Note on sets of first return

Valérie Berthé1, Clelia De Felice3, Francesco Dolce2, Dominique Perrin2, Christophe Reutenauer4, Giuseppina Rindone2

1CNRS, Université Paris 7, 2Université Paris Est, 3Università degli Studi di Salerno, 4Université du Québec à Montréal

May 11, 2014

Contents

1 Introduction 1

2 Preliminaries 2

2.1 Recurrent sets ........................................... 2

2.2 Tree sets .............................................. 3

2.3 Sets of first right return words .......................... 3

3 Main result 5

3.1 Stallings foldings of Rauzy graphs ...................... 6

3.2 Types of Rauzy graphs ................................. 7

3.3 The finite case ........................................ 8

3.4 The infinite case ...................................... 9

3.5 Examples ............................................ 12

Abstract

We prove that for an alphabet $A$ with three letters, the set of first return to a given word in a uniformly recurrent set satisfying the tree condition is a basis of the free group on $A$.

1 Introduction

Tree sets are symbolic dynamical systems subject to a restriction on the possible extensions of a given word. These sets are introduced in [2] as a common generalization of Sturmian sets and of regular interval exchange sets.

We prove in this note that, for an alphabet $A$ with three letters, the set of first return to a given word in a uniformly recurrent set satisfying the tree condition is a basis of the free group on $A$. 

1
In the first section, we prove some preliminary results concerning bispecial words. We define extension graphs and tree sets. In Section 3 we state our main result (Theorem 7). The proof uses a classification of possible Rauzy graphs. We end we some examples.

2 Preliminaries

In this section, we first recall some definitions concerning words. We give the definition of recurrent and uniformly recurrent sets of words. (see [1] for a more detailed presentation). We also give the definitions and basic properties of tree sets. We prove some preliminary properties of sets of first return words.

2.1 Recurrent sets

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 \( A^* \) the set of all words on \( A \). We denote by \( 1 \) or by \( \varepsilon \) the empty word.

For a set \( X \) of words and a word \( u \), we denote \( u^{-1}X = \{v \in A^* | uv \in X\} \) the right residual of \( X \) with respect to \( u \). This notation should not be confused with the notation for the inverse in the free group on \( A \).

A set of words is said to be factorial if it contains the factors of its elements.

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

\[
L(w) = \{a \in A | aw \in F\} \\
R(w) = \{a \in A | wa \in F\} \\
E(w) = \{(a, b) \in A \times A | awb \in F\}
\]

and further

\[
\ell(w) = \text{Card}(L(w)), \quad r(w) = \text{Card}(R(w)), \quad e(w) = \text{Card}(E(w)).
\]

A word \( w \) is right-extendable if \( r(w) > 0 \), left-extendable if \( \ell(w) > 0 \) and biextendable if \( e(w) > 0 \). A factorial set \( F \) is called right-essential (resp. left-essential, resp. biessential) if every word in \( F \) is right-extendable (resp. left-extendable, resp. biextendable).

A word is left-special if \( \ell(w) \geq 2 \), right-special if \( r(w) \geq 2 \), bispecial if it is both left-special and right-special.

A set of words \( F \) is recurrent if it is factorial and if for every \( u, w \in F \) there is a \( v \in F \) such that \( uvw \in F \). A recurrent set \( F \neq \{1\} \) is biessential.

A set of words \( F \) is said to be uniformly recurrent if it is right-essential and if, for any word \( u \in F \), there exists an integer \( n \geq 1 \) such that \( u^nw \in F \). A uniformly recurrent set is recurrent.
2.2 Tree sets

Let $F$ be a biessential set. For $w \in F$, we consider the extension graph of $w$ as the undirected set $G(w)$ on the set of vertices which is the disjoint union of $L(w)$ and $R(w)$ with edges the pairs $(a, b) \in E(w)$.

Recall that an undirected graph is a tree if it is connected and acyclic.

We say that $F$ is a tree set if it is biessential and if for every word $w \in F$, the graph $G(w)$ is a tree. A tree set has complexity $kn + 1$ with $k = \text{Card}(A \cap F) - 1$ (see [2], Proposition 3.2).

Let $F$ be a set of words. For $w \in F$, and $U, V \subseteq A^*$, let $U(w) = \{\ell \in U \mid \ell w \in F\}$ and let $V(w) = \{r \in V \mid wr \in F\}$. The generalized extension graph of $w$ relative to $U, V$ is the following undirected graph $G_{U,V}(w)$. The set of vertices is made of two disjoint copies of $U(w)$ and $V(w)$. The edges are the pairs $(\ell, r)$ for $\ell \in U(w)$ and $r \in V(w)$ such that $\ell wr \in F$. The extension graph $G(w)$ defined previously corresponds to the case where $U, V = A$.

**Example 1** Let $F$ be the Fibonacci set. Let $w = a$, $U = \{aa, ba, b\}$ and let $V = \{aa, ab, b\}$. The graph $G_{U,V}(w)$ is represented in Figure 1

![Figure 1: The graph $G_{U,V}(w)$.](image)

The following property is proved in [3]. It shows that in a tree set, not only the extension graphs but all generalized extension graphs are trees.

**Proposition 2** Let $F$ be a tree set. For any $w \in F$, any finite $F$-maximal suffix code $U \subseteq F$ and any finite $F$-maximal prefix code $V \subseteq F$, the generalized extension graph $G_{U,V}(w)$ is a tree.

Note that the statement also holds under the (apparently) weaker hypothesis that $U$ is a finite $w^{-1}$-maximal suffix code and $V$ a finite $w^{-1}$-$F$-maximal prefix code. Indeed, let then $U'$ be a finite $F$-maximal suffix code containing $U$ and $V'$ be a finite $F$-maximal prefix code containing $V$. Then $U'(w) = U(w)$ and $V'(w) = V(w)$. Thus $G_{U',V'}(w) = G_{U',V'}(w)$ and thus $G_{U,V}(w)$ is a tree.

2.3 Sets of first right return words

For $x \in F$, let $\Gamma_F(x) = \{r \in F \mid xr \in F \cap A^+x\}$ be the set of right return words to $x$ and let $R_F(x) = \Gamma_F(x) \setminus \Gamma_F(x)A^+$ be the set of first right return words to $x$. A word $x$ is unioccurrent in a word $y$ if there exist unique words $u, v$ such that $y = uxv$.

A set $F$ is periodic if it is the set of factors of $w^*$ for some word $w$. If $\text{Card}(A) \geq 2$, a set satisfying the tree condition is not periodic since otherwise its complexity would be bounded by a constant.
Proposition 3 Let $F$ be a recurrent set which is not periodic. Any word $x \in F$ is a factor of a bispecial word $y$. There is a unique shortest bispecial word $y$ containing $x$ and $x$ is unioccurrent in $y$. If $y = uxv$, one has $R_F(x) = vR_F(y)v^{-1}$.

Note that in the above statement, since all the elements of $R_F(y)$ end with $v$, the notation $v^{-1}$ can be indifferently interpreted as the residual or the inverse of $v$ in the free group.

Let $F$ be a factorial set. The Rauzy graph of order $n \geq 1$ is the following labeled graph $G_n$. Its vertices are the words in the set $F \cap A^n$. Its edges are the triples $(x, a, y)$ for all $x, y \in F \cap A^n$ and $a \in A$ such that $xa \in F \cap Ay$.

Observe that when $F$ is recurrent, any Rauzy graph is strongly connected.

Indeed, let $u, w \in F \cap A^n$. Since $F$ is recurrent, there is a $v \in F$ such that $uvw \in F$. Then there is a path in $G_n$ from $u$ to $w$ labeled $uv$.

Lemma 4 Let $F$ be a recurrent set. If $F$ is not periodic, any word of $F$ is a prefix of a right-special word and a suffix of a left-special word.

Proof. Let $x \in F$ be of length $n$. If $F$ is not periodic, there is a vertex of the Rauzy graph $G_n$ which is the origin of more than one edge (otherwise, the graph $G_n$ is reduced to a cycle and $F$ is periodic). Thus there is a word $y$ of length $n$ which is right-special. Thus the label $v$ of path from $x$ to $y$ is such that $xv$ is right-special.

The proof that $x$ is a suffix of a left-special word is symmetrical. ■

Lemma 5 Let $F$ be a recurrent set. If $F$ is not periodic, any word in $F$ is a factor of a bispecial word. More precisely, if $u$ is of minimal length such that $ux$ is left-special and $v$ of minimal length such that $xv$ is right-special, then $y = uxv$ is a bispecial word containing $x$ of minimal length and $x$ is unioccurrent in $y$.

Proof. Let us show by induction on the length of a prefix $v'$ of $v$ that $auxv' \in F$ for any $a \in L(ux)$. This is true if $v'$ is empty. Otherwise, set $v' = v'b$ with $b \in A$. For any $a \in L(ux)$, we have $auxv'' \in F$ by induction hypothesis. On the other hand, by the minimality of $v$, $xv''$ is not right-special and thus $auxv''$ is not right-special. This forces $auxv''b \in F$, which proves the property for $v'$. This implies that $uxv$ is left-special. The proof that $uxv$ is right-special is symmetrical. If $y = pxs$ is a bispecial word containing $x$, then $px$ is left-special and $xs$ is right-special. Thus $|p| \geq |u|$ and $|s| \geq |v|$ and thus $|y| \geq |uxv|$. Finally, assume that $x$ has a second occurrence in $y$. Then $y = u'xv'$ and either $|u'| < |u|$ or $|v'| < |v|$. Both options are impossible and thus $x$ is unioccurrent in $y$. ■

Proof of Proposition 3 We may assume $\text{Card}(A) \geq 2$. The first assertion is a consequence of Lemma 5.

For the second assertion, let $u$ be of minimal length such that $ux$ is left-special and $v$ of minimal length such that $xv$ is left-special. By Lemma 5 $y = uxv$ is a bispecial word containing $x$ of minimal length and $x$ is unioccurrent in $y$. Since $u$ is of minimal length, any extension of $x$ to the left is comparable...
with $u$ for the suffix order. In the same way any extension of $x$ to the right is comparable with $v$ for the prefix order. Thus there cannot be another bispecial word $u'xv'$ of the same length since otherwise $|u'| < |u|$, contradicting the hypothesis on $u$ since $u'x$ is left-special or $|v'| < |v|$, contradicting the hypothesis on $v$. Assume that $r \in R_F(x)$. Set $xr = sx$. Then $u$ is comparable with $s$ for the prefix order. Similarly $v$ is comparable with $r$ for the prefix order. This implies that $u$ is a suffix of $s$ and $v$ is a prefix of $r$ since otherwise $x$ would have a second occurrence in $y$.

Figure 2: First returns to $x$ and $y$.

Set $s = s'u$ and $r = vr'$. Then

$$yv^{-1}rv = u xv v^{-1}rv = u xv v'v = u xv = usvx = us'u xv = us'y.$$  

This implies that $v^{-1}rv$ is in $\Gamma_F(y)$ (see Figure 2). Since $r \in R_F(x)$, this implies actually that $v^{-1}rv \in R_F(y)$. The converse holds for the same reasons.

Example 6 Let $A = \{a, b\}$ and let $F$ be the Fibonacci set. The shortest bispecial word containing $x = aa$ is $y = abaaba$. One has $R_F(x) = \{baa, babaa\}$ and $R_F(y) = \{aba, baaba\}$. Thus $R_F(x) = baR_F(y)(ba)^{-1}$ as asserted in Proposition 3.

3 Main result

We will prove the following result.

Theorem 7 Let $A$ be three letter alphabet. Let $F$ be a uniformly recurrent set on the alphabet $A$ containing $A$ and satisfying the tree condition. For any $x \in F$, the set $R_F(x)$ is a basis of the free group on $A$.

The proof uses some preliminary results. Proposition 3 shows that it is enough to prove Theorem 7 for a bispecial word $x$.

Let $G$ be a graph labeled by words on the alphabet $A$ and let $E$ be its set of edges. We define $E^{-1} = \{(p, x^{-1}, q) \mid (q, x, p) \in E\}$. A generalized path in $G$ is a path in the graph having $E \cup E^{-1}$ as set of edges. Given a vertex $v$ of $G$, the set of labels of generalized paths in $G$ from $v$ to $v$ is a subgroup of the free group on $A$. It is called the group defined by the graph $G$ with respect to $v$. We will prove the following statement (which holds without hypothesis on the number of letters).
**Proposition 8** The group defined by any Rauzy graph of a tree set containing the alphabet $A$ with respect to any vertex is the free group on $A$.

We will prove by checking the possible types of Rauzy graphs the following result.

**Proposition 9** Let $F$ be a uniformly recurrent set on a three letter alphabet $A$, containing $A$ and satisfying the tree condition. For any bispecial word $x$, the set $R_F(x)$ has three elements and generates the group defined by the Rauzy graph of order $n = |x|$.

Together with Proposition 8, this implies that $R_F(x)$ is a basis of the free group on $A$ and thus, by Proposition 8, Theorem 7.

### 3.1 Stallings foldings of Rauzy graphs

A morpshism $\varphi$ from a labeled graph $G$ onto a labeled graph $H$ is a map from the set of vertices of $G$ onto the set of vertices of $H$ such that $(u, a, v)$ is an edge of $H$ if and only if there is an edge $(p, a, q)$ of $G$ such that $\varphi(p) = u$ and $\varphi(q) = v$. An isomorphism of labeled graphs is a bijective morphism.

The quotient of a labeled graph $G$ by an equivalence $\theta$, denoted $G/\theta$, is the graph with vertices the set of equivalence classes of $\theta$ and an edge from the class of $u$ to the class of $v$ labeled $a$ if there is an edge labeled $a$ from a vertex $u'$ equivalent to $u$ to a vertex $v'$ equivalent to $v$. The map from a vertex of $G$ to its equivalence class is a morphism from $G$ onto $G/\theta$.

We consider on the graph $G_n$ the equivalence $\theta_n$ formed by the pairs $(u, v)$ with $u = ax$, $v = bx$, $a, b \in L(x)$ such that there is a path from $a$ to $b$ in the graph $G(x)$ (and more precisely from the vertex corresponding to $a$ to the vertex corresponding to be in the copy corresponding to $L(x)$ in the bipartite graph $G(x)$).

**Proposition 10** If $F$ satisfies the tree condition, for each $n \geq 1$, the quotient of $G_n$ by the equivalence $\theta_n$ is isomorphic to $G_{n-1}$.

**Proof.** The map $\varphi : A^n \to A^{n-1}$ mapping a word of length $n$ to its suffix of length $n-1$ is clearly a morphism from $G_n$ onto $G_{n-1}$. If $u, v \in A^n$ are equivalent modulo $\theta_n$, then $\varphi(u) = \varphi(v)$. Thus there is a morphism $\psi$ from $G_n/\theta_n$ onto $G_{n-1}$. It is defined for any word $u \in F \cap A^n$ by $\psi(\bar{u}) = \varphi(u)$ where $\bar{u}$ denotes the class of $u$ modulo $\theta_n$. But since $F$ satisfies the tree condition, the class modulo $\theta_n$ of a word $ax$ of length $n$ has $\ell(x)$ elements, which is the same as the number of elements of $\varphi^{-1}(x)$. This shows that $\psi$ is an isomorphism. ■

Let $G$ be a labeled graph. A Stallings folding at vertex $v$ relative to letter $a$ of $G$ consists in identifying the edges coming into $v$ labeled $a$ and identifying their origins. A Stallings folding does not modify the group defined by the graph. **Proof of Proposition** The quotient $G_n/\theta_n$ can be obtained by a sequence of Stallings foldings from the graph $G_n$. Indeed, a Stallings folding at vertex $v$
identifies vertices which are equivalent modulo $\theta_n$. Conversely, consider $u = ax$ and $v = bx$, with $a, b \in A$ such that $a$ and $b$ (considered as elements of $L(x)$), are connected by a path in $G(x)$. Let $a_0, \ldots, a_k$ and $b_1, \cdots b_k$ with $a = a_0$ and $b = a_k$ be such that $(a_i, b_{i+1})$ for $0 \leq i \leq k - 1$ and $(a_i, b_i)$ for $1 \leq i \leq k$ are in $E(x)$. The successive Stallings foldings at $xb_1, xb_2, \ldots, xb_k$ identify the vertices $u = a_0x, a_1x, \ldots, a_kx = v$. Indeed, since $(a_i, b_{i+1}) \in F$, there are two edges labeled $b_{i+1}$ going out of $a_i x$ and $a_{i+1}x$ which end at $xb_{i+1}$. The Stallings folding identifies $a_i x$ and $a_{i+1}x$. The conclusion follows by induction.

Since the Stallings folding do not modify the group recognized, we deduce from Proposition 10 that the group defined by the Rauzy graph $G_n$ is the same as the group defined by the Rauzy graph $G_0$. Since $G_0$ is the graph with one vertex and with loops labeled by each of the letters, it defines the free group on $A$.

Example 11 Let $y$ be the infinite word obtained by decoding the Fibonacci word into blocks of length 2. Set $u = aa$, $v = ab$, $w = ba$. The graph $G_2$ is represented on the left of Figure 3. The classes of $\theta_2$ are indicated with colors.

![Figure 3: The Rauzy graphs $G_2$ and $G_1$ for the decoding of the Fibonacci word into blocks of length 2.](image)

The graph $G_1$ is represented on the right.

The following example shows that Proposition 10 is false for sets which do not satisfy the tree condition.

Example 12 Let $A = \{a, b, c\}$. The Chacon word on three letters is the fixpoint $x = f^\omega(a)$ of the morphism $f$ from $A^*$ into itself defined by $f(a) = aabc$, $f(b) = bc$ and $f(c) = abc$. Thus $x = aabcabcbabc\cdots$. The Chacon set is the set $F$ of factors of $x$. It is of complexity $2n + 1$ (see [4] Section 5.5.2). The Rauzy graph $G_1$ corresponding to the Chacon set is represented in Figure 4 on the left. The graph $G_1/\theta_1$ is represented on the right. It is not isomorphic to $G_0$ since it has two vertices instead of one.

3.2 Types of Rauzy graphs

Let $G$ be a graph labeled by words on an alphabet $A$. Call left-special (resp. right-special) a vertex of $G$ which has more than one predecessor (resp. successor). Similarly a vertex is bispecial if it is left and right-special. The type of $G$ is the graph $G'$ defined as follows. Its vertices are the left-special and right-special
vertices of $G$. There is an edge from $p$ to $q$ in $G'$ if there is a path $\pi$ from $p$ to $q$ in $G$ which uses between $p$ and $q$ only vertices which are not vertices of $G'$.

The label of this edge is the label of $\pi$.

The possible types of Rauzy graphs for a set of words on a three letter alphabet with complexity $2n + 1$ (and thus in particular for tree sets) have been described in [5]. There are 9 types of Rauzy graphs with a bispecial vertex $x$.

Seven of them are such that all cycles use this vertex. This implies that the set of paths of first returns to $x$ is finite and, for this reason we refer to this case as the finite case. The 7 graphs are represented in Figure 5.

The two remaining types of graphs are such that the set of paths of first returns to the bispecial vertex is infinite. We refer to this case as the infinite case. The two types of graphs are represented in Figure 6.

3.3 The finite case

We consider in turn the seven graphs.

Cases 1 to 5. In the first five cases, the proof is straightforward because for each path of first return to the bispecial vertex $x$, there is an edge which is only on this path. Since each edge is on an infinite path with its label in $F$, this implies that the labels of all paths of first return to the bispecial vertex $x$ are
in $F$. Thus Proposition \ref{prop:main} holds in these cases. We list in each of them the set $R(x)$.

| Case | $R(x)$ |
|------|---------|
| 1    | $\{u,v,w\}$ |
| 2    | $\{u, vw, vt\}$ |
| 3    | $\{u, vw, tw\}$ |
| 4    | $\{u, vwz, vtz\}$ |
| 5    | $\{ut, uzw, vz\}$ |

**Case 6.** Set $Z = \{uzw, uzt, vzt, vz\}$. We define

$$U = \{xu, xv\}y^{-1}, \quad V = \{w, t\}.$$ 

In view of applying Proposition \ref{prop:main} to the generalized extension graph $G_{U,V}(yz)$, we prove the following.

**Proposition 13** The set $U$ is an $F(yz)^{-1}$-maximal suffix code.

*Proof.* Each word $xu, xv$ contain only one occurrence of $x$. Thus $\{xu, xv\}$ is a suffix code and consequently also $\{xu, xv\}y^{-1}$. Let $s$ be such that $syz \in F$. Since the vertex $y$ can only be reached from $x$ by a path labeled $u$ or $v$, the word $sy$ is comparable for the suffix order with a word in $\{xu, xv\}$. Consequently, $s$ is comparable for the suffix order with a word in $U$. This shows that $U$ is an $F(yz)^{-1}$-maximal suffix code. $\blacksquare$

**Proposition 14** The set $V$ is a $(yz)^{-1}F$-maximal prefix code.

*Proof.* Since the words $w, t$ begin with distinct letters, the set $V$ is a prefix code. Let $p$ be such that $yzp \in F$. Then $p$ is prefix comparable with $t$ or $w$. This implies the conclusion. $\blacksquare$

We can now apply Proposition \ref{prop:main} to the graph $G_{U,V}(yz)$. Since this graph is a tree, it has three edges and this implies that $R(x)$ has three elements and generates the same group as $Z$. Thus Proposition \ref{prop:main} holds.

**Case 7.** Set $Z = \{uw, ut, vw, vt\}$. The proof is the same as previously using $U = \{xu, xv\}y^{-1}, V = \{w, t\}$ and considering the graph $G_{U,V}(y)$.

### 3.4 The infinite case

We consider in turn the two cases.
Case 1. Set $X = xvw^*$. In view of applying the strong tree condition to a graph of the form $G_{U,V}(yw^n)$, we prove the following.

**Proposition 15** Let $n \geq 0$ be such that $yw^n \in F$. Set $z = yw^n$ and $U = Xy^{-1} \cap Fz^{-1}$. The set $U$ is an $Fz^{-1}$-maximal suffix code.

**Proof.**
A word in $X$ has only one occurrence of $x$ which its prefix of length $n$. This implies that $X$ is a suffix code. Since any word in $X$ has $y$ as a suffix, the set $Xy^{-1}$ is also a suffix code and therefore also $U = Xy^{-1} \cap Fz^{-1}$.

Next, let $s$ be such that $sz \in F$. Then $sy$ is suffix comparable with a word beginning with $x$ and thus with a word in $xvw^*$. Thus $s$ is suffix comparable with a word in $Xy^{-1}$.

Set $Y = w^*t$. Symmetrically, we have

**Proposition 16** Let $n \geq 0$ be such that $yw^n \in F$. Set $z = yw^n$ and $V = Y \cap z^{-1}F$. The set $V$ is a $z^{-1}F$-maximal prefix code.

**Proof.** Since $w,t$ have distinct initial letters, the set $\{w,t\}$ is a prefix code and consequently $Y$ is a prefix code. Thus $V$ is a prefix code.

If $zp \in F$, then $zp$ is prefix comparable with a word ending in $x$ and thus with a word in $Y$.

**Proposition 17** We have $R_F(x) = \{u,vw^n t, vw^{n+1} t\}$ for some $n \geq 0$.

**Proof.** Since $F$ is uniformly recurrent, the set $R_F(x)$ is finite. Moreover, there is an $n \geq 1$ such that $vw^n t \in R_F(x)$. Indeed otherwise the word $w$ would not be a factor of $\Gamma(x)$.

Let us show that the set $R_F(x) \cap vw^* t$ cannot be reduced to one element. Indeed, assume that $R_F(x) \cap vw^* t$ is reduced to the word $vw^n t$. As we have just seen, we have $n \geq 1$. Consider the graph $G_{U,V}(y)$ with $U = Xy^{-1} \cap Fy^{-1}$ and $V = Y \cap y^{-1}F$. By Proposition 15 the set $U$ is an $Fy^{-1}$-maximal suffix code. By Proposition 16 $V$ is an $y^{-1}F$-maximal prefix code. Since $F$ is a tree set, by Proposition 2 the generalized extension graph $G = G_{U,V}(y)$ is a tree. But, since $\{xyy^{-1}, \ldots, xvw^n y^{-1}\} \subset U$ and $\{t, \ldots, w^n t\} \subset V$, and since
\[ xvy^{-1} \]
\[ \vdots \]
\[ xvw^n y^{-1} \]
\[ t \]
\[ \vdots \]
\[ (w^n t) \]

Figure 7: The graph \( G_{U,V}(y) \) when \( R_F(x) = \{vw^n t\} \).

\[ n \geq 1, \text{ the graph } G \text{ has at least four vertices (see Figure 7). Since it has only two edges, it is not connected, a contradiction.} \]

We have thus proved that \( R_F(x) \cap vw^* t \) has at least two elements. Set \( E = \{ i \geq 0 \mid vw^i t \in R_F(x) \} \). Let \( m, n \) with \( m < n \) be the two largest elements of the set \( E \). We show that \( n = m + 1 \). Set \( k = n - m \). Assume by contradiction that \( k \geq 2 \). Set \( z = yw^m \) and consider the graph \( G = G_{U,V}(z) \) with \( U = Xy^{-1} \cap Fz^{-1} \) and \( V = Y \cap z^{-1} F \). By Proposition 15 the set \( U \)
\[ xvy^{-1} \]
\[ \vdots \]
\[ xvw^k y^{-1} \]
\[ (w^k t) \]

Figure 8: The graph \( G_{U,V}(z) \) when \( vw^n t, vw^m t \in R_F(x) \).

is an \( Fz^{-1} \)-maximal suffix code and by Proposition 16 the set \( V \) is a \( z^{-1} F \)-maximal prefix code. Thus \( G \) is a tree. We have \{xvy^{−1},...,xvw^{k}y^{−1}\} ⊂ U and \{t, ..., w^{k}t\} ⊂ V. The vertices in the set \{xvy^{−1},xvw^{k}y^{−1}\} ⊂ U are only connected to vertices in the set \{t, w^{k}t\} ⊂ V (see Figure 8). Since \( k \geq 2 \), this implies that \( G \) is not connected, a contradiction.

We thus have proved that the two maximal elements of the set \( E \) are \( n, n + 1 \). We finally show that \( E = \{ n, n + 1 \} \). Consider indeed the graph \( G = G_{U,V}(y) \) with \( U = Xy^{-1} \cap Fy^{-1} \) and \( V = Y \cap y^{-1} F \). Any vertex \( u_{i} = xvw^{i}y^{-1} \in U \)
with \( 0 < i \leq n \) is connected to its two neighbours \( u_{i−1} \) and \( u_{i+1} \) by a path of length 2. Indeed, there are edges from \( u_{i−1} \) and \( u_{i} \) to \( w^{n−i+1}t \) and edges from \( u_{i} \) and \( u_{i+1} \) to \( v_{n−i}w^{n−i}t \) (see Figure 9). Similarly, there is an edge from any vertex \( v_{i} = w^{i}t \in V \) with \( 0 < i \leq n \) to its two neighbours. This shows that the graph \( G \) is connected. Therefore, the set \( E \) cannot contain any other element since it would create a cycle in the graph \( G \).

Set \( Z = u \cup vw^* t \).

**Corollary 18** The group generated by \( R_F(x) \) contains \( Z \).

**Proof.** This results from Proposition 17 and the fact that for all \( i \geq 1 \)
\[ vw^{i−1} t = vw^{i} t (vw^{i+1} t)^{−1} vw^{i} t, \quad vw^{i+2} t = vw^{i+1} t (vw^{i} t)^{−1} vw^{i+1} t \]
This implies that Proposition 9 is true in this case.

**Case 2.** The second case can be reduced to the previous one considering the equivalent graph represented in Figure 10. This graph is equivalent in the sense that the set of labels of paths of first return to \( x \) is the same.

![Figure 10: An equivalent graph for case 2.]

3.5 Examples

Set \( u = aa, v = ab, w = ba \) and \( B = \{u, v, w\} \). Let \( x \) be the infinite word on the alphabet \( B \) obtained by decoding the Fibonacci words in blocks of length 2. Let \( F \) be the set of factors of \( x \). There are bispecial words of length 1, 2, 3 and 5 (but none of length 4). The graphs \( G_1 \) and \( G_5 \) illustrate the infinite case while \( G_2 \) and \( G_3 \) illustrate the finite case.

The Rauzy graphs \( G_1 \) and \( G_5 \) are represented in Figure 11. They correspond both to the infinite case (case 1). We have \( R_F(v) = \{v, uwv, uwwv\} \) which

![Figure 11: The Rauzy graphs \( G_1 \) and \( G_5 \).]
corresponds to the case \( n = 0 \) in Proposition 17 and

\[ R(wvvuw) = \{ vvwv, wv(uwwv)vuw, wv(uwwv)^2vuw \} \]

which corresponds to the case \( n = 1 \).

The Rauzy graphs \( G_2 \) and \( G_3 \) are represented in Figure 12.

The word \( wv \) is bispecial. It corresponds to Case 6 of the finite case. One has \( R(wv) = \{ uwwv, vuvw, vuwwv \} \). The word \( vuw \) is bispecial and this time it is Case 5 of the finite case which is represented.

**References**

[1] Jean Berstel, Clelia De Felice, Dominique Perrin, Christophe Reutenauer, and Giuseppina Rindone. Bifix codes and Sturmian words. *J. Algebra*, 369:146–202, 2012.

[2] Valérie Berthé, Clelia De Felice, Francesco Dolce, Dominique Perrin, Christophe Reutenauer, and Giuseppina Rindone. Bifix codes and the finite index basis property. 2013. [http://arxiv.org/abs/1305.0127](http://arxiv.org/abs/1305.0127).

[3] Valérie Berthé, Clelia De Felice, Francesco Dolce, Dominique Perrin, Christophe Reutenauer, and Giuseppina Rindone. Bifix codes in acyclic sets. 2013. [http://arxiv.org/abs/1308.4260](http://arxiv.org/abs/1308.4260).

[4] N. Pythaes Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[5] Marie-Line Santini-Bouchard. Échanges de trois intervalles et suites minimales. *Theoret. Comput. Sci.*, 174(1-2):171–191, 1997.