n, k_n)_{n \geq 1} \in G(c)^\infty$. Note that by the definition of the edges we have $S^{k_n}\sigma[a_n \cdot b_n] \subset [a_{n-1} \cdot b_{n-1}]$ for all $n \geq 2$. Therefore, by (5), we have

$$S^{r_1}\sigma[a_1 \cdot b_1] \supset S^{r_2}\sigma^2[a_2 \cdot b_2] \supset S^{r_3}\sigma^3[a_3 \cdot b_3] \supset \dots$$

Furthermore, none of these sets is empty since $a_nb_n \in \mathfrak{L}_2(\sigma)$ for all $n \geq 1$. Hence,

$$\bigcap_{n \geq 1} S^{r_n}\sigma^n[a_n \cdot b_n] \neq \emptyset$$

and for each \tilde{w} contained in this intersection we obviously have $\Gamma_c(\tilde{w}) = (a_nb_n, k_n)_{n \geq 1}$. $\Gamma_c^{-1}(U) = O$ follows directly from the considerations after Lemma 4.4.

Continuity follows from the observation that U is an open neighbourhood of $\Gamma_c(\tilde{w})$, which can be made arbitrarily small by choosing n_0 sufficiently large, and O is an open neighbourhood of \tilde{w} . \square

Note that Γ_c is not necessarily an injective map. Section 6 is devoted exclusively to this problem. For now we will introduce a condition that will turn out to be sufficient for injectivity.

Definition 5.3. (Traceability) Let $G(c)$ be a coding graph. The infinite walk $(a_nb_n, k_n)_{n \geq 1} \in G(c)^\infty$ is called:

- *left-traceable* if $|\sigma(a_n)| + k_n > 1$ for infinitely many indices n ;
- *right-traceable* if $|\sigma(b_n)| - k_n > 1$ for infinitely many indices n ;
- *traceable* if it is both left-traceable and right-traceable.

A coding prescription is called *traceable* if all elements of $G(c)^\infty$ are traceable.

Observe that there exists no sequence that is neither left-traceable nor right-traceable. This will follow from the proof of Proposition 5.5. The following lemma shows that the traceability of a coding prescription is sufficient for the injectivity of Γ_c .

LEMMA 5.4. *Let σ be a primitive substitution over an alphabet \mathcal{A} , c a coding prescription and $\tilde{w} \in \tilde{\Omega}_\sigma$. If $\Gamma_c(\tilde{w})$ is traceable then $\Gamma_c(\tilde{w}') = \Gamma_c(\tilde{w})$ implies $\tilde{w}' = \tilde{w}$.*

Proof. Let $\Gamma_c(\tilde{w}) = (a_nb_n, k_n)_{n \geq 1}$, $a_0b_0 := \tilde{w}_0$ and define $A^{(n)}$, $B^{(n)}$ for $n \geq 0$ as in (7). We have already noticed that

$$\tilde{w} \in S^{r_n}\sigma^n[a_n \cdot b_n] \subseteq C_n = [\sigma^{n-1}(A^{(n-1)}) \dots A^{(0)}a_0 \cdot b_0B^{(0)} \dots \sigma^{n-1}(B^{(n-1)})]$$

for all $n \geq 1$.

We now show the lemma indirectly. Suppose that there existed a $\tilde{w}' \in \tilde{\Omega}_\sigma$ different from \tilde{w} with $\Gamma_c(\tilde{w}') = (a_n b_n, k_n)_{n \geq 1}$, too. Then there exists a $\delta > 0$ such that $d(\tilde{w}, \tilde{w}') = \delta$. Now observe that traceability implies that $|A^{(n)}| \geq 1$ and $|B^{(n)}| \geq 1$ for infinitely many indices. Thus, we can find an index n_0 with

$$\exp(-\min\{|\sigma^{n_0-1}(A^{(n_0-1)}) \cdots A^{(0)} a_0|, |b_0 B^{(0)} \cdots \sigma^{n_0-1}(B^{(n_0-1)})|\}) < \delta.$$

This immediately shows that $\tilde{w} \notin C_{n_0}$ and, hence, $\tilde{w}' \notin S^{r_{n_0}} \sigma^{n_0}[a_{n_0} \cdot b_{n_0}]$. This contradicts Theorem 5.2. \square

We see that for $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ traceable the coding is uniquely reversible and \tilde{w} can be represented by

$$\begin{aligned} \tilde{w} &= \bigcap_{n \geq 1} C_n = \lim_{n \rightarrow \infty} [\sigma^{n-1}(A^{(n-1)}) \cdots A^{(0)} a_0 \cdot b_0 B^{(0)} \cdots \sigma^{n-1}(B^{(n-1)})] \\ &= \lim_{n \rightarrow \infty} S^{r_n}[\sigma^n(a_n) \cdot \sigma^n(b_n)]. \end{aligned} \quad (8)$$

In particular, Γ_c is injective on $\tilde{\Omega}_\sigma$ if c is a traceable coding prescription.

Note that we can easily code each substitution dynamical system in a traceable and, hence, injective way. In particular, let σ be a primitive substitution over the alphabet \mathcal{A} . We can find an integer n such that $|\sigma^n(a)| \geq 3$ for all $a \in \mathcal{A}$. Note that $\tilde{\Omega}_\sigma = \tilde{\Omega}_{\sigma^n}$. By the choice of n we can find a coding prescription c with respect to σ^n such that for each $a \in \mathcal{A}$ neither $|\sigma^n(a)| - 1$ nor $1 - |\sigma^n(a)|$ is contained in $c(a)$. Thus, all elements of $G(c)^\infty$ are clearly left- and right-traceable and, hence, c is traceable. The induced coding Γ_c is injective and all elements of $\tilde{\Omega}_\sigma$ can be recovered consistently by (8).

In §6 we will give sufficient and necessary conditions for a coding prescription to be traceable. Furthermore, we will see that traceability is not necessary for a coding to be injective. However, we will precisely analyse for which non-traceable coding prescriptions the coding Γ_c is injective.

The rest of the section is devoted to the set of periodic points \tilde{P} of a substitution dynamical system $\tilde{\Omega}_\sigma$ and the corresponding Γ_c -images. We will need several formalisms concerning \tilde{P} from §3.3.

PROPOSITION 5.5. *Let σ be a primitive substitution over an alphabet \mathcal{A} and c a coding prescription with respect to σ . If $\langle \tilde{z}^{(0)}, \dots, \tilde{z}^{(p-1)} \rangle$ is a cycle of $\sigma|_{\tilde{P}}$ then, for each $m \in \{0, \dots, p-1\}$, $\Gamma_c(\tilde{z}^{(m)})$ is purely periodic with*

$$\Gamma_c(\tilde{z}^{(m)}) = (\tilde{z}_0^{(m-n)}, 0)_{n \geq 1}$$

where we consider the upper indices modulo p .

On the other hand, if $\Gamma_c(\tilde{w}) = (a_n b_n, 0)_{n \geq 1}$ then $\Gamma_c(\tilde{w})$ is purely periodic and \tilde{w} is a periodic point.

Proof. Note that $0 \in c(ab)$ for all $ab \in \mathcal{L}_2(\sigma)$, independently of the actual coding prescription c . Thus, for each $m \in \{0, \dots, p-1\}$, we have (modulo p) $\Theta(\tilde{z}^{(m)}) = \tilde{z}^{(m-1)}$ and $\varphi(\tilde{z}^{(m)}) = 0$. With this observation the first part of the proposition follows directly from the definition of Γ_c .

Now suppose that $\Gamma_c(\tilde{w}) = (a_n b_n, 0)_{n \geq 1}$. By Theorem 5.2 we know that $\Gamma_c(\tilde{w})$ is an infinite walk on the coding graph $G(c)$. Note that each vertex ab of this graph has exactly one incoming edge whose second entry equals 0 (namely, the edge $(ab, 0)$ that has its origin in $\sigma(a)_{-1}\sigma(b)_1$). Since $G(c)$ is finite we immediately see that $(a_n b_n, 0)_{n \geq 1}$ is purely periodic with period, say, p ($p \leq |\mathcal{L}_2(\sigma)|$).

Observe that $\Gamma_c(\tilde{w})$ is traceable. Indeed, suppose $\Gamma_c(\tilde{w})$ were not left-traceable (for right-traceability the argumentation is analogous). By the periodicity of $\Gamma_c(\tilde{w})$ this implies that $|\sigma(a_n)| = 1$ for all $n \in \{0, \dots, p-1\}$. Moreover, we have $\sigma(a_n)^n = a_n$, which contradicts the definition of a primitive substitution.

Now, by (iii) of Lemma 4.4, we see that $\Gamma_c(\tilde{w}) = \Gamma_c(\Theta^p(\tilde{w}))$, by Lemma 5.4 the traceability of $\Gamma_c(\tilde{w})$ implies that $\tilde{w} = \Theta^p(\tilde{w})$ and by the shape of $\Gamma_c(\tilde{w})$ we conclude that $\varphi(\Theta^{n-1}(w)) = 0$ for all $n \geq 1$. This immediately yields $\tilde{w} = \sigma^p(\tilde{w})$, showing that $\tilde{w} \in \tilde{P}$. \square

Note that Proposition 5.5 does not depend on the particular coding prescription c . We see that for a periodic points \tilde{z} the code $\Gamma_c(\tilde{z})$ is traceable and, hence, uniquely reversible in the sense of (8). Furthermore, Γ_c respects the cyclic structure of $\sigma|_{\tilde{P}}$, that is, a cycle of $\sigma|_{\tilde{P}}$ uniquely corresponds to a cycle of $G(c)$ of equal length.

6. When is the coding map injective?

In the present section we study the problem of injectivity of the coding Γ_c . We will see that the structure of the periodic points plays a major role. In the previous section we presented the concept of traceability that has already led to preliminary results. In particular, for a traceable coding prescription c , Γ_c is completely injective. The first main theorem of this section is a complete characterization of the traceable coding prescriptions with respect to a given substitution σ .

THEOREM 6.1. *Let σ be a primitive substitution and c a coding prescription. c is traceable if and only if:*

- (i) *for each cycle $\langle u^{(0)}, \dots, u^{(p-1)} \rangle$ of $\sigma|_L$*

$$\text{there exists } i \in \{0, \dots, p-1\} : |\sigma(u_{-1}^{(i)})| - 1 \notin c(u_{-1}^{(i)}) \quad \text{and} \quad (9)$$

- (ii) *for each cycle $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ of $\sigma|_R$*

$$\text{there exists } j \in \{0, \dots, q-1\} : 1 - |\sigma(v_1^{(j)})| \notin c(v_1^{(j)}). \quad (10)$$

But traceability is not necessary for Γ_c to be injective. The second important theorem here exactly shows under which conditions a not traceable coding prescription provides an injective coding.

THEOREM 6.2. *Let σ be a primitive substitution and c a non-traceable coding prescription. Γ_c is completely injective on $\tilde{\Omega}_\sigma$ if and only if:*

- (i) *for each cycle $\langle u^{(0)}, \dots, u^{(p-1)} \rangle$ of $\sigma|_L$ that does not satisfy (9) there exists a uniquely determined cycle $\langle \tilde{z}^{(0)}, \dots, \tilde{z}^{(p-1)} \rangle$ of $\sigma|_{\tilde{P}}$ with equal period such that $\pi^{(l)}(\tilde{z}^{(i)}) = u^{(i)}$ for all $i \in \{0, \dots, p-1\}$; and*

- (ii) for each cycle $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ of $\sigma|_R$ that does not satisfy (10) there exists a uniquely determined cycle $\langle \tilde{z}^{(0)}, \dots, \tilde{z}^{(q-1)} \rangle$ of $\sigma|_{\tilde{P}}$ with equal period such that $\pi^{(r)}(\tilde{z}^{(j)}) = v^{(j)}$ for all $j \in \{0, \dots, q-1\}$.

We see that the injectivity of Γ_c strongly depends on the injectivity of $\pi^{(l)}$ and $\pi^{(r)}$ on \tilde{P} . One immediately deduces the following corollary.

COROLLARY 6.3. *For a primitive substitution σ each coding prescription yields an injective coding of $\tilde{\Omega}_\sigma$ if and only if $|\tilde{P}| = |L| = |R|$, that is, $\pi^{(l)}$ and $\pi^{(r)}$ act completely injectively on \tilde{P} .*

To prove these results we need several technical lemmas. Furthermore, Proposition 6.9 will provide a complete description of the elements of a substitution dynamical system that (possibly) yield not traceable codes. The first lemma shows that not traceable sequences always have a purely periodic tail.

LEMMA 6.4. *A not traceable sequence $(a_n b_n, k_n)_{n \geq 1} \in G(c)^\infty$ is ultimately periodic.*

Proof. Observe that the sequence is either not left-traceable or not right-traceable. Thus, there exists an index $n_0 \geq 1$ such that $k_n = 1 - |\sigma(a_n)|$ for all $n \geq n_0$ or $k_n = |\sigma(b_n)| - 1$ for all $n \geq n_0$, respectively. By the definition of $G(c)$ a vertex ab has for each $k \in c(ab)$ exactly one incoming edge (ab, k) . Thus, each vertex ab has at most one incoming edge $(ab, 1 - |\sigma(a)|)$ and, analogously, at most one incoming edge $(ab, |\sigma(b)| - 1)$. Since $G(c)$ is finite this shows that $(a_n b_n, k_n)_{n \geq n_0}$ is periodic. \square

The next two lemmas show that for particular coding prescriptions the cycles of $\sigma|_L$ ($\sigma|_R$) give rise to purely periodic, not right-traceable (not left-traceable) codes.

LEMMA 6.5. *Let c be a coding prescription with respect to a primitive substitution σ , $\langle u^{(0)}, \dots, u^{(p-1)} \rangle$ a cycle of $\sigma|_L$ and suppose that $|\sigma(u_{-1}^{(i)})| - 1 \in c(u_{-1}^{(i)})$ for all $i \in \{0, \dots, p-1\}$. For each $\tilde{z} \in \tilde{P}$ with $\pi^{(l)}(\tilde{z}) = u^{(i)}$ for some $i \in \{0, \dots, p-1\}$,*

$$\Gamma_c(S^{-1}\tilde{z}) = (u_{[-2, -1]}^{(i-n)}, |\sigma(u_{-1}^{(i-n)})| - 1)_{n \geq 1}$$

(upper indices are considered modulo p), that is, $\Gamma_c(S^{-1}\tilde{z})$ is purely periodic and not right-traceable.

Proof. Consider the upper indices modulo p for the rest of the proof.

First we claim that for each $\tilde{z} \in \tilde{P}$ with $\pi^{(l)}(\tilde{z}) = u^{(i)}$ for some $i \in \{0, \dots, p-1\}$ there exists a $\tilde{z}' \in \tilde{P}$ with $\Theta(S^{-1}\tilde{z}) = S^{-1}\tilde{z}'$, $\pi^{(l)}(\tilde{z}') = u^{(i-1)}$ and $\varphi(S^{-1}\tilde{z}) = |\sigma(u_{-1}^{(i-1)})| - 1$. Indeed, since $\tilde{z} \in \tilde{P}$ there exists a (uniquely determined) $\tilde{z}' \in \tilde{P}$ with

$$\sigma(\tilde{z}') = \sigma(\pi^{(l)}(\tilde{z}')) \cdot \sigma(\pi^{(r)}(\tilde{z}')) = \pi^{(l)}(\tilde{z}) \cdot \pi^{(r)}(\tilde{z}) = \tilde{z}$$

and, by assumption, $\pi^{(l)}(\tilde{z}') = u^{(i-1)}$. We have $|\sigma(u_{-1}^{(i-1)})| - 1 \in c((S^{-1}\tilde{z}')_0) = c(u_{[-2, -1]}^{(i-1)})$ and using (3), we see that

$$S^{|\sigma(u_{-1}^{(i-1)})| - 1} \sigma(S^{-1}\tilde{z}') = S^{|\sigma(u_{-1}^{(i-1)})| - 1} \circ S^{-|\sigma(u_{-1}^{(i-1)})|} \sigma(\tilde{z}') = S^{-1}\tilde{z},$$

which shows the claim.

Now, successive application of this idea and observation of (iii) of Lemma 4.4 immediately completes the proof. \square

Similarly, we obtain an analogous result for the cycles of $\sigma|_R$.

LEMMA 6.6. *Let c be a coding prescription with respect to a primitive substitution σ , $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ a cycle of $\sigma|_R$ and suppose that $1 - |\sigma(v_1^{(j)})| \in c(v_1^{(j)})$ for all $j \in \{0, \dots, q-1\}$. For each $\tilde{z} \in \tilde{P}$ with $\pi^{(r)}(\tilde{z}) = v^{(j)}$ for some $j \in \{0, \dots, q-1\}$,*

$$\Gamma_c(S\tilde{z}) = (v_{[1,2]}^{(j-n)}, 1 - |\sigma(v_1^{(j-n)})|)_{n \geq 1}$$

(upper indices are considered modulo q), that is, $\Gamma_c(S\tilde{z})$ is purely periodic and not left-traceable.

Proof. The proof is analogous to the proof of Lemma 6.5. \square

We also have contrary statements.

LEMMA 6.7. *Let σ be a primitive substitution, c a coding prescription with respect to σ , and $\tilde{w} \in \tilde{\Omega}_\sigma$ such that $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ is not left-traceable and purely periodic with period q . Then $S^{-1}\tilde{w} \in \tilde{P}$ and there exists a uniquely determined cycle $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ of $\sigma|_R$ such that $a_n b_n = v_{[1,2]}^{(q-n)}$ for all $n \in \{1, \dots, q\}$ and $\pi^{(r)}(S^{-1}\tilde{w}) = v^{(0)}$.*

Proof. The fact that $\Gamma_c(\tilde{w})$ is not left-traceable and purely periodic implies that $k_n = 1 - |\sigma(a_n)|$ for all $n \in \mathbb{N}$. Consider the sequence $(\Theta^n(\tilde{w}))_{n \geq 0}$. Define $u^{(n)} := \pi^{(l)}(S^{-1}\Theta^n(\tilde{w}))$ and $v^{(n)} := \pi^{(r)}(S\Theta^n(\tilde{w}))$ for $n \geq 0$. Therefore, $\Theta^n(\tilde{w}) = u^{(n)} a_n \cdot b_n v^{(n)}$ (where $a_0 b_0 := \tilde{w}_0$).

We now claim that $\Theta(S^{-1}\Theta^n(\tilde{w})) = S^{-1}\Theta^{n+1}(\tilde{w})$ and $\varphi(S^{-1}\Theta^n(\tilde{w})) = 0$ for all $n \geq 0$. To show this we first note that $S^{1-|\sigma(a_n)|}\sigma(\Theta^{n+1}(\tilde{w})) = \Theta^n(\tilde{w})$ and then calculate

$$\begin{aligned} \sigma(S^{-1}\Theta^{n+1}(\tilde{w})) &= \sigma(u^{(n)}) \cdot \sigma(a_n) \sigma(b_n) \sigma(v^{(n)}) \\ &= S^{-|\sigma(a_n)|} (\sigma(u^{(n)}) \sigma(a_n) \cdot \sigma(b_n) \sigma(v^{(n)})) \\ &= S^{-1} \circ S^{1-|\sigma(a_n)|} \sigma(\Theta^{n+1}(\tilde{w})) = S^{-1}\Theta^n(\tilde{w}). \end{aligned}$$

$0 \in c(ab)$ for all $ab \in \mathfrak{L}_2(\sigma)$ and, thus, we have established the claim.

By the definition of Γ_c we immediately obtain that

$$\Gamma_c(S^{-1}\tilde{w}) = (u_{-1}^{(n)} a_n, 0)_{n \geq 1}.$$

Therefore, by Proposition 5.5, $S^{-1}\tilde{w} \in \tilde{P}$ and $\pi^{(r)}(S^{-1}\Theta^n(\tilde{w})) = a_n b_n v^{(n)} \in R$ for all $n \geq 0$. As $a_q b_q = a_0 b_0$, we have $\pi^{(r)}(S^{-1}\tilde{w}) = \pi^{(r)}(S^{-1}\Theta^q(\tilde{w}))$. Thus, setting $v^{(n)} := \pi^{(r)}(S^{-1}\Theta^{q-n}(\tilde{w}))$ for $n \in \{0, \dots, q-1\}$, we obtain the cycle $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ of $\sigma|_R$ which is uniquely determined and $v^{(q)} = v^{(0)} = \pi^{(r)}(S^{-1}\tilde{w})$. \square

For not left-traceable sequences an analogous result holds.

LEMMA 6.8. *Let σ be a primitive substitution, c a coding prescription with respect to σ , and $\tilde{w} \in \tilde{\Omega}_\sigma$ such that $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ is not right-traceable and purely periodic with period p . Then $S\tilde{w} \in \tilde{P}$ and there exists a uniquely determined cycle $\langle u^{(0)}, \dots, u^{(p-1)} \rangle$ of $\sigma|_L$ such that $a_n b_n = u_{[-2,-1]}^{(p-n)}$ for all $n \in \{1, \dots, p\}$ and $\pi^{(l)}(S^{-1}\tilde{w}) = u^{(p)}$.*

Proof. The lemma can be proved similarly to Lemma 6.7. \square

Summarizing the assertions of the previous lemmas, we see that each not traceable sequence in $G(c)$ ends up in a not traceable purely periodic tail. Each not left-traceable tail corresponds to a cycle of $\sigma|_R$ and each not right-traceable tail corresponds to a cycle of $\sigma|_L$. Together with the result from Proposition 5.5 we obtain a quite complete picture of how the cyclic structure of the periodic points is reflected in the code. The number of ways for a not traceable sequence to end up is bounded by $2|A|$, where the exact number depends on the particular coding prescription. This gives a characterization of the elements of a substitution dynamical system that yield non-traceable codes.

PROPOSITION 6.9. *Let σ be a primitive substitution, c a non-traceable coding prescription and $\tilde{w} \in \tilde{\Omega}_\sigma$.*

- *If $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ is not left-traceable then there exists an index $n_0 \geq 0$ such that $(a_n b_n, k_n)_{n \geq n_0+1}$ is purely periodic and $S^{-m} \tilde{w} \in \tilde{P}$ for some $m \geq 1$.*
- *If $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ is not right-traceable then there exists an index $n_0 \geq 0$ such that $(a_n b_n, k_n)_{n \geq n_0+1}$ is purely periodic and $S^m \tilde{w} \in \tilde{P}$ for some $m \geq 1$.*

Proof. Suppose that $\Gamma_c(\tilde{w}) = (a_n b_n, k_n)_{n \geq 1}$ is not left-traceable. Then, by Lemma 6.4, $\Gamma_c(\tilde{w})$ is ultimately periodic and, hence, there exists an index $n_0 \geq 0$ such that $(a_n b_n, k_n)_{n \geq n_0+1}$ is purely periodic and, of course, not left-traceable. If $n_0 = 0$ then we already know that $S^{-1} \tilde{w} \in \tilde{P}$ by Lemma 6.7. Otherwise we use (3) and (5) to calculate

$$\begin{aligned} \tilde{w} &= S^{r_{n_0}} \sigma^{n_0}(\Theta^{n_0}(\tilde{w})) = S^{r_{n_0}} \circ S^{|\sigma^{n_0}(a_{n_0})|} \sigma^{n_0}(S^{-1} \Theta^{n_0}(\tilde{w})) \\ &= S^{r_{n_0} + |\sigma^{n_0}(a_{n_0})|} \sigma^{n_0}(S^{-1} \Theta^{n_0}(\tilde{w})). \end{aligned}$$

Set $m := r_{n_0} + |\sigma^{n_0}(a_{n_0})|$ and note that $m \geq 1$. Furthermore, by (iii) of Lemma 4.4 and Lemma 6.7, we have $S^{-1} \Theta^{n_0}(\tilde{w}) \in \tilde{P}$. This yields

$$S^{-m} \tilde{w} = \sigma^{n_0}(S^{-1} \Theta^{n_0}(\tilde{w})) \in \tilde{P}.$$

For $\Gamma_c(\tilde{w})$ not right-traceable the proof is analogous. □

We are now in a position to prove the two theorems stated at the beginning of this section.

Proof of Theorem 6.1. The \Rightarrow part of the proof is done indirectly by means of Lemmas 6.5 and 6.6.

Turning to the \Rightarrow part, if c is not traceable then $G(c)^\infty$ contains either a not left-traceable or a not right-traceable sequence. Suppose first the existence of a sequence that is not left-traceable. By Lemma 6.4 this sequence is ultimately periodic. Thus, $G(c)^\infty$ also contains a purely periodic sequence $(a_n b_n, k_n)_{n \geq 1}$, $(a_n b_n, k_n) = (a_{n+q} b_{n+q}, k_{n+q})$ for all $n \geq 1$, which is also not left-traceable. Lemma 6.7 shows that there exists a uniquely determined cycle $\langle v^{(0)}, \dots, v^{(q-1)} \rangle$ of $\sigma|_R$ such that $a_n b_n = v_{[1,2]}^{(q-n)}$ for all $n \in \{1, \dots, q\}$. Since $(a_n b_n, k_n)_{n \geq 1}$ is not left-traceable we necessarily have $1 - |\sigma(a_n)| \in c(a_n)$ for all $n \in \{1, \dots, q\}$, contradicting (ii). Similarly, the assumption that $G(c)^\infty$ contains a not right-traceable sequence yields a contradiction to (i). □

Proof of Theorem 6.2. This is a direct consequence of Lemmas 6.5–6.8. □

We see that the only elements of a substitution dynamical system that possibly yield not traceable codes lie within the shift orbits of the periodic points. Since \tilde{P} is finite and the shift orbits are countable, these elements form a set of measure 0. For them, the coding is finite-to-one, depending on the fibres of $\pi^{(l)}$ and $\pi^{(r)}$. However, Γ_c is measure-theoretically bijective for every coding prescription c .

7. How does the shift map interact with Γ_c ?

Finally, we will see that coding prescriptions of a special shape induce a partial ordering on the edges of the coding graph that extends to a partial ordering on the infinite paths (cf. [Ver81]). This partial ordering interacts with the shift map of the underlying substitution dynamical system.

Definition 7.1. (Continuous coding prescription) Let σ be a primitive substitution over the alphabet \mathcal{A} . A coding prescription c with respect to σ is called *continuous* if for each $a \in \mathcal{A}$ the set $c(a)$ consists of consecutive integers.

Note that this definition immediately implies that for every $ab \in \mathfrak{L}_2(\sigma)$ the set $c(ab)$ is also a set of consecutive integers. For a continuous coding prescription the desubstitution $(\Theta(\tilde{w}), \varphi(\tilde{w}))$ has a nice behaviour with respect to the shift map.

LEMMA 7.2. Let c be a continuous coding prescription with respect to a primitive substitution σ , $\tilde{w} \in \tilde{\Omega}_\sigma$ and $(u \cdot v, k) := (\Theta(\tilde{w}), \varphi(\tilde{w}))$. Then

$$(\Theta(S\tilde{w}), \varphi(S\tilde{w})) = \begin{cases} (u \cdot v, k + 1) & \text{if } k + 1 \in c(u_{-1}v_1), \\ (S(u \cdot v), k + 1 - |\sigma(v_1)|) & \text{otherwise;} \end{cases} \quad (11)$$

$$(\Theta(S^{-1}\tilde{w}), \varphi(S^{-1}\tilde{w})) = \begin{cases} (u \cdot v, k - 1) & \text{if } k - 1 \in c(u_{-1}v_1), \\ (S^{-1}(u \cdot v), k - 1 + |\sigma(u_{-1})|) & \text{otherwise.} \end{cases} \quad (12)$$

Proof. We start by showing (11). Since $\tilde{w} = S^k \sigma(u \cdot v)$, we have $S\tilde{w} = S^{k+1} \sigma(u \cdot v)$. If $k + 1 \in c(u_{-1}v_1)$ then obviously $(\Theta(S\tilde{w}), \varphi(S\tilde{w})) = (u \cdot v, k + 1)$ by the uniqueness in Theorem 4.3.

Contrarily, suppose that $k + 1 \notin c(u_{-1}v_1)$. Observe that

$$S\tilde{w} = S^{k+1} \sigma(u \cdot v) = S^{k+1-|\sigma(v_1)|} \sigma(S(u \cdot v)),$$

where we have used (3). Since c is continuous and $0 \in c(ab)$ for each $ab \in \mathfrak{L}_2(\sigma)$ we immediately see that $k \geq 0$ and, hence, $k \in c(v_1)$, $|\sigma(v_1)| \geq k + 1 \notin c(v_1)$.

By the basic properties of coding prescriptions (Definition 4.2) we conclude that $0 \geq k + 1 - |\sigma(v_1)| \in c(v_1)$ and, therefore, $k + 1 - |\sigma(v_1)| \in c(v_1v_2)$. Again, by the uniqueness in Theorem 4.3, we see that $(\Theta(S\tilde{w}), \varphi(S\tilde{w})) = (S(u \cdot v), k - |\sigma(v_1)| + 1)$.

The proof of (12) is analogous. \square

Observe that for non-continuous coding prescriptions the lemma does not hold in general. Of course, the first part of the assertions (for $\varphi(\tilde{w}) + 1 \in c(\Theta(\tilde{w})_0)$ and $\varphi(\tilde{w}) - 1 \in c(\Theta(\tilde{w})_0)$, respectively) is still true but problems occur in the contrary case. More precisely, it may happen that $\Theta(S\tilde{w}) = S^{-1}\Theta(\tilde{w})$ (and, analogously, $\Theta(S^{-1}\tilde{w}) = S\Theta(\tilde{w})$). For the rest of the section we will concentrate on continuous coding prescriptions.

We define a partial ordering $<$ on \mathcal{R} by

$$(ab, k) < (a'b', k') \iff ab = a'b', k < k'.$$

We call (ab, k) minimal (maximal) if, for all $(ab, k') \in \mathcal{R}$, we have $k' \geq k$ ($k' \leq k$). Note that (depending on σ and the concrete coding prescription) an element (ab, k) of \mathcal{R} can be both minimal and maximal. In this case we necessarily have $k = 0$.

Analogously, we call an edge of $G(c)$ minimal (maximal) if it is labelled by a minimal (maximal) element of \mathcal{R} . By the definition of $G(c)$ we immediately see that each vertex ab has exactly one minimal and one maximal incoming edge. For a non-maximal element (ab, k) define the predecessor $T(ab, k)$ to be the (unique) element $(ab, k + 1)$. Similarly, for a non-minimal element (ab, k) , we set $T^{-1}(ab, d) := (ab, k - 1)$.

We can extend our partial ordering to $G(c)^\infty$:

$$(a_nb_n, k_n)_{n \geq 1} < (a'_nb'_n, k'_n)_{n \geq 1} \iff \text{there exists } n_0 : (a_nb_n, k_n) = (a'_nb'_n, k'_n) \\ \text{for } n > n_0 \text{ and } (a_{n_0}b_{n_0}, k_{n_0}) < (a'_{n_0}b'_{n_0}, k'_{n_0}).$$

$(a_nb_n, k_n)_{n \geq 1}$ is said to be minimal (maximal) if (a_nb_n, k_n) is minimal (maximal) for all $n \geq 1$.

The successor map T also extends the non-maximal elements of $G(c)^\infty$. In particular, consider an infinite non-maximal path $(a_nb_n, k_n)_{n \geq 1}$. Then we can find an index $n_0 \geq 1$ such that $(a_{n_0}b_{n_0}, k_{n_0})$ is not maximal while (a_nb_n, k_n) is maximal for all $n < n_0$. Define

$$T(a_nb_n, k_n)_{n \geq 1} := (a'_1b'_1, k'_1), \dots, (a'_{n_0-1}b'_{n_0-1}, k'_{n_0-1}), \\ T(a_{n_0}b_{n_0}, k_{n_0}), (a_{n_0+1}b_{n_0+1}, k_{n_0+1}), (a_{n_0+2}b_{n_0+2}, k_{n_0+2}), \dots$$

where $(a'_1b'_1, k'_1), \dots, (a'_{n_0-1}b'_{n_0-1}, k'_{n_0-1})$ is the uniquely determined path of length $n_0 - 1$ that consists of minimal edges only and ends in the initial vertex of the edge $T(a_{n_0}b_{n_0}, k_{n_0})$. Note that the uniqueness of the path follows from the observation that each vertex has exactly one incoming edge that is minimal.

Analogously, for a non-minimal path $(a_nb_n, k_n)_{n \geq 1}$ with $n_0 \geq 1$ the least index such that $(a_{n_0}b_{n_0}, k_{n_0})$ is not minimal, we have

$$T^{-1}(a_nb_n, k_n)_{n \geq 1} := (a'_1b'_1, k'_1), \dots, (a'_{n_0-1}b'_{n_0-1}, k'_{n_0-1}), \\ T^{-1}(a_{n_0}b_{n_0}, k_{n_0}), (a_{n_0+1}b_{n_0+1}, k_{n_0+1}), (a_{n_0+2}b_{n_0+2}, k_{n_0+2}), \dots$$

where $(a'_1b'_1, k'_1), \dots, (a'_{n_0-1}b'_{n_0-1}, k'_{n_0-1})$ is the unique path of length $n_0 - 1$ that consists of maximal edges only and ends in the initial vertex of the edge $T^{-1}(a_{n_0}b_{n_0}, k_{n_0})$.

THEOREM 7.3. *Let c be a continuous coding prescription with respect to a primitive substitution σ and $\tilde{w} \in \tilde{\Omega}_\sigma$. If $\Gamma_c(\tilde{w})$ is not maximal then $\Gamma_c(S\tilde{w}) = T\Gamma_c(\tilde{w})$, otherwise $\Gamma_c(S\tilde{w})$ is minimal.*

Proof. Let (ab, k) be the first term of $\Gamma_c(\tilde{w})$. Note that (ab, k) is maximal if and only if $k + 1 \notin c(ab)$. Suppose first that (ab, k) is not maximal. In this case, by (11), we have $\Theta(S\tilde{w}) = \Theta(\tilde{w})$ and $\varphi(S\tilde{w}) = k + 1$. Thus, $\Gamma_c(S\tilde{w})$ starts with $(ab, k + 1) = T(ab, k)$ and $\Gamma_c(S\tilde{w}) = T\Gamma_c(\tilde{w})$.

Now suppose that (ab, k) is maximal and let $(a'b', k')$ be the first term of $\Gamma_c(S\tilde{w})$. We claim that $(a'b', k')$ is minimal. From (11) we deduce that $a' = b$ and $k' = k + 1 - |\sigma(b)| \leq 0$. If $(a'b', k')$ were not minimal there would exist an element $(a'b'', k'') \in \mathcal{R}$ with $-|\sigma(a'')| = -|\sigma(b)| < k'' < k'$. By the definition of coding prescriptions this would imply that $0 < k'' + |\sigma(b)| \notin c(b)$. But this contradicts the continuity of c since $0 < k'' + |\sigma(b)| < k$ and $0 \in c(b)$ as well as $k \in c(b)$.

Now the theorem follows directly from (iii) of Lemma 4.4 and Lemma 7.2. \square

Analogously one can prove that $\Gamma_c(S^{-1}\tilde{w}) = T^{-1}(\Gamma_c(\tilde{w}))$ for $\Gamma_c(\tilde{w})$ not minimal, otherwise $\Gamma_c(S^{-1}\tilde{w})$ is maximal.

We summarize the previous results of this section in a theorem.

THEOREM 7.4. *Let σ be a primitive substitution and c a continuous coding prescription. Γ_c conjugates the shift map S on $\tilde{\Omega}_\sigma$ with the successor map T on $G(c)^\infty$ (provided that the latter one is defined). In other words, (Ω_σ, S) and $(G(c)^\infty, T)$ are topologically conjugated dynamical systems.*

We want to finish the section by showing some connections between minimality (maximality) of elements of $G(c)^\infty$ and traceability. Indeed, the following lemma shows that non-traceable and purely periodic element of $G(c)$ are maximal or minimal.

LEMMA 7.5. *Let σ be a primitive substitution and c a continuous coding prescription. A not traceable sequence of $G(\sigma, c)^\infty$ is minimal or maximal if and only if it is purely periodic.*

Proof. Since each vertex of $G(\sigma, c)^\infty$ has exactly one maximal and one minimal incoming edge we immediately see that a minimal or maximal sequence is purely periodic.

Now suppose that $(a_nb_n, k_n)_{n \geq 1} \in G(c)^\infty$ is not traceable and purely periodic. By Definition 5.3 we either have $k_n = 1 - |\sigma(a_n)|$ for all $n \geq 1$ (not left-traceable) or $k_n = |\sigma(b_n)| - 1$ for all $n \geq 1$ (not right-traceable). Obviously $(a_nb_n, k_n)_{n \geq 1}$ is minimal in the first case and maximal in the second case. \square

As a consequence of this lemma we obtain an equivalence in Proposition 6.9 for a continuous coding prescription.

COROLLARY 7.6. *Let σ be a primitive substitution and c a continuous coding prescription.*

$\Gamma_c(\tilde{w})$ is not right-traceable if and only if $S^m \tilde{w} \in \tilde{P}$ for some $m \geq 1$ and $\pi^{(l)}(S^m \tilde{w})$ is contained in a cycle of L that does not satisfy (9).

$\Gamma_c(\tilde{w})$ is not left-traceable if and only if $S^m \tilde{w} \in \tilde{P}$ for some $m \leq -1$ and $\pi^{(r)}(S^m \tilde{w})$ is contained in a cycle of R that does not satisfy (10).

8. Examples

We demonstrate the major results and techniques of the paper on some examples.

Example 8.1. Let $\sigma : 1 \mapsto 312, 2 \mapsto 32, 3 \mapsto 13$. For this substitution we have $\mathfrak{L}_2(\sigma) = \{12, 13, 21, 23, 31, 32, 33\}$. σ provides four periodic points:

$$\tilde{z}^{(1)} = u^{(1)} \cdot v^{(1)}, \quad \tilde{z}^{(2)} = u^{(1)} \cdot v^{(2)}, \quad \tilde{z}^{(3)} = u^{(2)} \cdot v^{(1)}, \quad \tilde{z}^{(4)} = u^{(2)} \cdot v^{(2)}$$

with

$$\begin{aligned} u^{(1)} &= \dots 1213312131332, & v^{(1)} &= 1331232312133121313 \dots, \\ u^{(2)} &= \dots 2133123231213, & v^{(2)} &= 3121313312321332133 \dots \end{aligned}$$

We have $\sigma(\tilde{z}^{(1)}) = \tilde{z}^{(2)}$, $\sigma(\tilde{z}^{(2)}) = \tilde{z}^{(1)}$, $\sigma(\tilde{z}^{(3)}) = \tilde{z}^{(4)}$ and $\sigma(\tilde{z}^{(4)}) = \tilde{z}^{(3)}$. Thus, the cyclic structure of $\sigma|_L$, $\sigma|_R$ and $\sigma|_{\tilde{P}}$ is given by

$$\sigma|_L = \langle u^{(1)} \rangle \langle u^{(2)} \rangle, \quad \sigma|_R = \langle v^{(1)}, v^{(2)} \rangle, \quad \sigma|_{\tilde{P}} = \langle \tilde{z}^{(1)}, \tilde{z}^{(2)} \rangle \langle \tilde{z}^{(3)}, \tilde{z}^{(4)} \rangle.$$

The following diagram shows the action of $\pi^{(l)}$ and $\pi^{(r)}$ on the cycles. In particular, $\pi^{(l)}$ as well as $\pi^{(r)}$ are not injective on \tilde{P} .

$$\begin{array}{ccccc} \sigma|_L & & \sigma|_{\tilde{P}} & & \sigma|_R \\ \langle u^{(1)} \rangle & \xleftarrow{\pi^{(l)}} & \langle \tilde{z}^{(1)}, \tilde{z}^{(2)} \rangle & \xrightarrow{\pi^{(r)}} & \langle v^{(1)}, v^{(2)} \rangle \\ \langle u^{(2)} \rangle & \xleftarrow{\pi^{(l)}} & \langle \tilde{z}^{(3)}, \tilde{z}^{(4)} \rangle & \xrightarrow{\pi^{(r)}} & \end{array}$$

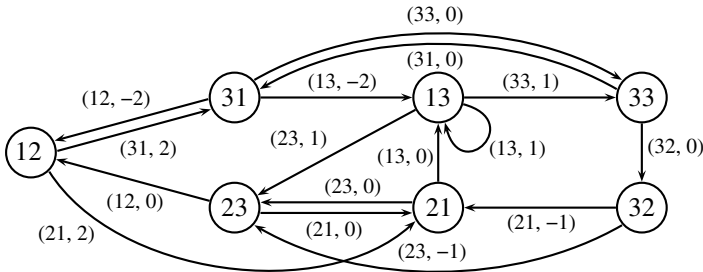
We choose the coding prescription

$$c : 1 \mapsto \{-2, 0, 2\}, \quad 2 \mapsto \{-1, 0\}, \quad 3 \mapsto \{0, 1\}$$

and want to determine the corresponding coding graph. The set of vertices is given by $\mathcal{L}_2(\sigma)$ and does not depend on the chosen coding prescription. The easiest way to construct the graph is to calculate the incoming edges for each vertex. We show this explicitly for the vertex 13. Note that $c(13) = \{-2, 0, 1\}$. From this we conclude that 13 has three incoming edges. We apply σ on 13 (the vertical bar divides $\sigma(1)$ and $\sigma(3)$).

$$\sigma(13) = \begin{array}{ccccccc} & -2 & & -1 & & 0 & & 1 \\ 3 & & 1 & & 2 & | & 1 & & 3 \end{array}$$

The superscript indices that appear in $c(13)$ indicate the particular origins of the edges. The first edge, corresponding to $-2 \in c(13)$, starts at 31 and is labelled by $(13, -2)$. The second edge starts at 21 (index 0) and is labelled by $(13, 0)$. The third edge has its origin in 13 (it is a self-loop) and the label $(13, 1)$. In this way it is not difficult to obtain the entire coding graph $G(c)$.



We now calculate $\Gamma_c(S\tilde{z}^{(2)})$. Observe that

$$S^{-2}\sigma(S\tilde{z}^{(1)}) = S^{-2} \circ S^{|\sigma(1)|} \sigma(\tilde{z}^{(1)}) = S\tilde{z}^{(2)}.$$

Since $(S\tilde{z}^{(1)})_0 = 13$ and $-2 \in c(13)$ we have $\Theta(S\tilde{z}^{(2)}) = S\tilde{z}^{(1)}$ and $\varphi(S\tilde{z}^{(2)}) = -2$ by Theorem 4.3. Hence, the first term of $\Gamma_c(S\tilde{z}^{(2)})$ equals $(13, -2)$. One can now easily

confirm that $\Theta^2(S\tilde{z}^{(2)}) = \Theta(S\tilde{z}^{(1)}) = \tilde{z}^{(2)}$ and $\varphi(S\tilde{z}^{(1)}) = 1$. We obtain $(23, 1)$ as second term of $\Gamma_c(S\tilde{z}^{(2)})$. Therefore, the infinite walk of the coding graph that corresponds to $\Gamma_c(S\tilde{z}^{(2)})$ starts at 31 (since $(S\tilde{z}^{(2)})_0 = 31$) and leads via 13 to 23.

Since $\Theta^2(S\tilde{z}^{(2)}) = \tilde{z}^{(2)} \in \tilde{P}$ we can use Proposition 5.5 and Lemma 4.4 to finish our calculation. By observing the cyclic structure of $\sigma|_{\tilde{P}}$ we obtain $\Gamma_c(\tilde{z}^{(2)}) = ((21, 0), (23, 0))^\infty$ and, finally,

$$\Gamma_c(S\tilde{z}^{(2)}) = (13, -2), (23, 1), ((21, 0), (23, 0))^\infty.$$

The other cycle of $G(c)$ that corresponds to a cycle of $\sigma|_{\tilde{P}}$ also has period 2 and is labelled by $(33, 0)$ and $(31, 0)$. In particular, we have

$$\Gamma_c(\tilde{z}^{(3)}) = ((33, 0), (31, 0))^\infty \quad \text{and} \quad \Gamma_c(\tilde{z}^{(4)}) = ((31, 0), (33, 0))^\infty.$$

Each of the latter codes is traceable and, hence, unique. We can use (8) to obtain a representation for the initial elements. For example, $\tilde{z}^{(3)}$ can be represented as limit

$$\tilde{z}^{(3)} = \bigcap_{n \geq 1} [A^{(n)} \cdot B^{(n)}] \quad \text{with} \quad A^{(n)} = \sigma^n(3), \quad B^{(n)} = \begin{cases} \sigma^n(1) & n \equiv 0 \pmod{2}, \\ \sigma^n(3) & n \equiv 1 \pmod{2}. \end{cases}$$

To analyse the traceability of c we use Theorem 6.1 and look at the cyclic structure of $\sigma|_L$ and $\sigma|_R$. We see that $\langle u^{(2)} \rangle$ is the only cycle that does not satisfy the conditions of the theorem. Hence, c is not traceable and the only not traceable purely periodic tail that occurs in $G(c)^\infty$ is $((13, 1))^\infty$. Indeed, using (7), we obtain $C_n = [A^{(n)} \cdot B^{(n)}]$ with $B^{(n)} = 3$ for all $n \geq 1$ (the sequence $(A^{(n)})_{n \geq 1}$ has unbounded length). This confirms that $((13, 1))^\infty$ is not right-traceable.

By Theorem 6.2, Γ_c is not injective since $\pi^{(l)}(\tilde{z}^{(3)}) = \pi^{(l)}(\tilde{z}^{(4)}) = u^{(2)}$. Hence,

$$\Gamma_c(S^{-1}\tilde{z}^{(3)}) = \Gamma_c(S^{-1}\tilde{z}^{(4)}) = ((13, 1))^\infty.$$

Due to Proposition 6.9, each not traceable code has the same periodic tail $((13, 1))^\infty$ and is produced by elements $S^{-m}\tilde{z}^{(3)}$ and $S^{-m}\tilde{z}^{(4)}$ with $m \geq 1$.

c is not continuous. We have already calculated $\Gamma_c(\tilde{z}^{(2)})$ and $\Gamma_c(S\tilde{z}^{(2)})$. Now consider $\Gamma_c(S^2\tilde{z}^{(2)})$. It is not difficult to see that

$$\Gamma_c(S^2\tilde{z}^{(2)}) = (21, 2), ((23, 0), (21, 0))^\infty.$$

Thus, $\Gamma_c(\tilde{z}^{(2)}) < \Gamma_c(S^2\tilde{z}^{(2)}) < \Gamma_c(S\tilde{z}^{(2)})$ which shows that the partial ordering $<$ does not respect the ordering induced by the shift map here.

Example 8.2. Let $\sigma : 1 \mapsto 32, 2 \mapsto 1, 3 \mapsto 13$. The substitution possesses the left-periodic sequences $L = \{u^{(1)}, u^{(2)}, u^{(3)}\}$ and the right-periodic sequences $R = \{v^{(1)}, v^{(2)}\}$ with

$$\begin{aligned} u^{(1)} &= \dots 313213131, & u^{(2)} &= \dots 213321332, & u^{(3)} &= \dots 321313213, \\ v^{(1)} &= 131321313 \dots, & v^{(2)} &= 321332131 \dots \end{aligned}$$

We have $\mathcal{L}_2(\sigma) = \{13, 21, 31, 32, 33\}$. Thus,

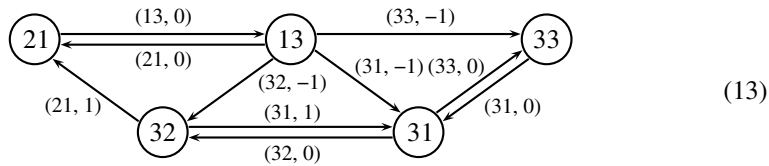
$$\tilde{P} = \{u^{(1)} \cdot v^{(2)}, u^{(2)} \cdot v^{(1)}, u^{(3)} \cdot v^{(1)}, u^{(3)} \cdot v^{(2)}\}.$$

Observe that σ induces two not admissible periodic points. Indeed, $\sigma^2(u^{(1)} \cdot v^{(1)}) = u^{(1)} \cdot v^{(1)}$ and $\sigma^2(u^{(2)} \cdot v^{(2)}) = u^{(2)} \cdot v^{(2)}$ but neither $u^{(1)} \cdot v^{(1)}$ nor $u^{(2)} \cdot v^{(2)}$ is contained in \tilde{P} . The following diagram shows the cyclic structure of $\sigma|_L$, $\sigma|_R$ and $\sigma|_{\tilde{P}}$ and the action of $\pi^{(l)}$ and $\pi^{(r)}$:

$$\begin{array}{ccccc}
 \sigma|_L & & \sigma|_{\tilde{P}} & & \sigma|_R \\
 \langle u^{(1)}, u^{(2)} \rangle & \xleftarrow{\pi^{(l)}} & \langle u^{(1)} \cdot v^{(2)}, u^{(2)} \cdot v^{(1)} \rangle & \xrightarrow{\pi^{(r)}} & \langle v^{(1)}, v^{(2)} \rangle \\
 \langle u^{(3)} \rangle & \xleftarrow{\pi^{(l)}} & \langle u^{(3)} \cdot v^{(1)}, u^{(3)} \cdot v^{(2)} \rangle & \xrightarrow{\pi^{(r)}} &
 \end{array}$$

There are four possible coding prescriptions, but we can easily see that none of them is traceable. In order to code all elements of $\tilde{\Omega}_\sigma$ in a traceable way we would have to consider higher powers of σ (since $\tilde{\Omega}_\sigma = \tilde{\Omega}_{\sigma^n}$ for all $n \geq 1$). In particular, one can easily verify that there exist traceable coding prescriptions with respect to σ^2 .

However, as $\pi^{(l)}$ acts injectively on the periodic points that form the first cycle of $\sigma|_{\tilde{P}}$ we conclude (using Theorem 6.2) that the coding prescription $c : 1 \mapsto \{0, 1\}$, $2 \mapsto \{0\}$, $3 \mapsto \{-1, 0\}$ yields an injective coding. The corresponding coding graph $G(c)$ is depicted in (13). Note that the choice of c maximizes the number of edges here.



Our coding prescription c is continuous. We quickly demonstrate the statement of Theorem 7.3. Consider the periodic point $\tilde{z} := u^{(1)} \cdot v^{(2)}$. Using Proposition 5.5 and the coding graph (13) we see that $\Gamma_c(\tilde{z}) = ((21, 0), (13, 0))^\infty$. The sequence is minimal since neither the vertex 21 nor the vertex 13 has an incoming edge that is smaller than (21, 0) and (13, 0), respectively, with respect to $<$. We easily obtain

$$T\Gamma_c(\tilde{z}) = \Gamma_c(S\tilde{z}) = (21, 1), ((13, 0), (21, 0))^\infty,$$

$$T^2\Gamma_c(\tilde{z}) = \Gamma_c(S^2\tilde{z}) = (13, 0), (32, -1), (21, 1), ((13, 0), (21, 0))^\infty,$$

$$T^3\Gamma_c(\tilde{z}) = \Gamma_c(S^3\tilde{z}) = (31, -1), (32, 0), (21, 1), ((13, 0), (21, 0))^\infty,$$

$$T^4\Gamma_c(\tilde{z}) = \Gamma_c(S^4\tilde{z}) = (31, 0), (32, 0), (21, 1), ((13, 0), (21, 0))^\infty,$$

$$T^5\Gamma_c(\tilde{z}) = \Gamma_c(S^5\tilde{z}) = (13, 0), (21, 0), (13, 0), (32, -1), (21, 1), ((13, 0), (21, 0))^\infty,$$

and so on.

Example 8.3. Consider the substitution $\sigma : 1 \mapsto 121, 2 \mapsto 3, 3 \mapsto 1$. σ provides four different coding prescriptions c_1, c_2, c_3, c_4 given by

$$\begin{aligned}
 c_1(1) &= \{-2, -1, 0\}, & c_2(1) &= \{-1, 0, 1\}, & c_3(1) &= \{0, 1, 2\}, & c_4(1) &= \{-2, 0, 2\}, \\
 c_1(2) &= c_1(3) = c_2(2) = c_2(3) = c_3(2) = c_3(3) = c_4(2) = c_4(3) = \{0\}.
 \end{aligned}$$

The only periodic point of $\tilde{\Omega}_\sigma$ is

$$\tilde{z} := \dots 2111213121 \cdot 1213121112 \dots$$

Thus, c_1 , c_3 and c_4 are not traceable. However, all coding prescription induce a completely injective coding. The coding prescription c_4 is the only one that is not continuous.

We want to concentrate on this last coding prescription c_4 . For the induced coding we have

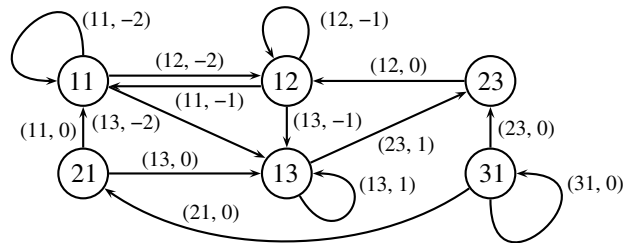
$$\Gamma_c(S\tilde{z}) = ((12, -2))^\infty$$

not left-traceable by Lemma 6.8. On the other hand,

$$\Gamma_c(S^2\tilde{z}) = (11, 2), ((11, 0))^\infty,$$

which is traceable. This shows that the equivalence in Corollary 7.6 does not hold in general for not continuous coding prescriptions.

Example 8.4. Consider $\sigma : 1 \mapsto 112, 2 \mapsto 3, 3 \mapsto 13$ and the (continuous) coding prescription $c : 1 \mapsto \{-2, -1, 0\}, 2 \mapsto \{0\}, 3 \mapsto \{0, 1\}$. The coding graph for this example is given by



and we see that it is not strongly connected. In fact, it consists of two strongly connected components. The vertex 21 is not contained in any of them. The (traceable) infinite walk $((31, 0))^\infty$ is both minimal and maximal.

Acknowledgement. The research was supported by the Brazilian FAPESP, Proc. 2009/07744-0.

REFERENCES

- [BS05] V. Berthé and A. Siegel. Tilings associated with beta-numeration and substitutions. *Integers* **5**(3:A2) (2005), 46 (electronic).
- [CS01a] V. Canterini and A. Siegel. Automate des préfixes-suffixes associé à une substitution primitive. *J. Théor. Nombres Bordeaux* **13**(2) (2001), 353–369.
- [CS01b] V. Canterini and A. Siegel. Geometric representation of substitutions of Pisot type. *Trans. Amer. Math. Soc.* **353**(12) (2001), 5121–5144 (electronic).
- [DHS99] F. Durand, B. Host and C. Skau. Substitutional dynamical systems, Bratteli diagrams and dimension groups. *Ergod. Th. & Dynam. Sys.* **19**(4) (1999), 953–993.
- [Fog02] N. P. Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics (Lecture Notes in Mathematics, 1794)*. Eds. V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. Springer, Berlin, 2002.
- [For97] A. H. Forrest. K -groups associated with substitution minimal systems. *Israel J. Math.* **98** (1997), 101–139.
- [HL86] T. Harju and M. Linna. On the periodicity of morphisms on free monoids. *RAIRO Inf. Théor. Appl.* **20** (1986), 47–54.

- [HZ01] C. Holton and L. Q. Zamboni. Directed graphs and substitutions. *Theory Comput. Syst.* **34**(6) (2001), 545–564.
- [LM95] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.
- [Mos92] B. Mossé. Puissances de mots et reconnaissabilité des points fixes d’une substitution. *Theoret. Comput. Sci.* **99**(2) (1992), 327–334.
- [Mos96] B. Mossé. Recognizability of substitutions and complexity of automatic sequences. *Bull. Soc. Math. France* **124**(2) (1996), 329–346.
- [Pan86] J.-J. Pansiot. Decidability of periodicity for infinite words. *RAIRO Inf. Théor. Appl.* **20** (1986), 43–46.
- [Que10] M. Queffélec. *Substitution Dynamical Systems: Spectral Analysis (Lecture Notes in Mathematics, 1294)*. 2nd edn. Springer, Dordrecht, 2010.
- [Ver81] A. M. Vershik. Uniform algebraic approximation of shift and multiplication operators. *Dokl. Akad. Nauk SSSR* **259**(3) (1981), 526–529.
- [VL92] A. M. Vershik and A. N. Livshits. Adic models of ergodic transformations, spectral theory, substitutions, and related topics. *Representation Theory and Dynamical Systems (Advances in Soviet Mathematics, 9)*. American Mathematical Society, Providence, RI, 1992, pp. 185–204.