Theory Comput. Systems **34**, 545–564 (2001)
DOI: 10.1007/s00224-001-1038-y

Theory of Computing Systems

Directed Graphs and Substitutions*

C. Holton[1] and L. Q. Zamboni[2]

[1]Department of Mathematics, University of California,
Berkeley, CA 94720-3840, USA
cholton@math.berkeley.edu

[2]Department of Mathematics, P.O. Box 311430, University of North Texas,
Denton, TX 76203-1430, USA
luca@unt.edu

Abstract. A substitution naturally determines a directed graph with an ordering of the edges incident at each vertex. We describe a simple method by which any primitive substitution can be modified (without materially changing the bi-infinite fixed points of the substitution) so that points in the substitution minimal shift are in bijective correspondence with one-sided infinite paths on its associated directed graph. Using this correspondence, we show that primitive substitutive sequences in the substitution minimal shift are precisely those sequences represented by eventually periodic paths. We use directed graphs to show that all measures of cylinders in a substitution minimal shift lie in a finite union of geometric sequences, confirming a conjecture of Boshernitzan. Our methods also yield sufficient conditions for a geometric realization of a primitive substitution to be "almost injective."

1. Introduction

Every substitution has an associated prefix automaton, a directed graph whose vertices are the letters of the substitution's alphabet and whose edges are labeled with the strict prefixes of the images of the letters under the substitution. In this paper we consider the directed graph obtained by reversing the directed edges in the prefix automaton. This graph is more natural from the point of view of iterated function systems.

* The second author was supported in part by NSF Grant INT-9726708 and a grant from the Texas Advanced Research Program.

Vershik introduced the notion of an adic transformation defined on the infinite path space of a Bratteli diagram as a means of studying operator algebras. Livshitz outlined a connection between substitution minimal systems and stationary adic transformations, the underlying spaces of which are essentially the infinite path spaces on the associated directed graphs: every substitution minimal shift is metrically isomorphic to a stationary adic transformation [24]. Forrest [18] and Durand et al. [15] proved that every stationary proper minimal Bratteli-Vershik system is topologically conjugate to either a substitution minimal system or a stationary odometer system.

In this paper we describe a simple but completely general way of modifying a primitive substitution so that the substitution minimal shift is homeomorphic to the space of one-sided infinite paths on its directed graph. The induced dynamical system on the infinite path space of the graph is just the continuous extension of the Vershik map. Other constructions of this kind exist, for example, [15], [18], and [19].

A primitive substitutive sequence is, by definition, an image under a morphism of a fixed point of a primitive substitution (see [13] and [21]). Directed graphs give a new way to characterize primitive substitutive sequences. As part of the characterization, we show that a point in the minimal shift arising from a primitive substitution is primitive substitutive if and only if it is represented by an eventually periodic path in the graph associated to the substitution. We also show that all primitive substitutive sequences in any minimal shift can be completely described by a single directed graph.

As another application of the directed graph construction, we establish the following conjecture of Boshernitzan: the measures of cylinders in a substitution minimal shift lie in a finite union of geometric sequences [6]. This is an unsurprising reflection of the self-similarity inherent in substitution dynamical systems.

Many substitutions admit geometric realizations equipped with piecewise translation mappings (see [20]). Ei and Ito gave a characterization of those substitutions on two letters whose geometric realization is an interval exchange [16]. In such cases the substitution minimal system is a symbolic representation for the interval exchange transformation. It is natural to ask which substitution dynamical systems admit an “almost injective” geometric realization in the sense of Queffélec [31]. In the last section we use the graph directed construction to obtain some partial results.

2. Substitutions

By an *alphabet* we mean a finite nonempty set of symbols. We refer to the members of an alphabet as *letters*. A *word* of length k in the alphabet $\mathcal{A}$ is an expression $u = u_0u_1 \cdots u_{k-1}$ where each u_j is a letter of $\mathcal{A}$. We write $|u|$ for the length of a word u. It is convenient to define the *empty word* ε, having length 0 and satisfying the concatenation rule $\varepsilon u = u = u\varepsilon$ for every word u. Let $\mathcal{A}^+$ be the set of all words of positive length in the alphabet $\mathcal{A}$ and set $\mathcal{A}^* = \mathcal{A}^+ \cup \{\varepsilon\}$.

A *morphism* defined on alphabet $\mathcal{A}$ is a map $\pi\colon \mathcal{A} \to \mathcal{B}^*$, where $\mathcal{B}$ is an alphabet. The length of π is defined to be $|\pi| = \max\{|\pi(a)|\colon a \in \mathcal{A}\}$. We say π is a *letter-to-letter morphism* if $\pi(\mathcal{A}) \subset \mathcal{B}$, and π is called *nonerasing* if $\varepsilon \notin \pi(\mathcal{A})$. A morphism $\pi\colon \mathcal{A} \to \mathcal{B}^*$ determines by concatenation a map $\mathcal{A}^* \to \mathcal{B}^*$, also denoted π. A *substitution* is a nonerasing morphism $\mathcal{A} \to \mathcal{A}^*$. A substitution ζ on alphabet $\mathcal{A}$ determines a map

$\mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ which we also denote by ζ. With this slight abuse of notation we define ζ^n for $n \geq 0$ on $\mathcal{A}$, $\mathcal{A}^+$ and $\mathcal{A}^{\mathbb{Z}}$ in the obvious way.

Associated to a substitution ζ on $\mathcal{A}$ is the $\mathcal{A} \times \mathcal{A}$ matrix M_ζ with abth entry equal to the number of occurrences of the letter a in the word $\zeta(b)$. The substitution ζ is *primitive* if there exists a positive integer n such that for each letter $a \in \mathcal{A}$ the word $\zeta^n(a)$ contains every letter of $\mathcal{A}$. Equivalently, ζ is primitive if and only if M_ζ is a primitive matrix, i.e., if some positive power of M_ζ has all entries strictly positive. If ζ is primitive, then

$$\forall a \in \mathcal{A}, \qquad |\zeta^m(a)| \to \infty \qquad \text{as} \quad m \to \infty. \tag{1}$$

A sequence $\omega \in \mathcal{A}^{\mathbb{Z}}$ is a *periodic point* of ζ if for some positive integer n we have $\zeta^n(\omega) = \omega$. Every primitive substitution has a periodic point. Denote by $\mathrm{Per}(\zeta)$ the set of periodic points of ζ.

For $w \in \mathcal{A}^+$ and $0 \leq m \leq n < |w|$ we write $w_{[m,n]}$ for the word $w_m w_{m+1} \cdots w_n$ (such words are called *factors* of w). Define $w_{[m,n)}$, $w_{(m,n]}$ and $w_{(m,n)}$ analogously, with the convention that $w_\emptyset = \varepsilon = \zeta(\varepsilon)$. Similarly, if $\omega \in \mathcal{A}^{\mathbb{Z}}$ and $m < n$, then we set $\omega_{[m,n]} = \omega_m \omega_{m+1} \cdots \omega_n$ and use the obvious interpretations for $\omega_{[m,n)}$, $\omega_{(m,n]}$ and $\omega_{(m,n)}$.

The topology of $\mathcal{A}^{\mathbb{Z}}$ is metrizable as the countable product of the discrete space $\mathcal{A}$. Specifically, we define a metric d on $\mathcal{A}^{\mathbb{Z}}$ by setting $d(\omega, \upsilon) = e^{-N}$ if N is the greatest integer for which $\omega_{(-N,N)} = \upsilon_{(-N,N)}$.

The *language* of a substitution ζ is the set $\mathcal{L}(\zeta)$ of all factors of the words $\zeta^m(a)$, $m \in \mathbb{N}$ and $a \in \mathcal{A}$. The subshift of $\mathcal{A}^{\mathbb{Z}}$ spanned by the language $\mathcal{L}(\zeta)$ is

$$X(\zeta) = \{\omega \in \mathcal{A}^{\mathbb{Z}} \colon \omega_{[-j,j]} \in \mathcal{L}(\zeta) \text{ for all } j \in \mathbb{N}\}.$$

It follows from (1) that if ζ is primitive, then $\mathcal{L}(\zeta)$ is an infinite set and $X(\zeta)$ is nonempty. The *cylinder* determined by a word w is $[w] = \{\omega \in X(\zeta) \colon \omega_{[0,|w|)} = w\}$. The *shift map* T on $X(\zeta)$ is defined by $(T\omega)_j = \omega_{j+1}$. It is easy to see that $X(\zeta)$ is a closed subset of $\mathcal{A}^{\mathbb{Z}}$, T is a homeomorphism $X(\zeta) \to X(\zeta)$, and if ζ is primitive, then the dynamical system $(X(\zeta), T)$ is minimal, i.e., every point has a dense orbit under the shift map (see [31]).

3. Directed Graphs Associated with a Substitution

A *directed graph* G consists of a finite nonempty set of vertices $\mathcal{V} = \mathcal{V}(G)$ and a finite set $\mathcal{E} = \mathcal{E}(G)$ of edges together with maps $\mathtt{i}, \mathtt{t} \colon \mathcal{E} \to \mathcal{V}$. A *walk* or *path* on G is a finite or infinite sequence of edges $e_l \in \mathcal{E}$ such that $\mathtt{i}(e_{l+1}) = \mathtt{t}(e_l)$ for all relevant indices l. A finite walk $(e_l)_{l=1}^k$ starts at its *initial vertex* $\mathtt{i}(e_1)$ and ends at its *terminal vertex* $\mathtt{t}(e_k)$. We say that G is *strongly connected* if for every ordered pair of vertices (a, b) there is a finite walk on G starting at a and ending at b. A strongly connected directed graph is said to be *aperiodic* if there are two walks of relatively prime lengths with the same initial vertex and the same terminal vertex.

The *adjacency matrix* of G is the $\mathcal{V} \times \mathcal{V}$ matrix $M = M(G)$ with abth entry M_{ab} equal to the number of edges starting at vertex a and ending at vertex b. It is easy to see that $(M^n)_{ab}$ is the number of paths of length n starting at a and ending at b. Thus, G is strongly connected if and only if M is irreducible, and G is aperiodic if and only if M is primitive. (Recall that a nonnegative matrix M is irreducible if for all indices a, b there is a positive integer n such that $(M^n)_{ab} > 0$.)

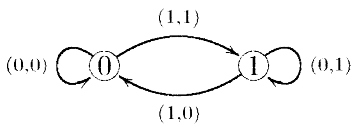

Fig. 1. Directed graph associated to the substitution in Example 3.1.

Denote by G^∞ the space of (one-sided) infinite walks on G, viewed as a subspace of the countable product of the discrete edge set. With this topology G^∞ is compact.

Let ζ be a substitution on an alphabet $\mathcal{A}$. We associate to ζ a directed graph $G = G(\zeta)$ with vertex set $\mathcal{V}(G) = \mathcal{A}$ and the edge set $\mathcal{E}(G) = \{(j, a)\colon\ a \in \mathcal{A} \text{ and } 0 \le j < |\zeta(a)|\}$, where $\mathtt{i}(j, a) = \zeta(a)_j$ and $\mathtt{t}(j, a) = a$. These graphs were independently studied by Canterini and Siegel in [8]. The adjacency matrix $M(G)$ is equal to the incidence matrix M_ζ. Thus ζ is primitive if and only if $G(\zeta)$ is aperiodic.

Example 3.1. The Thue–Morse substitution has alphabet $\{0, 1\}$ and $\zeta(0) = 01$, $\zeta(1) = 10$. The associated graph G is shown in Figure 1.

Our construction naturally determines a partial ordering $\prec$ on $\mathcal{E}$ in which edges are comparable if and only if they have the same terminal vertex. Specifically, for $(j_1, a_1), (j_2, a_2) \in \mathcal{E}$,

$$(j_1, a_1) \prec (j_2, a_2) \quad\Longleftrightarrow\quad a_1 = a_2 \text{ and } j_1 < j_2. \tag{2}$$

Conversely, a directed graph G having at least one edge ending at each vertex, together with an ordering $\prec$ on the set of edges ending at each vertex, determines a substitution ζ_G on alphabet $\mathcal{V}(G)$ given by $\zeta_G(a) = \mathtt{i}(e_{1,a})\mathtt{i}(e_{2,a})\cdots\mathtt{i}(e_{k_a,a})$, where $e_{1,a} \prec e_{2,a} \prec \cdots \prec e_{k_a,a}$ are the edges ending at $a \in \mathcal{V}(G)$. It is readily verified that $\zeta_{G(\zeta)} = \zeta$ and $G(\zeta_G) = G$.

Suppose ζ is a primitive substitution on alphabet $\mathcal{A}$ and G is the associated directed graph with partial ordering $\prec$ on the edge set given by (2). We also denote by $\prec$ the partial ordering on G^∞ defined by

$$(e_j)_{j=1}^\infty \prec (f_j)_{j=1}^\infty \quad\Longleftrightarrow\quad \exists k[(e_k \prec f_k) \text{ and } \forall j > k(e_j = f_j)].$$

An element of G^∞ is *minimal* (*maximal*) if it is minimal (maximal) with respect to this partial ordering. Note that for each vertex a and for each positive integer n there is a unique walk of length n ending at a and consisting entirely of minimal edges (ditto for maximal). Consequently,

Proposition 3.2. *Minimal and maximal elements always exist and are periodic* (*as sequences of edges*), *and there are, at most,* $|\mathcal{A}|$ *of each.*

Every nonmaximal element has a unique successor, and we write S for the successor map $G^\infty\backslash\{\text{maximal walks}\} \to G^\infty$. That is, if $(j_l, a_l)_{l=1}^\infty \in G^\infty$ is not maximal and k is the least index for which (j_k, a_k) is not a maximal edge, then

$$S((j_l, a_l)_{l=1}^{\infty}) = (0, \zeta^{k-2}(a)_0)(0, \zeta^{k-3}(a)_0) \\ \cdots (0, a)(j_k + 1, a_k)(j_{k+1}, a_{k+1})(j_{k+2}, a_{k+2}) \cdots, \tag{3}$$

where $a = \zeta(a_k)_{j_k+1}$.

4. Words and Walks

Fix a primitive substitution ζ on an alphabet $\mathcal{A}$ such that $X = X(\zeta)$ is infinite. Our first order of business is to give a method for coding points of X by infinite walks on $G = G(\zeta)$. The coding is suggested by the following lemma which states that all primitive substitutions have the unique admissable decoding (UAD) property of [25].

Theorem 4.1. *Let $\omega \in X$. Then there is a unique pair $(s(\omega), m(\omega)) \in X \times \mathbb{Z}$ such that $\omega = T^{m(\omega)}\zeta(s(\omega))$ and $0 \leq m(\omega) < |\zeta(s(\omega)_0)|$. Moreover, the map $\omega \mapsto (s(\omega), m(\omega))$ is continuous.*

The idea for the proof is probably due to Martin [26]. The result can also be inferred from work of Mossé in [30]. See [15] for more details.

Corollary 4.2. *The map $\zeta\colon X \to \zeta(X)$ is one-to-one.*

We record a useful fact whose proof is straightforward.

Lemma 4.3. *If $(j_l, a_l)_{l=1}^{k}$ is a walk on G and $\upsilon \in X$ with $\upsilon_0 = a_k$, then*

$$(T^{j_1} \circ \zeta \circ T^{j_2} \circ \zeta \circ \cdots \circ T^{j_k} \circ \zeta)(\upsilon) = T^{m_k}\zeta^k(\upsilon),$$

where $m_k = \sum_{l=1}^{k} |\zeta^{l-1}(\zeta(a_l)_{[0,j_l)})| < |\zeta^l(a_l)|$.

We define a map $H\colon X \to G^{\infty}$, which we call the *coding map* of ζ, by

$$H(\omega) = (m(s^{l-1}(\omega)), s^l(\omega)_0)_{l=1}^{\infty}.$$

Thus $H(\omega)$ starts at vertex a if and only if $\omega \in [a]$.

Proposition 4.4. *The coding map $H\colon X \to G^{\infty}$ is a continuous surjection satisfying*:

(a) $H(\zeta(\omega)) = (0, \zeta(\omega)_0)H(\omega)$.
(b) *$H \circ T = S \circ H$ whenever the latter is defined.*
(c) *$H(\omega)$ is minimal if and only if $\omega \in \mathrm{Per}(\zeta)$.*
(d) *$H(\omega)$ is maximal if and only if $T\omega \in \mathrm{Per}(\zeta)$.*
(e) *The infinite walk $H(\omega)$ begins in $(j_1, a_1) \cdots (j_k, a_k)$ if and only if*

$$\omega \in (T^{j_1} \circ \zeta \circ T^{j_2} \circ \zeta \circ \cdots \circ T^{j_k} \circ \zeta)([a_k]).$$

Proof. The image of H is dense in G^∞, for if $(j_1, a_1)(j_2, a_2)\cdots(j_k, a_k)$ is a finite walk on G and $\upsilon \in X$ is any point with $\upsilon_0 = a_k$, then $\omega := (T^{j_1} \circ \zeta \circ T^{j_2} \circ \zeta \circ \cdots \circ T^{j_k} \circ \zeta)(\upsilon)$ is a point of X and $H(\omega)$ begins in $(j_1, a_1)(j_2, a_2)\cdots(j_k, a_k)$. Continuity follows from Theorem 4.1 and surjectivity from compactness of X.

Assertions (a) and (e) follow immediately from the definition of H. To prove (b) suppose $\omega \in X$ and $H(\omega) = (j_l, a_l)_{l=1}^{\infty}$ is not maximal. Let k be the least index for which (j_k, a_k) is not maximal. Let υ be the unique point of X for which $\upsilon_0 = a_k$ and

$$\omega = (T^{j_1} \circ \zeta \circ T^{j_2} \circ \zeta \circ \cdots \circ T^{j_k} \circ \zeta)(\upsilon).$$

For any $\varpi \in X$ we have

$$(T \circ T^{|\zeta(\varpi_0)|-1} \circ \zeta)(\varpi) = \zeta(T\varpi), \tag{4}$$

and repeated application of this identity yields

$$\begin{aligned} T\omega &= (T \circ T^{j_1} \circ \zeta \circ T^{j_2} \circ \zeta \circ \cdots \circ T^{j_k} \circ \zeta)(\upsilon) \\ &= (\zeta^{k-1} \circ T^{j_k+1} \circ \zeta)(\upsilon). \end{aligned}$$

Putting $a = \zeta(a_k)_{j_k+1}$ we see from (a) that

$$\begin{aligned} H(T\omega) = {}& (0, \zeta^{k-2}(a)_0)(0, \zeta^{k-3}(a)_0) \\ & \cdots (0, \zeta(a)_0)(0, a)(j_k + 1, a_k)(j_{k+1}, a_{k+1})(j_{k+2}, a_{k+2})\cdots, \end{aligned}$$

which is, according to (3), the successor of $H(\omega)$.

To prove (c) we observe that $H(\omega)$ is minimal if and only if ω is in the image of ζ^k for all k. Assertion (c) now follows from the equality

$$\bigcap_{k\geq 0} \zeta^k(X) = \mathrm{Per}(\zeta). \tag{5}$$

To prove (d) we note that by (4) if $H(\omega)$ is maximal, then $T\omega$ is in the image of ζ^k for every positive integer k; by (5), $T\omega$ must be in $\mathrm{Per}(\zeta)$. Conversely, if $H(\omega)$ is not maximal, then (b) implies that $H(T\omega)$ is a successor (the successor of $H(\omega)$) and hence not minimal, whence it follows from (c) that $T\omega$ is not in $\mathrm{Per}(\zeta)$. $\square$

We wish to extend S to a continuous map on all of G^∞ so that H becomes a conjugacy. However, the map H need not be one-to-one. To see how this difficulty can arise, suppose ζ has a pair of distinct periodic points ω, υ with $\omega_{[0,\infty)} = \upsilon_{[0,\infty)}$. (The Thue–Morse substitution has two such pairs.) A substitution acts as a permutation on its set of periodic points, so there is a positive integer k such that $\zeta^k(\omega) = \omega$ and $\zeta^k(\upsilon) = \upsilon$. This implies for all $l > 0$,

$$s^l(\omega) = \zeta^{(k-l) \bmod k}(\omega) \quad \text{and} \quad m^{l-1}(\omega) = 0$$

and the analogous statement for υ. Thus

$$s^l(\omega)_0 = \zeta^{(k-l) \bmod k}(\omega)_0 = \zeta^{(k-l) \bmod k}(\upsilon)_0 = s^l(\upsilon)_0,$$

whence $H(\omega) = H(\upsilon)$.

A pair as in the preceeding paragraph exists precisely when there are more maximal elements than minimal elements in G^∞. This is the only thing to check; it can be shown that H is bijective if and only if

$$|\{\text{minimal elements of } G^\infty\}| = |\{\text{maximal elements of } G^\infty\}|.$$

To extend the successor map S we modify the substitution so that the resulting coding map H is one-to-one. Let $\mathcal{A}^{[3]} = \mathcal{A}^{[3]}(\zeta)$ be the subset of $\mathcal{L}(\zeta)$ consisting of the words of length 3. We define a substitution $\zeta_{[3]}$ on $\mathcal{A}^{[3]}$ by setting for $(abc) \in \mathcal{A}^{[3]}$,

$$\zeta((abc))_j = (\zeta(abc)_{[|\zeta(a)|+j-1, |\zeta(a)|+j+1]}), \qquad 0 \le j < |\zeta(b)|.$$

Note $|\zeta_{[3]}((abc))| = |\zeta(b)|$. Write $T_{[3]}$ for the shift map on $X(\zeta_{[3]})$. For $\omega \in X(\zeta)$ denote by $\omega^{[3]}$ the sequence in the letters of $\mathcal{A}^{[3]}$ with jth entry $(\omega^{[3]})_j = (\omega_{j-1}\omega_j\omega_{j+1})$. It is easy to see that

Proposition 4.5. *With the above notation, $\zeta_{[3]}$ is primitive and $\omega \mapsto \omega^{[3]}$ is a conjugacy $X(\zeta) \to X(\zeta_{[3]})$. Moreover, $(\zeta(\omega))^{[3]} = \zeta_{[3]}(\omega^{[3]})$ and $[(abc)] = (T[abc])^{[3]}$.*

Denote by $H_{[3]}\colon X(\zeta_{[3]}) \to G(\zeta_{[3]})^\infty$ the coding map of $\zeta_{[3]}$. Let $(j_l, (a_l b_l c_l))_{l=1}^\infty \in G(\zeta_{[3]})^\infty$ and set $m_k = \sum_{l=1}^k |\zeta^{l-1}(\zeta(a_l)_{[0,j_l)})|$. By Lemma 4.3 and (e) of Proposition 4.4, if $H_{[3]}(\omega^{[3]}) = (j_l, (a_l b_l c_l))_{l=1}^\infty$, then $\omega^{[3]} \in T_{[3]}^{m_k}\zeta_{[3]}^k([(a_k b_k c_k)])$, so, by Proposition 4.5, $\omega \in T^{m_k}\zeta^k(T[a_k b_k c_k])$. Any $\upsilon \in T^{m_k}\zeta^k(T[a_k b_k c_k])$ must satisfy $\upsilon_{[-|\zeta^k(a_k)|-m_k, |\zeta^k(b_k)|-m_k+|\zeta^k(c_k)|]} = \zeta^k(a_k b_k c_k)$, and therefore

$$\operatorname{diam} T^{m_k}\zeta^k(T[a_k b_k c_k]) \le \exp(-\min(|\zeta^k(a_k)|, |\zeta^k(c_k)|)),$$

where diam means diameter in the metric d. This together with (e) of Proposition 4.4 implies

$$H_{[3]}^{-1}((j_l, a_l b_l c_l)_{l=1}^\infty) \subset \bigcap_{k=1}^\infty (T^{j_1} \circ \zeta_{[3]} \circ \cdots \circ T^{j_k} \circ \zeta_{[3]})([(a_k b_k c_k)]) = \{\omega^{[3]}\}.$$

Thus, $H_{[3]}$ is seen to be a homeomorphism and the successor map $S_{[3]}$, initially defined on the nonmaximal members of $G(\zeta_{[3]})^\infty$, may be extended to a continuous homeomorphism (also denoted $S_{[3]}$) on all of $G(\zeta_{[3]})^\infty$. In this way, $S_{[3]}$ sends each maximal path in $G(\zeta_{[3]})^\infty$ onto a minimal path. We summarize this discussion as follows.

Proposition 4.6. *If ζ is a primitive substitution, then there exists a primitive substitution τ and a conjugacy $\varphi\colon X(\zeta) \to X(\tau)$ such that $\varphi \circ \zeta = \tau \circ \varphi$ and the coding map $H\colon X(\tau) \to G(\tau)^\infty$ is a conjugacy.*

For the remainder of the paper we assume that ζ is a primitive substitution on the alphabet $\mathcal{A}$ and that $X = X(\zeta)$ is an infinite set conjugate to $G = G(\zeta)$ via the coding map H.

5. Characterizing Left Special Sequences

A point $\omega \in X$ is called *left-special* if there exists $\upsilon \in X$ such that $\omega_{[0,\infty)} = \upsilon_{[0,\infty)}$ and $\omega_{-1} \neq \upsilon_{-1}$. Queffélec showed in [31] that the set of left-special points of X is finite.

Proposition 5.1. *If $\omega \in X$ is left-special, then $H(\omega)$ is eventually periodic.*

Proof. Let $\mathcal{S}$ be the set of ordered pairs $(\omega, \upsilon) \in X \times X$ such that $\omega_{[0,\infty)} = \upsilon_{[0,\infty)}$ and $\omega_{-1} \neq \upsilon_{-1}$. It follows from Queffélec's result that $\mathcal{S}$ is finite. For $(\omega, \upsilon) \in \mathcal{S}$ we have $\zeta(\omega)_{[0,\infty)} = \zeta(\upsilon)_{[0,\infty)}$ and $\zeta(\omega) \neq \zeta(\upsilon)$ by Theorem 4.1, so there must be a nonnegative integer $L(\omega, \upsilon)$ such that

$$(T^{-L(\omega,\upsilon)}\zeta(\omega), T^{-L(\omega,\upsilon)}\zeta(\upsilon)) \in \mathcal{S}.$$

We will show that each point of $\mathcal{S}$ has at most one preimage under the map $\mathcal{S} \to \mathcal{S}$ given by

$$(\omega, \upsilon) \mapsto (T^{-L(\omega,\upsilon)}\zeta(\omega), T^{-L(\omega,\upsilon)}\zeta(\upsilon)),$$

and hence this map is a permutation. For suppose (ω', υ') is a preimage of (ω, υ). Let m be the least nonnegative integer such that $|\zeta(\omega'_{[-m,0)})| \geq L(\omega', \upsilon')$. Then

$$T^{|\zeta(\omega'_{[-m,0)})|-L(\omega',\upsilon')} \circ \zeta(T^{-m}\omega') = T^{-L(\omega',\upsilon')} \circ \zeta(\omega') = \omega$$

and $0 \leq |\zeta(\omega'_{[0,m)})| - L < |\zeta(\omega'_{-m})|$. By the uniqueness assertion of Theorem 4.1 we have $s(\omega) = T^{-m}(\omega')$. Likewise, υ' is in the shift orbit of $s(\upsilon)$. Since X has no periodic points there can be only one pair of integers (j, k) for which $(T^j(s(\omega)), T^k(s(\upsilon))) \in \mathcal{S}$, thus the preimage (ω', υ') is unique.

We now fix a left-special $\omega \in X$ and let υ be such that $(\omega, \upsilon) \in \mathcal{S}$. It follows that for some integers $n > 0$ and $N \geq 0$ we have

$$(T^{-N}\zeta^n(\omega), T^{-N}\zeta^n(\upsilon)) = (\omega, \upsilon).$$

From (a) and (b) of Proposition 4.4 we obtain

$$S^{-N}((0, \zeta^n(\omega)_0)(0, \zeta^{n-1}(\omega)_0) \cdots (0, \zeta(\omega)_0)H(\omega)) = H(\omega). \tag{6}$$

If $H(\omega)$ is in the orbit of a minimal or maximal element, then $H(\omega)$ is eventually periodic by Proposition 3.2. If $H(\omega)$ is not in the S-orbit of a minimal or maximal element of G^∞, then $H(\omega)$ and $S^N(H(\omega))$ differ in only finitely many coordinates; this together with (6) shows that $H(\omega)$ is an eventually periodic sequence of edges. $\square$

Remark 5.2. Related to $\mathcal{S}$ is the number $N(\zeta)$ defined as follows. Let $\approx$ be the equivalence relation on X given by

$$\omega \approx \upsilon \iff \exists m \exists n\colon\ \omega_{[m,\infty)} = \upsilon_{[n,\infty)}.$$

The $\approx$ equivalence class $\langle\omega\rangle$ of ω is a union of finitely many orbits, namely those of ω and all points υ such that $(T^m\omega, \upsilon) \in \mathcal{S}$ for some m. Define

$$N(\zeta) := \sum_{\langle\omega\rangle} (\text{\# of orbits in } \langle\omega\rangle - 1)\,.$$

This sum is finite as only finitely many equivalence classes contain more than one orbit, but no sharp bound is known. We conjecture that $N(\zeta) \leq |\mathcal{A}|(|\mathcal{A}| - 1)$. The conjecture holds when $|\mathcal{A}| = 2$, and in this case there is a simple algorithm for finding all pairs of $\mathcal{S}$. The question seems to be open for $|\mathcal{A}| \geq 3$.

6. Primitive Substitutive Sequences

In [13] Durand adapted the notion of automatic sequences to the context of primitive substitutions: a one-sided infinite sequence ω is called *primitive substitutive* if and only if it is the image of a fixed point of a primitive substitution under a letter-to-letter morphism. Durand [13] showed that primitive substitutive sequences are precisely those minimal sequences which admit only finitely many derived sequences on prefixes (see Theorem 6.5 below). This description was generalized by Holton and Zamboni in [21] to obtain a characterization of *primitive substitutive subshifts*. In this section we extend the notion of primitive substitutive to bi-infinite sequences in the natural way, and derive an alternative characterization in terms of the associated directed graph. A two-sided infinite sequence is called *primitive substitutive* if it is the image of a fixed point of a primitive substitution under a letter-to-letter morphism. The main result of this section is

Theorem 6.1. *Let ζ be a primitive substitution, and let π be a morphism taking $X(\zeta)$ into an infinite minimal shift space Y. Then, for every $\omega \in X(\zeta)$, the sequence $\pi(\omega)$ is primitive substitutive if and only if $H(\omega)$ is eventually periodic.*

We begin with a few technical remarks. Durand showed that the hypothesis that the morphism be letter-to-letter in the definition of primitive substitutive may be relaxed (see Proposition 3.1 of [13]) to arbitrary nonerasing morphisms. Durand's proof extends immediately to the bi-infinite case:

Lemma 6.2. *Let $\omega \in \mathcal{A}^{\mathbb{Z}}$ be primitive substitutive. If π is a morphism such that $|\pi(a)| \geq 1$ for some letter a occurring in ω, then $\cdots\pi(\omega_{-2})\pi(\omega_{-1})\cdot\pi(\omega_0)\pi(\omega_1)\cdots$ is primitive substitutive.*

The next lemma characterizes the sequences in X represented by eventually periodic paths in G^∞. The proof is straightforward and is left to the reader.

Lemma 6.3. *Let $\omega \in X$. Then $H(\omega)$ is eventually periodic with period n if and only if $\zeta^n(\omega)$ and ω lie in the same T-orbit.*

Our proof of Theorem 6.1 requires several lemmas. We begin with some terminology adapted from [13]. Let Y be a minimal shift space and let u be a factor of some (hence every) sequence of Y. A *return word* to u in Y is a nonempty word w such that

- wu is a factor of some sequence of Y,
- $(wu)_{[j,j+|u|)} = u$ if and only if $j \in \{0, |w|\}$.

Denote by $\mathcal{R}_u = \mathcal{R}_u(Y)$ the set of return words to u in Y. It follows from minimality of Y that $\mathcal{R}_u$ is finite. Durand showed that primitive substitutive shifts are *linearly recurrent*:

Lemma 6.4 [13, Theorem 4.5]. *If Y is infinite and contains a primitive substitutive sequence, then there exists a positive constant C such that for all u, every return word $w \in \mathcal{R}_u(Y)$ satisfies $C^{-1}|u| < |w| < C|u|$.*

Fix, for each u, a bijection $\Theta_u\colon \mathcal{R}_u \to \{1, 2, \ldots, |\mathcal{R}_u|\}$. If $\upsilon \in Y$ with $\upsilon_{[0,|u|)} = u$, then υ can be written uniquely as a concatenation:

$$\upsilon = \cdots w_{-2}w_{-1} \cdot w_0 w_1 w_2 \cdots, \qquad \text{each} \quad w_j \in \mathcal{R}_u,$$

and we obtain the *derived sequence* of υ with respect to u:

$$\mathcal{D}_u(\upsilon) = \cdots \Theta_u(w_{-2})\Theta_u(w_{-1}) \cdot \Theta_u(w_0)\Theta_u(w_1) \cdots.$$

Derived sequences are unique up to permutation of their alphabets, and Durand's result holds for bi-infinite sequences:

Theorem 6.5 [13, Theorem 2.5]. *Let Y be a minimal shift space. A sequence $\upsilon \in Y$ is primitive substitutive if and only if $\{\mathcal{D}_{\upsilon_{[0,n)}}(\upsilon)\colon n > 0\}$ is finite.*

Remark 6.6. In case $\upsilon \in Y$ has only finitely many derived sequences, Durand's proof gives a method for constructing a primitive substitution ζ fixing a point $\omega \in X$ and a letter-to-letter morphism π mapping X onto Y with $\pi(\omega) = \upsilon$. In view of Theorem 6.1 this is all one needs to identify all primitive substitutive sequences in Y.

A bi-infinite sequence ω is primitive substitutive if and only if $T(\omega)$ is primitive substitutive (see Lemma 2 in [22]). We introduce the notion of *derived orbit* to reflect this fact. Let Y be a minimal shift and let u, $\mathcal{R}_u$ and Θ_u be as above. If $\upsilon \in Y$, then (by minimality of Y) there is a point υ' in the shift orbit of υ such that $\upsilon'_{[0,|u|)} = u$, and we set $\mathcal{O}_u(\upsilon)$ equal to the shift orbit of the sequence $\mathcal{D}_u(\upsilon')$. Clearly, the derived orbit $\mathcal{O}_u(\upsilon)$ depends only on the orbit of υ.

The following lemma, together with Lemmas 6.2 and 6.3, establishes the "if" part of Theorem 6.1.

Lemma 6.7. *If $\upsilon \in X(\zeta)$ is such that $\zeta(\upsilon)$ lies in the shift orbit of υ, then υ is primitive substitutive.*

Proof. Replacing υ with $T^m(\upsilon)$ if necessary and using Lemma 4.3 we may assume that $\upsilon = T^k\zeta(\upsilon)$ for some integer $k \in [0, |\zeta(\upsilon_0)|)$. If $k = 0$ or $k = |\zeta(\upsilon_0)| - 1$, then

υ or $T\upsilon$ is fixed by ζ, in which cases υ is primitive substitutive. Thus we consider $0 < k < |\zeta(\upsilon_0)| - 1$. Put $a = \upsilon_0$, let $\mathcal{R}_a$ be the set of return words to a in X and let $\Theta_a\colon \mathcal{R}_a \to \{1, 2, \ldots, |\mathcal{R}_a|\}$ be a bijection. If $w \in \mathcal{R}_a$, then $\zeta(wa)_{[k,|\zeta(w)|+k)}a$ is a factor of υ which begins and ends in a (since $w_0 = a = \zeta(a)_k$) and therefore $\zeta(wa)_{[k,|\zeta(w)|+k)}$ can be written uniquely as a concatenation of return words. We may thus define a substitution τ on $\{1, 2, \ldots, |\mathcal{R}_a|\}$ in the following way: if $b \in \{1, 2, \ldots, |\mathcal{R}_a|\}$, then we may write

$$\zeta(\Theta_a^{-1}(b)a)_{[k,|\zeta(\Theta_a^{-1}(b))|+k)} = r_0 r_1 \cdots r_n, \qquad \text{each} \quad r_j \in \mathcal{R}_a,$$

and we set

$$\tau(b) = \Theta_a(r_0)\Theta_a(r_1)\cdots\Theta_a(r_n).$$

We leave it to the reader to verify that τ is a primitive substitution fixing $\omega = \mathcal{D}_a(\upsilon)$. The proof is completed by invoking Lemma 6.2 with $\pi = \Theta_a^{-1}$. □

We now turn our attention to proving the "only if" part of Theorem 6.1. We show

$$\begin{aligned} &\pi(\omega) \text{ is primitive substitutive} \\ &\quad \Rightarrow \omega \text{ is primitive substitutive} \\ &\quad \Rightarrow \omega \text{ has only finitely many derived orbits} \\ &\quad \Rightarrow \exists l > 0\colon \ \omega \text{ and } \zeta^l(\omega) \text{ are in the same shift orbit.} \end{aligned} \tag{7}$$

We begin by paraphrasing a result from [21].

Theorem 6.8 [21, Theorem 1.3]. *If Y is a minimal shift space containing a primitive substitutive sequence, then there is a finite set of disjoint minimal shift spaces whose union contains every derived sequence of every sequence of Y. Each of these shift spaces contains a primitive substitutive sequence.*

We say a function f is *bounded-to-one* if there is a positive integer L such that each point in the range of f has at most L preimages. To establish the first implication of (7) we need the following key fact.

Proposition 6.9. *If Y is a minimal shift space containing a primitive substitutive sequence and π is a morphism defined on the alphabet of Y such that $\pi(Y)$ is infinite, then $\pi\colon Y \to \pi(Y)$ is bounded-to-one.*

Proof. The hypotheses imply that Y is infinite. Let C be as in Lemma 6.4, so that every word of length k in the language of Y occurs in every word of length kC in the language of Y. Since some letter in the alphabet of Y is not sent by π to the empty word, we must have $\lfloor C^{-1}|u|\rfloor < |\pi(u)| < |\pi||u|$ for every word u in the language of Y.

Suppose that $\omega^{(1)}, \omega^{(2)}, \ldots, \omega^{(n)}$ are distinct sequences of Y, all sent to the same sequence by π, such that $\pi(\omega_0^{(j)}) \neq \varepsilon$. For all large enough m the words $\omega^{(j)}_{[-m,m]}$, $j = 1, 2, \ldots, n$, are distinct. Fix such $m > C$. For any two of the words $\pi(\omega^{(j)}_{[-m,0)})$,

$j = 1, 2, \ldots, n$, one must be a suffix of the other; let $w^{(l)}$ be the shortest of these words. Likewise $w^{(r)} :=$ the shortest of the words $\pi(\omega^{(j)}_{[0,m]})$, $j = 1, 2, \ldots, n$, is a prefix of each of these words. Put $w = w^{(l)}w^{(r)}$. Then $|w| > 2\lfloor C^{-1}m\rfloor > C^{-1}m$.

Now let u be a word in the language of Y of length $(2m+1)C$. Then $\pi(u)$ is a word of length $\leq |\pi|(2m+1)$. Each of the words $\omega^{(j)}_{[-m,m]}$, $j = 1, 2, \ldots, n$, must occur in u, and if $u_{[r,r+2m]} = \omega^{(j)}_{[-m,m]}$, then $w^{(l)}$ is a suffix of $\pi(u_{[0,r+m)})$ and $w^{(r)}$ is a prefix of $\pi(u_{[r+m,r+2m]})$. Since $\pi(\omega^{(j)}_0) \neq \varepsilon$ there must be at least n occurrences of the word w in $\pi(u)$. Let Z be the union of the first $|\pi|$ shift images of $\pi(Y)$. Then Z is an infinite minimal shift space which by Lemma 6.2 contains a primitive substitutive sequence. Denote by C' the return constant for Z guaranteed by Lemma 6.4. Then $\pi(u)$ contains at least $n-1$ return words to w in Z and by Lemma 6.4 we have $|\pi|(2m+1) > (n-1)(C')^{-1}C^{-1}m$. Thus $n \leq 1 + 2CC'|\pi|$.

Finally, if $\omega \in Y$ and k is the least nonnegative integer for which $\pi(\omega_k) \neq \varepsilon$, then $k < C$ and if ν is the kth shift image of ω, then $\pi(\omega) = \pi(\nu)$. It follows that $\pi\colon Y \to \pi(Y)$ is at most $C(1 + 2CC'|\pi|)$-to-one. □

Remark 6.10. Proposition 6.9 can be seen as a consequence of a very general result of Durand [14]: Let (X, T) be a linearly recurrent subshift. There exists a constant L such that for every factor map $\varphi\colon (X, T) \to (Y, T)$, where Y is a nonperiodic subshift, and all $y \in Y$, we have $\#\varphi^{-1}(\{y\}) \leq L$.

The next proposition is a partial converse of Lemma 6.2.

Proposition 6.11. *Let Y be a minimal shift space containing a primitive substitutive sequence and suppose π is a morphism such that $\pi(Y)$ is infinite. If $\varpi \in Y$ is such that $\pi(\varpi)$ is primitive substitutive, then ϖ is primitive substitutive.*

Proof. Let C, Z and C' be as in the proof of Proposition 6.9. For all sufficiently large n, if $w \in \mathcal{R}_{\varpi_{[0,n)}}(Y)$, then $\pi(w)$ may be written uniquely as a concatenation of elements of $\mathcal{R}_{\pi(\varpi_{[0,n)})}(Z)$. This defines a morphism $\varphi_n\colon \{1, 2, \ldots, |\mathcal{R}_{\varpi_{[0,n)}}(Y)|\} \to \{1, 2, \ldots, |\mathcal{R}_{\pi(\varpi)_{[0,n)}}(Z)|\}$ having the property that

$$\varphi_n(\mathcal{D}_{\varpi_{[0,n)}}(\varpi)) = \mathcal{D}_{\pi(\varpi)_{[0,n)}}(\pi(\varpi)).$$

By Lemma 6.4 $|u| < C'|w|$ for any $w \in \mathcal{R}_u(Z)$. This implies that $|\varphi_n(b)| < CC'|\pi|n/\lfloor C^{-1}n\rfloor$ for any $b \in \{1, 2, \ldots, |\mathcal{R}_{\varpi_{[0,n)}}(Y)|\}$, from which we obtain the crude bound $|\varphi_n| < 2C^2C'|\pi|$, valid for $n \geq C$. It follows from the bound on the $|\varphi_n|$ that there are only finitely many distinct maps φ_n.

By Durand's theorem, there are only finitely many different derived sequences $\mathcal{D}_{\pi(\varpi)_{[0,n)}}(\pi(\varpi))$ and each is primitive substitutive. By Theorem 6.8, there is a finite disjoint collection of minimal shift spaces $Y_1, Y_2, \ldots, Y_k$, each containing a primitive substitutive sequence, with the property that any derived sequence of a sequence in Y is contained in one of them. If $\mathcal{D}_{\varpi_{[0,n)}}(\varpi) \in Y_{l(n)}$, then φ_n is defined on $Y_{l(n)}$ and takes it to an infinite subset of the closure of the shift orbit of $\mathcal{D}_{\pi(\varpi)_{[0,n)}}(\pi(\varpi))$. By Proposition 6.9, there are only finitely many sequences in $Y_{l(n)}$ which are sent to $\mathcal{D}_{\pi(\varpi)_{[0,n)}}(\pi(\varpi))$ by φ_n, and $\mathcal{D}_{\varpi_{[0,n)}}(\varpi)$ is one of them. Thus we see that there can be only finitely many distinct sequences $\mathcal{D}_{\varpi_{[0,n)}}(\varpi)$. □

Lemma 6.12. *A sequence υ belonging to a minimal shift space Y is primitive substitutive if and only if $\{\mathcal{O}_u(\upsilon)\colon\ u$ a factor of $\upsilon\}$ is finite.*

Proof. Suppose first that υ is primitive substitutive. We will construct a finite collection of sequences which contains a member of each derived orbit. Let u be any factor of υ and let $n = n(u)$ be the least positive integer for which u is a factor of $\upsilon_{[0,n)}$. Then $\upsilon_{[0,n)}$ ends in u and $n < (C+1)|u|$ where C is the constant from Lemma 6.4. If w is a return word to $\upsilon_{[0,n)}$, then $w\upsilon_{[0,n)}$ begins in $\upsilon_{[0,n)}$ and therefore $(w\upsilon_{[0,n)})_{[n-|u|,|w|+n)}$ begins and ends in u. Thus $(w\upsilon_{[0,n)})_{[n-|u|,|w|+n-|u|)}$ can be written uniquely as a concatenation of return words to u,

$$(w\upsilon_{[0,n)})_{[n-|u|,|w|+n-|u|)} = r_{u,w,0}r_{u,w,1}\cdots r_{u,w,m(u,w)}, \qquad \text{each} \quad r_{u,w,j} \in \mathcal{R}_u.$$

It follows from Lemma 6.4 that $m(u,w) < C^2(C+1)+1$.

Consider the morphism $\varphi_u\colon \{1,2,\ldots,|\mathcal{R}_{\upsilon_{[0,n(u))}}|\} \to \{1,2,\ldots,|\mathcal{R}_u|\}$ given by

$$\varphi_u(b) = \Theta_u(r_{u,w,0})\Theta_u(r_{u,w,1})\cdots\Theta_u(r_{u,w,m(u,w)}), \qquad w := \Theta_{\upsilon_{[0,n(u))}}(b).$$

One checks easily that $\varphi_u(\mathcal{D}_{\upsilon_{[0,n(u))}}(\upsilon))$ is in $\mathcal{O}_u(\upsilon)$. It follows from the bound on the $m(u,w)$ that there are only finitely many different possible maps φ_u, and by Durand's theorem there are only finitely many distinct derived sequences $\mathcal{D}_{\upsilon_{[0,n)}}(\upsilon)$. Thus, $\{\varphi_u(\mathcal{D}_{\upsilon_{[0,n(u))}}(\upsilon))\colon u$ a factor of $\upsilon\}$ is a finite set containing a representative of each derived orbit.

Now suppose $\{\mathcal{O}_u(\upsilon)\colon\ u$ a factor of $\upsilon\}$ is finite. We may assume that υ is not periodic. Then we can find positive integers $n_1 < n_2$ for which $\mathcal{D}_{\upsilon_{[0,n_1)}}(\upsilon) \in \mathcal{O}_{\upsilon_{[0,n_2)}}(\upsilon)$ and every return word to $\upsilon_{[0,n_2)}$ is longer than the longest return word to $\upsilon_{[0,n_1)}$. Every return word to $\upsilon_{[0,n_2)}$ can be written uniquely as a concatenation of return words to $\upsilon_{[0,n_1)}$ and thus we may define a substitution $\tau\colon \{1,2,\ldots,|\mathcal{R}_{\upsilon_{[0,n_2)}}|\} \to \{1,2,\ldots,|\mathcal{R}_{\upsilon_{[0,n_2)}}|\}$ by putting $\tau = \Theta_{\upsilon_{[0,n_1)}} \circ \Theta^{-1}_{\upsilon_{[0,n_2)}}$. It is easy to verify that τ is primitive, and if m is the integer for which $(\mathcal{D}_{\upsilon_{[0,n_1)}}(\upsilon))_j = (\mathcal{D}_{\upsilon_{[0,n_2)}}(\upsilon))_{m+j}$ holds for all j, then

$$\tau(\mathcal{D}_{\upsilon_{[0,n_2)}}(\upsilon))_j = (\mathcal{D}_{\upsilon_{[0,n_2)}}(\upsilon))_{j+m} \qquad \text{for all } j.$$

Lemma 6.7 asserts that $\mathcal{D}_{\upsilon_{[0,n_2)}}(\upsilon)$ is primitive substitutive, and an application of Lemma 6.2 with $\pi = \Theta^{-1}_{\upsilon_{[0,n_2)}}$ completes the proof that υ is primitive substitutive. □

We recall a well-known inequality (see [31]). Since ζ is primitive the matrix M_ζ is also primitive and the Perron–Frobenius theorem guarantees a positive eigenvalue θ strictly greater in absolute value than any other eigenvalue of M_ζ.

Lemma 6.13. *There exist positive constants C_1, C_2 such that $C_1|w|\theta^n < |\zeta^n(w)| < C_2|w|\theta^n$ for every word w and for every positive integer n.*

The next lemma completes the proof of Theorem 6.1.

Lemma 6.14. *A sequence $\omega \in X(\zeta)$ has only finitely many distinct derived orbits if and only if there exists a positive integer l such that $\zeta^l(\omega)$ is in the shift orbit of ω.*

Proof. If $\zeta^l(\omega)$ is in the shift orbit of ω, then by Lemma 6.7 ω is primitive substitutive, whence it has only finitely many distinct derived orbits, by Lemma 6.12.

Now suppose ω is primitive substitutive. Let u be any factor of ω and consider the array

$$\begin{array}{llll} \mathcal{O}_u(\omega) & & & \\ \mathcal{O}_{\zeta(u)}(\omega) & \mathcal{O}_{\zeta(u)}(\zeta(\omega)) & & \\ \mathcal{O}_{\zeta^2(u)}(\omega) & \mathcal{O}_{\zeta^2(u)}(\zeta(\omega)) & \mathcal{O}_{\zeta^2(u)}(\zeta^2(\omega)) & \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

There are only finitely many distinct entries $\mathcal{O}_{\zeta^n(u)}(\omega)$ in the first column, by Lemma 6.12.

If $w \in \mathcal{R}_u$, then wu begins in u and is a factor of ω. It follows that $\zeta^k(wu)$ is a factor of ω which begins and ends in $\zeta^k(u)$, and thus $\zeta^k(w)$ may be written uniquely as a concatenation of words of $\mathcal{R}_{\zeta^k(u)}$. This induces a morphism $\varphi_k\colon \{1, 2, \ldots, |\mathcal{R}_u|\} \to \{1, 2, \ldots, |\mathcal{R}_{\zeta^k(u)}|\}$, defined for $b \in \{1, 2, \ldots, |\mathcal{R}_u|\}$ by

$$\varphi_k(b) = \Theta_{\zeta^k(u)}(r_{k,b,0})\Theta_{\zeta^k(u)}(r_{k,b,1}) \cdots \Theta_{\zeta^k(u)}(r_{k,b,l(k,b)}),$$

where

$$\zeta^k(\Theta_u^{-1}(b)) = r_{k,b,0}r_{k,b,1} \cdots r_{k,b,l(k,b)}, \qquad \text{each} \quad r_{k,b,j} \in \mathcal{R}_{\zeta^k(u)},$$

is the unique representation of $\zeta^k(\Theta_u^{-1}(b))$ as a concatenation of return words to $\zeta^k(u)$. One may verify that for any $\upsilon \in X(\zeta)$,

$$\varphi_k(\mathcal{O}_u(\upsilon)) \subset \mathcal{O}_{\zeta^k(u)}(\zeta^k(\upsilon)).$$

Thus, for each $m \geq n \geq 0$ we have

$$\varphi_n(\mathcal{O}_u(\zeta^{m-n}(\omega))) \subset \mathcal{O}_{\zeta^n(u)}(\zeta^m(\omega)).$$

By Lemmas 6.4 and 6.13 we have $|\varphi_k(b)| < C_2C/C_1$ for each $b \in \{1, 2, \ldots, |\mathcal{R}_u|\}$, independent of k and u. This means there are only finitely many distinct morphisms φ_k.

It follows that the entire array has only finitely many distinct entries. In particular, for some $m, n, l > 0$ we must have $\mathcal{O}_{\zeta^n(u)}(\zeta^m(\omega)) = \mathcal{O}_{\zeta^n(u)}(\zeta^{m+l}(\omega))$ from which we see that $\zeta^m(\omega)$ and $\zeta^{m+l}(\omega)$ lie in the same shift orbit. Thus, $\zeta^l(\omega)$ is in the shift orbit of ω. $\square$

7. Measures of Cylinders

It is well known that if ζ is primitive, then $X(\zeta)$ is uniquely ergodic [31]. Denote by ν the T-invariant probability measure on X. By the Perron–Frobenius theorem, $M(G)$ admits a simple positive eigenvalue θ, greater in absolute value than any other eigenvalue of $M(G)$, and a unique strictly positive (right) eigenvector $(p_a)_{a\in\mathcal{A}}$, normalized so that $\sum_{a\in\mathcal{A}} p_a = 1$. For a finite walk $e_1e_2\cdots e_k$ denote by $[e_1e_2\cdots e_k]$ the cylinder set $\{(f_j)_{j=1}^{\infty} \in G^{\infty}\colon\ f_j = e_j \text{ for all } 1 \le j \le k\}$, and set

$$\mu([e_1e_2\cdots e_k]) = \theta^{-k} p_{\mathtt{t}(e_k)}. \tag{8}$$

One readily verifies the consistency requirement

$$\sum_{e\in\mathcal{E},\, \mathtt{i}(e)=\mathtt{t}(e_k)} \mu([e_1e_2\cdots e_k e]) = \mu([e_1e_2\cdots e_k]),$$

so by Kolmogorov's theorem μ extends to a probability measure on G^{∞}.

Proposition 7.1. *The measure μ is S-invariant and $\mu \circ H = \nu$.*

Proof. The first assertion is proved by considering sets of the form $[e_1e_2\cdots e_k] \subset G^{\infty}$ while the second follows from unique ergodicity. $\square$

The following was conjectured by Boshernitzan [6].

Theorem 7.2. *The measures of all cylinders in X lie in a finite union of geometric sequences.*

Proof. Fix positive constants C_1, C_2 as in Lemma 6.13 and let $K > 0$ be such that for each $\varpi \in X$ and each word $w \in \mathcal{L}(\zeta)$ the minimum gap between successive occurrences of w in ϖ is at least $K|w|$. The value C^{-1} of Lemma 6.4 will suffice for K. Let $w \in \mathcal{L}(\zeta)$ and set $n = \lceil(\log|w| - \log C_1)/\log\theta\rceil$. Then for every letter a we have $|w| < |\zeta^n(a)| < C_2\theta|w|/C_1$.

If $u \in \mathcal{L}(\zeta)$, then $H([u])$ is compact and open in G^{∞} and hence is a finite union of cylinders. Thus there is a positive integer m such that for every word u of length 2 one can write $H([u])$ as a union of cylinders of the form $[e_1e_2\cdots e_m]$. Our choice of n implies, for every word u of length 2 and for every integer $\ell \in [0, |\zeta^n(u_0)|)$, either $T^{\ell}(\zeta^n([u])) \subset [w]$ or $T^{\ell}(\zeta^n([u])) \cap [w] = \emptyset$. Our choice of K ensures that at most $\lfloor C_2\theta/C_1K\rfloor$ of these ℓ satisfy $T^{\ell}(\zeta^n([u])) \subset [w]$.

Now if $(j_1, a_1)\cdots(j_{n+m}, a_{n+m})$ is a walk on G, then $T^{j_n+1}\circ\zeta\circ\cdots\circ T^{j_{m+n}}([a_{n+m}]) \subset [u]$ for some word u of length 2 and if $[(j_1, a_1)\cdots(j_{n+m}, a_{n+m})]$ meets $H([w])$, then

$\ell := \sum_{l=1}^{n} |\zeta^{l-1}(\zeta(a_l)_{[0,j_l)})|$ is one of those few integers for which $T^{\ell}(\zeta^n([u])) \subset [w]$. Thus for each walk $e_{n+1}\cdots e_{n+m}$ there are at most $\lfloor C_2\theta/C_1K\rfloor$ choices of $e_1e_2\cdots e_n$ such that $[e_1e_2\cdots e_{n+m}] \cap H([w]) \neq \emptyset$, and $[e_1e_2\cdots e_{n+m}] \subset H([w])$ whenever this intersection is nonempty. Setting L equal to the number of walks on G of length m this shows that $H(\omega)$ is a union of no more than $\lfloor C_2\theta L/C_1K\rfloor$ cylinders of the form $[e_1\cdots e_{n+m}]$, and we can use (8) and Proposition 7.1 to write

$$\nu([w]) = \theta^{-(n+m)} \sum_{a\in\mathcal{A}} t_a p_a,$$

where each $t_a \in \mathbb{Z}$ and $0 \le t_{abc} \le \lfloor C_2\theta L/C_1K\rfloor$. Thus the measure of every cylinder in X is in

$$\bigcup_{n\ge 0} \theta^{-n} \left\{ \sum_{a\in\mathcal{A}} t_a p_a \colon\ t_a \in \mathbb{Z},\ 0 \le t_a \le \left\lfloor \frac{C_2\theta L}{C_1K} \right\rfloor \right\}. \qquad \square$$

8. Geometric Realizations of Substitutions

In 1982 Rauzy showed that the subshift (X, T) generated by the *Tribonacci substitution* $1 \mapsto 12$, $2 \mapsto 13$, $3 \mapsto 1$ is a natural coding of a rotation on the two-dimensional torus, i.e., is metrically isomorphic to an exchange of three fractal domains on a compact set in $\mathbb{R}^2$, each domain being translated by the same vector modulo a lattice [32]. This example, also studied in great detail by Messaoudi in [28] and [29] and Ito and Kimura in [23], prompted a general interest in the question of which substitutions admit a geometric realization as in the case of Tribonacci. This question was made precise by Queffélec in [31]. Since then many partial results have been obtained for various types of substitutions (see [2]–[5], [7], [9], [11], [16], [17], [20], and [33] to name just a few).

By a *complex geometric realization* [20] of a primitive substitution ζ on an alphabet $\mathcal{A}$ we mean a continuous function $h\colon\ X(\zeta) \to \mathbb{C}$, a nonzero complex number β and a nonzero vector $v = (v_a)_{a\in\mathcal{A}} \in \mathbb{C}^{|\mathcal{A}|}$ such that

- $h \circ \zeta = \beta h$ and
- $h(T\omega) = h(\omega) + v_{\omega_0}$ for all $\omega \in X(\zeta)$.

One can deduce from the definition that $|\beta| < 1$, $vM_\zeta = \beta v$ and $h(\omega) = 0$ for each $\omega \in \mathrm{Per}(\zeta)$ (see [20]). In [20] we showed that every nonzero eigenvector of M_ζ corresponding to a nonzero eigenvalue of modulus strictly less than 1 determines a geometric realization of ζ.

Suppose that (h, β, v) is a geometric realization of ζ. Consider the *edge weights*

$$\rho(j, a) = v_{\zeta(a)_0} + v_{\zeta(a)_1} + \cdots + v_{\zeta(a)_{j-1}}$$

for $(j, a) \in \mathcal{E}$. Let $\omega \in X$ and $H(\omega) = (j_l, a_l)_{l=1}^{\infty}$. If $k > 0$ and ϖ is the unique point of X for which $\varpi_0 = a_k$ and $\omega = (T^{j_1} \circ \zeta \circ \ldots T^{j_k} \circ \zeta)(\varpi)$, then

$$h(\omega) = \beta^k h(\varpi) + \sum_{l=1}^{k} \beta^{l-1}\rho(j_l, a_l). \tag{9}$$

Proposition 8.1. *The function* $\psi\colon G^\infty \to \mathbb{C}$ *given by*

$$\psi((j_l, a_l)_{l=1}^\infty) = \sum_{l=1}^\infty \beta^{l-1}\rho(j_l, a_l)$$

is continuous and $\psi \circ H = h$.

Proof. Continuity follows from the fact that $|\beta| < 1$. The second assertion follows from (9). ☐

For each $w \in \mathcal{L}(\zeta)$ put $\Omega_w = h([w])$, and for each finite walk $e_1 e_2 \cdots e_k$ on G set $J_{e_1\cdots e_k} = \psi([e_1 e_2 \cdots e_k])$.

The function ψ above is related to the graph-directed construction of Mauldin and Williams in [27]. Define *edge maps* $\psi_{(j,a)}\colon \Omega_a \to \Omega_{\mathtt{i}(j,a)}$ by setting $\psi_{(j,a)}(z) = \beta z + \rho(j, a)$. Then

$$J_{e_1\cdots e_k} = (\psi_{e_1} \circ \cdots \circ \psi_{e_k})(\Omega_{\mathtt{t}(e_k)}). \tag{10}$$

The Mauldin and Williams construction requires that the edge maps be defined on compact subsets of $\mathbb{R}^n$ with nonempty interiors and have nonoverlapping images. Although these conditions need not be satisfied by our edge maps, the methods can still be applied to give an upper bound on the Hausdorff dimension of Ω.

Proposition 8.2. *The Hausdorff dimension of* Ω *is no greater than* $-\log\theta/\log\beta$.

Proof. We see from (10) that $\operatorname{diam}(J_{e_1\cdots e_k}) \le \beta^k \operatorname{diam}(\Omega_{\mathtt{t}(e_k)})$. Set $\alpha = -\log\theta/\log\beta$. For each positive integer k we have

$$\begin{aligned}
\sum(\operatorname{diam} J_{e_1\cdots e_k})^\alpha &\le \sum(\beta^k \operatorname{diam}\Omega_{\mathtt{t}(e_k)})^\alpha \\
&= \sum \theta^{-k}(\operatorname{diam}\Omega_{\mathtt{t}(e_k)})^\alpha \\
&= \sum \mu([e_1 e_2 \cdots e_k])(\operatorname{diam}\Omega_{\mathtt{t}(e_k)})^\alpha / p_{\mathtt{t}(e_k)} \\
&\le \sup\{(\operatorname{diam}\Omega_a)^\alpha / p_a\colon a \in \mathcal{A}\},
\end{aligned}$$

where the sums are over all walks $e_1 e_2 \cdots e_k$ on G. This shows $\mathcal{H}^\alpha(\Omega) < \infty$ as required. ☐

We are interested in a condition which guarantees that the map h is one-to-one off a set of ν-measure zero. The remainder of the section is devoted to some partial results.

Suppose $\mathcal{H}^\alpha(\Omega) > 0$. For each $a \in \mathcal{A}$, $[a] = \bigcup\{H^{-1}([e])\colon\ e \in \mathcal{E},\ \mathtt{i}(e) = a\}$ and thus

$$\Omega_a = \bigcup\{J_e\colon\ e \in \mathcal{E},\ \mathtt{i}(e) = a\}, \tag{11}$$

and therefore

$$\begin{aligned}\mathcal{H}^\alpha(\Omega_a) &\leq \sum_{e\in\mathcal{E},\ \mathtt{i}(e)=a} \mathcal{H}^\alpha(J_e)\\ &= \sum_{e\in\mathcal{E},\ \mathtt{i}(e)=a} \mathcal{H}^\alpha(\beta\Omega_{\mathtt{t}(e)} + \rho(e))\\ &= \theta^{-1} \sum_{e\in\mathcal{E},\ \mathtt{i}(e)=a} \mathcal{H}^\alpha(\Omega_{\mathtt{t}(e)})\\ &= \theta^{-1} \sum_{b\in\mathcal{A}} M_{ab}\mathcal{H}^\alpha(\Omega_b). \end{aligned}\tag{12}$$

Writing $\mathbf{h}$ for the vector whose ath entry is $\mathcal{H}^\alpha(\Omega_a)$, we have $M\mathbf{h} \geq \theta\mathbf{h}$. This implies that in fact $M\mathbf{h} = \theta\mathbf{h}$, so we must also have equality in the first line of (12). It follows from this and (11) that if e and e' are distinct edges with $\mathtt{i}(e) = \mathtt{i}(e')$, then $\mathcal{H}^\alpha(J_e \cap J_{e'}) = 0$. A straightforward generalization of this argument to longer paths yields:

Proposition 8.3. *If $\mathcal{H}^\alpha(\Omega) > 0$ and $w \in \mathcal{L}(\zeta)$, then the restriction of h to $[w]$ is one-to-one off a set of ν-measure zero. Moreover, the restriction of $\mathcal{H}^\alpha$ to $h([w])$ is a multiple of $\nu \circ \left(h|_{[w]}\right)^{-1}$.*

Remark 8.4. We have no general method for verifying the hypothesis of Proposition 8.3.

Conjecture 8.5. *If $\theta|\beta| < 1$, then $\mathcal{H}^\alpha(\Omega) > 0$. If $\beta \in \mathbb{C}\backslash\mathbb{R}$ and $\theta|\beta|^2 < 1$, then $\mathcal{H}^\alpha(\Omega) > 0$.*

We do not know whether Proposition 8.3 implies that h is one-to-one off a set of ν-measure zero whenever $\mathcal{H}^\alpha(\Omega) > 0$. One partial result is the following.

Suppose a, b are vertices of G and N, n are nonnegative integers such that

(A1) $N < \min(|\zeta^n(a)|, |\zeta^n(b)|)$,
(A2) $\sum_{j=0}^{N-1} v_{\zeta^n(a)_j} = \sum_{j=0}^{N-1} v_{\zeta^n(b)_j}$ and
(A3) $\zeta^n(a)_N = \zeta^n(b)_N$.

Then there exist walks $e_1 \cdots e_n,\ f_1 \cdots f_n$, starting at $\zeta^n(a)_N$ and ending at a, b, respectively, for which

$$\sum_{j=1}^{n} \beta^{(j-1)}\rho(e_j) = \sum_{j=1}^{n} \beta^{(j-1)}\rho(f_j).$$

We have

$$J_{e_1\cdots e_n} = \beta^n \Omega_a + \sum_{j=1}^{n} \beta^{(j-1)} \rho(e_j) \quad \text{and} \quad J_{f_1\cdots f_n} = \beta^n \Omega_b + \sum_{j=1}^{n} \beta^{(j-1)} \rho(f_j),$$

and the comments preceding Proposition 8.3 show

$$\mathcal{H}^\alpha(J_{e_1\cdots e_n} \cap J_{f_1\cdots f_n}) = 0$$

from which we deduce $\mathcal{H}^\alpha(\Omega_a \cap \Omega_b) = 0$.

Thus, by Proposition 8.3, if $\mathcal{H}^\alpha(\Omega) > 0$ and for each pair a, b of vertices there exist nonnegative integers N, n satisfying (A1)–(A3) above, then h is one-to-one off a set of ν-measure zero.

There is a simple combinatorial criterion on substitutions which guarantees (A1)–(A3) will be satisfied for each a, b. The substitution ζ is said to have the *coincidence property* if for each pair $\omega, \varpi \in \mathrm{Per}(\zeta)$ there is an integer n for which $\omega_n = \varpi_n$ and the number of occurrences of each letter of the alphabet in $\omega_{[0,n)}$ is the same as that in $\varpi_{[0,n)}$. This condition was originally introduced by Dekking in [10] in the context of constant length substitutions. Arnoux and Ito [4] generalized it to arbitrary substitutions.

Proposition 8.6. *If $\mathcal{H}^\alpha(\Omega) > 0$ and ζ has the coincidence property, then h is one-to-one off a set of ν-measure zero.*

Acknowledgments

We are grateful to B. Host and F. Durand for their helpful comments and suggestions. We also thank A. Forrest for useful e-mail exchanges. Finally, we wish to thank the referee for suggestions and for pointing out a simpler proof of Theorem 7.2.

References

[1] J. P. Allouche, E. Cateland, W. J. Gilbert, H.-O. Peitgen, J. O. Shallit, G. Skordev, Automatic maps in exotic numeration systems, *Theory Comput. Systems*, **30**(3) (1997), 285–331.

[2] P. Arnoux, Un exemple de semi-conjugaison entre un échange d'intervalles et une translation sur le tore, *Bull. Soc. Math. France*, **116**(4) (1988), 489–500.

[3] P. Arnoux, V. Berthé, S. Ito, Discrete planes, $\mathbb{Z}^2$-actions, Jacobi–Perron algorithm and substitutions, preprint (1999).

[4] P. Arnoux, S. Ito, Pisot substitutions and Rauzy fractals, *Journées Montoises d'Informatique Théorique Bull. Belg. Math. Soc. Simon Stevin*, **8** (2001), 181–207.

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

[6] M. Boshernitzan, Private communication.

[7] M. Boshernitzan, I. Kornfeld, Interval translation mappings, *Ergodic Theory Dynamical Systems*, **15** (1995), 821–832.

[8] V. Canterini, A. Siegel, Automate des préfixes-suffixes associé à une substitution primitive, preprint (1999). To appear in *J. Th. Nombres Bordeaux*.

[9] V. Canterini, A. Siegel, Geometric representations of Pisot substitutions, *Trans. Amer. Math. Soc.*, **353** (2001), 5121–5144..

[10] F. M. Dekking, The spectrum of dynamical systems arising from substitutions of constant length, *Z. Wahrsch. Verw. Gebiete*, **41** (1978), 221–239.
[11] F. M. Dekking, Recurrent sets, *Adv. in Math.*, **44** (1982), 78–104.
[12] J. M. Dumont, A. Thomas, Systèmes de numération et fonctions fractales relatifs aux substitutions, *Theoret. Comput. Sci.*, **65** (1989), 153–169.
[13] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math.*, **179**(1–3) (1998), 89–101.
[14] F. Durand, Linearly recurrent subshifts have a finite number of non-periodic subshift factors, *Ergodic Theory Dynamical Systems*, **20**(4) (2000), 1061–1078.
[15] F. Durand, B. Host, C. Skau, Substitutional dynamical systems, Bratteli diagrams and dimension groups, *Ergodic Theory Dynamical Systems* **18**(4) (1999), 953–993.
[16] H. Ei, S. Ito, Decomposition theorem on invertible substitutions, *Osaka J. Math.*, **35**(4) (1998), 821–834.
[17] S. Ferenczi, Les transformations de Chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité 2n+1, *Bull. Soc. Math. France*, **123**(2) (1995), 271–292.
[18] A. H. Forrest, K-Groups associated with substitution minimal systems, *Israel J. Math.*, **98** (1997), 101–139.
[19] R. Herman, I. F. Putnam, C. Skau, Ordered Bratteli diagrams, dimension groups and topological dynamics, *Internat. J. Math.*, **3**(6) (1992), 827–864.
[20] C. Holton, L. Q. Zamboni, Geometric realizations of substitutions, *Bull. Soc. Math. France*, **126** (1998), 149–179.
[21] C. Holton, L. Q. Zamboni, Descendants of primitive substitutions, *Theory Comput. Systems*, **32** (1999), 133–157.
[22] C. Holton, L. Q. Zamboni, Iteration of maps by primitive substitutive sequences, in *Dynamical Systems* (Luminy-Marseille, 1998), World Science, River Edge, NJ, 2000, pp. 137–143.
[23] S. Ito, M. Kimura, On Rauzy fractal, *Japan J. Indust. Appl. Math.*, **8** (1991), 461–486.
[24] A. N. Livshitz, A sufficient condition for weak mixing of substitutions and stationary adic transformations, *Mat. Zametki*, **44**(6) (1988), 920–925.
[25] A. N. Livshitz, A. M. Vershik, Adic models of ergodic transformations, spectral theory, substitutions and related topics, *Adv. Soviet Math.*, **9** (1992), 185–204.
[26] J. C. Martin, Substitution minimal flows, *Amer. J. Math.*, **93** (1971), 503–526.
[27] R. D. Mauldin, S. C. Williams, Hausdorff dimension in graph directed constructions, *Trans. Amer. Math. Soc.*, **309** (1988), 811–829.
[28] A. Messaoudi, Propriétés arithmétiques et dynamiques du fractal de Rauzy, *J. Th. Nombres Bordeaux*, **10** (1998), 135–162.
[29] A. Messaoudi, Frontière du fractal de Rauzy et système de numération complexe, *Acta Arith.*, **95** (2000), 195–224.
[30] B. Mossé, Puissances de mots et reconnaissabilité des points fixes d'une substitution, *Theoret. Comput. Sci.*, **99** (1992), 327–334.
[31] M. Queffélec, *Substitution Dynamical Systems—Spectral Analysis*, Lecture Notes in Mathematics #1294, Springer-Verlag, Berlin, 1987.
[32] G. Rauzy, Nombres algébriques et substitutions, *Bull. Soc. Math. France*, **110** (1982), 147–178.
[33] B. Solomyak, Substitutions, adic transformations, and beta-expansions, *Contemp. Math.*, **135** (1992), 361–372.

Received January 24, 2000, *and in final form August* 9, 2001. *Online publication November* 23, 2001.