Rigidity and substitutive tree words

V. Berthé, F. Dolce, F. Durand, J. Leroy, D. Perrin

April 10, 2017

Abstract

Tree words are infinite words that are defined in terms of extension graphs that describe the left and right extensions of their factors; tree words are such that their extension graphs are trees. This class of words includes classical families of words such as Sturmian and episturmian words, codings of interval exchanges or else Arnoux-Rauzy words. We investigate here the properties of substitutive tree words and prove some rigidity properties, that is, algebraic properties on the set of substitutions that fix a tree word.

1 Introduction

Tree words are infinite words that are defined in terms of extension graphs that describe the left and right extensions of their factors. Extension graphs can be roughly described as follows: given an infinite word $x$, and given a finite factor $w$ of $x$, one puts an edge between the letters $a$ and $b$ if $awb$ is a factor of $x$. Tree words are such that all their extension graphs are trees. For precise definitions, see Section 2.1 and 2.2. This class of words with linear factor complexity includes classical families of words such as Sturmian and episturmian words, codings of interval exchanges or else Arnoux-Rauzy words. They have been introduced in [7] and studied in several papers (as, for instance, [1, 8, 9, 7]). These words have striking combinatorial, ergodic and algebraic properties. This includes in particular algebraic properties of their return words [7], and of maximal bifix codes defined with respect to their languages [8, 9].

We investigate here the properties of substitutive tree words and prove some rigidity properties. Rigidity has to do with the algebraic properties of the monoid of substitutions that fix a tree word: an infinite word generated by a substitution is rigid if all the substitutions which generate this word are powers of a unique substitution. In the present paper, we concentrate
on the iterative stabilizer according to the terminology of [23]: we focus on non-erasing morphisms and on infinite words generated by iterating a substitution.

There exist numerous specific results on the two-letter case concerning rigidity (see [34, 35] and also [3]). It is indeed well known that Sturmian words generated by substitutions are rigid [34, 35]. The situation is more contrasted as soon as the size of the alphabet increases. For instance, over a ternary alphabet, the monoid of morphisms generating a given infinite word by iteration can be infinitely generated, even when the word is generated by iterating an invertible primitive morphism (see [10, 23]).

Our main results are the following. We provide a characterization of substitutive primitive tree words in terms of $S$-adic expansions and tame substitutions (Theorem 3.3). An $S$-adic expansion corresponds to an infinite composition of substitutions of the form $\sigma_1 \circ \cdots \circ \sigma_n$, and tame substitutions are elementary substitutions that thus extend to free group automorphisms (see Section 2.3 for definitions). We prove that if an infinite tree word is fixed by two primitive substitutions, then they coincide up to a power, and there exists a given substitution $\theta$ such that all the primitive substitutions of the stabilizer of conjugate up to powers to $\theta$ (Theorem 4.1). We also prove that tree words cannot be generated by constant length substitutions (Corollary 5.1).

Our proofs rely on the notion of return words and on the so-called Return Theorem [7, Theorem 4.5] that states that for every infinite tree word defined over the alphabet $A$, the set of (first right) return words is a basis of the free group generated by the alphabet $A$.

Let us briefly sketch the contents of this paper. We recall in Section 2 the first basic definitions that are required, such as the notions of extension graphs and tree words, return words, stabilizers, etc. We provide in Section 3 a characterization of substitutive tree words in terms of derived sequences and $S$-adic expansions. Rigidity properties are considered in Section 4. Lastly, we prove that recurrent tree words cannot have rational continuous eigenvalues in Section 5. We conclude this paper with several questions in Section 3.2.

2 Basic definitions

2.1 Words, extensions and subshifts

Let $A$ be a finite nonempty alphabet. All words considered below, unless stated explicitly, are supposed to be on the alphabet $A$. We denote by $\varepsilon$ the
empty word of the free monoid $A^*$, by $A^+$ the free semigroup and by $A^\mathbb{N}$ the set of infinite words over $A$. The Parikh vector of a word $w \in A^*$ is the vector in $\mathbb{N}^{\text{Card } A}$ whose coordinates are equal to the number of occurrences on letters in $w$, i.e., its $i$-th entry is equal to $|w|i$, where $|w|i$ stands for the number of occurrences of the letter $i$ in $w$. The notation $|w|$ stands for the length of $w$.

We say that a word $u$ is a factor of a word $w$ if there exists words $p,s$ such that $w = pvs$. If $p = \varepsilon$ (resp. $s = \varepsilon$) we say that $u$ is a prefix (resp. suffix) of $w$.

Let $F$ be a set of words on the alphabet $A$. For $w \in F$, we denote

$$L(w) = \{ a \in A \mid aw \in F \},$$
$$\ell(w) = \text{Card}(L(w)),$$
$$R(w) = \{ a \in A \mid wa \in F \},$$
$$r(w) = \text{Card}(R(w)),$$
$$B(w) = \{ (a,b) \in A \times A \mid awb \in F \},$$
$$b(w) = \text{Card}(B(w)).$$

Let $F$ be a set of words. For a word $w \in F$, we consider an undirected bipartite graph $E(w)$ called its extension graph in $F$ and defined as follows: its set of vertices is the disjoint union of $L(w)$ and $R(w)$ and its edges are the pairs $(a,b) \in B(w)$. For an illustration, see Example 2.1 below.

**Example 2.1** Let $F$ be a set of words on the alphabet $\{a,b\}^*$ having as factors of length less then 4 the set $\{\varepsilon, a, b, ab, ba, aab, aba, bab\}$. The extension graphs of the empty word and of the two letters are represented in Figure 1.

\[\begin{align*}
\mathcal{E}(\varepsilon) & \quad \mathcal{E}(a) & \quad \mathcal{E}(b) \\
\begin{array}{c}
\begin{array}{ccc}
\varepsilon & a & a \\
\wedge & \wedge & \wedge \\
b & b & b
\end{array}
\end{array} & \begin{array}{c}
\begin{array}{ccc}
a & a & a \\
\wedge & \wedge & \wedge \\
b & b & b
\end{array}
\end{array} & \begin{array}{c}
\begin{array}{ccc}
a & a \\
\wedge & \wedge \\
b & b
\end{array}
\end{array}
\end{align*}\]

Figure 1: The extension graphs of $\varepsilon$ (on the left), $a$ (on the center) and $b$ (on the right).

A set of words $F$ is factorial if it contains the factors of all its elements. It is biessential if it is factorial and if for all $w \in F$, one has $r(w) \geq 1$ and $\ell(w) \geq 1$. It is recurrent if for every $u,v \in F$ there exists a word $w \in F$ such that $uwv \in F$. A uniformly recurrent set is a biessential set of words of $A^*$ such that for any word $u \in F$, there exists an integer $n \geq 1$ such that
$u$ is a factor of every word of $F$ of length $n$. Every uniformly recurrent set is recurrent while the opposite is in general not true.

A word $w \in F$ is called right-special if $r(w) \geq 2$. It is called left-special if $\ell(w) \geq 2$. It is called bispecial if it is both right- and left-special. For a word $w \in F$, let

$$m(w) = b(w) - \ell(w) - r(w) + 1.$$ 

We say that $w$ is neutral (resp. weak, resp. strong) if $m(w) = 0$ (resp. $m(w) \leq 0$, resp. $m(w) \geq 0$). A factorial set $F$ is said to be neutral (resp. weak, resp. strong) if any word of $F$ is neutral (resp. neutral of weak, resp. neutral or strong). Note that our definition of neutral set corresponds to the one of neutral set of characteristic 1 in [13]. We will work here with a subclass of the family of neutral sets, namely tree sets introduced in Section 2.2. But before defining them, we also introduce notions corresponding to infinite words and shifts.

An infinite word in $A^\mathbb{N}$ is said to be uniformly recurrent if the set of its factors is uniformly recurrent. In other words, an infinite word $x = (x_n)_n$ is uniformly recurrent if every word occurring in $u$ occurs in an infinite number of positions with bounded gaps, that is, if for every factor $w$, there exists $s$ such that for every $n$, $w$ is a factor of $x_n \ldots x_{n+s-1}$. The set of factors $F(x)$ of an infinite word $x$ is called its language.

The mapping $S$ acting on sets of infinite words is the (one-sided) shift $S$ acting on $A^\mathbb{N}$:

$$S ((x_n)_{n \in \mathbb{N}}) = (x_{n+1})_{n \in \mathbb{N}}.$$ 

A subshift is a couple $(X, S)$ where $X$ is a closed shift invariant subset of some $A^\mathbb{N}$. A subshift is said to be minimal if it admits no non-trivial closed and shift-invariant subset. If $X$ is a subshift, then its language $F(X)$ is defined as the set of factors of elements of $X$. Given a biessential set of words $F \subset A^*$, there exists a unique subshift $X$ such that $F(X) = F$.

### 2.2 Tree sets

The following notions has been introduced in [7]. We say that a set of words $F$ is a tree set if it is biessential and if for every word $w \in F$, the graph $\mathcal{E}(w)$ is a tree (this corresponds to the definition of tree set of characteristic 1 in [13]). Note that a biessential set $F$ is a tree set if and only if the graph $\mathcal{E}(w)$ is a tree for every bispecial word $w$. Indeed the extension graph associated to every non-bispecial word is trivially a tree.

If the extension graph $\mathcal{E}(w)$ of $w$ is a tree, then $m(w) = 0$. Thus $w$ is neutral. Note that if $\mathcal{E}(w)$ is acyclic, one has $m(w) = 1 - c$ where $c$ is
the number of connected components of the graph $\mathcal{E}(w)$. Sturmian sets and regular interval exchange sets are example of tree sets (see [7]).

The sequence $(p_n)_{n \geq 0}$ with $p_n = \text{Card}(F \cap A^n)$ is called the factor complexity of $F$. Set $k = \text{Card}(F \cap A) - 1$. The factor complexity of a neutral set $F$ is equal to $kn + 1$ (see [7]). Since a tree set is neutral, we deduce that the factor complexity of a tree set is also $kn + 1$ (see [7]).

The following result shows that in neutral sets (and thus in tree sets) the notion of recurrence and uniformly recurrence coincide.

**Proposition 2.1** (Corollary 5.3 [13]) A recurrent neutral set is uniformly recurrent.

We similarly define a tree word as an infinite word $x$ such that its language $F(x)$ is a tree set, and a tree subshift as a subshift $(X, S)$ such that $F(X)$ is a tree set.

### 2.3 Morphisms and free group

A morphism $\sigma : A^* \to B^*$ is a monoid morphism from $A^*$ into $B^*$. We consider here exclusively non-erasing morphisms, that is, morphisms such that the image of every element in $A^+$ belongs to $B^+$. When $B = A$, such a morphism is a substitution. If there exists a letter $a \in A$ such that the word $\sigma(a)$ begins with $a$ and if $|\sigma^n(a)|$ tends to infinity with $n$, there exists a unique infinite word denoted $\sigma^\omega(a)$ which has all words $\sigma^n(a)$ as prefixes. Such an infinite word is called a fixpoint of the substitution $\sigma$.

A substitution $\sigma : A^* \to A^*$ is called primitive if there is a positive integer $k$ such that for all $a, b \in A$, the letter $b$ appears in $\sigma^k(a)$. If $\sigma$ is a primitive substitution, there exists a power $\sigma^k$ that admits a fixpoint, and the set of factors of any fixpoint of $\sigma$ (or of some power of $\sigma$) is uniformly recurrent (see for example [21, Proposition 1.2.3]).

An infinite word $x$ over the alphabet $A$ is said to be substitutive if there exist a substitution $\sigma$ over an alphabet $B$ with a fixpoint $y = \sigma^\omega(b)$, for some $b \in B$, and a morphism $\tau : B^* \to A^*$, such that $\tau(y) = x$. It is said substitutive primitive when $\sigma$ is primitive.

The incidence matrix (also called substitution matrix) of a substitution $\sigma$ defined over the alphabet $A$ is the square matrix of size the cardinality of $A$ such that its entry $(i, j)$, for $(i, j) \in A$, counts the number $|\sigma(j)|_i$ of occurrences of the letter $i$ in $\sigma(j)$.

The subshift $(X_\sigma, S)$ generated by a primitive substitution $\sigma$ over $A$ is the set of infinite words $x$ such that any word $w$ in the language $F(x)$ is
a factor of some $\sigma^n(a)$, for some $a \in A$ and some positive integer $n$. It is minimal.

**Example 2.2** Let $\sigma_F$ be the Fibonacci morphism defined over the alphabet \{a, b\} by $\sigma_F(a) = ab$ and $\sigma_F(b) = a$. The morphism $\sigma_F$ is a primitive substitution and the uniformly recurrent infinite word $\sigma_F^\omega(a) = abababaabaab \cdots$ is called the Fibonacci word. The incidence matrix of $\sigma_F$ is

$$
\begin{pmatrix}
1 & 1 \\
1 & 0
\end{pmatrix}
$$

The set of factors of the Fibonacci word, called the Fibonacci set, is a tree set (see, e.g., [7]) and the extensions graphs of $\varepsilon, a, b$ are shown in Figure 1.

We denote by $F_A$ the free group on the alphabet $A$. It is the set of all words on the alphabet $A \cup A^{-1}$ which are reduced, in the sense that they do not have any factor $aa^{-1}$ or $a^{-1}a$ for $a \in A$. A morphism $\sigma$ from $A^*$ to $A^*$ can be extended to a morphism from $F_A$ to $F_A$ by defining $\sigma(a^{-1}) = (\sigma(a))^{-1}$ for all $a \in A \cup A^{-1}$. A morphism $\sigma$ of the free monoid is said to be invertible if when extended to a morphism of the free group it is an automorphism, that is, there exists a morphism $\sigma^{-1}$ from $F_A$ to $F_A$ such that $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = Id$. In particular, if $\sigma$ is such that there exists $n$ such that $\sigma^n$ is an automorphism, then $\sigma$ is itself an automorphism.

An automorphism $\alpha$ of the free group on $A$ is positive if $\alpha(a)$ belongs to $A^+$ for every $a \in A$. We say that a positive automorphism of the free group on $A$ is tame if it belongs to the submonoid generated by the permutations of $A$ and the automorphisms $\alpha_{a,b}, \tilde{\alpha}_{a,b}$ defined for $a, b \in A$ with $a \neq b$ by

$$
\alpha_{a,b}(c) = \begin{cases} 
ab & \text{if } c = a, \\
\quad c & \text{otherwise}
\end{cases} \quad \text{and} \quad \tilde{\alpha}_{a,b}(c) = \begin{cases} 
ba & \text{if } c = a, \\
\quad c & \text{otherwise}.
\end{cases}
$$

Thus $\alpha_{a,b}$ places a letter $b$ after each $a$ and $\tilde{\alpha}_{a,b}$ places a letter $b$ before each $a$, without modifying the other letters. The above automorphisms and the permutations of $A$ are called the elementary positive automorphisms on $A$. We let $S_e$ denote the set of elementary positive automorphisms. This is a subset of the set of Nielsen’s transformations (see e.g. [28]). A substitution that extends as a positive automorphism that is tame is said to be a tame substitution.

The monoid of tame substitutions contains strictly the monoid of episturmian morphisms (also called Arnoux-Rauzy substitutions). Note that the case of a two-letter alphabet corresponds to the Stumian case. The monoid
of episturmian morphisms has been thoroughly investigated e.g. in [33]. It is generated by the permutations together with the set of automorphisms
\[ \psi, \tilde{\psi}, \] defined for \( a \in A \) by
\[ \psi_a(c) = \begin{cases} ac & \text{if } c \neq a, \\ a & \text{if } c = a. \end{cases} \]
\[ \tilde{\psi}_{a,b}(c) = \begin{cases} ca & \text{if } c \neq a, \\ a & \text{if } c = a. \end{cases} \]

The submonoid of the monoid of episturmian morphisms generated by the permutations together with the set of automorphisms \( \psi_a \), defined for \( a \in A \) (that is, no \( \tilde{\psi}_b \) is allowed) is called the monoid of epistandard morphisms. When all letters of the alphabet \( A \) occur, it is said to be strict.

Note also that the monoid of tame automorphisms is strictly included in the monoid of positive automorphisms. Indeed the monoid of positive automorphisms on an alphabet containing at least three letters is not finitely generated [33, 36, 38]. However invertible substitutions over a binary alphabet are exactly the Sturmian substitutions (see, e.g. [29, 37]), and the monoid of all invertible substitutions (i.e., the Sturmian monoid) is finitely generated.

Tame substitutions are closely related to tree words such as shown in Section 3 where a characterization of substitutive tree words is provided. But, not every tame substitution admits as a fixed point a tree word (see [8, Example 5.25]). The situation is thus more contrasted that in the Sturmian case (or in the episturmian case) where every Sturmian substitution generates a Sturmian word (see e.g. [27]). Note however that it is decidable whether the language of a primitive aperiodic substitution is a tree set (see [12]).

### 2.4 Stabilizers

The stabilizer of an infinite word \( x \in A^\mathbb{N} \), denoted by \( \text{Stab}(x) \), is the monoid of substitutions \( \sigma \) defined on the alphabet \( A \) that satisfy \( \sigma(x) = x \). Words that have a cyclic stabilizer are called rigid\(^1\).

We can also define the stabilizer of a shift. We will restrict to the minimal case. Let \( (X, S) \) be a minimal subshift. For a primitive substitution \( \sigma \), recall that \( X_\sigma \) stands for the shift generated by \( \sigma \). The primitive stabilizer of a shift \( X \) is the set of all primitive substitutions \( \sigma \) defined on the alphabet \( A \) that satisfy \( X_\sigma = X \). In other words, these are the primitive substitutions whose language coincides with the language of \( X \). A subshift \( (X, S) \) is said rigid if its stabilizer is cyclic.

\(^1\) Note that rigidity has nothing to do with the ergodic notion of rigidity.
Note that we concentrate here on the iterative stabilizer according to the terminology of [23]. Results on the possible growth of elements of the stabilizer are provided in [10] and [19]. It is shown in particular that polynomial and exponential growth cannot co-exist in the stabilizer for aperiodic words.

Words generated by Sturmian substitutions are rigid (see [35]). It is proved in [23] that fixpoints of strict epistandard morphisms are rigid. Epistandard morphisms belong to the monoid generated by the permutations together with the set of automorphisms \( \psi_a \), defined for \( a \in A \) (that is, no \( \tilde{\psi}_b \) is allowed). When all letters of the alphabet \( A \) occur, it is said to be strict. It is an open question to know whether fix points of episturmian morphisms are rigid.

Note also that the question of the existence of non-negative integers such that \( \sigma^n = \tau^p \) is decidable (see ref Pansiot 1981b page 536 du Allouche-Shallit).

2.5 Return words

Let \( F \subset A^* \) be a set of words. For \( w \in F \), let

\[
\Gamma_F(w) = \{ x \in F \mid wx \in F \cap A^+ w \} \quad \text{and} \quad \Gamma'_F(w) = \{ x \in F \mid xw \in F \cap wA^+ \}
\]

be respectively the set of right return words and of left return words to \( w \). If \( F \) is recurrent, the sets \( \Gamma_F(w) \) and \( \Gamma'_F(w) \) are nonempty. Let

\[
\mathcal{R}_F(w) = \Gamma_F(w) \setminus \Gamma_F(w)A^+ \quad \text{and} \quad \mathcal{R}'_F(w) = \Gamma'_F(w) \setminus A^+ \Gamma'_F(w)
\]

be respectively the set of first right return words and the set of first left return words to \( w \). Note that \( w \mathcal{R}_F(w) = \mathcal{R}'_F(w)w \). Note also that a recurrent set \( S \) is uniformly recurrent if and only if the set \( \mathcal{R}_S(w) \) is finite for any \( w \in S \).

Return words will play a crucial role in the following. The following theorem is proved in [7] and it is referred as the Return Theorem.

**Theorem 2.1 (Theorem 4.5 [7])** Let \( F \) be a recurrent tree set containing the alphabet \( A \). Then for any \( w \in S \), the set \( \mathcal{R}_F(w) \) is a basis of the free group \( \mathbb{F}_A \) on \( A \). Similarly, for any \( w \in S \), the set \( \mathcal{R}'_F(w) \) is a basis of the free group on \( A \).

3 Substitutive tree words

We first recall some basic definitions concerning \( S \)-adic representations in terms of return words. For more on \( S \)-adic words, see e.g. [2, 20, 24, 25, 26].
Let \( S \) be a set of morphisms. An infinite word \( x \in A^\mathbb{N} \) is said to be \( S \)-adic if there is a sequence of morphisms \( s = (\sigma_n : A_{n+1}^* \to A_n^*)_{n \in \mathbb{N}} \in S^\mathbb{N} \) and a sequence of letters \( a = (a_n \in A_n)_{n \geq 1} \) such that \( x = \lim_{n \to +\infty} \sigma_0 \sigma_1 \cdots \sigma_n(a_{n+1}) \). The pair \((s, a)\) is called an \( S \)-adic representation of \( x \) and the sequence \( s \) a directive sequence of \( x \). The pair \((s, a)\) is eventually periodic if there exists \( N, n \in \mathbb{N} \) such that \( \sigma_{m+n} = \sigma_m \) and \( a_{m+n} = a_m \) for all \( m \geq N \). The pair \((s, a)\) is said to be primitive whenever the directive sequence \( s \) is primitive, i.e., for all \( r \geq 0 \), there exists \( r' > r \) such that all letters of \( A_r \) occur in all images \( \sigma_r \sigma_{r+1} \cdots \sigma_r(a) \), for all \( a \in A_{r+1} \). Observe that if \( x \) has a primitive \( S \)-adic representation then \( x \) is uniformly recurrent (see, e.g., [16]). If \( X \) is a minimal subshift and if \( x \in X \) admits an \( S \)-adic representation \((s, a)\), then we say that \((s, a)\) is also an \( S \)-adic representation of \( x \).

Return words provide \( S \)-adic representations of uniformly recurrent words. Indeed, let \( x = (x_n)_{n \in \mathbb{N}} \in A^\mathbb{N} \) be a uniformly recurrent word. We consider left return word with respect to the first letter \( x_0 \) of \( x \) (note that such a return word starts with \( x_0 \)). There thus exists a unique sequence \((w_n)_{n \in \mathbb{N}} \in (\mathcal{R}(x_0))^\mathbb{N} \) such that \( x = w_0 w_1 w_2 \cdots \). Let \( R \) be the alphabet \( \{1, 2, \ldots, \text{Card}(\mathcal{R}(x_0))\} \) and consider the return morphism \( \lambda : R^* \to A^* \) such that \( \lambda(R) = \mathcal{R}(x_0) \) and \( \lambda(i) \) is the \( i \)-th return word occurring in \((w_n)_{n \in \mathbb{N}}\) for all \( i \). More precisely, \( \lambda \) is such that for all \( i \in R \), if \( m \) is the smallest integer such that \( w_m = \lambda_i \), then \( \{w_0, w_1, \ldots, w_{m-1}\} = \{\lambda(j) \mid 1 \leq j < i\} \).

The derived sequence of \( x \) is the unique sequence \( \mathcal{D}(x) \in R^\mathbb{N} \) such that \( \lambda(\mathcal{D}(x)) = x \). We recursively define \( \mathcal{D}^n(x), \lambda_n \) and \( R_n \) by \( \mathcal{D}^0(x) = x, R_0 = A \) and \( \mathcal{D}^n(x) = \mathcal{D}^{n-1}(x) \in R_n^\mathbb{N} \) with \( \lambda_n : R_{n+1}^* \to A^* \) the return morphism of \( \mathcal{D}^n(x) \).

Let us consider the set of morphisms \( \Lambda = \{\lambda_n \mid n \in \mathbb{N}\} \). We denote \( \lambda = (\lambda_n)_{n \in \mathbb{N}} \) and \( 1 = (1)_{n \geq 1} \) the constant sequence that takes values 1. The pair \((\lambda, 1)\) is clearly primitive (see also [8, Proposition 5.22]). We call it the \( \Lambda \)-adic representation of \( x \). For all \( n \geq 1 \), we set \( \theta_n = \lambda_0 \lambda_1 \cdots \lambda_{n-1} : R_n^* \to A^* \). Thus we have \( \theta_n(\mathcal{D}^n(x)) = x \).

The \( \Lambda \)-adic representation allows the following formulation for characterizing primitive substitutive words.

**Theorem 3.1 ([15])** A uniformly recurrent word \( x \in A^\mathbb{N} \) is primitive substitutive if and only if the set of its derived sequence \( \{\mathcal{D}^n(x) \mid n \in \mathbb{N}\} \) is finite.

The next result implies that any recurrent tree words admits a primitive \( S_c \)-adic representation (see also [8, Theorem 5.23]). The main property used below is that in a recurrent tree set the set of return words of a given
word forms a basis of the free group ([7, Theorem 4.5]). We recall that $\mathcal{S}_e$ stands for the set of elementary positive automorphisms such as defined in Section 2.3.

**Theorem 3.2** Let $x \in A^\mathbb{N}$ be a recurrent tree word over the alphabet $A = \{1, \ldots, d\}$ and let $(\lambda, 1)$ be its $\Lambda$-adic representation.

1. For all $n \in \mathbb{N}$, $D^n(x) \in A^\mathbb{N}$ is a recurrent tree word.

2. For all $n \in \mathbb{N}$, the morphism $\lambda_n : A^* \to A^*$ extends to a tame automorphism of $\mathbb{F}_A$.

In other words, the $\Lambda$-adic representation of $x$ is an $\mathcal{S}_e$-adic representation.

**Proof.** Item 1 follows from [8, Theorems 5.13]. Item 2 follows from [7, Theorem 4.5] and [8, Theorem 5.19].\[\]

We now can state the main result of this section that provides a characterization of primitive substitutive tree words. A similar statement is known to hold for Sturmian words (see e.g. [5]). In this later case it can even be expressed in terms of ultimately periodic continued fraction and Ostrowski expansion,

**Theorem 3.3** A recurrent tree word is primitive substitutive if and only if it has an eventually periodic primitive $\mathcal{S}_e$-adic representation.

**Proof.** One easily checks that the condition is sufficient. The necessary part follows from Theorem 3.1: the set $\{D^n(x) \mid n \in \mathbb{N}\}$ being finite, there exists $m, n \in \mathbb{N}$, $m < n$ such that $D^m(x) = D^n(x)$. By construction of $\lambda$, this means that for all $k \in \mathbb{N}$, $D^{m+k}(x) = D^{n+k}(x)$ and $\lambda_{m+k} = \lambda_{n+k}$, i.e., $(\lambda, 1)$ is eventually periodic. We then apply Theorem 3.2.\[\]

Observe that if the $\Lambda$-adic representation of a tree word $x$ is purely periodic, then $x$ is the fixpoint of a primitive substitution. The converse is not true such as illustrated by the substitution $\sigma : 0 \mapsto 010, 1 \mapsto 10$ and its fix point $x = \sigma^\omega(0)$. Indeed, one has $x = \lambda_0\lambda_1(\lambda_2)^\omega$ where $\lambda_0 : 0 \mapsto 01, 1 \mapsto 0$, $\lambda_1 : 0 \mapsto 0, 1 \mapsto 01$, $\lambda_2 : 0 \mapsto 01, 1 \mapsto 011$.

4 **Stabilizers of tree words**

The following theorem states that one has a weak form rigidity for tree words or tree subshifts together with a structure theorem for the stabilizer
Theorem 4.1 Let $x$ be a tree word. Primitive substitutions in the stabilizer $\text{Stab}(x)$ of $x$ coincide up to powers. More precisely, if $x$ is a fixpoint of both $\sigma$ and $\tau$ primitive substitutions, then there exist $i, j \geq 1$ such that $\tau^i = \sigma^j$.

Let $x$ be a recurrent substitutive tree word. There is a primitive tame substitution $\theta$ such that any primitive substitution $\sigma \in \text{Stab}(x)$ has a power that is conjugate to a power of $\theta$, that is, there exist a positive automorphism $\tau$ such that $\sigma^i = \tau \theta^j \tau^{-1}$, for some $i, j$.

In particular, if $x$ is a tree word, any primitive substitution in $\text{Stab}(x)$ extends to an automorphism of the free group and is a tame substitution.

Note that the first statement implies that the Perron-Frobenius eigenvalues of $\sigma$ and $\tau$ are multiplicatively dependent, which is also a consequence of Cobham’s theorem [18].

Proof. The first statement is a direct consequence of [17, Corollary 22]. Indeed, given any finite word $w$, the set of Parikh vectors of returns words to $w$ generates $\mathbb{Z}^d$ by the Return theorem (see Theorem 2.1).

Now, let $x$ be a recurrent tree word. If $x$ is primitive substitutive, then the set $\{D^n(x) \mid n \in \mathbb{N}\}$ is finite by Theorem 3.1. Let $(\lambda, 1)$ be its $\Lambda$-adic representation, and, $k, l \in \mathbb{N}$, $k < l$, be such that $D^k(x) = D^l(x)$ and all derived sequences in $\{D^n(x) \mid 0 \leq n < l\}$ are pairwise distinct. Let $\theta$ be the morphism such that $\theta_l = \theta_k \theta$, i.e., $\theta = \lambda_k \lambda_{k+1} \cdots \lambda_{l-1}$. Since $((\lambda_n)_{n \geq k}, 1)$ is a primitive $\Lambda$-adic representation of $D^k(x)$, $\theta$ is a primitive substitution having $D^k(x)$ for fixed point.

Let $\sigma$ be a primitive substitution in $\text{Stab}(x)$. There exists a primitive substitution $\sigma_k$ satisfying $\sigma \theta_k = \theta_k \sigma_k$ and having $D^k(x)$ for fixed point, by [15, Proposition 5.1]. By Theorem 3.2 and by the first statement, we have $\sigma^j_k = \theta^j$ for some $i, j \geq 1$. We thus get $\sigma^i = \theta_k \theta^j \theta_k^{-1}$, which finishes the proof of the second statement.

We deduce that if $\sigma$ is a primitive element of $\text{Stab}(x)$, then it is invertible. By [8, Theorem 5.19], it is thus a tame substitution.

5 Tree sets and continuous eigenvalues

We prove in this section that aperiodic minimal tree subshifts cannot be generated by substitutions of constant length (see Corollary 5.1 below). We
provide here a spectral proof and prove the more general result that aperiodic minimal tree subshifts cannot have rational eigenvalues. Let us start by recalling some definitions.

Let \((X, S)\) be a subshift. We say that \((X, S)\) is totally minimal whenever \((X, S^n)\) is minimal for all \(n\).

A cyclic partition of \((X, S)\) is a partition \(X = \bigcup_{i=1}^{m} X_i\) in closed subsets such that \(X_{i+1} = S(X_i)\), for \(1 \leq i \leq m - 1\), and \(S(X_m) = X_1\). Note that the elements \(X_i\) are thus clopen sets.

A topological eigenvalue of \((X, S)\) is a complex number \(\lambda\) such that there exists a nonzero continuous function \(f: X \to \mathbb{C}\) satisfying \(f \circ S = \lambda f\); \(f\) is called a continuous eigenfunction associated with \(\lambda\). A continuous eigenvalue of the form \(\exp(2i\pi k/n)\), for some \(n \geq 2\) and some integer \(k\), is said to be a rational continuous eigenvalue.

**Example 5.1** Let \(\sigma_{TM}\) be the Thue-Morse substitution defined on \(\{a, b\}^*\) by \(\sigma_{TM}(a) = ab\), \(\sigma_{TM}(b) = ba\). Let \(X_1 = \sigma[a] \cup \sigma[b]\). The partition \((X_1, SX_1)\) is a cyclic partition of the subshift \(X_{TM}\) generated by \(\sigma_{TM}\). Indeed, one checks that \(S^2 X_1 = X_1\), and moreover that \(X_1\) and \(SX_1\) are disjoint, by recognizability of the Thue-Morse substitution [30, 31].

The eigenvalue \(i = \exp(i\pi)\) is a continuous eigenvalue. Indeed, consider the function \(f\) that maps every element of \([a]\) to the constant value \(i\), and every element of \(S[a]\) to the constant value \(1\). One has \(f(Sx) = if(x)\) for every element \(x\) of \(X_{TM}\).

The following result is part of the folklore of topological dynamical systems (see e.g. [32]). It shows that these notions are intimately related.

**Lemma 5.1** Let \((X, S)\) be a minimal subshift. The following are equivalent.

1. \((X, S)\) has a cyclic partition \(X_1, \ldots, X_n\) for some \(n \geq 2\);
2. \(\exp(2i\pi/n)\) is a continuous eigenvalue of \((X, S)\) for some \(n \geq 2\);
3. \((X, S)\) is not totally minimal.

**Proof.** It is clear that (3) is equivalent to (1).

Let us prove that (1) implies (2). Let \(f\) be defined as the constant function taking the value \(\exp(2i\pi k/n)\) on \(X_k\). One checks that \(f\) is a continuous eigenfunction associated with the eigenvalue \(\exp(2i\pi/n)\).

Conversely, let us prove that (2) implies (1). Let \(\exp(2i\pi/n)\) be a continuous eigenvalue of \((X, S)\), and let \(f\) be a continuous eigenfunction for this eigenvalue. One can suppose \(f(x) = 1\) for some \(x \in X\). One has
\( f(S^nx) = f(x) = 1 \). By minimality, every element \( y \) of \( X \) can be written as 
\( y = \lim S^{n_i}(x) \), for some non-decreasing subsequence \((n_i)\). There exists \( k \) 
with \( 0 \leq k \leq n - 1 \) such that infinitely many \( n_i \) are congruent to \( k \) modulo \( n \). By continuity of \( f \), one has 
\( y = \exp(2i\pi k/n)f(x) = \exp(2i\pi k/n) \). Hence 
\( f(X) = \{\exp(2i\pi k/n) \mid 0 \leq k \leq n - 1\} \). Let \( X_k = f^{-1}(\{\exp(2i\pi k/n)\}) \), for 
\( 0 \leq k \leq n - 1 \). This defines a cyclic partition.

We now prove that minimal tree subshifts are totally minimal. We use 
below the fact that the set of factors of a given length over the alphabet 
\( A \) is a code, as well as the properties of stability of tree sets by maximal 
bifix decoding, and the fact that recurrent tree sets are in fact uniformly 
recurrent. Recall that a set \( X \subset A^+ \) of nonempty words over the alphabet 
\( A \) is a code if the relation 
\[
x_1 \cdots x_n = y_1 \cdots y_m
\]
with \( n, m \geq 1 \) and \( x_1, \ldots, x_n, y_1, \ldots, y_m \in X \) implies \( n = m \) and \( x_i = y_i \) 
for \( i = 1, \ldots, n \). A coding morphism for a code \( X \subset A^+ \) is a morphism 
\( f: B^* \to A^* \) which maps bijectively \( B \) onto \( X \).

**Proposition 5.1** A minimal tree subshift is totally minimal.

**Proof.** Let \((X, S)\) be a minimal tree subshift. Let \( n \geq 2 \). We consider 
the code \( G \) made of the factors of the language of \( X \) of length \( n \), that 
is, \( G = F(X) \cap A^n \). Let \( f: B \to G \) be a coding morphism for \( G \). Let 
\( H = f^{-1}(F(X)) \) and let \((Y, T)\) be the subshift defined by \( F(Y) = H \), with 
\( T \) standing for the shift acting on \( Y \). One thus has \( Y \subset B^n \), with \( B \) being 
in bijection with the set \( G \) of factors of length \( n \) of \( X \). Then \( f \) extends to an 
isomorphism from \((Y, S)\) onto \((X, S^n)\). Since \( G \) is an \( F(X) \)-maximal bifix 
code, the set \( H \) is a recurrent tree set by [8, Theorem 6.1]. By Proposition 
2.1 (i.e., by [13, Corollary 5.3]) this implies that \((Y, T)\) is minimal, which 
yields that \((X, S^n)\) is minimal.

We say a minimal subshift is periodic whenever it is finite. Otherwise it 
is said to be aperiodic. We now can state the main result of this section.

**Theorem 5.1** Let \((X, S)\) be an aperiodic minimal tree subshift. Then it 
admits no rational continuous eigenvalue.

**Proof.** This is a direct consequence of Proposition 5.1 together with Lemma 
5.1.
Example 5.2 Consider the Tribonacci substitution $\sigma_T$ defined on $\{a, b, c\}^*$ that satisfies $\sigma_T(a) = ab$, $\sigma_T(b) = ac$, $\sigma_T(c) = a$. Let $X_T$ be the shift generated by $\sigma_T$. The subshift $X_T$ is an aperiodic tree subshift by [1]. It thus admits no rational continuous eigenvalue. Let $\tau: \{a, b\}^* \to \{a, b, c\}^*$ be the morphism defined by $\tau(a) = aa$, $\tau(b) = ab$, $\tau(c) = ac$. Let $X$ be the subshift generated by $\tau \circ \sigma_T$. Let $X_1 = [a]$. The partition $([a], S[a])$ is a cyclic partition of $X$. Indeed, one checks that $S^2[a] = [a]$. The shift $(X, S)$ admits $i = \exp(i\pi)$ as a rational continuous eigenvalue. The shift $(X, S)$ is thus not a tree shift.

We recall that a shift generated by a primitive constant length substitution admits rational continuous eigenvalues (see e.g. [11]). A subshift $X$ is said to be Toeplitz if any element $x = (x_n)_n \in X$ satisfies the following: for all $n$, there exists a positive integer $p$ such that $x_n = x_{n+kp}$, for all $k$. For more on Toeplitz shifts, see e.g. [14]. Toeplitz shifts are also known to have rational continuous eigenvalues.

Corollary 5.1 Let $(X, S)$ be an aperiodic minimal tree subshift. Then, it can neither be generated by a primitive constant-length substitution, nor a Toeplitz shift.

6 Concluding questions

Tree words generalize Sturmian and episturmian words. They are also defined combinatorially but they exhibit a larger diversity of behaviors. Let us focus on the comparison between tree words and episturmian words for what concerns substitutions.

It is proved in [23] that fixed points of epistandard morphisms are rigid. We are not able here to answer the question asked in [23] on the rigidity of strict episturmian words, even if Theorem 4.1 provides some elements of answer, but we extend this question to the general framework of tree words. We then ask the following: are uniformly recurrent tree words or minimal tree subshifts rigid?

By analogy with episturmian morphisms, we introduce the monoid of \textit{standard tame substitutions} as the monoid generated by the permutations of $A$ and the automorphisms $\alpha_{a,b}$, for $a, b \in A$ with $a \neq b$, and the monoid of \textit{antistandard tame substitutions} as the monoid generated by the permutations of $A$ and the automorphisms $\tilde{\alpha}_{a,b}$, for $a, b \in A$ with $a \neq b$. There is a simple characterization of elements of the monoid of standard tame substitutions which extends [33, Lemma 2.4]. A tame substitution is standard.
(resp. antistandard) if and only if the set of the first (resp. last) letters of the images of letters in $A$ is equal to $A$.

One notable difference with the episturmian case is that the tame automorphism $\alpha_{a,b}$ is not rotationally conjugate to $\tilde{\alpha}_{a,b}$, whereas $\tilde{\psi}_a$ is conjugate to $\psi_a$. We say that two substitutions $\sigma$ and $\rho$ over the finite alphabet $A$ are rotationally conjugate if $\sigma = \gamma_w \circ \rho$ for some $w \in A^*$, where $\gamma_w$ is the inner automorphism of $F_A$ defined by $\gamma_w(x) = wxw^{-1}$, for all $x \in F_A$. The substitutions $\sigma$ and $\rho$ thus belong to the same outer class, they are obtained by action of an inner automorphism.

Another difference comes from the fact that not every fixpoint of a tame substitution is a tree set.

Several questions occur naturally. Do fixpoints of standard tame substitutions play the same role for tree words as in the episturmian case, in particular with respect to left special factors? Is this notion relevant in the present context? What can be said on tame substitutions that have the same incidence matrix? What can be said when one exchanges $\tilde{\alpha}_{a,b}$ with $\tilde{\alpha}_{a,b}$ in a decomposition of a tame substitution? Can rotational conjugation (by composition by inner automorphism) be seen on the decomposition by elementary morphisms of a substitution that preserves a tree set?

Let $x$ be a uniformly recurrent tree set. If $x$ is a fixpoint of a substitution, is this substitution primitive? in other words, is any nontrivial element of $\text{Stab}(x)$ primitive? Recall that any sequence which is substitutive with respect to a substitution with polynomial growth and a substitution with exponential growth is ultimately periodic (see [19]).

Can one characterize in terms of the $S_e$-directive sequence of a tree word (see Theorem 3.2) when the stabilizer of a recurrent tree word is nontrivial? More generally, can one characterize among substitutive tree words the tree words that are fixpoints of substitutions? For the two-letter Sturmian case, see e.g. [4] and the references therein.

The very fact that the family of tree subshifts encompasses Arnoux-Rauzy subshifts shows the diversity of spectral behaviors of minimal tree subshifts. Indeed, by [6], Arnoux-Rauzy subshifts can have continuous eigenvalues or not; the same holds for measure-theoretic ones. Note that if we focus on tree subshifts generated by primitive substitutions, it is known that measure-theoretical and continuous eigenvalues are the same [22]. One way to tackle spectral questions concerning tree subshifts is to interpret the $S_e$-adic representation in terms of continued fractions and understand the underlying convergence.
References


