International Journal of Foundations of Computer Science
Vol. 23, No. 5 (2012) 1067–1083

DOI: 10.1142/S012905411240045X

ALMOST RICH WORDS AS MORPHIC IMAGES OF RICH WORDS

EDITA PELANTOVÁ
Department of Mathematics, FNSPE, Czech Technical University in Prague
Trojanova 13, 120 00 Praha 2, Czech Republic
edita.pelantova@fjfi.cvut.cz

ŠTĚPÁN STAROSTA
Department of Mathematics, FNSPE and Department of Applied Mathematics
FIT, Czech Technical University in Prague
Trojanova 13, 120 00 Praha 2, Czech Republic
stepan.starosta@fit.cvut.cz

Received 28 October 2011
Accepted 16 February 2012
Communicated by Alberto Leporati

We focus on Θ-rich and almost Θ-rich words over a finite alphabet $\mathcal{A}$, where Θ is an involutive antimorphism over $\mathcal{A}^*$. We show that any recurrent almost Θ-rich word $\mathbf{u}$ is an image of a recurrent Θ'-rich word under a suitable morphism, where Θ' is also an involutive antimorphism. Moreover, if the word $\mathbf{u}$ is uniformly recurrent, we show that Θ' can be set to the reversal mapping. We also treat one special case of almost Θ-rich words: we show that every Θ-standard word with seed is an image of an Arnoux-Rauzy word.

Keywords: Palindrome; pseudopalindrome; palindromic defect; richness.

1. Introduction

In this paper we deal with infinite words over a finite alphabet $\mathcal{A}$. Given a word $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ we are interested whether its language is saturated, in a certain sense, by generalized palindromes, here called Θ-palindromes. We use the symbol Θ for an *involutive antimorphism*, i.e., a mapping $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ such that $\Theta^2 = \mathrm{Id}$ and $\Theta(uv) = \Theta(v)\Theta(u)$ for all $u, v \in \mathcal{A}^*$. Fixed points of Θ are called Θ-*palindromes*. The notion of Θ-palindrome seems to have appeared independently on several places in the literature, see [13, 17].

A strong impulse for study of palindromes came from outside of mathematics. Physicists discovered a role of classical palindromes in the description of the spectrum of Schrödinger operators with aperiodic potentials, see [16]. In genetics, the so-called Watson-Crick palindromes play, for instance, an important role in the description of unwanted bindings of nucleotides in a

DNA strand (see [18]). In our terminology, the Watson-Crick palindromes are Θ-palindromes where the involutive antimorphism Θ acts on a quaternary alphabet and has no fixed point of length one.

The most common antimorphism used in combinatorics on words is the *reversal* mapping. We denote it by R. The reversal mapping assigns to every word $w = w_1 w_2 \dots w_n$ its mirror image $R(w) = w_n w_{n-1} \dots w_1$. In the case $w = R(w)$, we sometimes say that w is a *palindrome* or classical palindrome instead of R-palindrome.

The set of distinct Θ-palindromes occurring in a finite word w is denoted by $\mathrm{Pal}_\Theta(w)$. Since the empty word ε is a Θ-palindrome for any Θ, we have a simple lower bound $\#\mathrm{Pal}_\Theta(w) \geq 1$.

In 2001, Droubay *et al.* gave in [14] an upper bound for the reversal mapping R. They deduced that $\#\mathrm{Pal}_R(w) \leq |w| + 1$, where $|w|$ denotes the length of the word w. In [5], Blondin Massé *et al.* studied involutive antimorphisms with no fixed points of length 1. For such Θ they decreased the upper bound, in particular, they showed that $\#\mathrm{Pal}_\Theta(w) \leq |w|$ for all non-empty word w. In [23], the upper bound is more precise. The following estimate is valid for any involutive antimorphism Θ:

$$\#\mathrm{Pal}_\Theta(w) \leq |w| + 1 - \gamma_\Theta(w), \tag{1}$$

where $\gamma_\Theta(w) := \#\big\{\{a, \Theta(a)\} \mid a \in \mathcal{A}, a \text{ occurs in } w, \text{and } a \neq \Theta(a)\big\}$. Let us illustrate the mapping γ_Θ on the following example.

Example. Suppose $\mathcal{A} = \{0, 1, 2\}$ and Θ is the antimorphism determined by $0 \mapsto 1, 1 \mapsto 0, 2 \mapsto 2$. We have

$$\gamma_\Theta(\varepsilon) = 0, \quad \gamma_\Theta(0) = 1, \quad \gamma_\Theta(2) = 0, \quad \gamma_\Theta(00) = \gamma_\Theta(01) = \gamma_\Theta(02) = \gamma_\Theta(20) = 1.$$

Let us note that if $\Theta = R$, then $g_R(w) = 0$ for any finite word w, and the upper bound in (1) is the same as for classical palindromes.

In [6], the authors deal with the case $\Theta = R$ and with words for which the equality in (1) holds. They call such words *full*. According to the terminology for classical palindromes introduced in [15] and for Θ-palindromes in [23], we say that a finite word w is Θ-*rich* if the equality in (1) holds. An infinite word $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is Θ-*rich* if every factor $w \in \mathcal{L}(\mathbf{u})$ is Θ-rich, where by $\mathcal{L}(\mathbf{u})$ we denote the set of all factors of $\mathbf{u}$ called the *language of* $\mathbf{u}$. In [6], the authors introduce the *palindromic defect* of a finite word w as the difference between the upper bound $|w| + 1$ and the actual number of distinct palindromic factors. We define analogously the Θ-*palindromic defect* of w as

$$D_\Theta(w) := |w| + 1 - \gamma_\Theta(w) - \#\mathrm{Pal}_\Theta(w).$$

We define for an infinite word $\mathbf{u}$ its Θ-*palindromic defect* as

$$D_\Theta(\mathbf{u}) = \sup\{D_\Theta(w) \mid w \in \mathcal{L}(\mathbf{u})\},$$

which is again a generalization of the case of classical palindromes introduced in [6]. Words with finite Θ-palindromic defect are referred to as *almost* Θ-*rich*.

The notion of almost richness for classical palindromes was introduced and studied in [15].

In [12], it is shown that rich words (i.e. R-rich words) can be characterized using an inequality shown in [2] for infinite words with languages closed under reversal. Results of both mentioned papers were generalized in [23] for an arbitrary involutive antimorphism. In particular, it is shown that if an infinite word has its language closed under Θ, the following inequality holds

$$\mathcal{C}(n+1) - \mathcal{C}(n) + 2 \geq \mathcal{P}_\Theta(n) + \mathcal{P}_\Theta(n+1) \text{ for all } n \geq 1, \tag{2}$$

where $\mathcal{C}(n)$ is the *factor complexity* defined by $\mathcal{C}(n) := \#\{w \in \mathcal{L}(\mathbf{u}) \mid n = |w|\}$ and $\mathcal{P}_\Theta(n)$ is the Θ-*palindromic complexity* defined by $\mathcal{P}_\Theta(n) := \#\{w \in \mathcal{L}(\mathbf{u}) \mid w = \Theta(w) \text{ and } n = |w|\}$. Let us denote by $T_\Theta(n)$ the difference between the left side and the right side in (2), i.e., the quantity

$$T_\Theta(n) := \mathcal{C}(n+1) - \mathcal{C}(n) + 2 - \mathcal{P}_\Theta(n+1) - \mathcal{P}_\Theta(n).$$

This quantity decides about Θ-richness: in [23], it is also shown that an infinite word with language closed under Θ is Θ-rich if and only if

$$T_\Theta(n) = 0 \text{ for all } n \geq 1.$$

The list of infinite words which are R-rich is quite extensive. See for instance [2, 9, 11, 15]. Examples of Θ-rich words can be found in the class of words called Θ-episturmian words. A condition when such a word is Θ-rich can be found in [23]. In [1], the authors also deal with Θ-episturmian words (they are called pseudopalindromic in the paper). However, the result of Theorem 2 in [1] is valid only for the subset of Θ-rich Θ-episturmian words, not for all Θ-episturmian words as stated in the paper.

Fewer examples of words with finite nonzero palindromic defect are known. Periodic words with finite nonzero R-defect can be found in [6], aperiodic ones are studied in [15] and [4]. To our knowledge, examples of words with $0 < D_\Theta(\mathbf{u}) < +\infty$ and $\Theta \neq R$ have not yet been explicitly exhibited. As we will show, such examples are Θ-standard words with seed defined in [10] and thus also their subset, standard Θ-episturmian words, which can be constructed from standard episturmian words, see [8].

The main aim of this paper is to show that among words with finite Θ-palindromic defect, Θ-rich words, i.e. words with $D_\Theta(\mathbf{u}) = 0$, play an important role. We will prove the following theorems.

Theorem 1. *Let $\Theta_1 : \mathcal{A}^* \mapsto \mathcal{A}^*$ be an involutive antimorphism. If $\mathbf{u} \in \mathcal{A}^\mathbb{N}$ is a recurrent infinite word such that $D_{\Theta_1}(\mathbf{u}) < +\infty$, then there exist an involutive antimorphism $\Theta_2 : \mathcal{B}^* \mapsto \mathcal{B}^*$, a morphism $\varphi : \mathcal{B}^* \mapsto \mathcal{A}^*$ and an infinite recurrent word $\mathbf{v} \in \mathcal{B}^\mathbb{N}$ such that*

$$\mathbf{u} = \varphi(\mathbf{v}) \text{ and } \mathbf{v} \text{ is } \Theta_2\text{-rich.}$$

A stronger statement can be shown if uniform recurrency is assumed.

Theorem 2. *Let $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ be an involutive antimorphism. If $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is a uniformly recurrent infinite word such that $D_\Theta(\mathbf{u}) < +\infty$, then there exist a morphism $\varphi : \mathcal{B}^* \mapsto \mathcal{A}^*$ and an infinite uniformly recurrent word $\mathbf{v} \in \mathcal{B}^{\mathbb{N}}$ such that*

$$\mathbf{u} = \varphi(\mathbf{v}) \text{ and } \mathbf{v} \text{ is } R\text{-rich.}$$

One can conclude that rich words, using the classical notion of a palindrome, play somewhat a more important role than Θ-rich words for an arbitrary $\Theta \neq R$.

The proofs of the two stated theorems do not provide any relation between the size of the alphabet $\mathcal{B}$ of the word $\mathbf{v}$ and the size of the original alphabet $\mathcal{A}$. Nevertheless, one can find a bound on the size of $\mathcal{B}$ using the factor complexity of $\mathbf{u}$, see Corollary 8. The following theorem is a special case of the last theorem. In this case, the size of $\mathcal{B}$ can be bounded by the size of the alphabet $\mathcal{A}$ and the word $\mathbf{v}$ is more specific, namely it is Arnoux-Rauzy. Let us recall that an infinite word $\mathbf{v}$ is an Arnoux-Rauzy word if for every n we have $\mathcal{C}(n) = (\#\mathcal{A} - 1)n + 1$ and there is exactly one factor $w \in \mathcal{L}(\mathbf{v})$ of length n which can be extended to the left in more than one way, i.e., is left special. Ternary Arnoux-Rauzy words were first mentioned in [22].

Theorem 3. *If $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ is an involutive antimorphism and $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is a Θ-standard word with seed, then there exist an Arnoux-Rauzy word $\mathbf{v} \in \mathcal{B}^{\mathbb{N}}$ and a morphism $\varphi : \mathcal{B}^* \mapsto \mathcal{A}^*$ such that*

$$\mathbf{u} = \varphi(\mathbf{v}) \text{ and } \#\mathcal{B} \leq \#\mathcal{A}.$$

One of the reviewers of this paper pointed out to us that the last theorem is in fact a particular case of Theorem 1 in [7]. We keep it here with a proof for the sake of completeness in the context of Θ-richness.

All three mentioned theorems present an almost Θ_1-rich word as an image of a Θ_2-rich word by a suitable morphism. The opposite question when a morphic image of a Θ_1-rich word is almost Θ_2-rich is not tackled here. In [15], a type of morphisms preserving the set of almost R-rich words is studied.

2. Properties of Words with Finite Θ-Defect

Let $\mathcal{A}$ be an *alphabet*: a finite set of symbols called *letters*. We say that w is a *finite word* if $w = w_0 w_1 \cdots w_n$ where $w_i \in \mathcal{A}$ for all i such that $0 \leq i \leq n$. The *length* of w is denoted $|w|$ and equals $n + 1$. The set of all finite words is the free monoid $\mathcal{A}^*$ which includes the *empty word* ε. We consider mainly infinite words $\mathbf{u} = (u_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$. A finite word is a *factor* of a finite or infinite word $\mathbf{u}$ if there exists an index j such that $w = u_j u_{j+1} \cdots u_{j+n}$. The index j is called an *occurrence* of w in $\mathbf{u}$. We deal mainly with infinite words having their language $\mathcal{L}(\mathbf{u})$ *closed under a given involutive antimorphism* Θ, i.e., for any factor $w \in \mathcal{L}(\mathbf{u})$ we have $\Theta(w) \in \mathcal{L}(\mathbf{u})$.

We say that a finite word p is a *prefix* of a (finite or infinite) word v if there exists a word s such that $v = ps$. A finite word s is a *suffix* of a finite word v if $v = ps$ for some $p \in \mathcal{A}^*$.

If each factor of $\mathbf{u}$ has infinitely many occurrences in $\mathbf{u}$, the infinite word $\mathbf{u}$ is said to be *recurrent*. It is easy to see that if the language of $\mathbf{u}$ is closed under Θ, then $\mathbf{u}$ is recurrent. For a recurrent infinite word $\mathbf{u}$, we may define the notion of a *complete return word* of any $w \in \mathcal{L}(\mathbf{u})$. It is a factor $v \in \mathcal{L}(\mathbf{u})$ such that w is a prefix and a suffix of v and w occurs in v exactly twice. By a *return word* of a factor w we mean a word $q \in \mathcal{L}(\mathbf{u})$ such that qw is a complete return word of w. If every factor w of a recurrent word $\mathbf{u}$ has only finitely many return words, then the infinite word $\mathbf{u}$ is called *uniformly recurrent*.

An important role for the description of languages closed under Θ is played by the so-called *super reduced Rauzy graphs* $G_n(\mathbf{u})$, introduced in [12]. Before defining them, we introduce some necessary notions.

We say that a factor $w \in \mathcal{L}(\mathbf{u})$ is *left special* (LS) if w has at least two left extensions, i.e., if there exist two letters $a, b \in \mathcal{A}$, $a \neq b$, such that $aw, bw \in \mathcal{L}(\mathbf{u})$. A *right special* (RS) factor is defined analogously. A *special* factor is a factor which is RS or LS. If a factor is LS and RS, we refer to it as *bispecial*. The fact that $\mathcal{L}(\mathbf{u})$ is closed under Θ assures the following relation: a factor w is LS if and only if the factor $\Theta(w)$ is RS.

An *n-simple path* e is a factor of $\mathbf{u}$ of length at least $n+1$ such that the only special factors of length n occurring in e are its prefix and suffix of length n. If w is the prefix of e of length n and v is the suffix of e of length n, we say that the n-simple path e begins with w and ends with v. We denote by $G_n(\mathbf{u})$ an undirected graph whose set of vertices is formed by unordered pairs $\{w, \Theta(w)\}$ such that $w \in \mathcal{L}(\mathbf{u})$, $|w| = n$, and w is RS or LS. We connect two vertices $\{w, \Theta(w)\}$ and $\{v, \Theta(v)\}$ by an unordered pair $\{e, \Theta(e)\}$ if e or $\Theta(e)$ is an n-simple path beginning with w or $\Theta(w)$ and ending with v or $\Theta(v)$. Note that the graph $G_n(\mathbf{u})$ may have multiple edges and loops.

As first shown for classical palindromes in [12], the super reduced Rauzy graph $G_n(\mathbf{u})$ can be used to detect equality in (2). Let us cite Corollary 7 from [23].

Proposition 4. *If $n \in \mathbb{N}$ and $\mathcal{L}(\mathbf{u})$ is closed under Θ, then $T_\Theta(n) = 0$ if and only if*

(1) *all n-simple paths forming a loop in $G_n(\mathbf{u})$ are Θ-palindromes and*
(2) *the graph obtained from $G_n(\mathbf{u})$ by removing all loops is a tree.*

In [4], various properties are shown for words with finite R-palindromic defect. These properties and their proofs are valid even if we replace the antimorphism R by an arbitrary Θ.

Proposition 5. *If $\mathbf{u}$ is a recurrent infinite word such that $D_\Theta(\mathbf{u}) < +\infty$, then there exists a positive integer H such that $\mathbf{u}$ has the following Properties:*

(i) *every prefix of* $\mathbf{u}$ *longer than* H *has a unioccurrent* Θ*-palindromic suffix;*
(ii) $\mathcal{L}(\mathbf{u})$ *is closed under* Θ*;*
(iii) *for any factor* $w \in \mathcal{L}(\mathbf{u})$ *such that* $|w| > H$*, occurrences of* w *and* $\Theta(w)$ *in the word* $\mathbf{u}$ *alternate;*
(iv) *for any* $w \in \mathcal{L}(\mathbf{u})$ *such that* $|w| > H$*, every factor* $v \in \mathcal{L}(\mathbf{u})$ *beginning with* w*, ending with* $\Theta(w)$*, and with no other occurrences of* w *or* $\Theta(w)$ *is a* Θ*-palindrome;*
(v) $T_\Theta(n) = 0$ *for any integer* $n > H$*.*

In the case of the reversal mapping and $D_R(\mathbf{u}) = 0$, the listed properties are generalizations of properties already shown in [15] and [12]. In this case we have $H = 0$.

The main difference for an arbitrary Θ with comparison to R is that there can be non-Θ-palindromic letters. However, this can be dealt with by a good choice of the constant H. To prove Properties (i), (iii) and (iv), one can follow the proofs for $\Theta = R$ in [4]. Property (ii) generalizes Proposition 4.4. in [15]. Proof of Property (v) is the most intricate and is in fact a special case of a more general result from [21] where we deal with words closed under all elements of a finite group generated by involutive antimorphisms. Therefore, we give only a sketch of the proof.

Sketch of the proof.

(i): Take a prefix p of $\mathbf{u}$ such that all letters of $\mathcal{A}$ occur in it. Denote by p' the word such that $p = p'a$ with $a \in \mathcal{A}$. Since all letters occur in p, we have $\gamma_\Theta(p') = \gamma_\Theta(p)$. Moreover, suppose p does not have a unioccurrent Θ-palindromic suffix. If p has no Θ-palindromic suffix, then we have $\#\mathrm{Pal}_\Theta(p') = \#\mathrm{Pal}_\Theta(p)$. If p has a Θ-palindromic suffix, then again we have $\#\mathrm{Pal}_\Theta(p') = \#\mathrm{Pal}_\Theta(p)$ as the suffix is not unioccurrent. Thus, $D_\Theta(p) > D_\Theta(p')$ which can happen for only finitely many prefixes p because of $D_\Theta(\mathbf{u}) < +\infty$.

Let us denote by q such a prefix of $\mathbf{u}$ that $D_\Theta(\mathbf{u}) = D_\Theta(q)$. It is enough to set $H := \max\{|p|, |q|\}$.

(ii): Suppose that w is a factor of $\mathbf{u}$ such that $\Theta(w) \notin \mathcal{L}(\mathbf{u})$. Since $\mathbf{u}$ is recurrent, we can find two consecutive occurrences i and j of the factor w such that $i, j > H$ and $i < j$. Denote p the prefix of $\mathbf{u}$ ending with w occurring at j, i.e., $|p| = j + |w|$. Since $|p| > H$, there exists a unioccurrent Θ-palindromic suffix of p. Denote s to be such a suffix. If $|s| \leq |w|$, then s is a factor of w and thus occurs at least twice in p — a contradiction with the unioccurrence of s. If $|s| > |w|$, then w is a factor of s which is a Θ-palindrome and thus contains $\Theta(w)$ as well — a contradiction with the assumption that $\Theta(w) \notin \mathcal{L}(\mathbf{u})$.

(iii): Suppose $w \in \mathcal{L}(\mathbf{u})$ such that $\Theta(w) \neq w$ and $|w| > H$. Take a factor v such that it contains exactly 2 occurrences of w and no occurrence $\Theta(w)$. It is clear that $|v| > H$. Take the shortest prefix p of $\mathbf{u}$ such that it contains exactly one occurrence of v or $\Theta(v)$. It is easy to see that p does not have a unioccurrent Θ-palindromic suffix which is in contradiction with Property (i).

(iv): Suppose $w \in \mathcal{L}(\mathbf{u})$ such that $|w| > H$. Take a factor v, with $|v| > |w|$, such that it contains w as its prefix, $\Theta(w)$ as its suffix, and no other occurrences of w or $\Theta(w)$. Take the shortest prefix p of $\mathbf{u}$ such that it contains exactly one occurrence of v or $\Theta(v)$. Since p has a unioccurrent Θ-palindromic suffix, one can deduce that v is a Θ-palindrome.

(v): According to Property (ii) the language $\mathcal{L}(\mathbf{u})$ is closed under Θ. In order to use Proposition 4, we need to prove that for all $n > H$

(1) all n-simple paths forming a loop in $G_n(\mathbf{u})$ are Θ-palindromes and
(2) the graph obtained from $G_n(\mathbf{u})$ by removing all loops is a tree.

To prove this, we use a more general result from [21]. The claim follows from Property (iv) and Lemma 19 in [21] applied to the group $G = \{\Theta, \mathrm{Id}\}$. □

As already mentioned, the first property stated in the previous proposition, in fact, characterizes words with finite Θ-defect. We do not know whether this is the case of the Properties (iii), (iv) and (v). If we restrict our attention to uniformly recurrent words, only then can we show several characterizations of words with finite Θ-defect. The next proposition states two of them that we use in what follows.

Proposition 6. *If $\mathbf{u}$ is a uniformly recurrent infinite word with language closed under Θ, then the following statements are equivalent:*

(i) $D_\Theta(\mathbf{u}) < +\infty$;
(ii) *there exists a positive integer H such that for any $w \in \mathcal{L}(\mathbf{u})$, $|w| > H$, the longest Θ-palindromic suffix of w is unioccurrent in w;*
(iii) *there exists a positive integer K such that for any Θ-palindrome $w \in \mathcal{L}(\mathbf{u})$ of length $|w| \geq K$, all complete return words of w are Θ-palindromes.*

The provided proposition was proved in [4] for the special case $\Theta = R$. We give again a sketch of the proof using more general results from [21].

Sketch of the proof.
We use results of [21] for words closed under all elements of the group $G = \{\Theta, \mathrm{Id}\}$. According to Theorem 31 in [21], finite G-defect (in our case it coincides with Θ-defect as defined above) is equivalent with almost G-richness of $\mathbf{u}$. Using Lemmas 16 and 19 from [21], we obtain equivalence with Property (ii). Lemmas 22 and 24 from [21] imply equivalence of Properties (ii) and (iii). □

A Θ-standard word with seed is an infinite word defined using Θ-palindromic closure, for details see [10]. Construction of such a word $\mathbf{u}$ guarantees that $\mathbf{u}$ is uniformly recurrent (cf. Proposition 3.5. in [10]). The authors of [10] showed (Proposition 4.8) that any complete return word of a sufficiently long Θ-palindromic factor is a Θ-palindrome as well. Therefore, Θ-standard words with seed serve as an example of almost Θ-rich words.

Corollary 7. *Let* $\mathbf{u}$ *be a* Θ*-standard word with seed. Then* $D_\Theta(\mathbf{u}) < +\infty$.

3. Proofs

In this section we give proofs of all three theorems stated in the introduction. Although Theorem 2 seems to be only a refinement of Theorem 1, constructions of the morphisms φ in their proofs differ substantially because of stronger properties that we can exploit for a uniformly recurrent word.

As already mentioned, the list of infinite words with finite but nonzero Θ-defect is very modest. Moreover, if Conjecture 1 from [5] holds (and we believe so), no fixed point of a primitive substitution has finite nonzero defect. Therefore, it is very difficult to demonstrate validity of our Theorems 1 and 2 on reasonably described examples of almost rich words. Instead of this, we accompany our proofs with graphical presentation of the crucial idea for construction of suitable morphisms.

Our theorems convey that one method of construction of almost rich words is via morphic images of rich words. From this point of view, it would be of great importance to characterize morphisms under which the set of almost rich words is invariant.

Proof of Theorem 1. Recall that according to Proposition 5 Property (ii) the language $\mathcal{L}(\mathbf{u})$ is closed under Θ_1.

Suppose $\mathbf{u}$ is eventually periodic. Since $\mathcal{L}(\mathbf{u})$ is closed under Θ_1, $\mathbf{u}$ is recurrent. This implies that $\mathbf{u}$ is purely periodic (see Proposition 4.3.2. in [19]). Any purely periodic word is a morphic image of a word $\mathbf{v}$ over a one-letter alphabet under the morphism which assigns to this letter a word w such that $\mathbf{u} = www\ldots = w^\omega$. Therefore, we may assume without loss of generality that $\mathbf{u}$ is not eventually periodic.

Since $D_{\Theta_1}(\mathbf{u}) < +\infty$, according to Propositions 4 and 5, there exists an integer $H \in \mathbb{N}$ such that

(1) for any $w \in \mathcal{L}(\mathbf{u})$, $|w| > H$, occurrences of w and $\Theta_1(w)$ alternate;
(2) for any $w \in \mathcal{L}(\mathbf{u})$, $|w| > H$, every factor beginning with w, ending with $\Theta_1(w)$ and with no other occurrences of w or $\Theta_1(w)$ is a Θ_1-palindrome;
(3) for any $n \geq H$, every loop in $G_n(\mathbf{u})$ is a Θ_1-palindrome and the graph obtained from $G_n(\mathbf{u})$ by removing all loops is a tree.

Fix $n > H$. If an edge $\{b, \Theta_1(b)\}$ in $G_n(\mathbf{u})$ is a loop, then, according to Property 3, we have $b = \Theta_1(b)$. If the edge $\{b, \Theta_1(b)\}$ connects two distinct vertices $\{w_1, \Theta_1(w_1)\}$ and $\{w_2, \Theta_1(w_2)\}$, then there exist exactly two n-simple paths b and $\Theta_1(b)$ such that without loss of generality the n-simple path b begins with w_1

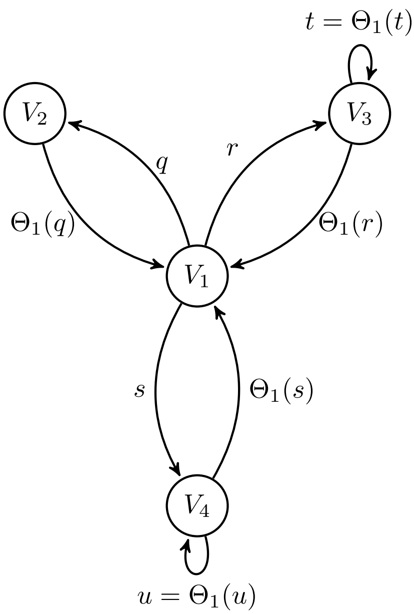

Fig. 1. Example of a graph $G_n(\mathbf{u})$ where an edge $\{e, \Theta_1(e)\}$ is drawn as two directed edges e and $\Theta_1(e)$. V_i denotes the vertex $\{w_i, \Theta_1(w_i)\}$ for a special factor w_i. In this case, there are exactly 8 n-simple paths and thus the alphabet $\mathcal{B}$ consists of 8 letters, i.e., $\mathcal{B} = \{[q], [\Theta_1(q)], [r], [\Theta_1(r)], [s], [\Theta_1(s)], [t], [u]\}$.

and ends with w_2 and the n-simple path $\Theta_1(b)$ begins with $\Theta_1(w_2)$ and ends with $\Theta_1(w_1)$.

We assign to every n-simple path b a new symbol $[b]$, i.e., we define the alphabet $\mathcal{B}$ as

$$\mathcal{B} := \{[b] \mid b \in \mathcal{L}(\mathbf{u}) \text{ is an } n\text{-simple path}\}.$$

See Figure 1 for an example of construction of $\mathcal{B}$.

We define on $\mathcal{B}$ an involutive antimorphism $\Theta_2 : \mathcal{B}^* \mapsto \mathcal{B}^*$ in the following way:

$$\Theta_2([b]) := [\Theta_1(b)].$$

We are now going to construct a suitable infinite word $\mathbf{v} \in \mathcal{B}^{\mathbb{N}}$. Let $(s_i)_{i\in\mathbb{N}}$ denote a strictly increasing sequence of indices such that s_i is an occurrence of a LS or RS factor of length n and every LS and RS factor of length n occurs at some index s_i. We define $\mathbf{v} = (v_i)_{i\in\mathbb{N}}$ by the formula

$$v_i = [b] \quad \text{if} \quad b = u_{s_i}u_{s_i+1}u_{s_i+2}\dots u_{s_{i+1}+n-1}.$$

This construction can be done for any $n > H$. Since infinitely many prefixes of $\mathbf{u}$ are LS or RS factors, we can choose $n > H$ such that the prefix of $\mathbf{u}$ of length n is LS or RS, i.e., $s_0 = 0$.

According to Proposition 12 in [23], to prove that $\mathbf{v}$ is Θ_2-rich we need to show the following:

(i) for every non-empty factor $w \in \mathcal{L}(\mathbf{v})$, any factor v beginning with w and ending with $\Theta_2(w)$, with no other occurrences of w or $\Theta_2(w)$, is a Θ_2-palindrome;

(ii) for every letter $[b] \in \mathcal{B}$ such that $[b] \neq \Theta_2([b])$, the occurrences of $[b]$ and $\Theta_2([b])$ in the word $\mathbf{v}$ alternate.

Let us first verify (i). Let e and f be factors of $\mathbf{v}$ such that e is a prefix of f and $\Theta_2(e)$ is a suffix of f and there are no other occurrences of e or $\Theta_2(e)$ in f. In that case there exist integers $r \leq k$ such that $f = [b_1][b_2]\dots[b_k]$ and $e = [b_1][b_2]\dots[b_r]$. The case $r = k$ is trivial. Suppose $r < k$. Since $\mathbf{v}$ is defined as a coding of consecutive occurrences of n-simple paths in $\mathbf{u}$, factor f codes a certain segment of the word $\mathbf{u}$. Let us denote that segment by $F = u_j \dots u_l$ where $j = s_t$ for some $t \in \mathbb{N}$ and $l = s_{t+k-1} + n - 1$. Factor e codes in the same way a factor $E = u_j \dots u_h$ where $h = s_{t+r-1} + n - 1$.

Due to the definition of Θ_2, the fact that e is a prefix of f and $\Theta_2(e)$ is a suffix of f ensures that E is a prefix of F and $\Theta(E)$ is a suffix of F. Suppose f is not a Θ_2-palindrome. This implies that F is not a Θ_1-palindrome which contradicts Property 3.

Let us now verify (ii). Consider $[b] \in \mathcal{B}$ such that $[b] \neq \Theta_2([b])$. Moving along the infinite word $\mathbf{u} = u_0u_1u_2\dots$ from the left to the right with a window of width n corresponds to a walk in the graph $G_n(\mathbf{u})$. The pair b and $\Theta_1(b)$ of n-simple paths in $\mathbf{u}$ represents an edge in $G_n(\mathbf{u})$ connecting two distinct vertices. Moreover, moving along the n-simple path b and moving along $\Theta_1(b)$ can be viewed as traversing that edge in opposite directions. Since the graph obtained from $G_n(\mathbf{u})$ by removing all loops is a tree, the only way to traverse an edge is alternately in one direction and in the other. Thus, the occurrences of letters $[b]$ and $\Theta_2([b])$ in $\mathbf{v}$ alternate.

We have shown that $\mathbf{v}$ is Θ_2-rich. It is now obvious how to define a morphism $\varphi : \mathcal{B}^* \mapsto \mathcal{A}^*$. If an n-simple path b equals $b = u_{s_i}u_{s_i+1}\dots u_{s_{i+1}+n-1}$, then we set $\varphi([b]) := u_{s_i}u_{s_i+1}\dots u_{s_{i+1}-1}$. □

Corollary 8. *Let $\mathcal{B}$ be the alphabet given by Theorem* 1. *If $\mathbf{u}$ is eventually periodic, then $\#\mathcal{B} = 1$. If $\mathbf{u}$ is aperiodic, then*

$$\#\mathcal{B} \leq 3(\mathcal{C}(n+1) - \mathcal{C}(n)),$$

where n is the integer from the proof of Theorem 1.

Proof. If $\mathbf{u}$ is eventually periodic, then the claim follows from the previous proof.

Suppose $\mathbf{u}$ is aperiodic. The size of the alphabet $\mathcal{B}$ defined in the previous proof equals the number of n-simple paths. Since any n-simple path starts in a special factor, we have

$$\#\mathcal{B} = \sum_{w \text{ is LS or RS}} \#\mathrm{Rext(w)},$$

where $\mathrm{Rext(w)} = \{\mathrm{a} \in \mathcal{A} \mid \mathrm{wa} \in \mathcal{L}(\mathbf{u})\}$. Similarly, we denote $\mathrm{Lext(w)} = \{\mathrm{a} \in \mathcal{A} \mid \mathrm{aw} \in \mathcal{L}(\mathbf{u})\}$. Because of

$$\mathcal{C}(n+1) - \mathcal{C}(n) = \sum_{w \in \mathcal{L}(\mathbf{u})} (\#\mathrm{Rext(w)} - 1) = \sum_{\mathrm{w} \in \mathcal{L}(\mathbf{u})} (\#\mathrm{Lext(w)} - 1),$$

we can estimate

$$\#\mathcal{B} \le \mathcal{C}(n+1) - \mathcal{C}(n) + \#\{w \in \mathcal{L}(\mathbf{u}) \mid w \text{ is RS}\} + \#\{w \in \mathcal{L}(\mathbf{u}) \mid w \text{ is LS}\}.$$

Trivially,

$$\#\{w \in \mathcal{L}(\mathbf{u}) \mid w \text{ is RS}\} \le \sum_{w \in \mathcal{L}(\mathbf{u})} (\#\mathrm{Rext(w)} - 1) = \mathcal{C}(\mathrm{n}+1) - \mathcal{C}(\mathrm{n})$$

and

$$\#\{w \in \mathcal{L}(\mathbf{u}) \mid w \text{ is LS}\} \le \sum_{w \in \mathcal{L}(\mathbf{u})} (\#\mathrm{Lext(w)} - 1) = \mathcal{C}(\mathrm{n}+1) - \mathcal{C}(\mathrm{n}).$$

□

Proof of Theorem 2. Recall again that according to Proposition 5 Property (ii) the language $\mathcal{L}(\mathbf{u})$ is closed under Θ.

Next, we show that infinitely many Θ-palindromes are also prefixes of $\mathbf{u}$. Consider an integer H whose existence is guaranteed by Proposition 5 and denote by w a prefix of $\mathbf{u}$ longer than H. Since occurrences of factors w and $\Theta(w)$ in $\mathbf{u}$ alternate, according to the same proposition, the prefix of $\mathbf{u}$ ending with the first occurrence of $\Theta(w)$ is a Θ-palindrome.

Let us denote by p a Θ-palindromic prefix of $\mathbf{u}$ of length $|p| > K$ where K is the constant from Proposition 6. All complete return words of p are Θ-palindromes. Since $\mathbf{u}$ is uniformly recurrent, there exist only finite number of complete return words to p. Let $r^{(1)}, r^{(2)}, \ldots, r^{(M)}$ be the list of all these complete return words. Any complete return word $r^{(i)}$ has the form $q^{(i)}p = r^{(i)}$ for some factor $q^{(i)}$, usually called a return word of p. See Figure 2. Since $r^{(i)}$ and p are Θ-palindromes, we have

$$p\Theta(q^{(i)}) = q^{(i)}p \quad \text{for any return word} \quad q^{(i)}. \tag{3}$$

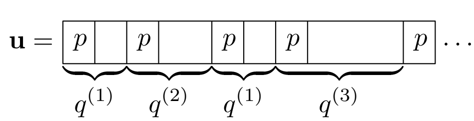

Fig. 2. Example of occurrences of return words $q^{(i)}$ of a palindrome p in the word $\mathbf{u}$.

Let us define a new alphabet $\mathcal{B} = \{1, 2, \ldots, M\}$ and morphism $\varphi : \mathcal{B}^* \to \mathcal{A}^*$ by the prescription

$$\varphi(i) = q^{(i)}, \quad \text{for} \;\; i = 1, 2, \ldots, M\,.$$

First, we shall check the validity of the relation

$$\Theta\big(\varphi(w)p\big) = \varphi\big(R(w)\big)p \quad \text{for any} \;\; w \in \mathcal{B}^*\,. \tag{4}$$

Let $w = i_1 i_2 \ldots i_n$. Then $\Theta\big(\varphi(i_1 i_2 \ldots i_n)p\big)$ is equal to

$$\Theta(p)\Theta\big(\varphi(i_n)\big)\Theta\big(\varphi(i_{n-1})\big)\ldots\Theta\big(\varphi(i_1)\big) = p\Theta\big(q^{(i_n)}\big)\Theta\big(q^{(i_{n-1})}\big)\ldots\Theta\big(q^{(i_1)}\big)$$

and we may apply repeatedly n times the equality (3) to rewrite the right side as

$$q^{(i_n)}q^{(i_{n-1})}\ldots q^{(i_1)}p = \varphi(i_n)\varphi(i_{n-1})\ldots\varphi(i_1)p = \varphi\big(R(i_1 i_2 \ldots i_n)\big)p.$$

This proves the relation (4).

An important property of the morphism φ is its injectivity. Indeed, in accordance with the definition, the number of occurrences of the factor p in $\varphi(w)p$ equals to the number of letters in w plus one. Moreover, each occurrence of p in $\varphi(w)p$ indicates a beginning of an image of a letter under φ. Therefore, $\varphi(w)p = \varphi(v)p$ necessarily implies $w = v$.

Let us finally define the word $\mathbf{v}$. As p is a prefix of $\mathbf{u}$, the word $\mathbf{u}$ can be written as a concatenation of return words $q^{(i)}$ and thus we can determine a sequence $\mathbf{v} = (v_n) \in \mathcal{B}^{\mathbb{N}}$ such that

$$\mathbf{u} = q^{(v_0)}q^{(v_1)}q^{(v_2)}\ldots$$

Directly from the definition of $\mathbf{v}$ we have $\mathbf{u} = \varphi(\mathbf{v})$. See Figure 3.

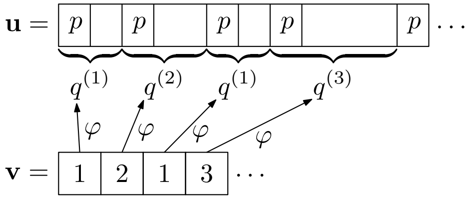

Fig. 3. Idea for the definition of the morphism φ.

Since $\mathbf{u}$ is uniformly recurrent, the word $\mathbf{v}$ is uniformly recurrent as well. To prove that $\mathbf{v}$ is an R-rich word, we shall show that any complete return word of any R-palindrome in the word $\mathbf{v}$ is an R-palindrome as well. According to Theorem 2.14 in [15], this implies the R-richness of $\mathbf{v}$.

If s is an R-palindrome in $\mathbf{v}$ and w its complete return word, then $\varphi(w)p$ has precisely two occurrences of the factor $\varphi(s)p$. Since s is an R-palindrome, we have according to the equation (4) that $\varphi(s)p$ is a Θ-palindrome of length $|\varphi(s)p| \geq |p| > K$.

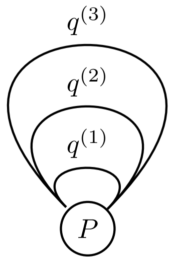

Fig. 4. Example of graph $G_{|p|}(\mathbf{u})$. There is only one vertex $P = \{p, \Theta(p)\}$ and all $|p|$-simple paths are loops.

Therefore, $\varphi(w)p$ is a complete return word of a long enough Θ-palindrome and according to our assumption $\varphi(w)p$ is a Θ-palindrome as well. Thus, by using (4) we have

$$\varphi(w)p = \Theta\big(\varphi(w)p\big) = \varphi\big(R(w)\big)p$$

and injectivity of φ gives $w = R(w)$, as we claimed. □

Theorem 6.1 in [8] states that every standard Θ-episturmian word is an image of a standard episturmian word. Again, the role of R can be perceived as more important. Also, compared to Theorem 2, it may be seen as a special case since Θ-episturmian words, according to Corollary 7, have finite Θ-defect.

Proof of Theorem 3.

If $\mathbf{u}$ is periodic, then the claim is trivial. Suppose $\mathbf{u}$ is aperiodic.

We are going to repeat the proof of Theorem 2 with a more specific choice of p. Theorem 4.4 in [10] implies that there exists $L \in \mathbb{N}$ such that any LS factor of $\mathbf{u}$ longer than L is a prefix of $\mathbf{u}$. Without loss of generality, we may assume that the constant L is already chosen in such a way that all prefixes of $\mathbf{u}$ longer than L have the same left extensions. Let us denote their number by M. According to the same theorem, infinitely many prefixes of $\mathbf{u}$ are Θ-palindromes and thus bispecial factors as well.

According to Corollary 7, $\mathbf{u}$ has finite Θ-palindromic defect. Let K be the constant from Proposition 6. Altogether, there exists a bispecial factor p, $|p| > \max\{L, K\}$, such that it is a prefix of $\mathbf{u}$ and a Θ-palindrome. Since p is longer than K, all complete return words to p are Θ-palindromes. As p is the unique left special factor of length $|p|$ in $\mathbf{u}$, its return words (i.e., complete return words after erasing the suffix p) end with distinct letters. This means that there are exactly M return words of p, denoted again by $q^{(i)}$. Let us recall that by M we denoted the number of left extensions of some factor, therefore $M \leq \#\mathcal{A}$. See Figure 4.

The construction of the word $\mathbf{v}$ and the definition of the morphism φ over the alphabet $\mathcal{B} = \{1, 2, \ldots, M\}$ can be done in exactly the same way as in the proof of Theorem 2. It remains to show that $\mathbf{v}$ is an Arnoux-Rauzy word.

According to Theorem 2 we know that $\mathbf{v}$ is R-rich and uniformly recurrent. Applying Property (ii) from Proposition 5 we deduce that the language $\mathcal{L}(\mathbf{v})$ is closed under reversal.

Suppose there exist $v, w \in \mathcal{L}(\mathbf{v})$, two LS factors such that $|v| = |w|$ and $v \neq w$. Since the words $q^{(i)}$ end with distinct letters, it is clear that $\varphi(w)p$ is a LS factor of $\mathbf{u}$ and it has the same number of left extensions as w. The same holds for $\varphi(v)p$. Since both these factors have their length greater than or equal to $|p| > L$ and are both LS, one must be a prefix of the other. Let without loss of generality $\varphi(w)p$ is a prefix of $\varphi(v)p$, i.e., $\varphi(v)p = \varphi(ww')p$. The injectivity of φ implies $w' = \varepsilon$ and thus $v = w$ — a contradiction. □

Remark 9. *Note that the proof of Theorem* 3 *is in fact a combination of methods used in the preceding proofs of Theorems* 1 *and* 2 *in the sense that the set of complete return words* $r^{(i)}$ *of the factor* p *and the set of* $|p|$*-simple paths in* $\mathbf{u}$ *coincide.*

4. Open Problems and Remarks on Graphs Hidden in the Structure of Rich Words

In this section, we list some related open problems and remarks.

The role of uniform recurrence

All presented results concern infinite words whose language is closed under one involutive antimorphism. In particular, we proved that any uniformly recurrent Θ-rich word $\mathbf{u}$ is a morphic image of an R-rich word, or equivalently, that the reversal mapping R is more important than other involutive antimorphisms. The question whether this statement is valid even in case when $\mathbf{u}$ is not uniformly recurrent is still open. Infinite words whose languages are invariant under more antimorphisms are not treated at all in the paper. The famous Thue-Morse word is one such word. Recently the authors proved that aperiodic words with a larger group of symmetries cannot be Θ-rich for any antimorphisms Θ from the group. Therefore, a new definition of richness which respects all symmetries present in an infinite word is suggested, see [20]. This definition is based on the notion of the graph of symmetries of a given infinite word. The super reduced graph is its special case, when the group of symmetries consists just from Θ and $\Theta^2 = \mathrm{Id}$. As already metioned, in [21], we generalize most of the known characterizations of richness with respect to all symmetries, including the notion of defect. Despite many other equivalent descriptions of (almost) rich words, we have chosen for generalization the property of the super reduced graphs, since it seems to be crucial for deducing all other characterization of (almost) rich words.

Graphs hidden in the structure of rich words

The main tool for proving Theorem 1 was the notion of the super reduced Rauzy graph $G_n(\mathbf{u})$. Richness of $\mathbf{u}$ implies for all $n \in \mathbb{N}$ that the graph obtained from

$G_n(\mathbf{u})$ by removing loops is a tree. In [3], other types of graphs are exploited for characterizations of richness. To any bispecial factor w of an infinite word $\mathbf{u}$ with language closed under reversal, we assign a graph $G_w = (V_w, E_w)$. Its definition differs for palindromic and non-palindromic factors w.

If w is non-palindromic then G_w is bipartite with the set of vertices $\{aw \mid aw \in \mathcal{L}(\mathbf{u})\} \cup \{wb \mid wb \in \mathcal{L}(\mathbf{u})\}$ and the set of edges $E_w = \{awb \mid awb \in \mathcal{L}(\mathbf{u})\}$.

If w is palindromic, then $V_w = \{aw \mid aw \in \mathcal{L}(\mathbf{u})\}$ and two vertices aw and bw are connected with an edge if awb belongs to $\mathcal{L}(\mathbf{u})$. In this case, the graph G_w may have loops.

As follows from the proof of Theorem 11 in [3],

- an infinite word $\mathbf{u}$ is rich if and only if for any bispecial factor $w \in \mathcal{L}(\mathbf{u})$ the graph obtained from G_w by removing loops is a tree;
- an infinite word $\mathbf{u}$ is almost rich if and only if there exists a constant M such that for any bispecial factor $w \in \mathcal{L}(\mathbf{u})$ of length $|w| \geq M$ the graph obtained from G_w by removing loops is a tree.

The property of a graph G_w "to be a tree" can be reformulated by value of bilateral order of w.

Morphisms fixing rich words

As we have already mentioned in the introduction, we do not discuss here the question of morphisms preserving the set of (almost) rich words. A description of such morphisms can be a tough problem. Our scepticism is supported by Example 5.6. from [4], in which we construct a primitive morphism ψ and two rich words $\mathbf{u}$ and $\mathbf{v}$, such that the palindromic defect of $\psi(\mathbf{u})$ is infinite and the palindromic defect of $\psi(\mathbf{v})$ is zero.

We believe that a characterization of (almost) richness by the above mentioned graphs G_w assigned to bispecial factors may serve to identify morphisms φ which do not preserve the tree structure of G_w after applying φ on w.

Defect of aperiodic fixed points of primitive morphisms

Almost rich words which are preserved by a primitive morphism, i.e. fixed points of primitive substitutions with finite defect, are studied in the article [5]. Its authors conjectured:

Conjecture 10. *Let* $\mathbf{u}$ *be a fixed point of a primitive morphism* φ*. If the defect is such that* $0 < D(\mathbf{u}) < \infty$*, then* $\mathbf{u}$ *is periodic.*

In [5], the authors conclude that the conjecture holds for a special class of morphisms (see Section 6 in [5]). A full verification of the conjecture remains an open question.

Acknowledgments

We would like to thank all the reviewers of this paper for their fruitful remarks. We acknowledge financial support by the Czech Science Foundation grant GAČR 201/09/0584, by the grants MSM6840770039 and LC06002 of the Ministry of Education, Youth, and Sports of the Czech Republic, and by the grant of the Grant Agency of the Czech Technical University in Prague grant No. SGS11/162/OHK4/3T/14.

References

[1] V. Anne, L. Q. Zamboni, and I. Zorca. Palindromes and pseudo-palindromes in episturmian and pseudo-palindromic infinite words. In S. Brlek and C. Reutenauer, editors, *Words 2005 5th International Conference on Words*, volume 36, pages 91–100. LaCIM, 2005.

[2] P. Baláži, Z. Masáková, and E. Pelantová. Factor versus palindromic complexity of uniformly recurrent infinite words. *Theoret. Comput. Sci.*, 380(3):266–275, 2007.

[3] L. Balková, E. Pelantová, and Š. Starosta. Sturmian jungle (or garden?) on multiliteral alphabets. *RAIRO-Theoret. Inf. Appl.*, 44:443–470, 2010.

[4] L. Balková, E. Pelantová, and Š. Starosta. Infinite words with finite defect. *Adv. in Appl. Math.*, 47(3):562–574, 2011.

[5] A. Blondin Massé, S. Brlek, A. Garon, and S. Labbé. Combinatorial properties of f-palindromes in the Thue-Morse sequence. *Pure Math. Appl.*, 19(2-3):39–52, 2008.

[6] S. Brlek, S. Hamel, M. Nivat, and C. Reutenauer. On the palindromic complexity of infinite words. *Internat. J. Found. Comput.*, 15(2):293–306, 2004.

[7] M. Bucci and A. De Luca. On a family of morphic images of Arnoux-Rauzy words. In *LATA '09: Proceedings of the 3rd International Conference on Language and Automata Theory and Applications*, pages 259–266, Berlin, Heidelberg, 2009. Springer-Verlag.

[8] M. Bucci, A. de Luca, and A. De Luca. Characteristic morphisms of generalized episturmian words. *Theoret. Comput. Sci.*, 410:2840–2859, 2009.

[9] M. Bucci, A. de Luca, and A. De Luca. Rich and periodic-like words. In *Proceedings of the 13th International Conference on Developments in Language Theory*, pages 145–155, Berlin, Heidelberg, 2009. Springer-Verlag.

[10] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni. On different generalizations of episturmian words. *Theoret. Comput. Sci.*, 393(1-3):23–36, 2008.

[11] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni. On theta-episturmian words. *European J. Combin.*, 30(2):473–479, 2009.

[12] M. Bucci, A. De Luca, A. Glen, and L. Q. Zamboni. A connection between palindromic and factor complexity using return words. *Adv. in Appl. Math.*, 42(1):60–74, 2009.

[13] A. de Luca and A. De Luca. Pseudopalindrome closure operators in free monoids. *Theoret. Comput. Sci.*, 362(1–3):282–300, 2006.

[14] X. Droubay, J. Justin, and G. Pirillo. Episturmian words and some constructions of de Luca and Rauzy. *Theoret. Comput. Sci.*, 255(1-2):539–553, 2001.

[15] A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni. Palindromic richness. *European J. Combin.*, 30(2):510–531, 2009.

[16] A. Hof, O. Knill, and B. Simon. Singular continuous spectrum for palindromic Schrödinger operators. *Comm. Math. Phys.*, 174:149–159, 1995.

[17] L. Kari and K. Mahalingam. Watson-Crick conjugate and commutative words. In M. Garzon and H. Yan, editors, *DNA Computing*, volume 4848 of *Lecture Notes in Computer Science*, pages 273–283. Springer-Verlag, Berlin, Heidelberg, 2008.

[18] L. Kari and K. Mahalingam. Watson-Crick palindromes in DNA computing. *Nat. Comput.*, 9:297–316, 2010.

[19] M. Lothaire. *Algebraic combinatorics on words*. Number 90 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2002.

[20] E. Pelantová and Š. Starosta. Languages invariant under more symmetries: overlapping factors versus palindromic richness. *to appear in Discrete Math., preprint available at http://arxiv.org/abs/1103.4051*, 2011.

[21] E. Pelantová and Š. Starosta. Palindromic richness and coxeter groups. *preprint available at http://arxiv.org/abs/1108.3042*, 2011.

[22] G. Rauzy. Suites à termes dans un alphabet fini. *Séminaire de Théorie des Nombres de Bordeaux*, Anné 1982–1983(exposé 25), 1983.

[23] Š. Starosta. On theta-palindromic richness. *Theoret. Comput. Sci.*, 412(12-14): 1111–1121, 2011.