

## On Morphisms Preserving Palindromic Richness

Francesco Dolce, Edita Pelantová\*

Czech Technical University in Prague

160 00 Prague 6 - Dejvice, Czech Republic

[dolcefra@fit.cvut.cz](mailto:dolcefra@fit.cvut.cz), [edita.pelantova@fffi.cvut.cz](mailto:edita.pelantova@fffi.cvut.cz)

---

**Abstract.** It is known that each word of length  $n$  contains at most  $n + 1$  distinct palindromes. A finite *rich word* is a word with maximal number of palindromic factors. The definition of palindromic richness can be naturally extended to infinite words. Sturmian words and Rote complementary symmetric sequences form two classes of binary rich words, while episturmian words and words coding symmetric  $d$ -interval exchange transformations give us other examples on larger alphabets. In this paper we look for morphisms of the free monoid, which allow us to construct new rich words from already known rich words. We focus on morphisms in Class  $P_{ret}$ . This class contains morphisms injective on the alphabet and satisfying a particular palindromicity property: for every morphism  $\varphi$  in the class there exists a palindrome  $w$  such that  $\varphi(a)w$  is a first complete return word to  $w$  for each letter  $a$ . We characterize  $P_{ret}$  morphisms which preserve richness over a binary alphabet. We also study marked  $P_{ret}$  morphisms acting on alphabets with more letters. In particular we show that every Arnoux-Rauzy morphism is conjugated to a morphism in Class  $P_{ret}$  and that it preserves richness.

**Keywords:** Palindromic richness, morphisms, bispecial factors, Class  $P_{ret}$ , Class  $P$ .

### 1. Introduction

Palindromes are words that coincide when read left to right and right to left. In natural languages we can find several known, and sometimes amusing, examples of palindromes. In the theory of formal languages, questions related to palindromes are usually elegantly formulated but difficult to solve. An

---

\*Address for correspondence: Czech Technical University in Prague, 160 00 Prague 6 - Dejvice, Czech Republic.

example of this is the famous HKS conjecture, whose first vague formulation appears in 1995 in [24], and that still remains unsolved in its general case. According to this conjecture, if the language of a fixed point of a primitive morphism contains infinite many palindromes, then the language coincides with the language of a fixed point of a morphism from Class  $P$  (see Definition 2 for the definition of Class  $P$  and [27] for a more detailed explanation of the HKS conjecture). Another recent conjecture is the one about palindromic length of aperiodic words, defined in [19], which still resists many attempts to prove it.

This article is devoted to words rich in palindromes. The definition of richness is a natural consequence of a simple observation, made by Droubay, Justin and Pirillo in [15], that each word of length  $n$  contains at most  $n + 1$  distinct palindromes, counting also the empty word  $\varepsilon$  which is considered to be a palindrome of length zero. A word of length  $n$  is said to be *rich* in palindromes if  $n + 1$  distinct palindromes occur in it. For example, *ananas* (Czech and Italian for "pineapple") is rich in taste but also in palindromes since it contains the 7 palindromes:  $\varepsilon, a, n, s, ana, nan, anana$ . Another rich word is *pizza*, since it contains the palindromes  $\varepsilon, a, i, p, z, zz$ . A counterexample is given by the word *hawaiianpizza* of length 13 that only contains 12 palindromes, namely  $\varepsilon, a, h, i, n, p, w, z, ii, zz, awa, aia$ , and thus is not rich in palindromes (nor in taste, according to the first author).

Every factor of a rich word is rich as well. Therefore the set of all rich words over a given finite alphabet is a so-called *factorial language*. Its factor complexity is known to be subexponential (see [34]) and superpolynomial (see [22]). Nevertheless the gap between the best known upper and lower bounds on number of rich words of length  $n$  over a fixed alphabet is still huge. An efficient algorithm to enumerate rich words over a fixed alphabet up to a given length is provided in [33]. To improve the lower bound requires to construct more rich words. In this paper we look for rewriting rules (formally morphisms of the free monoid), which allow to construct new rich words from already known rich words. An example of such a rewriting rule is the so-called *Fibonacci morphism*: 0 rewrites to 01 and 1 rewrites to 0. Applying the Fibonacci morphism to a rich word, say 0001110, we get a new rich word, i.e., 01010100001.

Because of the result of J. Vesti on extendability of rich words (see [36]), it is enough to apply rewriting rules only to infinite words, i.e., to sequences of the form  $\mathbf{u} = u_0u_1u_2\cdots$ , where  $u_i$  is an element of a finite alphabet  $\mathcal{A}$  for each  $i \in \mathbb{N}$ . The definition of palindromic richness can be naturally extended to infinite words. We say that  $\mathbf{u}$  is rich in palindromes if all its finite prefixes are rich in palindromes. The most studied class of binary words, namely Sturmian words, consists of words rich in palindromes. Therefore rich words can be considered as one of the possible generalizations of Sturmian words. The Rote complementary symmetric sequences form another class of binary rich words. Except for these two classes, only several singular examples of binary rich words are known (e.g., the period doubling word). On  $d$ -ary alphabets, with  $d \geq 3$ , two disjoint classes of rich words were found: the episturmian words and the words coding  $d$ -interval exchange transformation with the symmetric permutation of intervals. On ternary alphabets a further class of rich words is described in [2].

Here we study morphisms in Class  $P_{ret}$ . A morphism  $\varphi$  is in this class if it is injective and if there exists a palindrome  $w$  such that  $\varphi(a)w$  is a first complete return word to  $w$  for each letter  $a$ . This class was introduced in [4]. Harju, Vesti and Zamboni used these morphisms in [23] to prove a

weak version of the HKS conjecture for rich words. The palindromic richness of an infinite word can be equivalently described by a property of its bispecial factors, therefore we concentrate on bispecial factors in images of rich words under a morphism in Class  $P_{ret}$ . This enables us to characterize  $P_{ret}$  morphisms which preserve richness on a binary alphabet. On alphabets with more letters we obtain only partial results for the so-called marked morphisms. In particular we show that Arnoux-Rauzy morphisms are conjugated to morphisms in Class  $P_{ret}$  and preserve the set of all rich words.

The article is organized as follows. In Section 2 necessary definitions and tools for the study of richness are presented. In Section 3 we define Class  $P_{ret}$  morphisms and list some of their properties, including their relation to morphisms in Class  $P$ . Marked morphisms in Class  $P_{ret}$  are studied in Section 4. Section 5 is devoted to the proofs of our main results on morphisms preserving richness. These results are applied in Section 6 where we show how to find new rich words. Finally, some suggestions for further research are presented in the Section 7.

## 2. Preliminaries

### 2.1. Words

Let  $\mathcal{A}$  be a finite alphabet and  $\mathcal{A}^*$  the free monoid over  $\mathcal{A}$ . The elements of  $\mathcal{A}$  are called *letters* and the elements of  $\mathcal{A}^*$  are called (finite) *words*. The *length* of a word  $u = a_0a_1 \cdots a_{n-1}$ , with  $a_i \in \mathcal{A}$ , is the number  $|u| = n$ . The identity of the monoid is called the *empty word* and is denoted by  $\varepsilon$ . Note that  $\varepsilon$  is the only word of length 0. If a word  $u \in \mathcal{A}^*$  can be written as  $u = xwy$ , with  $x, w, y \in \mathcal{A}^*$ , we say that  $w$  is a *factor* of  $u$ ,  $x$  is a *prefix* of  $u$  and  $y$  is a *suffix* of  $u$ . If  $u = a_0a_1 \cdots a_n$  and  $w = a_i a_{i+1} \cdots a_j$  we say that  $i$  is an *occurrence* of  $w$  in  $u$ . If  $z = xy \in \mathcal{A}^*$ , then we denote  $x^{-1}z = y$  and  $zy^{-1} = x$ . Given a non-empty word  $u = a_0a_1 \cdots a_{n-1}$ , with  $a_i \in \mathcal{A}$ , we define its *mirror image*  $\bar{u}$  as  $\bar{u} = a_{n-1}a_{n-2} \cdots a_0$ . The mirror image of the empty word is  $\bar{\varepsilon} = \varepsilon$ . A word  $u$  is called a *palindrome* (or *palindromic*) if  $u = \bar{u}$ .

A (right) infinite word is an infinite sequence of letters  $\mathbf{u} = a_0a_1a_2 \cdots$ . We denote by  $\mathcal{A}^{\mathbb{N}}$  the set of infinite words over  $\mathcal{A}$ . In this paper we use bolder letters to denote infinite words. We can extend in a natural way to infinite words the notions of factor, prefix, suffix and occurrence. An infinite word  $\mathbf{u}$  is called *ultimately periodic* if  $\mathbf{u} = uv^\omega = uvvv \cdots$  for some  $u, v \in \mathcal{A}^*$ ,  $v \neq \varepsilon$ . If, moreover,  $u$  can be chosen to be empty, the word  $\mathbf{u}$  is said to be (purely) *periodic*. A word that is not ultimately periodic is called *aperiodic*. A word  $\mathbf{u}$  is called *recurrent* if every factor of  $\mathbf{u}$  appears at least twice (thus an infinite number of times). It is *uniformly recurrent* if these factors appear with bounded gaps, i.e., if for every factor  $u$  of  $\mathbf{u}$  there exists an integer  $n \geq 0$  such that all factors of length  $n$  of  $\mathbf{u}$  are of the form  $xuy$  for certain  $x, y \in \mathcal{A}^*$ .

The *language* of an infinite word  $\mathbf{u}$ , is the set  $(\mathbf{u}) \subset \mathcal{A}^*$  of its finite factors, i.e.,

$$(\mathbf{u}) = \left\{ u \in \mathcal{A}^* : \mathbf{u} = xuy \text{ with } x \in \mathcal{A}^*, y \in \mathcal{A}^{\mathbb{N}} \right\}.$$

We say that the language  $(\mathbf{u})$  is periodic (resp. aperiodic, recurrent, uniformly recurrent) if the word  $\mathbf{u}$  is periodic (resp. aperiodic, recurrent, uniformly recurrent). A language  $(\mathbf{u})$  is said to be *closed under reversal* if for every  $u \in (\mathbf{u})$  we have  $\bar{u} \in (\mathbf{u})$  as well. It is known that if  $(\mathbf{u})$  is closed under reversal, then  $\mathbf{u}$  is recurrent.

The set of *complete return words* to a word  $u \in (\mathbf{u})$  (with respect to the language  $(\mathbf{u})$ ) is the set  $\mathcal{CR}_{\mathbf{u}}(u)$  of words in  $(\mathbf{u})$  having exactly two factors equal to  $u$ , one as a proper prefix and the other as a proper suffix. It is known that  $\mathbf{u}$  is uniformly recurrent if and only if it is recurrent and for every word  $u \in (\mathbf{u})$  the set  $\mathcal{CR}_{\mathbf{u}}(u)$  is finite. If  $uw$  is a complete return word to  $u$ , then  $w$  is called a (right) *return word* to  $u$ . We denote by  $\mathcal{R}_{\mathbf{u}}(u)$  the set of return words to  $u$ . Clearly  $\#\mathcal{CR}_{\mathbf{u}}(u) = \#\mathcal{R}_{\mathbf{u}}(u)$ .

Let us consider an infinite word  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  and its language  $(\mathbf{u})$ . Given a word  $u$  we define the following subsets of  $(\mathbf{u})$ :

$$L_{\mathbf{u}}(u) = \{a \in \mathcal{A} : au \in (\mathbf{u})\}, \quad R_{\mathbf{u}}(u) = \{b \in \mathcal{A} : ub \in (\mathbf{u})\},$$

called respectively the set of left extensions and the set of right extensions of  $u$  in  $(\mathbf{u})$ , and the subset of  $(\mathbf{u}) \times (\mathbf{u})$

$$B_{\mathbf{u}}(u) = \{(a, b) \in \mathcal{A} \times \mathcal{A} : aub \in (\mathbf{u})\}$$

called the set of bi-extensions of  $u$  in  $(\mathbf{u})$  (see [7] for a more detailed explanation).

**Example 2.1.** Let us consider an infinite word  $\mathbf{u} \in \{0, 1\}^{\mathbb{N}}$  such that its language  $(\mathbf{u})$  contains the following words of length 2: 00, 01, 10 (but not 11); and the following words of length 3: 001, 010, 100, 101 (but not 000). Then one has

$$\begin{aligned} L_{\mathbf{u}}(\varepsilon) &= R_{\mathbf{u}}(\varepsilon) = \{0, 1\}, & B_{\mathbf{u}}(\varepsilon) &= \{(0, 0), (0, 1), (1, 0)\}, \\ L_{\mathbf{u}}(0) &= R_{\mathbf{u}}(0) = \{0, 1\}, & B_{\mathbf{u}}(0) &= \{(0, 1), (1, 0), (1, 1)\}, \\ L_{\mathbf{u}}(1) &= R_{\mathbf{u}}(1) = \{0\}, & B_{\mathbf{u}}(1) &= \{(0, 0)\}. \end{aligned}$$

A word  $u \in (\mathbf{u})$  is *left special* (resp. *right special*) if  $\#L_{\mathbf{u}}(u) \geq 2$  (resp.  $\#R_{\mathbf{u}}(u) \geq 2$ ). A *bispecial word* is a word that is both left special and right special. We denote by  $\mathcal{B}(\mathbf{u})$  the set of all bispecial factors of  $\mathbf{u}$ . We also define the *bilateral order* of a word  $u$  as

$$b_{\mathbf{u}}(u) = \#B_{\mathbf{u}}(u) - \#L_{\mathbf{u}}(u) - \#R_{\mathbf{u}}(u) + 1.$$

**Example 2.2.** Let  $\mathbf{u}$  be the infinite word defined in Example 2.1. Both the empty word  $\varepsilon$  and the factor 0 are in  $\mathcal{B}(\mathbf{u})$ , while the factor 1 is neither left special nor right special. Moreover, one can check that  $b_{\mathbf{u}}(\varepsilon) = b_{\mathbf{u}}(0) = b_{\mathbf{u}}(1) = 0$ .

An infinite word over an alphabet  $\mathcal{A}$  is *episturmian* if its language is closed under reversal and contains for each  $n$  at most one word of length  $n$  which is right special. An episturmian word is *strict episturmian*, or *Arnoux-Rauzy*, if it has exactly one right special word for each length and moreover each right special factor  $u$  is such that  $\#R_{\mathbf{u}}(u) = \#\mathcal{A}$ . Note that when  $\#\mathcal{A} = 2$  every episturmian word is also a strict episturmian word. Episturmian words over a binary alphabet are called *Sturmian* words.

## 2.2. Richness

In [15] it is proved that a finite word  $w$  has at most  $|w| + 1$  factors that are palindromes. The *defect*  $D(w)$  of a finite word  $w$  is defined as the difference between  $|w| + 1$  and the actual number of palindromes contained in  $w$ . Finite words with zero defects are called *rich words*. If  $u$  is a factor of  $w$ , then  $D(u) \leq D(w)$ . In particular, factors of a rich word are rich as well. Droubay, Justin and Pirillo demonstrated that a finite word  $u$  is rich if and only if every prefix  $v$  of  $u$  has the so-called Property Ju: the longest palindromic suffix of  $v$  is unioccurrent in  $v$ . A direct consequence of Property Ju is the following lemma (see [4, Lemma 4.4] for a proof).

### Lemma 2.3. ([4])

Let  $u, q \in \mathcal{A}^*$  with  $q \neq \bar{q}$ . Assume that  $u$  is rich.

1. If both  $q$  and  $\bar{q}$  occur in  $u$ , then occurrences of  $q$  and  $\bar{q}$  in  $u$  alternate, i.e., every complete return word to  $q$  in  $u$  contains  $\bar{q}$  and every complete return word to  $\bar{q}$  in  $u$  contains  $q$ .
2. If  $p$  is a factor of  $u$  such that  $q$  occurs in  $p$  exactly once as a prefix of  $p$  and  $\bar{q}$  occurs in  $p$  exactly once as a suffix of  $p$ , then  $p$  is a palindrome.

The following result gives us a more precise characterization of rich words in case of a binary alphabet.

### Lemma 2.4. ([31])

Let  $u$  be a finite word over  $\{0, 1\}$ . The word  $u$  is not rich if and only if there exists a non-palindromic word  $q$  such that  $0q0, 1q1, 0\bar{q}1$  and  $1\bar{q}0$  occur in  $u$ .

The definition of richness can be extended to infinite words. The *defect* of an infinite word  $\mathbf{u}$  is defined as the number (finite or infinite)

$$D(\mathbf{u}) = \sup \{D(w) : w \text{ is a prefix of } \mathbf{u}\}.$$

An infinite word  $\mathbf{u} = a_0a_1a_2\cdots$  is *rich* if for every  $n \in \mathbb{N}$  the prefix  $a_0a_1\cdots a_{n-1}$  of length  $n$  contains exactly  $n + 1$  different palindromes (see [15]). Infinite words with zero defect correspond exactly to rich words. Infinite words with finite (but not necessarily zero) defect are called *almost rich*. Such words have been studied in [21]. A simple test for richness in periodic words was provided by Brlek, Hamel, Nivat and Reutenauer in [11, Theorems 4 and 6].

### Theorem 1. ([11])

If  $\mathbf{u}$  is periodic and  $(\mathbf{u})$  contains infinitely many palindromes, then there exist two palindromes  $p, q$  such that  $\mathbf{u} = (pq)^\omega$ .

Moreover,  $\mathbf{u}$  is rich if its prefix of length  $|pq| + \max \left\{ \left\lfloor \frac{|q|-|p|}{3} \right\rfloor, \left\lfloor \frac{|p|-|q|}{3} \right\rfloor \right\}$  is rich.

To test richness in general infinite words we introduce the following notion. Given an infinite word  $\mathbf{u} \in \mathcal{A}^\mathbb{N}$  and a palindrome  $p \in (\mathbf{u})$ , we define the number of palindromic extensions of  $p$  in  $\mathbf{u}$  as

$$\#\text{Pext}_{\mathbf{u}}(p) = \#\{a \in \mathcal{A} : apa \in (\mathbf{u})\}.$$

The following relation between bispecial factors and richness in the language of an infinite word has been proved in [3, Theorem 11].

**Theorem 2. ([3])**

Let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  be such that  $(\mathbf{u})$  is closed under reversal. Then  $\mathbf{u}$  is rich if and only if every bispecial factor  $u \in (\mathbf{u})$  satisfies

$$b_{\mathbf{u}}(u) = \begin{cases} \#\text{Pext}_{\mathbf{u}}(u) - 1 & \text{if } u \text{ is a palindrome,} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Extendability of finite rich words into infinite rich words were studied by Vesti. He proved the following property (see [36, Propositions 2.11 and 2.12]).

**Proposition 1. ([36])**

Let  $u$  be a finite rich word. There exist an infinite aperiodic rich word and an infinite periodic rich word such that  $u$  is a factor of both of them.

**2.3. Morphisms**

Let  $\mathcal{A}, \mathcal{B}$  be two alphabets. A *morphism*  $\varphi : \mathcal{A}^* \rightarrow \mathcal{B}^*$  is a monoid morphism from  $\mathcal{A}^*$  to  $\mathcal{B}^*$ . A morphism  $\varphi : \mathcal{A}^* \rightarrow \mathcal{B}^*$  is called *non-erasing* if  $|\varphi(a)| > 0$  for all letters  $a \in \mathcal{A}$ .

Let us consider a morphism  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$ . If  $a \in \mathcal{A}$  is such that  $\varphi(a)$  begins with  $a$  and  $|\varphi^n(a)|$  tends to infinity with  $n$ , then the morphism is called a *substitution*. The unique word which has all words  $\varphi^n(a)$  as prefixes is called a *fixed point* of the substitution  $\varphi$  and is denoted by  $\varphi^\omega(a)$ .

A morphism  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  is called *primitive* if there exists an integer  $k$  such that for all  $a, b \in \mathcal{A}$ , the letter  $b$  appears as a factor in  $\varphi^k(a)$ . If  $\varphi$  is a primitive morphism, then the language  $(\mathbf{u})$  is uniformly recurrent for all fixed points  $\mathbf{u}$  of  $\varphi$  (see [32, Proposition 13]).

A morphism  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  is *right conjugated* to a morphism  $\psi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  (or is a *right conjugate* of  $\psi$ ) if there exists a word  $x \in \mathcal{A}^*$ , called the *conjugate word*, such that  $\varphi(a)x = x\psi(a)$  for each  $a \in \mathcal{A}$ . A *left conjugated* morphism is defined symmetrically. We say that  $\psi$  is the *rightmost conjugate* of  $\varphi$ , if  $\psi$  is right conjugated to  $\varphi$  and  $\psi$  is the only right conjugated to itself. In this case we denote such rightmost conjugate by  $\varphi_R$ . Analogously we define the *leftmost conjugate* of  $\varphi$  and denote it by  $\varphi_L$ .

If there exists no rightmost conjugate (resp. no leftmost conjugate) of  $\varphi$ , then  $\varphi$  is conjugated to itself by a non-empty conjugate word  $x$ . Such a morphism is called *cyclic* and there exists  $z \in \mathcal{A}$  such that  $\varphi(a) \in z^*$  for each letter  $a \in \mathcal{A}$ . A fixed point of a cyclic morphism is periodic (see [29]). A morphism that is not cyclic is called *acyclic*. Note that an acyclic morphism can have a periodic fixed point (see, e.g., the morphism  $\varphi_2$  in Example 3.1 later).

Over an alphabet  $\mathcal{A}$  we define the morphisms  $\psi_a$  and  $\bar{\psi}_a$  for each  $a \in \mathcal{A}$  as follows:

$$\psi_a : \begin{cases} a \mapsto a \\ b \mapsto ab \text{ if } b \neq a \end{cases} \quad \text{and} \quad \bar{\psi}_a : \begin{cases} a \mapsto a \\ b \mapsto ba \text{ if } b \neq a \end{cases} \quad (2)$$

These morphisms together with permutations of letters in  $\mathcal{A}$  are called *elementary Arnoux-Rauzy morphisms*. All elementary Arnoux-Rauzy morphisms generate a monoid of morphisms called the *Arnoux-Rauzy monoid*. A morphism from this monoid is said to be an *Arnoux-Rauzy morphism*. Arnoux-Rauzy morphisms are known to send episturmian words to episturmian words (see [25]).

**Example 2.5.** Let us consider the morphism  $\phi : \{0, 1\}^* \rightarrow \{0, 1\}^*$  defined by

$$\phi : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 0 \end{cases}$$

Such a morphism, called the *Fibonacci morphism*, is an Arnoux-Rauzy morphism since it can be obtained as the composition  $\phi = \psi_0 \circ \pi$ , where  $\pi = (1, 0)$  is the permutation interchanging the two letters. Its fixed point  $\mathbf{f} = \varphi^\omega(0)$  is called the *Fibonacci word*.

Similarly, the morphism  $\tau : \{0, 1, 2\}^* \rightarrow \{0, 1, 2\}^*$  defined by

$$\tau : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 02 \\ 2 \mapsto 0 \end{cases}$$

is an Arnoux-Rauzy morphism, called the *Tribonacci morphism*. It can be obtained as the composition  $\tau = \psi_0 \circ \pi'$ , where  $\pi' = (1, 2, 0)$ . Its fixed point  $\mathbf{t} = \tau^\omega(0)$  is called the *Tribonacci word*.

Fixed points of Arnoux-Rauzy morphisms are known to be rich.

A morphism  $\psi$  over the binary alphabet  $\mathcal{A} = \{0, 1\}$  is called a *standard Sturmian morphism* if it belongs to the monoid generated by  $\pi$  and  $\phi$ , where  $\pi$  is the permutation of the two letters and  $\phi$  is the Fibonacci morphism defined in Example 2.5.

### 3. Class $P_{ret}$

The basic set in which we will look for morphisms preserving richness is the so-called Class  $P_{ret}$ . This class was introduced in [4] to study relations between rich and almost rich words.

**Definition 1.** A morphism  $\varphi : \mathcal{A}^* \mapsto \mathcal{B}^*$  belongs to Class  $P_{ret}$ , if there exists a palindrome  $w$  such that

- $\varphi(a)w$  is a palindromic complete return word to  $w$  for each  $a \in \mathcal{A}$ ,
- $\varphi(a) \neq \varphi(b)$  for each pair  $a, b \in \mathcal{A}$ , with  $a \neq b$ .

Clearly, if  $\varphi \in P_{ret}$ , then  $\varphi$  is non-erasing. Indeed a complete return word to  $w$  is longer than  $w$  itself. In this paper we focus on  $P_{ret}$  morphisms for which  $\mathcal{A} = \mathcal{B}$ . In Proposition 2 we will prove that the palindrome  $w$  in Definition 1 is unique. Following [23], we call it the *marker* of the morphism  $\varphi \in P_{ret}$ .

**Example 3.1.** One can easily check that the following standard Sturmian morphisms

$$\varphi_1 : \begin{cases} 0 \mapsto 010 \\ 1 \mapsto 01001 \end{cases} \quad \text{and} \quad \varphi_2 : \begin{cases} 0 \mapsto 1110 \\ 1 \mapsto 1 \end{cases}$$

belong to Class  $P_{ret}$  with markers respectively  $w_1 = 010010$  and  $w_2 = 111$ .

**Example 3.2.** Let  $p, q, \ell$  be positive integers, with  $p \neq q$ . The morphism

$$\sigma : \begin{cases} 0 \mapsto 0^\ell 1^p \\ 1 \mapsto 0^\ell 1^q \end{cases}$$

belongs to Class  $P_{ret}$  and the associated marker is  $w = 0^\ell$ .

**Example 3.3.** For each  $a \in \mathcal{A}$ , the elementary Arnoux-Rauzy morphism  $\psi_a$  defined in (2) belongs to Class  $P_{ret}$  and  $w = a$  is its marker. A morphism which permutes letters of the alphabet, i.e., a permutation on  $\mathcal{A}$ , is in Class  $P_{ret}$  as well, the associated marker being  $w = \varepsilon$ . The morphism  $\overline{\psi}_a$  does not belong to Class  $P_{ret}$ , but is conjugated to  $\psi_a \in P_{ret}$  with conjugate word  $a$ .

**Corollary 1.** Each Arnoux-Rauzy morphism is conjugated to a morphism in Class  $P_{ret}$ .

**Proof:**

As proven in [23, Lemma 2.5], if a morphism  $\varphi$  is conjugated to  $\varphi'$  and  $\psi$  is conjugated to  $\psi'$ , then their composition  $\varphi \circ \psi$  is conjugated to  $\varphi' \circ \psi'$ . Using the previous example, every elementary Arnoux-Rauzy morphism is conjugated to a morphism in Class  $P_{ret}$ . As shown in [4], the Class  $P_{ret}$  is closed under composition and thus every composition of elementary Arnoux-Rauzy morphisms belongs to Class  $P_{ret}$ .  $\square$

Let us list some properties of morphisms in Class  $P_{ret}$ . They are stated in [4] as Remark 5.2 and Proposition 5.3.

**Lemma 1. ([4])**

Let  $\varphi, \sigma : \mathcal{A}^* \mapsto \mathcal{A}^*$  be morphisms in Class  $P_{ret}$ . Denote by  $w_\varphi$  and  $w_\sigma$  the markers associated to  $\varphi$  and  $\sigma$ , respectively. For every  $u, v \in \mathcal{A}^*$ :

1.  $\varphi(u) = \varphi(v)$  implies  $u = v$ ,
2.  $\overline{\varphi(u)w_\varphi} = \varphi(\overline{u})w_\varphi$ ,
3.  $u$  is a palindrome if and only if  $\varphi(u)w_\varphi$  is a palindrome,
4.  $\sigma \circ \varphi$  is a morphism in Class  $P_{ret}$  and its associated marker is  $\sigma(w_\varphi)w_\sigma$ .

**Example 3.4.** Let us consider the morphism  $\varphi_1$  and  $\varphi_2$  of Example 3.1.

One has  $\varphi_2(01)w_2 = \overline{11101111} = 11110111 = \varphi_2(10)111$ . Moreover one has that both  $010$  and  $\varphi_2(010)w_2 = 111011110111$  are palindromes. Finally one can check that  $\varphi_1 \circ \varphi_2$  is in Class  $P_{ret}$  with associated marker  $\varphi_1(111)010010$ .

Using Lemma 1, we deduce several auxiliary useful properties.

**Lemma 2.** Let  $\varphi : \mathcal{A}^* \mapsto \mathcal{A}^*$  be a morphism in Class  $P_{ret}$  with marker  $w$  and let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ .

1. If  $v \in \mathcal{A}^*$ , then  $w$  occurs in  $\varphi(v)w$  exactly  $|v| + 1$  times and  $w$  is a prefix of  $\varphi(v)w$ .
2. The set of return words to  $w$  in  $\varphi(\mathbf{u})$  is  $\mathcal{R}_{\varphi(\mathbf{u})}(w) = \{\varphi(a) : a \in \mathcal{A} \cap (\mathbf{u})\}$ .
3. If  $u \in (\mathbf{u})$ , then  $\varphi(u)w \in (\varphi(\mathbf{u}))$ .



**Proof:**

**Item (1).** We proceed by induction on  $|v|$ . The claim is clearly valid when  $v = \varepsilon$ . Assume now that  $v = v'a$ , with  $a \in \mathcal{A}$ . Then  $\varphi(v)w = \varphi(v')\varphi(a)w$ . By definition of  $P_{ret}$ , the factor  $w$  is a prefix of  $\varphi(a)w$  and it has 2 occurrences in  $\varphi(a)w$ . Therefore we can write  $\varphi(v)w = \varphi(v')ww^{-1}\varphi(a)w$ . By induction hypothesis,  $w$  is a prefix of  $\varphi(v')w$  and occurs  $|v'| + 1 = |v|$  times in  $\varphi(v')w$ . Consequently,  $w$  is a prefix of  $\varphi(v)w = \varphi(v')ww^{-1}\varphi(a)w$  and has  $|v| + 1$  occurrences in  $\varphi(v)w$ .

**Item (2).** Let  $\mathbf{u} = a_0a_1a_2 \cdots$ . First we show that for every  $i \in \mathbb{N}$ , the word  $\varphi(a_i a_{i+1} a_{i+2} \cdots)$  has  $w$  as a prefix. Indeed, the word  $\varphi(a_i a_{i+1} \cdots a_{i+k})$  is longer than  $|w|$  for some  $k \in \mathbb{N}$ . By Item (1), the word  $\varphi(a_i a_{i+1} \cdots a_{i+k})w$  has  $w$  as a prefix. Therefore,  $w$  is a prefix of  $\varphi(a_i a_{i+1} \cdots a_{i+k})$  as well.

Since  $w$  is a prefix both of  $\varphi(a_i a_{i+1} a_{i+2} \cdots)$  and  $\varphi(a_{i+1} a_{i+2} a_{i+3} \cdots)$ , the word  $\varphi(a_i)w$  is a prefix of  $\varphi(a_i a_{i+1} a_{i+2} \cdots)$ . By definition of Class  $P_{ret}$ ,  $w$  occurs twice in  $\varphi(a_i)w$ . This implies that  $\varphi(a_i)$  is a return word to  $w$  in  $\varphi(\mathbf{u})$ .

**Item (3).** Let  $u \in (\mathbf{u})$ . Since the language  $(\mathbf{u})$  is extendable and the morphism  $\varphi$  is non-erasing, there exist letters  $c_1, c_2, \dots, c_k \in \mathcal{A}$  such that  $uc_1c_2 \cdots c_k \in (\mathbf{u})$  and  $|\varphi(c_1c_2 \cdots c_k)| \geq |w|$ . Thus  $\varphi(uc_1c_2 \cdots c_k)$  is an element of  $(\varphi(\mathbf{u}))$ . It is enough to show that  $w$  is a prefix of  $\varphi(c_1c_2 \cdots c_k)$ . We know from Item (1) that  $w$  is a prefix of  $\varphi(c_1c_2 \cdots c_k)w$ . Since  $|\varphi(c_1c_2 \cdots c_k)| \geq |w|$ , the word  $\varphi(c_1c_2 \cdots c_k)$  starts with  $w$  as well.  $\square$

**Example 3.5.** Let us consider again the morphism  $\varphi_2$  from Examples 3.1 and 3.4. According to Lemma 2, the word  $w_2 = 111$  appears exactly three times in  $\varphi_2(01)111$ .

**Proposition 2.** Each morphism in Class  $P_{ret}$  has a unique marker.

**Proof:**

Let  $\varphi$  be a morphism in Class  $P_{ret}$ . Assume by contradiction that  $w$  and  $u$  are two distinct markers associated to  $\varphi$ . From Item (1) of Lemma 2 it follows that both  $w$  and  $u$  are prefixes of  $\varphi(v)$  for a long enough word  $v \in \mathcal{A}^*$ . Therefore  $|w| \neq |u|$ . Without loss of generality, let us assume that  $|w| > |u|$ . In particular,  $u$  is a non-empty prefix of  $w$ , i.e.,  $w = uz$  for some non-empty word  $z \in \mathcal{A}^*$ . Since both  $w$  and  $u$  are palindromes, we get  $uz = \bar{z}u$ . It is a well-known fact that this equations on words implies that there exist words  $x, y \in \mathcal{A}^*$  and  $i \in \mathbb{N}$  such that  $z = yx$ ,  $\bar{z} = xy$  and  $u = (xy)^i x$ . Since  $w = uz = (xy)^{i+1} x = xyu = uxy$ , the marker  $u$  has at least two occurrences in  $w$ . Let us discuss two cases.

- 1) There exists a letter  $a \in \mathcal{A}$  such that  $|\varphi(a)| \geq |w|$ . Then  $\varphi(a)$  has a prefix  $w$  and thus  $\varphi(a)u$  has at least three occurrences of  $u$  - a contradiction.
- 2) For all letters  $a \in \mathcal{A}$  one has  $|\varphi(a)| < |w|$ . Let us fix  $a \in \mathcal{A}$  and denote

$$\varphi(a)w = v_0v_1 \cdots v_{n-1}v_n \cdots v_{n+m-1},$$

where  $n = |\varphi(a)|$ ,  $m = |w|$  and  $v_i \in \mathcal{A}$  for all  $i$ . Then 0 and  $n$  are the two occurrences of  $w$  in  $\varphi(a)w$ . The form of  $u$  implies that 0,  $|xy|$ , and  $n$  are occurrences of  $u$  in  $\varphi(a)u$ . As the marker  $u$  occurs in  $\varphi(a)u$  exactly twice, and  $|z| > 0$ , we have  $|xy| = n$  and hence  $\varphi(a) = xy$ . Since this is true for every letter  $a \in \mathcal{A}$ , we have a contradiction with the injectivity of  $\varphi$ .  $\square$

**Corollary 2.** Let  $\varphi : \mathcal{A}^* \mapsto \mathcal{A}^*$  be a morphism in Class  $P_{ret}$  with marker  $w$ . Let  $\mathbf{u}, \mathbf{v} \in \mathcal{A}^{\mathbb{N}}$ , with  $\mathbf{u} = a_0a_1a_2\cdots$ ,  $\mathbf{v} = b_0b_1b_2\cdots$ , such that  $\mathbf{v} = \varphi(\mathbf{u})$ .

- If  $v \in (\mathbf{v})$  has  $w$  both as a prefix and as a suffix, then there exists a unique  $u \in \mathcal{A}^*$  such that  $v = \varphi(u)w$ . Moreover, if an index  $i$  is an occurrence of  $v$  in  $\mathbf{v}$ , then  $b_ib_{i+1}b_{i+2}\cdots = \varphi(a_ja_{j+1}a_{j+2}\cdots)$ , where the index  $j$  is an occurrence of  $u$  in  $\mathbf{u}$ .
- $(\mathbf{u})$  is closed under reversal if and only if  $(\mathbf{v})$  is closed under reversal.

**Proof:**

Let  $v'$  such that  $v = v'w$ . Then  $v'$  is a concatenation of return words to  $w$ . By Item (2) of Lemma 2, the factor  $v$  can be written as  $v = \varphi(u)w$  for a certain word  $u$ . Since  $\varphi(a) \neq \varphi(b)$  for every  $a \neq b$ , such a factor  $u$  is given uniquely. Moreover, Item (2) of Lemma 2 implies that  $w$  occurs in  $\mathbf{v}$  only as a prefix of  $\varphi(a_ja_{j+1}a_{j+2}\cdots)$  for some  $j$ .

To prove the second part of the corollary, we first assume that  $(\mathbf{u})$  is closed under reversal and let  $v \in (\varphi(\mathbf{u}))$ . Then, there exists a factor  $u \in (\mathbf{u})$  such that  $v$  occurs in  $\varphi(u)$ . By Item (3) of Lemma 2,  $\varphi(u)w \in (\varphi(\mathbf{u}))$ . Since  $\bar{u} \in (\mathbf{u})$ , it follows from Item (2) of Lemma 1 that the factor  $\overline{\varphi(u)w} = \varphi(\bar{u})w$  belongs to  $(\varphi(\mathbf{u}))$ . Obviously,  $\bar{v}$  occurs in  $\overline{\varphi(u)w}$  and thus belongs to  $(\varphi(\mathbf{u}))$  too.

Now, let us assume that  $(\varphi(\mathbf{u}))$  is closed under reversal and let  $u \in (\mathbf{u})$ . Then  $\varphi(u)w \in (\varphi(\mathbf{u}))$  and thus, using Item (2) of Lemma 1, we have  $\overline{\varphi(u)w} = \varphi(\bar{u})w \in (\mathbf{v})$ . By the first statement of this corollary,  $\bar{u}$  belongs to  $(\mathbf{u})$ .  $\square$

**Example 3.6.** Let us consider the morphism  $\phi$  of Example 2.5 and its fixed point  $\mathbf{f}$ . Such a morphism is in Class  $P_{ret}$  and its marker is  $w = 0$ . The word  $v = 010$  starts and ends with 0. The letter 0 is the unique word  $u \in \mathcal{L}(\mathbf{f})$  such that  $v = \phi(u)w$ .

**Remark 3.7.** If  $\varphi \in P_{ret}$ , then every factor of  $\varphi(\mathbf{u})$  can be decomposed (up to a short prefix and suffix) uniquely into concatenation of return words to  $w$ . Therefore, the synchronizing delay as defined in [13] is at most the length of a factor which does not contain  $w$ . In other words, the synchronizing delay of  $\varphi$  is at most  $\max\{|\varphi(a)w| : a \in \mathcal{A}\} - 2$ .

**Proposition 3.8.** Every morphism in Class  $P_{ret}$  is acyclic.

**Proof:**

Assume that  $\varphi$  is a cyclic morphism in Class  $P_{ret}$  with marker  $w$ . Then, there exist  $z \in \mathcal{A}^*$ , distinct letters  $a, b \in \mathcal{A}$  and positive integers  $m, n$ , with  $m < n$  such that  $\varphi(a) = z^m$  and  $\varphi(b) = z^n$ . Moreover,  $\varphi(a)w = z^mw$  and  $\varphi(b)w = z^nw$  are complete return words to  $w$ . In particular, there exists a non-empty word  $u \in \mathcal{A}^*$  such that  $z^mw = wu$ . Hence we have  $z^nw = z^{n-m}z^mw = z^{n-m}wu$ , which implies that  $w$  appears at least three times as a factor of  $\varphi(b)w$ , a contradiction.  $\square$

Let  $\text{Fst}(u)$  and  $\text{Lst}(u)$  denote respectively the first and the last letter of a non-empty word  $u$ . Given a non-erasing morphism  $\varphi$  we define the two maps on  $\mathcal{A}$

$$\rho_\varphi : a \mapsto \text{Lst}(\varphi_R(a)) \quad \text{and} \quad \lambda_\varphi : a \mapsto \text{Fst}(\varphi_L(a))$$

**Proposition 3.** Let  $\varphi$  be a morphism in Class  $P_{ret}$  and  $w$  be its marker.

1. If  $\psi$  is a right conjugate of  $\varphi$ , then  $\psi$  belongs to Class  $P_{ret}$ . In particular,  $\varphi_R \in P_{ret}$ .
2. If  $\varphi = \varphi_R$ , then the marker  $w$  is the conjugate word between  $\varphi_L$  and  $\varphi_R$ , i.e.,

$$\varphi_R(a)w = w\varphi_L(a) \quad \text{for all } a \in \mathcal{A}.$$

Moreover,  $\rho_\varphi(a) = \lambda_\varphi(a)$ .

**Proof:**

First note that, by Proposition 3.8, the morphism  $\varphi$  is acyclic and thus  $\varphi_R$  and  $\varphi_L$  are well defined.

**Item (1).** Let us assume that there exists a letter  $x \in \mathcal{A}$  such that  $\psi(a)x = x\varphi(a)$  for each  $a \in \mathcal{A}$ . Then  $x$  is a suffix of  $\varphi(a)$  for each  $a \in \mathcal{A}$ . Since  $\varphi(a)w$  is a palindrome, then  $wx$  is a prefix of  $\varphi(a)w$ . Thus  $\psi(a)xwx = x\varphi(a)wx$  is a palindrome as well and the palindrome  $p = xwx$  occurs in  $\psi(a)xwx$  exactly twice (otherwise we would have a third occurrence of  $w$  as a factor of  $\varphi(a)w$ ). Therefore,  $\psi \in P_{ret}$  and  $p = xwx$  is the associated marker of  $\psi$ . We have shown that  $\psi$  obtained from  $\varphi$  by right conjugation using one letter  $x$  belongs to Class  $P_{ret}$ . As each right conjugation can be obtained step by step by conjugations using one letter only, the assertion is proven.

**Item (2).** Since  $w$  is a prefix of  $\varphi(a)w$ , we can define a new morphism  $\xi$  by setting  $\xi(a) = w^{-1}\varphi(a)w$  for each  $a \in \mathcal{A}$ . Obviously,  $\xi$  is left conjugated to  $\varphi$  and  $w$  is the conjugate word. Since  $\varphi(a)w$  is a palindrome,  $\text{Fst}(\xi(a)) = \text{Lst}(\varphi(a))$  for each  $a \in \mathcal{A}$ . By our assumption,  $\varphi$  is rightmost conjugate and thus there exist two letters, say  $b_1$  and  $b_2$ , such that  $\text{Lst}(\varphi(b_1)) \neq \text{Lst}(\varphi(b_2))$ . Therefore,  $\text{Fst}(\xi(b_1)) \neq \text{Fst}(\xi(b_2))$  and thus  $\xi$  is the leftmost conjugate of  $\varphi$ .  $\square$

**Example 3.9.** Let us consider the morphism  $\sigma$  from Example 3.2 with  $\ell = p = 1$  and  $q = 3$ . It can be checked that this morphism is not Sturmian. Its rightmost conjugate and leftmost conjugate are, respectively, the morphisms

$$\sigma_R : \begin{cases} 0 \mapsto 10 \\ 1 \mapsto 1011 \end{cases} \quad \text{and} \quad \sigma_L : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 1101 \end{cases}.$$

The morphisms  $\sigma_R$  belongs to Class  $P_{ret}$  and its marker is 101. Moreover, for all  $a \in \{0, 1\}$  one has  $\sigma_R(a)101 = 101\sigma_L(a)$  and  $\rho_\sigma(a) = \lambda_\sigma(a)$ .

The following class was introduced by Hof, Knill and Simon in [24] in order to construct words containing infinitely many palindromic factors. Such words play an important role in the study of the spectrum of the Schrödinger operator associated with an aperiodic sequence.

**Definition 2.** A morphism  $\eta : \mathcal{A}^* \rightarrow \mathcal{A}^*$  belongs to Class  $P$ , if there exists a palindrome  $p \in \mathcal{A}^*$  such that  $\eta(a) = pq_a$  for each  $a \in \mathcal{A}$ , where  $q_a$  is a palindrome.

It is easy to see that every fixed point of a substitution in Class  $P$  contains infinitely many palindromes. Let us explain the relation between Class  $P$  and Class  $P_{ret}$ .

As demonstrated in [27, Lemma 20], an acyclic morphism  $\varphi$  is conjugated to a morphism in Class  $P$  if and only if  $\varphi_R(a) = \overline{\varphi_L(a)}$  for each  $a \in \mathcal{A}$ . If it is the case, the conjugate word  $w$  is a palindrome. Using this result, Proposition 3.8 and Item (2) of Proposition 3, we deduce that every morphism in Class  $P_{ret}$  is conjugated to an acyclic morphism in Class  $P$ .

The converse is not true. For example, the morphism given by

$$\varphi : \begin{cases} 0 \mapsto 010101 \\ 1 \mapsto 0101001 \end{cases}$$

belongs to Class  $P$  and is acyclic. In particular,

$$\varphi_R : \begin{cases} 0 \mapsto 010101 \\ 1 \mapsto 0101010 \end{cases}, \quad \varphi_L : \begin{cases} 0 \mapsto 101010 \\ 1 \mapsto 0101010 \end{cases}$$

and the conjugate word is  $w = 0101010$ . Nevertheless, in  $\varphi_R(0)w$  the word  $w$  occurs four times. Therefore, no conjugate of  $\varphi$  belongs to Class  $P_{ret}$ .

## 4. Marked morphisms

In our attempt to find morphisms preserving richness we will restrict to morphisms for which "desubstitution" is easy. We use marked morphisms. Our definition is slightly more general than the definition used by Frid in [18].

**Definition 3.** An acyclic morphism  $\varphi$  is *right marked* (resp. *left marked*) if the mappings  $\rho_\varphi$  and  $\lambda_\varphi$  are injective on  $\mathcal{A}$ .

A morphism is *marked* if it is both right marked and left marked.

A marked morphism is called *well-marked* if both  $\rho_\varphi$  and  $\lambda_\varphi$  are the identity map on  $\mathcal{A}$ .

The following proposition is a direct consequence of Item (2) in Proposition 3.

**Proposition 4.** Let  $\varphi$  be a morphism in Class  $P_{ret}$ . If  $\varphi$  is right marked, then it is left marked too.

**Lemma 3.** If  $\varphi \in P_{ret}$  is right marked, then there exists a positive integer  $k$  such that  $\varphi^k$  is well-marked.

**Proof:**

The statement follows simply from the following three facts:

- 1) The mappings  $\lambda_\varphi$  and  $\rho_\varphi$  are permutations on  $\mathcal{A}$ .
- 2) If  $\pi$  is a permutation on  $d$  elements, then for  $k = d!$ , the power  $\pi^k$  is the identity.
- 3) Class  $P_{ret}$  is closed under composition (see [4]).

□

**Example 4.1.** The morphism  $\sigma$  in Example 3.9 is well marked. On the other hand, the Tribonacci morphism  $\tau$  seen in Example 2.5, which is in Class  $P_{ret}$  with marker 0, is marked but not well marked. Indeed, one has  $\tau = \tau_R$  and

$$\tau_L : \begin{cases} 0 \mapsto 10 \\ 1 \mapsto 20 \\ 2 \mapsto 0 \end{cases} .$$

On the other hand one can check that the morphism  $\tau^3$  is well-marked. Indeed, one has

$$\tau_R^3 : \begin{cases} 0 \mapsto 0102010 \\ 1 \mapsto 010201 \\ 2 \mapsto 0102 \end{cases} \quad \text{and} \quad \tau_L^3 : \begin{cases} 0 \mapsto 0102010 \\ 1 \mapsto 102010 \\ 2 \mapsto 2010 \end{cases} .$$

**Lemma 4.** Let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  and  $\varphi \in P_{ret}$  be a marked morphism over  $\mathcal{A}$  with marker  $w$ , such that  $\varphi = \varphi_R$ . Let  $a, b \in \mathcal{A}$  such that  $b = \rho_\varphi(a)$ .

- The factor  $wb$  occurs in  $\varphi(\mathbf{u})$  only as a prefix of  $\varphi(a)w = w\varphi_L(a)$ .
- The factor  $bw$  occurs in  $\varphi(\mathbf{u})$  only as a suffix of  $\varphi(a)w = w\varphi_L(a)$ .

**Proof:**

Let us set  $\varphi(\mathbf{u}) = b_0b_1b_2 \cdots$  and let us suppose that  $i$  is an occurrence of the word  $wb$  in  $\varphi(\mathbf{u})$ . Let us consider a  $j > i$  such that  $b_ib_{i+1} \cdots b_{j-1}$  is a return word to  $w$  in  $\varphi(\mathbf{u})$ . By Item (2) of Lemma 2, the index  $i$  is an occurrence of  $\varphi(c)w$  for some  $c \in \mathcal{A}$ . Since  $\varphi(c)w = w\varphi_L(c)$  is a palindrome, then  $b = \lambda_\varphi(c)$ . By Item (2) of Proposition 3, we have  $b = \rho_\varphi(c)$ . Since  $\rho_\varphi$  is injective, we have  $c = a$ .

Now, let us assume that  $i - 1$  is an occurrence of the word  $bw$  in  $\varphi(\mathbf{u})$ . Since  $\varphi(\mathbf{u})$  has a prefix  $w$ , there exists an index  $j < i$  such that  $b_jb_{j+1} \cdots b_{i-1}$  is a return word to  $w$  in  $\varphi(\mathbf{u})$ . By the same argument as above,  $i - |\varphi(c)|$  is an occurrence of  $\varphi(c)w$  in  $\varphi(\mathbf{u})$  for some  $c \in \mathcal{A}$ . Obviously,  $b = \rho_\varphi(c)$  and thus  $c = a$ .  $\square$

**Example 4.2.** Let  $\tau = \tau_R$  and  $\mathbf{t}$  be as in Example 2.5 and let  $w_\tau = 0$  be the marker of  $\tau$ . Since  $\rho_\tau(1) = 2$ , it follows from Lemma 4 that 02 occurs in  $\mathbf{t}$  only as a prefix of 020 and that 20 occurs in  $\mathbf{t}$  only as a suffix of 020.

**Proposition 4.3.** Let  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  be a marked morphism in Class  $P_{ret}$  such that  $\varphi = \varphi_R$  and  $w$  be its marker. Let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  and  $a, b \in \mathcal{A}$ . Then

$$aub \in (\mathbf{u}) \iff \rho_\varphi(a) \varphi(u)w \rho_\varphi(b) \in (\varphi(\mathbf{u})). \quad (3)$$

**Proof:**

Let us first suppose that  $aub \in (\mathbf{u})$ . By Item (3) of Lemma 2,  $\varphi(a)\varphi(u)\varphi(b)w \in (\varphi(\mathbf{u}))$ . By definition  $\varphi(a)$  has  $\rho_\varphi(a)$  as a suffix. Since  $\varphi = \varphi_R$ , we can apply Item (2) of Proposition 3 to get  $\varphi(b)w = w\varphi_L(b)$ . Therefore,  $\varphi(b)w$  starts with  $w\lambda_\varphi(b) = w\rho_\varphi(b)$ . Hence,  $\varphi(a)\varphi(u)\varphi(b)w$  has a factor  $\rho_\varphi(a)\varphi(u)w\rho_\varphi(b)$ .

To show the opposite implication, let us assume that  $i - 1$  is an occurrence in  $\varphi(\mathbf{u})$  of the factor  $\rho_\varphi(a)\varphi(u)w\rho_\varphi(b) \in (\varphi(\mathbf{u}))$ . By the second assertion of Lemma 4,  $i - |\varphi(a)|$  is an occurrence of the factor  $\varphi(a)\varphi(u)w\rho_\varphi(b)$ . By the first assertion of the same lemma,  $i - |\varphi(a)|$  is an occurrence of the factor  $\varphi(a)\varphi(u)\varphi(b)w = \varphi(aub)w$ . Since  $w$  is a prefix and a suffix of  $\varphi(aub)w$ , Corollary 2 implies  $aub \in (\mathbf{u})$ .  $\square$

The next result describes the relation between the set  $\mathcal{B}(\mathbf{u})$  of bispecial factors in  $\mathbf{u}$  and the set  $\mathcal{B}(\varphi(\mathbf{u}))$  of bispecial factors in  $\varphi(\mathbf{u})$ , when  $\varphi \in P_{ret}$ . Given a permutation  $\pi$  on  $\mathcal{A}$  and a set  $S \subset \mathcal{A}$  we denote  $\pi(S) = \{\pi(s) : s \in S\}$ .

**Corollary 3.** Let  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  be a marked morphism in Class  $P_{ret}$  such that  $\varphi = \varphi_R$  and  $w$  be its marker. Let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ .

1. If  $v \in \mathcal{B}(\varphi(\mathbf{u}))$  and  $w$  is a factor of  $v$ , then there exists  $u \in \mathcal{B}(\mathbf{u})$  such that  $v = \varphi(u)w$ .
2. For every  $u \in (\mathbf{u})$  we have  $\rho_\varphi(L_{\mathbf{u}}(u)) = L_{\varphi(\mathbf{u})}(\varphi(u)w)$ ,  $\rho_\varphi(R_{\mathbf{u}}(u)) = R_{\varphi(\mathbf{u})}(\varphi(u)w)$ , and  $\rho_\varphi(B_{\mathbf{u}}(u)) = B_{\varphi(\mathbf{u})}(\varphi(u)w)$ .
3.  $u \in \mathcal{B}(\mathbf{u})$  if and only if  $\varphi(u)w \in \mathcal{B}(\varphi(\mathbf{u}))$ .
4.  $b_{\mathbf{u}}(u) = b_{\varphi(\mathbf{u})}(\varphi(u)w)$  for every  $u \in (\mathbf{u})$ .
5.  $\#Pext_{\mathbf{u}}(u) = \#Pext_{\varphi(\mathbf{u})}(\varphi(u)w)$  for every palindromic  $u \in (\mathbf{u})$ .

**Proof:**

**Item (1).** Let  $v$  be a bispecial factor in  $(\varphi(\mathbf{u}))$  and  $w$  a factor of  $v$ . There exist two letters  $b_1$  and  $b_2$ ,  $b_1 \neq b_2$  such that  $vb_1, vb_2 \in (\varphi(\mathbf{u}))$ . We show that  $w$  is a suffix of  $v$ . Assume the contrary, i.e., that  $v = ywx$ , with  $x, y \in \mathcal{A}^*$ , such that the factor  $w$  occurs in  $wx$  only once and  $x \neq \varepsilon$ . Let  $b$  be the first letter of  $x$ . By Lemma 4 and the fact that both  $wxb_1, wxb_2 \in (\varphi(\mathbf{u}))$ , we have that  $\varphi(a)w$  is a prefix of  $wx$ , where  $b = \rho_\varphi(a)$ , a contradiction with the definition of  $x$ . Therefore  $w$  is a suffix of  $v$ . By the same argument,  $w$  is also a prefix of  $v$ . Corollary 2 implies that  $v = \varphi(u)w$ .

**Item (2).** Proposition 4.3 says that  $(a, b)$  is a bi-extension of  $u$  in  $(\mathbf{u})$  if and only if  $(\rho_\varphi(a), \rho_\varphi(b))$  is a bi-extension of  $\varphi(u)w$  in  $(\varphi(\mathbf{u}))$ .

**Item (3).** A consequence of the previous item.

**Item (4).** A consequence of the previous item.

**Item (5).** A consequence of Equation (3) and Item (3) of Lemma 1.  $\square$

**Example 4.4.** Let  $\tau = \tau_R$ ,  $\mathbf{t}$  and  $w_\tau$  be as in Example 4.2. One has  $\rho_\tau = (1, 2, 0)$ . Let us consider the word  $u = 2010$ . One has

$$L_{\mathbf{t}}(u) = \{0\}, \quad R_{\mathbf{t}}(u) = \{0, 1, 2\}, \quad B_{\mathbf{t}}(u) = \{(0, 0), (0, 1), (0, 2)\}$$

and

$$L_{\mathbf{t}}(\tau(u)0) = \{1\}, \quad R_{\mathbf{t}}(\tau(u)0) = \{1, 2, 0\}, \quad B_{\mathbf{t}}(\tau(u)0) = \{(0, 1), (0, 2), (0, 0)\}.$$

**Remark 4.5.** Let  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  be recurrent and  $\varphi$  and  $\psi$  be two mutually conjugate morphisms over  $\mathcal{A}$ . Clearly, the languages of  $\varphi(\mathbf{u})$  and  $\psi(\mathbf{u})$  coincide. Since the palindromic richness can be seen as a property of a language and not of an infinite word itself, it is enough to examine richness for one of these languages.

**Corollary 4.** Let  $\varphi : \mathcal{A}^* \rightarrow \mathcal{A}^*$  be a marked morphism in Class  $P_{ret}$  and  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  such that its language is closed under reversal. If  $\varphi(\mathbf{u})$  is rich, then  $\mathbf{u}$  is rich.

**Proof:**

Let us recall that if a language is closed under reversal, then it is recurrent. By Remark 4.5, we may assume without loss of generality that  $\varphi = \varphi_R$ . Let  $w$  denote the marker associated to  $\varphi$ . By Corollary 2, the language of  $\varphi(\mathbf{u})$  is closed under reversal as well. Therefore, to demonstrate the richness of  $\mathbf{u}$  we may use Theorem 2. If  $u \in \mathcal{B}(\mathbf{u})$ , then by Item (1) of Corollary 3,  $\varphi(u)w \in \mathcal{B}(\varphi(\mathbf{u}))$ . Since  $\varphi(\mathbf{u})$  is rich, we have that  $b_{\varphi(\mathbf{u})}(\varphi(u)w) = \#\text{Pext}_{\varphi(\mathbf{u})}(\varphi(u)w) - 1$  if  $\varphi(u)w$  is a palindrome, and  $b_{\varphi(\mathbf{u})}(\varphi(u)w) = 0$  otherwise. Recall that by Lemma 1,  $u$  is a palindrome if and only if  $\varphi(u)w$  is a palindrome. Therefore, Items (4) and (5) of Corollary 3 imply that  $b_{\mathbf{u}}(u) = \#\text{Pext}_{\mathbf{u}}(u) - 1$  if  $u$  is a palindrome, and  $b_{\mathbf{u}}(u) = 0$  otherwise. Consequently,  $\mathbf{u}$  is rich.  $\square$

**Remark 4.6.** Let us point out that the elementary Arnoux-Rauzy morphism  $\psi_a$ , with  $a \in \mathcal{A}$ , belongs to Class  $P_{ret}$  and  $w = a$  is its associated marker. Moreover,  $\psi_a$  is well marked. Hence Lemma 5.3. in [21] is a special case of the previous corollary for  $\varphi = \psi_a$ .

## 5. Morphisms preserving richness

Our aim is to study the implication opposite to Corollary 4, i.e., for which morphism  $\varphi$  in Class  $P_{ret}$ , richness of  $\mathbf{u}$  forces richness of  $\varphi(\mathbf{u})$ . Behaviour of morphisms from this class may differ according to the infinite word we apply the morphism to. This is testified by the following example (see [4]).

**Example 5.1.** Let us consider the morphism  $\zeta : \{0, 1, 2\}^* \mapsto \{0, 1\}^*$  defined as

$$\zeta : \begin{cases} 0 \mapsto 0100 \\ 1 \mapsto 01011 \\ 2 \mapsto 010111 \end{cases} .$$

This morphism is in Class  $P_{ret}$ . It is proved in [4, Lemma 5.6] that  $\zeta$  maps a rich word (over the ternary alphabet  $\{0, 1, 2\}$ ) to a word with infinite palindromic defect (over the binary alphabet  $\{0, 1\}$ ). Moreover, the same morphism also maps the Tribonacci word (which is rich) to a rich word.

Starosta studied in [35] morphic images of episturmian words by morphisms of Class  $P_{ret}$ . It is known that episturmian words are rich (see, for instance, [20]).

**Theorem 3.** ([35]) Let  $\mathbf{u}$  be an episturmian word and  $\varphi : \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a morphism of Class  $P_{ret}$ . The word  $\varphi(\mathbf{u})$  is almost rich.

In [21, Corollary 6.3] it is proved that the image of each rich word by an Arnoux-Rauzy morphism is rich as well. Here we provide an alternative proof of this statement.

**Theorem 4.** ([21]) Let  $\varphi$  be an Arnoux-Rauzy morphism over the alphabet  $\mathcal{A}$  and  $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$  such that its language is closed under reversal. Then  $\mathbf{u}$  is rich if and only if  $\varphi(\mathbf{u})$  is rich.

**Proof:**

Since the Arnoux-Rauzy monoid is generated by elementary morphisms, let us prove the result for these morphisms.

Let us first consider the case  $\varphi = \psi_a$  for a certain letter  $a \in \mathcal{A}$ .

By Remark 4.6, if  $\psi_a(\mathbf{u})$  is rich, then  $\mathbf{u}$  is rich as well. To prove the opposite implication let us assume that  $\mathbf{u}$  is rich and let us prove that for all bispecial words in  $\mathcal{L}(\psi_a(\mathbf{u}))$  Equation (1) is satisfied. The empty word  $\varepsilon$  is clearly a palindrome. The only palindromic extension of the empty word in  $(\psi_a(\mathbf{u}))$  is  $aa$ , and  $b_{\psi_a(\mathbf{u})}(\varepsilon) = 1 - 1 = 0$ . The only bispecial factor of length 1 is  $a$ , that is clearly a palindrome. Using Items (4) and (5) of Corollary 3 we have:

$$\#\text{Pext}_{\psi_a(\mathbf{u})}(a) = \#\text{Pext}_{\mathbf{u}}(\varepsilon) = b_{\mathbf{u}}(\varepsilon) + 1 = b_{\psi_a(\mathbf{u})}(a) + 1.$$

Let us now consider a bispecial word  $v \in (\psi(\mathbf{u}))$  of length at least 2. Since the only left special (resp. right special) letter is  $a$ , we can write  $v = av'a$  for a certain word  $v'$ . Moreover  $av' = \psi_a(u)$  for a certain word  $u \in (\mathbf{u})$ , hence  $u = \psi_a(w)a$  and, by Item (3) of Corollary 3,  $u \in \mathcal{B}(\mathbf{u})$ . Using Item (3) of Lemma 1, we have that  $v$  is a palindrome if and only if  $u$  is a palindrome. Using the same reasoning as before we find that whenever  $v$  is a palindrome, then

$$\#\text{Pext}_{\psi_a(\mathbf{u})}(v) = \#\text{Pext}_{\mathbf{u}}(u) = b_{\mathbf{u}}(u) + 1 = b_{\psi_a(\mathbf{u})}(u) + 1,$$

while whenever  $v$  is not a palindrome, then

$$b_{\psi_a(\mathbf{u})}(v) = b_{\mathbf{u}}(u) = 0.$$

Thus, according to Theorem 2,  $\psi_a(\mathbf{u})$  is rich.

Let us now consider the case  $\varphi = \bar{\psi}_a$  for a certain letter  $a \in \mathcal{A}$ . We note that  $\bar{\psi}_a$  is conjugated to  $\psi_a$ . Due to Remark 4.5 we have that

$$\bar{\psi}_a(\mathbf{u}) \text{ is rich} \iff \psi_a(\mathbf{u}) \text{ is rich} \iff \mathbf{u} \text{ is rich}$$

Finally, let us consider a morphism permuting letters. Since the definition of richness is based on counting palindromes occurring in a factor of  $\mathbf{u}$ , permutations of letters do not change the number of palindromes and the statement of the theorem is trivially verified.  $\square$

In the case of the binary alphabet the set of morphisms in Class  $P_{ret}$  which preserve richness can be extended even more. Indeed, in a binary alphabet, every acyclic morphism  $\psi$  is marked and moreover a morphism in Class  $P_{ret}$  is acyclic by Proposition 3.8.



**Lemma 5.** Let  $w$  be the marker of a binary morphism  $\varphi = \varphi_R$  in Class  $P_{ret}$  and  $\mathbf{u} \in \{0, 1\}^{\mathbb{N}}$  be a recurrent word. Assume that  $\mathbf{u}$  and  $\varphi(01)w$  are rich. Each bispecial factor of  $\varphi(\mathbf{u})$  which does not contain the marker  $w$  satisfies Condition (1) of Theorem 2.

**Proof:**

Let  $v$  be a bispecial factor of  $\varphi(\mathbf{u})$  which does not contain  $w$  as a factor. By Lemma 2, we deduce that  $v$  is a factor of  $c^{-1}\varphi(0)wc^{-1}$  or  $c^{-1}\varphi(1)wc^{-1}$ , where  $c = \text{Fst}(w) = \text{Lst}(w)$ . Since  $\varphi(0)w$  and  $\varphi(1)w$  are palindromes, we have

$$xvy \in (\varphi(01)w) \iff y\bar{v}x \in (\varphi(01)w) \quad \text{for } x, y \in \{0, 1\}, \quad (4)$$

and

$$xvy \in (\varphi(01)w) \iff xvy \in (\varphi(\mathbf{u})) \quad \text{for } x, y \in \{0, 1\}. \quad (5)$$

Let us consider the set of bi-extensions  $\mathfrak{B} = \{xvy : (x, y) \in B_{\varphi(\mathbf{u})}(v)\}$ . By Equation (5), we have  $\mathfrak{B} \subset (\varphi(01)w)$ . Let us show that  $v$  satisfies condition (1) of Theorem 2, by considering the bilateral order of  $v$ . Note that, since we are in a binary alphabet, the only possible values for  $b_{\varphi(\mathbf{u})}(v)$  are 0, 1 and  $-1$ .

**Case  $b_{\varphi(\mathbf{u})}(v) = 0$**  In this case we have  $\#\mathfrak{B} = 3$ . If  $v$  is not a palindrome, then Equation (1) is trivially fulfilled. If  $v$  is a palindrome, then Equation (4) implies that  $\mathfrak{B} = \{ava, avb, bva\}$  for  $a, b \in \{0, 1\}$ , and  $a \neq b$ . Therefore,  $\#\text{Pext}_{\varphi(\mathbf{u})}(v) = 1$  and Equation (1) is satisfied.

**Case  $b_{\varphi(\mathbf{u})}(v) = 1$**  In this case we have  $\mathfrak{B} = \{0v1, 1v0, 0v0, 1v1\}$ . If  $v$  is a palindrome, then we have  $\#\text{Pext}_{\varphi(\mathbf{u})}(v) = 2$  and Equation (1) is fulfilled. If  $v$  is not a palindrome, then Equation (4) implies that  $0v0, 1v1, 0\bar{v}1$  and  $1\bar{v}0$  occur in  $\varphi(01)w$ . By Lemma 2.4, this is a contradiction with richness of  $\varphi(01)w$ . Therefore,  $b_{\varphi(\mathbf{u})}(v)$  cannot be positive for a non-palindromic  $v$ .

**Case  $b_{\varphi(\mathbf{u})}(v) = -1$**  In this case  $\#\mathfrak{B} = 2$ . First assume that  $\mathfrak{B} = \{0v1, 1v0\}$ . If  $v$  is a palindrome, then  $\#\text{Pext}_{\varphi(\mathbf{u})}(v) = 0$  and Equation (1) is fulfilled. If  $v$  is not a palindrome, then by Lemma 2.3 occurrences of  $v$  and  $\bar{v}$  alternate in  $c^{-1}\psi(01)wc^{-1}$  and occurrences of  $0v1$  and  $1\bar{v}0$  alternate in  $\varphi(01)w$ . Item (2) of Lemma 2.3 implies that  $0v1$  and  $1v0$  cannot occur simultaneously in  $\varphi(01)w$  which leads to a contradiction. Therefore,  $\mathfrak{B} \neq \{0v1, 1v0\}$  for a non-palindromic  $v$ . Moreover,  $\mathfrak{B} \neq \{0v0, 1v1\}$ . Indeed, by Item (2) of Lemma 2.3 we can not have simultaneous occurrences of  $0v0$  and  $1v1$  in  $\varphi(01)w$  for both palindromic and non-palindromic  $v$ . □

**Example 5.2.** Let  $\mathbf{f}$  be as in Example 2.5 and  $\varphi_1$  and  $w_1$  as in Example 3.1. Let  $\mathbf{v} = \varphi_1(\mathbf{f})$ . The word  $\varphi_1(01)w_1 = 11101111$  is rich. The only bispecial factors of  $\mathbf{v}$  which do not contain  $w_1$  as a factor are the palindromes  $\varepsilon, 1$  and  $11$ . One can check that  $b_{\mathbf{v}}(\varepsilon) = b_{\mathbf{v}}(1) = b_{\mathbf{v}}(11) = 0$  and that  $\#\text{Pext}_{\mathbf{v}}(\varepsilon) = \#\text{Pext}_{\mathbf{v}}(1) = \#\text{Pext}_{\mathbf{v}}(11) = 1$ .

**Lemma 6.** Let  $\mathbf{u}$  be a fixed point of a primitive morphism  $\varphi : \{0, 1\}^* \rightarrow \{0, 1\}^*$  conjugated to a morphism in Class  $P_{ret}$ . Then  $(\mathbf{u})$  is closed under reversal.



**Corollary 5.** Let  $\varphi$  be a binary morphism in Class  $P_{ret}$  and  $\mathbf{u} \in \{0, 1\}^{\mathbb{N}}$  be a non-unary recurrent word. If  $\varphi(\mathbf{u})$  is rich, then  $\varphi(\mathbf{v})$  is rich for every recurrent rich word  $\mathbf{v} \in \{0, 1\}^{\mathbb{N}}$ .

**Proof:**

Let  $w$  be the marker of  $\varphi_R$ . Since  $\mathbf{u}$  is non-unary recurrent word, then  $01 \in (\mathbf{u})$ . Remark 4.5 and Item (4) of Lemma 2 imply that  $\varphi_R(01)w$  is a factor of  $\mathbf{u}$  and thus  $\varphi_R(01)w$  is rich. The statement follows by Item (1) of the previous theorem.  $\square$

**Example 5.4.** Let  $\phi$  be as in Example 2.5 and  $\sigma$  as in Example 5.3. Since  $\mathbf{f} = \phi(\mathbf{f})$  is recurrent and rich, we have that

$$\phi(\sigma(0)^\omega) = 0100100001001000010000100001001000010010000100 \dots$$

is rich. Moreover, images of rich periodic words by  $\phi$  (or of any other Sturmian morphism) are rich as well.

## 6. Construction of new rich words

Theorems 4 and 5 enable us to construct new rich words from known ones. We give several examples.

On alphabets of cardinality  $d$ , two classes of infinite words are known to be rich:

- episturmian words and
- the so-called *d-iet words*, that is words coding  $d$ -interval exchange transformation under the symmetric permutation  $(d, d - 1, \dots, 2, 1)$ .

Applying an arbitrary Arnoux-Rauzy morphism to a  $d$ -iet word gives a new rich word which does not belong to any of the two classes mentioned above.

Although both our theorems consist in applying a richness-preserving morphism to infinite words, we can apply them to finite words as well. The construction of infinite words J. Vesti provided in the proof of Proposition 1 is based on repetitive application of palindromic closure. Therefore, the languages of infinite words he constructed are closed under reversal and thus we can apply our theorems to factors of these words to get new finite rich words.

In [22], Guo, Shallit and Shur gave a lower bound on the number of binary rich words of length  $n$ . They proved that every word of the form  $0^{a_1}1^{b_1}0^{a_2}1^{b_2} \dots 0^{a_k}1^{b_k}$ , where  $a_1 \leq a_2 \leq \dots \leq a_k$  and  $b_1 \leq b_2 \leq \dots \leq b_k$ , is rich and thus they counted words of this form of length  $n$ . Mapping any of these rich words through a Sturmian morphism or a morphism from Theorem 5 gives us a new rich word.

To give another type of examples of morphisms in Class  $P_{ret}$  which have rich fixed points let us recall the definition of derived words as introduced by Durand in [17].

**Definition 4.** Let  $\mathcal{R}_{\mathbf{u}}(x) = \{r_1, r_2, \dots, r_k\}$  be the set of return words to a factor  $x$  of a uniformly recurrent word  $\mathbf{u}$ . Let  $z$  be the shortest prefix of  $\mathbf{u}$  such that  $\mathbf{u}$  can be written in the form  $\mathbf{u} = z r_{d_0} r_{d_1} r_{d_2} \dots$ , where  $r_{d_j} \in \mathcal{R}_{\mathbf{u}}(x)$  for each  $j \in \mathbb{N}$ . The *derived word* of  $\mathbf{u}$  to the factor  $x$  is the infinite word  $\mathbf{d}_{\mathbf{u}}(x) = d_0 d_1 d_2 \dots$  over the alphabet of cardinality  $\#\mathcal{R}_{\mathbf{u}}(x) = k$ .

**Example 6.1.** Let us consider the binary morphism

$$\xi : \begin{cases} 0 \mapsto 11 \\ 1 \mapsto 10 \end{cases}$$

fixing the period doubling sequence

$$\mathbf{x} = 1011101010111011101 \dots$$

which is known to be rich (see [10]). The morphism  $\xi$  does not belong to  $P_{ret}$ , nevertheless it is in Class  $P$ . Let us find the derived word of  $\mathbf{x}$  to the prefix  $p = 1$ . This prefix has two return words, namely  $\mathbf{r} = 10$  and  $\mathbf{s} = 1$ . We thus have

$$\mathbf{x} = \underbrace{10}_{\mathbf{r}} \underbrace{11}_{\mathbf{s}} \underbrace{10}_{\mathbf{r}} \underbrace{10}_{\mathbf{r}} \underbrace{11}_{\mathbf{s}} \underbrace{10}_{\mathbf{r}} \underbrace{11}_{\mathbf{s}} \underbrace{10}_{\mathbf{r}} \underbrace{11}_{\mathbf{s}} \underbrace{10}_{\mathbf{r}} \underbrace{11}_{\mathbf{s}} \dots$$

and the derived word over the alphabet  $\{\mathbf{r}, \mathbf{s}\}$  is

$$\mathbf{d}_{\mathbf{x}}(p) = \mathbf{r} \mathbf{s} \mathbf{s} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{r} \mathbf{s} \mathbf{s} \mathbf{r} \mathbf{s} \mathbf{s} \mathbf{r} \mathbf{s} \dots$$

By a result of Durand ([17]), the derived word to a prefix is fixed by a substitution. Since  $\xi(\mathbf{r}) = 1011 = \mathbf{r} \mathbf{s} \mathbf{s}$  and  $\xi(\mathbf{s}) = 10 = \mathbf{r}$ , the substitution fixing  $\mathbf{d}_{\mathbf{x}}(p)$  is

$$\eta : \begin{cases} \mathbf{r} \mapsto \mathbf{r} \mathbf{s} \mathbf{s} \\ \mathbf{s} \mapsto \mathbf{r} \end{cases}.$$

One can easily check that  $\eta : \{\mathbf{r}, \mathbf{s}\}^* \rightarrow \{\mathbf{r}, \mathbf{s}\}^*$  is in Class  $P_{ret}$ , with the marker  $w = \mathbf{r}$ . It is well-marked and the word  $\eta(\mathbf{r} \mathbf{s}) w = \mathbf{r} \mathbf{s} \mathbf{s} \mathbf{r} \mathbf{r}$  is rich. Theorem 5 implies that the derived word to the prefix  $p = 1$  is rich as well.

What seen in the previous example is not surprising. Indeed in [23, Lemma 3.3] it is stated that the derived word to the first letter of a rich word is rich as well. The same result for a palindromic factor of a rich word is contained in the proof of [4, Theorem 5.5.].

**Lemma 7. ([23, 4])**

Let  $\mathbf{u}$  be a rich infinite word. The derived word  $\mathbf{d}_{\mathbf{u}}(x)$  to any palindromic factor  $x$  of  $\mathbf{u}$  is rich.

The assumption that  $x$  is palindromic in the previous lemma is necessary, as shown in the following remark.

**Remark 6.2.** In [30], the derived words to bispecial factors of complementary symmetric Rote sequences are described. The derived word to each bispecial factor is the coding of a three interval exchange transformation. If a bispecial factor  $x$  is palindromic, then the interval exchange transformation is given by the symmetric permutation  $(3, 2, 1)$ . If a bispecial factor  $x$  is non-palindromic, then the corresponding permutation is  $(2, 3, 1)$ . In the latter case the language of the derived word  $\mathbf{d}_{\mathbf{u}}(x)$  is not closed under reversal and contains only finitely many palindromes. Therefore,  $\mathbf{d}_{\mathbf{u}}(x)$  cannot be rich. In fact, the defect of  $\mathbf{d}_{\mathbf{u}}(x)$  is infinite.



$\text{Lst}(r_0) \neq \text{Lst}(r_2)$ . Since  $r_1r_0$  and  $r_2r_0$  are factors of  $\mathbf{v}$ , the above described construction gives two possible substitutions:

$$\alpha : \begin{cases} \mathbf{r} \mapsto \mathbf{r r s s r s s r s s} \\ \mathbf{s} \mapsto \mathbf{r} \end{cases} \quad \text{and} \quad \beta : \begin{cases} \mathbf{r} \mapsto \mathbf{r r s s} \\ \mathbf{s} \mapsto \mathbf{r} \end{cases}.$$

It is easy to verify that the dominant eigenvalue of  $M_\alpha$  is  $\Lambda_\alpha = 2 + \sqrt{10}$ , whereas  $\Lambda_\beta = 1 + \sqrt{3}$ . Therefore, fixed points of  $\eta$ ,  $\alpha$  and  $\beta$  are rich and differ substantially. We can continue applying the same construction to the fixed points of the new substitutions.

## 7. Comments and open questions

Theorem 5 gives a simple criterion to decide whether a fixed point of a binary primitive morphism  $\varphi$  in Class  $P_{ret}$  is rich: namely, one has to check richness of the finite word  $\varphi_R(01)w$ . It would be helpful to find a similar criterion for morphisms over larger alphabets starting, for instance, with marked ternary morphisms in Class  $P_{ret}$ . The main technical obstacle is the absence of a ternary analogue of Lemma 2.4.

On ternary alphabet, every word with language closed under reversal with factor complexity  $\mathcal{C}(n) = 2n + 1$ , i.e., having exactly  $2n + 1$  factors of length  $n$  for every integer  $n \geq 0$ , is rich (see [2]). This can be generalized to the class of *dendric sets*, i.e., factorial sets such that the extension graph of every word in the set is acyclic and connected. Indeed, all dendric sets closed under reversal are rich (see [9]). It would be useful to characterize all dendric sets which are closed under reversal.

A morphism over an alphabet  $\mathcal{A}$  is called *tame* if it belongs to the submonoid generated by the permutations of letters of  $\mathcal{A}$  and the morphisms  $\alpha_{a,b}, \tilde{\alpha}_{a,b}$  defined for  $a, b \in \mathcal{A}$  with  $a \neq b$  by

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

(see, for instance [8], where the definition is given for morphisms of the free group instead of the free monoid). All Arnoux-Rauzy morphisms are tame, but the opposite is not true (the two classes coincide only when  $\#\mathcal{A} \leq 2$ ). As an example we can consider the morphism

$$\psi = \pi_{(c a b)} \circ \alpha_{a,b} \circ \tilde{\alpha}_{a,b} \begin{cases} \mathbf{a} \mapsto \mathbf{aca} \\ \mathbf{b} \mapsto \mathbf{a} \\ \mathbf{c} \mapsto \mathbf{b} \end{cases}$$

on the ternary alphabet  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ . This morphism is tame but not Arnoux-Rauzy. It also belongs to Class  $P$  and therefore sends a language closed under reversal to a language closed under reversal.

However not all tame morphisms preserve richness. An interesting question would be to characterize tame morphisms that map rich words to rich words.

Baranwal and Shallit in [5] and [6] looked for infinite rich words over  $\mathcal{A}$  with the minimal critical exponent. They found lower bounds on the minimum value for cardinality of the alphabet  $k = 2, 3, 4, 5$ . In [14], Curie, Mol and Rampersad proved that for  $k = 2$  the suggested bound is the best possible. Solving some of the previous questions about richness, may help determining the optimal lower bound for  $k \geq 3$ .

## Acknowledgement

The research received funding from the Ministry of Education, Youth and Sports of the Czech Republic through the projects CZ.02.1.01/0.0/0.0/16\_019/0000765 and CZ.02.1.01/0.0/0.0/16\_019/0000778.

## References

- [1] Arnoux P, Rauzy G. *Représentation géométrique de suites de complexité  $2n + 1$* , Bulletin de la Société Mathématique de France, 1991. 119(2):199–215. doi:10.24033/bsmf.2164.
- [2] L. Balková, E. Pelantová, Š. Starosta, *Palindromes in infinite ternary words*, RAIRO - Theoretical Informatics and Applications, 2009. 43(4):687–702. doi:10.1051/ita/2009016.
- [3] Balková L, Pelantová E, Starosta Š. *Sturmian Jungle (or Garden?) on Multiliteral Alphabets*, RAIRO - Theoretical Informatics and Applications, 2010. 44(4):443–470. doi:10.1051/ita/2011002.
- [4] Balková L, Pelantová E, Starosta Š. *Infinite Words with Finite Defect*, Advances in Applied Mathematics, 2011. 47(3):562–574. doi:10.1016/j.aam.2010.11.006.
- [5] Baranwal AR. *Decision Algorithms for Ostrowski-Automatic Sequences*, Master Thesis, University of Waterloo (2020). ID: 229124363.
- [6] Baranwal AR, Shallit J. *Repetitions in infinite palindrome-rich words*, in: R. Mercas and D. Reidenbach (eds.), Proceedings WORDS 2019, Lecture Notes in Computer Science, vol. 11682, Springer, 2019 pp. 93–105. doi:10.1007/978-3-030-28796-2\_7.
- [7] Berthé V, De Felice C, Dolce F, Leroy J, Perrin D, Reutenauer C, Rindone G. *Acyclic, connected and tree sets*, Monatshefte für Mathematik, 2015. 176:521–550. doi:10.1007/s00605-014-0721-4.
- [8] Berthé V, De Felice C, Dolce F, Leroy J, Perrin D, Reutenauer C, Rindone G. *Maximal bifix decoding*, Discrete Mathematics, 2015. 338(5):725–742. doi:10.1016/j.disc.2014.12.010.
- [9] Berthé V, De Felice C, Delecroix V, Dolce F, Leroy J, Perrin D, Reutenauer C, Rindone G. *Specular sets*, Theoretical Computer Science, 2017. 684:3–28. doi:10.1016/j.tcs.2017.03.001.
- [10] Blondin Massé A, Brlek S, Labbé S, Vuillon L. *Palindromic complexity of codings of rotations*, Theoretical Computer Science, 2011. 412(46):6455–6463. doi:10.1016/j.tcs.2011.08.007.
- [11] Brlek S, Hamel S, Nivat M, Reutenauer C. *On the palindromic complexity of infinite words*, in: J. Berstel, J. Karhumäki, D. Perrin (Eds.), Combinatorics on Words with Applications, International Journal of Foundations of Computer Science, 2004. 15(2):293–306. doi:10.1142/S012905410400242X.
- [12] Bucci M, De Luca A, Glen A, Zamboni LQ. *A connection between palindromic and factor complexity using return words*, Advances in Applied Mathematics. 2009. 42(1):60–74. doi:10.1016/j.aam.2008.03.005.
- [13] Cassaigne J. *An algorithm to test if a given circular HDOL-language avoids a pattern*, in: IFIP World Computer Congress'94, Elsevier, 1994. pp. 459–464. ID:17440661.

- [14] Currie JD, Mol L, Rampersad N. *The repetition threshold for binary rich words*, Discrete Mathematics & Theoretical Computer Science, 2020. vol. 22 no 1. doi:10.23638/DMTCS-22-1-6.
- [15] Droubay X, Justin J, Pirillo G. *Episturmian words and some constructions of de Luca and Rauzy*, Theoretical Computer Science, 2001. 255(1-2):539–553. doi:10.1016/S0304-3975(99)00320-5.
- [16] Durand F. *A generalization of Cobham Theorem*, Theory of Computing Systems, 1998. 31:169–185. Id: hal-00303322.
- [17] Durand F. *A characterization of substitutive sequences using return words*, Discrete Mathematics, 1998. 179(1-3):89–101. doi:10.1016/S0012-365X(97)00029-0.
- [18] Frid A. *Applying a uniform marked morphism to a word*, Discrete Mathematics and Theoretical Computer Science, 1999. 3(3):125–139. doi:10.46298/dmtcs.255.
- [19] Frid A, Puzynina S, Zamboni LQ. *On palindromic factorization of words*, Advances in Applied Mathematics, 2013. 50(5) 737–748. doi:10.1016/j.aam.2013.01.002.
- [20] Glen A, Justin J. *Episturmian words: a survey*, RAIRO - Theoretical Informatics and Applications, 2009. 43(3)403–442. doi:10.1051/ita/2009003.
- [21] Glen A, Justin J, Widmer S, Zamboni LQ. *Palindromic richness*, European Journal of Combinatorics, 2009. 30(2):510–531. doi:10.1016/j.ejc.2008.04.006.
- [22] Guo C, Shallit J, Shur AM. *Palindromic rich words and run-length encodings*, Information Processing Letters, 2016. 116(12):735–738. doi:10.1016/j.ipl.2016.07.001.
- [23] Harju T, Vesti J, Zamboni LQ. *On a question of Hof, Knill and Simon on palindromic substitutive systems*, Monatshefte für Mathematik, 2016. 179:379–388. Id: hal-01829318.
- [24] Hof A, Knill O, Simon B. *Singular continuous spectrum for palindromic Schrödinger operators*, Communications in mathematical physics, 1995. 174(1):149–159. doi:10.1007/BF02099468.
- [25] Justin J, Pirillo G. *Episturmian words and episturmian morphisms*, Theoretical Computer Science, 2002. 276(1-2):281–313. doi:10.1016/S0304-3975(01)00207-9.
- [26] Klouda K. *Bispecial factors in circular non-pushy DOL languages*, Theoretical Computer Science, 2012. 445(3):63–74. doi:10.1016/j.tcs.2012.05.007.
- [27] Labbé S, Pelantová E. *Palindromic sequences generated from marked morphisms*, European Journal of Combinatorics, 2016. 51:200–214. doi:10.1016/j.ejc.2015.05.006.
- [28] Meyer CD. Matrix Analysis and Applied Linear Algebra, SIAM, 2000. ISBN-13:978-0898714548, 10:0898714540.
- [29] Lothaire M. *Algebraic Combinatorics on Words*, Cambridge University Press, 2002. ISBN-13:978-0521180719, 10:0521180716.
- [30] Medková K, Pelantová E, Vuillon L. *Derived sequences of complementary symmetric Rote sequences*, RAIRO - Theoretical Informatics and Applications, 2019. 53(3-4):125–151. doi:10.1051/ita/2019004.
- [31] Pelantová E, Starosta Š. *On Words with the Zero Palindromic Defect*, In: Brlek S., Dolce F., Reutenauer C., Vandomme É. (eds), Combinatorics on Words. WORDS 2017, Lecture Notes in Computer Science, vol. 10432. Springer, Cham, 2017. doi:10.1007/978-3-319-66396-8\_7.
- [32] Pytheas Fogg N. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of Lecture Notes in Mathematics. Springer-Verlag (2002). Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. ISBN-13:978-3540441410, 10:3540441417.



- [33] Rubinchik M, Shur AM. *EERTREE: An efficient data structure for processing palindromes in strings*, European J. Combin., 2018. 68:249–265. doi:10.1016/j.ejc.2017.07.021.
- [34] Rukavicka J. *On the number of Rich Words*, In E. Charlier, J. Leroy, M. Rigo (Eds), DLT 2017, Lecture Notes in Computer Science, vol 10396, Springer (2017), 345–352.
- [35] Starosta Š. *Morphic images of episturmian words having finite palindromic defect*, European Journal of Combinatorics, 2016. 51:359–371.
- [36] Vesti J. *Extensions of rich words*, Theoretical Computer Science, 2014. 548(4): 14–24. doi:10.1016/j.tcs.2014.06.033.