České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Czech Technical University in Prague
Faculty of Nuclear Sciences and Physical Engineering

Ing. Štěpán Starosta, Ph.D.

Role palindromického defektu při zkoumání jazyků pevných bodů morfismů

Role of palindromic defect in the study of languages of fixed points of morphisms

Summary

Combinatorics on Words is a research domain providing tools to study discrete structures that can be modelled using sequences over finite sets of symbols. We give a short overview of Combinatorics on Words, focusing on the notion of fixed points of morphisms and on palindromic defect of languages of infinite words.

Palindromic defect is a measure of how much palindromes are missing in a language closed under taking mirror images. We reveal its role in the study of such symmetric languages that are languages of fixed points of morphisms. We present the so-called "zero defect conjecture" stating that a fixed point of a primitive morphism which is closed under taking mirror images has either zero palindromic defect (i.e., no palindromes are missing) or its palindromic defect is infinite. We shortly present a proof of the conjecture for the class of marked primitive morphisms and we reveal a connection to another conjecture stating that if an infinite word, fixed by a primitive morphism, contains infinitely many palindromes, then the language of the word may be generated by a morphism in a very special form, so-called class P. We also give briefly describe an application in algorithmic study of languages generated by morphism.

Souhrn

Kombinatorika na slovech je oblast výzkumu, která poskytuje nástroje ke studiu diskrétních struktur, které mohou být modelovatelné pomocí posloupností nad konečnými množinami. Krátce představíme tuto oblast a zaměříme se na pojmy pevný bod morfismu a palindromický defekt jazyku nekonečného slova.

Palindromický defekt je míra absence palindromů v jazyku, který je uzavřený na braní zrcadlových obrazů svých prvků. Ukážeme roli palindromického defektu při studiu těchto symetrických jazyků, které jsou vygenerované morfismy. Představíme takzvanou „zero defect conjecture“, tedy domněnka nulového defektu, která říká, že pevný bod primitivního morfismu, jehož jazyk je uzavřený na zrcadlení, má palindromický defekt nula (nechybí žádné palindromy) nebo nekonečno. Uvedeme krátký důkaz tvrzení této domněnky pro třídu takzvaných „marked“ primitivních morfismů a představíme souvislost s další domněnkou, která říká, že jazyk nekonečného slova, které je pevným bodem morfismu a obsahující nekonečně mnoho palindromů, může být vygenerován morfismem ve speciálním tvaru, tzv. třídě P. Krátce popíšeme aplikaci při algoritmickém studiu jazyků pevných bodů morfismů.

Klíčová slova

palindrom
morfické slovo
palindromický jazyk
palindromický defekt
domněnka nulového defektu
domněnka třídy P

Keywords

palindrome
morphic words
palindromic language
palindromic defect
zero defect conjecture
class P conjecture

Contents

1 Introduction

We present several problems and results in Combinatorics on Words concerning fixed points of morphisms and their languages. We start by an overview of the investigation of the number of missing palindromes in a language of an infinite word possessing certain symmetries. This number is called the palindromic defect and some of the techniques used for its study have applications in the investigation of other problems, which are seemingly unrelated. We also give an application in algorithmic analysis of infinite words that are fixed points of a morphism. We start with a brief introduction to Combinatorics on Words and the used notation.

1.1 On Combinatorics on Words

The very beginning of Combinatorics on Words is mostly attributed to Axel Thue and his articles published from 1906 till 1914. Axel Thue gave birth to a systematic study of objects called words: finite or infinite sequences of elements from a finite set called alphabet. The reader may refer to [5, 46, 38] for translations of Thue's papers and comments on his results.

The systematic study continued with the appearance of probably the most famous infinite word, the *Thue–Morse word*, in the article of Marston Morse in 1921 [33]. Axel Thue studied the same word in different context, which is the reason for the name of the word. Let us note that the word Thue–Morse word is also sometimes called Prouhet–Thue–Morse since is appeared already in 1851 in [39] by Eugène Prouhet.

The next stepping stone in the history of Combinatorics on Words is the article [34] of Marston Morse and Gustav Hedlund from 1940. Their work includes the study of another famous infinite words called *Sturmian words* in the honour of the famous mathematician Jacques Charles François Sturm. A Sturmian word is an infinite word over a two-letter alphabet having factor complexity $n + 1$, that is, for each n the number of total distinct contiguous subsequences of length n found in the word is equal to $n + 1$.

Such finite contiguous subsequence is called a *factor*, thus the name *factor complexity* since it is one of the basic measures of chaos (or order) of an infinite word. Factor complexity has the following essential property: if the factor complexity of an infinite word is bounded, then the word is (eventually) periodic. The converse is also true and we may deduce that Sturmian words are binary words having the least possible factor complexity so that they are not periodic.

Let us illustrate these notions on the example of the so-called Fibonacci word $\mathbf{f}$. The word $\mathbf{f}$ is a binary word, i.e., it is an infinite sequence over two symbols, 0 and 1. It may be defined as follows. First, we set $q_1 = 0$. To obtain q_2, we apply the rewriting rule $0 \mapsto 01$ and $1 \mapsto 0$ to q_1. We obtain the word $q_2 = 01$. We apply again the rule to q_2 and obtain $q_3 = 010$. We repeat the procedure and obtain

$$\begin{aligned} q_1 &= 0, & q_4 &= 01001, \\ q_2 &= 01, & q_5 &= 01001010, \\ q_3 &= 010, & q_6 &= 0100101001001. \end{aligned}$$

Each word q_i is a prefix of q_{i+1} and the length of q_i is strictly increasing. Thus, there is a unique infinite word over $\{0, 1\}$ having each q_i as its prefix, and it is the Fibonacci word $\mathbf{f}$.

After the mentioned works, the field of Combinatorics on Words has been growing steadily. The reader may refer to [6] for an overview of early progress in the area. The steady growth of the domain is underlined by collective publications containing overview of results in Combinatorics on Words and closely related domains and various monographs.

The first item on the list of such publications is the book *Combinatorics on Words*, first published in 1983, written by a collective of authors under the pseudonym M. Lothaire [28]. Two more books by M. Lothaire were published later, *Algebraic Combinatorics on Words* in 2002 [29] and *Applied Combinatorics on Words* in 2005 [30].

The growth of Combinatorics of Words may be also seen in its increasing connection to other domains. *Substitutions in Dynamics, Arithmetics and Combinatorics* published in 2002 [17] is a basic reference for the connection of Combinatorics of Words and Symbolic Dynamics. The publication *Combinatorics, Automata, and Number Theory* of 2010 [7] contains useful results interconnecting the domain in the title of the publication. The strong connection to Automata, Theory of Codes and Formal Languages may be also observed in the following books [22, 31, 41, 42].

Besides the mentioned domains, Combinatorics on Words finds its application in many other domains. Let us name some of them: Algebra, Logic, Music Theory, Stringology, and Biology. The reader may refer to [23] for an overview of some mentioned applications.

1.2 Notations and definitions

Let $\mathcal{A}$ be a finite set, called an *alphabet*. Its elements are called *letters*. A *finite word* w is an element of $\mathcal{A}^n$ for $n \in \mathbb{N}$. The *length* of w is n and is denoted $|w|$. The set of all finite words over $\mathcal{A}$ is denoted $\mathcal{A}^*$. An *infinite word* over $\mathcal{A}$ is an infinite sequence of letters from $\mathcal{A}$.

A finite word w is a *factor* of a finite or infinite word v if there exist words p and s such that v is a concatenation of p, w, and s, denoted $v = pws$. The word p is said to be a *prefix* and s a *suffix* of v. The set of all factors of a word $\mathbf{u}$ is the *language of* $\mathbf{u}$ and is denoted $\mathcal{L}(\mathbf{u})$. All factors of $\mathbf{u}$ of length n are denoted by $\mathcal{L}_n(\mathbf{u})$.

An *occurrence* of $w = w_0w_1 \cdots w_{n-1} \in \mathcal{A}^n$ in a word $v = v_0v_1v_2 \ldots$ is an index i such that $v_i \cdots v_{i+n-1} = w$. A factor w is *unioccurrent* in v if there is exactly one occurrence of w in v. A *complete return word* of a factor w (in v) is a factor f (of v) containing exactly two occurrences of w such that w is its prefix and also its suffix. For instance, the word 010011010 is a complete return word of 010.

The *reversal* or mirror mapping assigns to a word $w \in \mathcal{A}^*$ the word $\mathrm{R}(w)$ with the letters reversed, i.e.,

$$\mathrm{R}(w) = w_{n-1}w_{n-2} \cdots w_1w_0 \quad \text{where } w = w_0w_1 \cdots w_{n-1} \in \mathcal{A}^n.$$

A word is *palindrome* if $w = \mathrm{R}(w)$. We say that a language $\mathcal{L} \subset \mathcal{A}^*$ is *closed under reversal* if for all $w \in \mathcal{L}$ we have $\mathrm{R}(w) \in \mathcal{L}$.

Given an infinite word $\mathbf{u}$, its *factor complexity* $\mathcal{C}_{\mathbf{u}}(n)$ is the count of its factors of length n:

$$\mathcal{C}_{\mathbf{u}}(n) = \#\mathcal{L}_n(\mathbf{u}) \quad \text{for all } n \in \mathbb{N}.$$

Let $\mathrm{Pal}(\mathbf{u})$ be the set of all palindromic factors of the infinite word $\mathbf{u}$. The *palindromic complexity* $\mathcal{P}_{\mathbf{u}}(n)$ of $\mathbf{u}$ is given by

$$\mathcal{P}_{\mathbf{u}}(n) = \#(\mathcal{L}_n(\mathbf{u}) \cap \mathcal{P}(\mathbf{u})) \quad \text{for all } n \in \mathbb{N}.$$

We omit the subscript $\mathbf{u}$ if there is no confusion.

Example 1. Thus, for the Fibonacci word we have $\mathcal{P}_{\mathbf{f}}(n) = 1$ if n id odd and $\mathcal{P}_{\mathbf{f}}(n) = 2$ if n is even. In fact, the palindromic complexity of all Sturmian words is the same.

2 Palindromic defect

In [16], given a finite word $w \in \mathcal{A}^*$, the authors investigate the set of all its palindromic factors, denoted $\mathrm{Pal}(w)$, and give the following upper bound on $\#\mathrm{Pal}(w)$:

$$\#\mathrm{Pal}(w) \leq |w| + 1.$$

Note that the *empty word* ε, the unique word of length 0, is an element of $\mathrm{Pal}(w)$ for all w.

For instance, we have $\mathrm{Pal}(011) = \{\varepsilon, 0, 1, 11\}$. Thus, for the word 011, the upper bound on $\#\mathrm{Pal}(011)$ is attained.

The difference of the upper bound and the actual number of palindromic factors is the *palindromic defect* of w, see [10]. It is denoted $D(w)$. We have

$$D(w) = |w| + 1 - \#\mathrm{Pal}(w).$$

A basic property of the palindromic defect is that $D(w) \geq 0$ for any w and $D(v) \leq D(w)$ for any factor v of w. The properties of palindromic defect allow for a natural extension to infinite words:

$$D(\mathbf{u}) = \sup\{D(w) \colon w \in \mathcal{L}(\mathbf{u})\}.$$

One can say that it measures the number of "missing" palindromic factors in the given word.

The language $\mathcal{L}(\mathbf{u})$ may possess more symmetries besides R. By a symmetry we mean an involutory antimorphism. In such a case, the notion of palindromic defect may be generalized to the so-called *G-palindromic defect* where G is a group generated by the present involutory antimorphisms that fix the language, see [35, 36].

3 Fixed points of morphisms and their properties

Morphisms are an important tool as they allow to generate infinite words and their languages. For instance, they are the main object in Lindenmayer systems, or L-systems, which were originally proposed to model plant growth (see for instance [43]).

A *morphism* φ is a mapping $\mathcal{A}^* \to \mathcal{B}^*$ where $\mathcal{A}$ and $\mathcal{B}$ are alphabets such that $\forall v, w \in \mathcal{A}^*$ we have $\varphi(vw) = \varphi(v)\varphi(w)$ (it is a homomorphism of the monoids $\mathcal{A}^*$ and $\mathcal{B}^*$). Its action is extended to $\mathcal{A}^{\mathbb{N}}$: if $\mathbf{u} = u_0 u_1 u_2 \ldots \in \mathcal{A}^{\mathbb{N}}$ with $u_i \in \mathcal{A}$, then

$$\varphi(\mathbf{u}) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots \in \mathcal{B}^{\mathbb{N}}.$$

If φ is an endomorphism of $\mathcal{A}^*$, we may find its fixed point, i.e., a word $\mathbf{u}$ such that $\varphi(\mathbf{u}) = \mathbf{u}$. We are interested mainly in the case of $\mathbf{u}$ being infinite. A morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ is *primitive* if for every $a, b \in \mathcal{A}$ there exists an integer k such that b occurs in $\varphi^k(a)$.

Two morphisms $\varphi, \psi : \mathcal{A}^* \to \mathcal{B}^*$ are *conjugate* if there exists a word $w \in \mathcal{B}^*$ such that

$$\forall a \in \mathcal{A}, \varphi(a)w = w\psi(a) \quad \text{or} \quad \forall a \in \mathcal{A}, w\varphi(a) = \psi(a)w.$$

If φ is primitive, then the languages of fixed points of φ and ψ are the same.

A morphism $\psi : \mathcal{A}^* \to \mathcal{B}^*$ is of class P if $\psi(a) = pp_a$ for all $a \in \mathcal{A}$ where p and p_a are both palindromes (possibly empty). A morphism φ is of class P' if it is conjugate to a morphism of class P.

A morphism is *uniform* if the lengths of images of letters are the same.

The following examples illustrate the last few notions.

Example 2. Let $\varphi : \{a,b\}^* \to \{a,b\}^*$ be determined by $\varphi : \begin{array}{ccc} a & \mapsto & abab, \\ b & \mapsto & aab. \end{array}$ The fixed point of φ is

$$\mathbf{u} = \lim_{k\to+\infty} \varphi^k(a) = \underbrace{abab}_{\varphi(a)}\underbrace{aab}_{\varphi(b)}\underbrace{abab}_{\varphi(a)}\underbrace{aab}_{\varphi(b)}\underbrace{abab}_{\varphi(a)}\ldots$$

The morphism φ is of class P' since it is conjugate to ψ given by $\psi : \begin{array}{ccc} a & \mapsto & abab, \\ b & \mapsto & aba. \end{array}$ Indeed, we have $ab\varphi(a) = \psi(a)ab$ and $ab\varphi(b) = \psi(b)ab$. To see that ψ is of class P, i.e., it is of the form $a \mapsto pp_a$ and $b \mapsto pp_b$, it suffices to set $p = aba$, $p_a = b$ and $p_b = \varepsilon$. The fixed point of ψ is

$$\mathbf{v} = \lim_{k\to+\infty} \psi^k(a) = \underbrace{abab}_{\psi(a)}\underbrace{aba}_{\psi(b)}\underbrace{abab}_{\psi(a)}\underbrace{aba}_{\psi(b)}\underbrace{abab}_{\psi(a)}\ldots$$

We have $\mathcal{L}(\mathbf{u}) = \mathcal{L}(\mathbf{v})$.

Since $|\varphi(a)| \neq |\varphi(b)|$, the morphism φ is not uniform.

Example 3. The two already mentioned famous examples of infinite words, the Thue-Morse word $\mathbf{t}$ and the Fibonacci word $\mathbf{f}$, are both fixed points of a morphism.

The word $\mathbf{t}$ is fixed by the morphism φ_{TM} determined by $\varphi_{TM}(0) = 01$ and $\varphi_{TM}(1) = 10$. Note that this uniform morphism in fact has two fixed points, one being the other one after replacing 0 with 1 and 1 with 0. The word $\mathbf{t}$ as given above is the fixed points starting in 0.

The word $\mathbf{f}$ is fixed by the morphism φ_F defined by $\varphi_F(0) = 01$ and $\varphi_F(1) = 0$.

An (infinite) fixed point of a morphism of class P' clearly contains infinitely many palindromes which is one motivation for this notion. Class P is introduced in [20] in the context of discrete Schrödinger operators.

4 The study of palindromic defect

4.1 Characterizations of words with finite defect

We start by giving some of the known characterizations of words having finite palindromic defect.

Theorem 4. *For an infinite word* $\mathbf{u}$ *with language closed under reversal the following statements are equivalent:*

1. $D(\mathbf{u})$ *is finite ([10]);*
2. *there exists an integer* P *such that any prefix of* $\mathbf{u}$ *longer than* P *has a unioccurrent longest palindromic suffix ([16, 2]);*
3. *there exists an integer* N *such that for any palindromic factor of* $\mathbf{u}$ *having length at least* N*, every its complete return word is a palindrome ([2, 35]);*
4. *there exists an integer* N *such that for any factor* w *of* $\mathbf{u}$ *having length at least* N*, every factor of* $\mathbf{u}$ *that contains* w *only as its prefix and* $\mathrm{R}(w)$ *only as its suffix is a palindrome ([2, 35]);*
5. *there exists an integer* N *such that for each* $n \geq N$ *we have* $\mathcal{C}(n+1) - \mathcal{C}(n) + 2 = \mathcal{P}(n) + \mathcal{P}(n+1)$ *([3]).*

Sturmian words have palindromic defect equal to 0. It follows for instance for that fact that property 5 of the last theorem is satisfied for $N = 0$.

The following property is conjectured in [11] and shown in [3].

Theorem 5. *Let* $\mathbf{u}$ *be an infinite word with language closed under reversal. We have* $\sum_{n=0}^{+\infty} \left(\mathcal{C}_{\mathbf{u}}(n+1) + \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)\right) = 2D(\mathbf{u})$.

Besides these general properties, many examples of words with zero or finite palindromic defect were found:

- In [36, 13], another characterizations of words with zero palindromic defect are given.
- In [12], the relation of words with zero palindromic defect to so-called periodic-like words is exhibited.
- Links to another class of words, trapezoidal words, are shown in [15].
- The number of all words with zero palindromic defect of a given length and other properties are investigated in [48, 18].
- Words coding symmetric interval exchange transformations have zero palindromic defect by [1].
- In [9], the authors show that words coding rotation on the unit circle with respect to partition consisting of two intervals have zero palindromic defect.

- In [40], the author show a connection of words having zero palindromic defect with Burrows–Wheeler transform.
- In [45], we show that morphic images of episturmian words, a know class of words with zero palindromic defect, produces a word with finite palindromic defect.
- The articles [44, 21, 37] exhibit more examples of words with finite palindromic defect (along with some examples of words with finite generalized palindromic defect).

4.2 Zero defect conjecture

We now focus on words that are fixed by a morphism with the assumption that their language is closed under reversal. The main motivation to study their palindromic defect is the following conjecture.

Conjecture 6 (Zero defect conjecture [8])**.** *Let* $\mathbf{u}$ *be an aperiodic fixed point of a primitive morphism having its language closed under reversal. We have* $D(\mathbf{u}) = 0$ *or* $D(\mathbf{u}) = +\infty$.

The Thue–Morse word $\mathbf{t}$ and the Fibonacci word $\mathbf{f}$ are examples of aperiodic fixed points of a primitive morphism (see Example 3) having their language closed under reversal. We have $D(\mathbf{f}) = 0$ and $D(\mathbf{t}) = +\infty$.

Counterexamples to the conjecture were given in [14, 4]. Thus, the current statement of the conjecture is not true. However, there still might some refinement of the current statement that is valid as there are many witnesses and the found counterexamples seem to have some specific properties. Indeed, in [27] we prove that the conjecture is true for a special class of morphisms. A morphism φ is *marked* if there exists two morphisms φ_1 and φ_2, both being conjugate to φ, such that

$$\{\text{last letter of } \varphi_1(a)\colon a \in \mathcal{A}\} = \{\text{first letter of } \varphi_2(a)\colon a \in \mathcal{A}\} = \mathcal{A}.$$

In other words, the set of the last letters of the images of letters by φ_1 is the whole alphabet $\mathcal{A}$ and the set of the first letters of the images of letters by φ_2 is also the whole alphabet $\mathcal{A}$.

For instance, $\varphi = \varphi_{TM} : 0 \mapsto 01, 1 \mapsto 10$ is marked (here $\varphi = \varphi_1 = \varphi_2$). For $\varphi = \varphi_F : 0 \mapsto 01, 1 \mapsto 0$ we have $\varphi = \varphi_1$ and $\varphi_2 : 0 \mapsto 10, 1 \mapsto 0$. Thus, φ_F is also marked.

In [27] we show the following theorems:

Theorem 7. *Let* φ *be a primitive marked morphism and let* $\mathbf{u}$ *be its fixed point with finite palindromic defect. If all complete return words of all letters in* $\mathbf{u}$ *are palindromes or* φ *is conjugate to a morphism distinct from* φ *itself, then* $D(\mathbf{u}) = 0$.

In fact, any non-trivial morphism on binary alphabet is marked. Moreover, the binary alphabet allows for all of the assumptions to be dropped:

Theorem 8. *If $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is a fixed point of a primitive morphism over binary alphabet and $D(\mathbf{u}) < +\infty$, then $D(\mathbf{u}) = 0$ or $\mathbf{u}$ is periodic.*

We thus confirm that for a large class of fixed points of morphisms, their palindromic defect is either zero or infinite.

Let us outline the proof of Theorem 7:

1. We proceed by contradiction and assume that $\mathbf{u}$ has finite nonzero palindromic defect and is not periodic.
2. Under these assumptions, we show that there exists a non-empty factor q with a specific property: a graph[1] describing the extensions of q in $\mathcal{L}(\mathbf{u})$ contains a cycle ([27, Theorem 26]).
3. Using the results of [24, 26] and the fact that the morphism is marked, we find an infinite sequence of factors having the same property as the factor q.
4. Based on the results of [3], we show that a word with finite palindromic defect may have only finitely many factors with the same property as the factors q.

Let us comment on the assumptions of Theorem 7. The case which does not satisfy the assumptions, i.e., the case of primitive marked morphisms such that the morphism is only conjugate to itself and its fixed point $\mathbf{u}$ contains a non-palindromic complete return word to a letter, remains an open question. In this case, we obtain the empty word as the factor q from the above proof outline. Consequently, the method of step 3 does not yield infinitely many factors (it repeatedly yields $q = \varepsilon$). Thus, the proof is not applicable to this case and probably it is not easily extendable to it.

4.3 Related results and applications

The mentioned proof of Theorem 7 reveals the most intrinsic role of palindromic defect in the investigation of fixed points of morphisms. The connection links the symmetries of the investigated language of a fixed point, that is, in our case, its closedness under reversal, with the property of having infinitely many palindromic factors.

One direction of this connections is trivial: If a fixed point of a primitive morphism contains infinitely many palindromes, then its language is closed under reversal. The non-trivial converse is shown in [26] for marked morphisms. The authors of the mentioned article tackle the following conjecture:

[1] We omit the exact definition in the proof sketch, see [27] for a definition.

Conjecture 9 (Class P conjecture [20]). *Let $\mathbf{u}$ be a fixed point of a primitive morphism φ containing infinitely many palindromic factors. There exists a morphism of class P' such that its fixed point has the same language as $\mathbf{u}$.*

The original statement of the conjecture in [20] is ambiguous and allows for more interpretations, see also [26] or [19]. The above given statement of Conjecture 9 follows from two results. First, for binary alphabet the question is solved in [47]: if a fixed point of a primitive morphism φ over a binary alphabet contains infinitely many palindromes, then φ or φ^2 is of class P'. Second, in [25], the author shows that if we restrict ourselves just to infinite words, not more general languages of fixed points, the answer is negative: there exists a word $\mathbf{w}$ over ternary alphabet which is a fixed point of a primitive morphism, containing infinitely many palindromic factors, and not being fixed by any morphism of class P'. However, the authors of [19] note that the language of the word $\mathbf{w}$ may indeed be generated by a morphism of class P.

At this moment only partial answers to Conjecture 9 are known: as already mentioned, the binary case is solved ([47]); for larger alphabets an affirmative answer is provided only for some special classes of morphisms.

In [32], we confirm the conjecture for morphisms fixing a codings a non-degenerate exchange of 3 intervals. In [26], the authors prove the validity of the conjecture for marked morphisms. Moreover, the author show that a power of the marked morphism itself is in class P'. The technique and results used in the proofs of the latter fact is crucial in showing the defect conjecture for marked morphisms in [27].

The mentioned results lead to the formulation of the following conjecture.

Conjecture 10. *Let $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ be a primitive morphism having a fixed point $\mathbf{u}$. Its language $\mathcal{L}(\mathbf{u})$ is closed under reversal if and only if it contains infinitely many palindromic factors.*

As stated above, the conjecture is true for marked primitive morphisms. A proof in full generality of this conjecture has applications in algorithmic analysis of the language of a given morphism. Specifically, it allows for an efficient test whether the language of a fixed point is closed under reversal. For marked morphisms, such an algorithm may be devised based on the already mentioned results of [26]: it suffices to check whether a given power of the marked morphism in question is in class P'. Moreover, the algorithm is efficient as for a marked morphism its correct power may be easily determined and checking if it is in class P' is a straightforward task. Thus, in the general case, the last conjecture is a first step to provide an efficient test for any primitive morphism.

References

[1] P. BALÁŽI, Z. MASÁKOVÁ, AND E. PELANTOVÁ, *Factor versus palindromic complexity of uniformly recurrent infinite words*, Theoret. Comput. Sci., 380 (2007), pp. 266–275.

[2] L. BALKOVÁ, E. PELANTOVÁ, AND Š. STAROSTA, *Infinite words with finite defect*, Adv. in Appl. Math., 47 (2011), pp. 562–574.

[3] ——, *Proof of the Brlek-Reutenauer conjecture*, Theoret. Comput. Sci., 475 (2013), pp. 120–125.

[4] B. BAŠIĆ, *On highly potential words*, Eur. J. Combin., 34 (2013), pp. 1028–1039.

[5] J. BERSTEL, *Axel Thue's papers on repetitions in words: a translation*, Pub. LaCIM, 20 (1995).

[6] J. BERSTEL AND D. PERRIN, *The origins of combinatorics on words*, Eur. J. Combin., 28 (2007), pp. 996–1022.

[7] V. BERTHÉ AND M. RIGO, eds., *Combinatorics, automata, and number theory.*, vol. 135 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2010.

[8] A. BLONDIN MASSÉ, S. BRLEK, A. GARON, AND S. LABBÉ, *Combinatorial properties of f-palindromes in the Thue-Morse sequence*, Pure Math. Appl., 19 (2008), pp. 39–52.

[9] A. BLONDIN MASSÉ, S. BRLEK, S. LABBÉ, AND L. VUILLON, *Palindromic complexity of codings of rotations*, Theoret. Comput. Sci., 412 (2011), pp. 6455–6463.

[10] S. BRLEK, S. HAMEL, M. NIVAT, AND C. REUTENAUER, *On the palindromic complexity of infinite words*, Int. J. Found. Comput. Sci., 15 (2004), pp. 293–306.

[11] S. BRLEK AND C. REUTENAUER, *Complexity and palindromic defect of infinite words*, Theoret. Comput. Sci., 412 (2011), pp. 493–497.

[12] 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, Berlin, Heidelberg, 2009, Springer-Verlag, pp. 145–155.

[13] M. BUCCI, A. DE LUCA, A. GLEN, AND L. Q. ZAMBONI, *A new characteristic property of rich words*, Theoret. Comput. Sci., 410 (2009), pp. 2860–2863.

[14] M. BUCCI AND E. VASLET, *Palindromic defect of pure morphic aperiodic words*, in Proceedings of the 14th Mons Days of Theoretical Computer Science, 2012.

[15] A. DE LUCA, A. GLEN, AND L. Q. ZAMBONI, *Rich, sturmian, and trapezoidal words*, Theoret. Comput. Sci., 407 (2008), pp. 569–573.

[16] X. DROUBAY, J. JUSTIN, AND G. PIRILLO, *Episturmian words and some constructions of de Luca and Rauzy*, Theoret. Comput. Sci., 255 (2001), pp. 539–553.

[17] N. P. FOGG, *Substitutions in Arithmetics, Dynamics and Combinatorics*, vol. 1794 of Lecture notes in mathematics, Springer, 1st ed., 2002.

[18] C. GUO, J. SHALLIT, AND A. M. SHUR, *Palindromic rich words and run-length encodings*, Inform. Process. Lett., 116 (2016), pp. 735–738.

[19] T. HARJU, J. VESTI, AND L. Q. ZAMBONI, *On a question of Hof, Knill and Simon on palindromic substitutive systems*, Monatsh. Math., 179 (2016), pp. 379–388.

[20] A. HOF, O. KNILL, AND B. SIMON, *Singular continuous spectrum for palindromic Schrödinger operators*, Comm. Math. Phys., 174 (1995), pp. 149–159.

[21] T. JAJCAYOVÁ, E. PELANTOVÁ, AND Š. STAROSTA, *Palindromic closures using multiple antimorphisms*, Theoret. Comput. Sci., 533 (2014), pp. 37–45.

[22] C. R. JEAN BERSTEL, DOMINIQUE PERRIN, *Codes and Automata*, Encyclopedia of Mathematics and its Applications 129, Cambridge University Press, 1 ed., 2009.

[23] J. KARHUMÄKI, *Combinatorics on words: A new challenging topic*, Tech. Rep. 645, Turku Centre for Computer Science, 2004.

[24] K. KLOUDA, *Bispecial factors in circular non-pushy D0L languages*, Theoret. Comput. Sci., 445 (2012), pp. 63–74.

[25] S. LABBÉ, *A counterexample to a question of Hof, Knill and Simon*, Electron. J. Combin., 21 (2014).

[26] S. LABBÉ AND E. PELANTOVÁ, *Palindromic sequences generated from marked morphisms*, Eur. J. Combin., 51 (2016), pp. 200–214.

[27] S. LABBÉ, E. PELANTOVÁ, AND Š. STAROSTA, *Zero defect conjecture*, submitted to Eur. J. Combin., (2016).

[28] M. LOTHAIRE, *Combinatorics on Words*, vol. 17 of Encyclopaedia of Mathematics and its Applications, Addison-Wesley, Reading, Mass., 1983. Reprinted in the *Cambridge Mathematical Library*, Cambridge University Press, Cambridge UK, 1997.

[29] ——, *Algebraic Combinatorics on Words*, no. 90 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2002.

[30] ——, *Applied Combinatorics on Words*, no. 105 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2005.

[31] C. MARTÍN-VIDE, V. MITRANA, AND G. PĂUN, eds., *Formal Languages and Applications*, vol. 148 of Studies in Fuzziness and Soft Computing, Springer Berlin Heidelberg, 2004.

[32] Z. MASÁKOVÁ, E. PELANTOVÁ, AND Š. STAROSTA, *Exchange of three intervals: substitutions and palindromicity*, Eur. J. Combin., 62 (2017), pp. 217–231.

[33] H. M. MORSE, *Recurrent geodesics on a surface of negative curvature*, T. Am. Math. Soc., 22 (1921), pp. 84–100.

[34] M. MORSE AND G. A. HEDLUND, *Symbolic dynamics II - Sturmian trajectories*, Amer. J. Math., 62 (1940), pp. 1–42.

[35] E. PELANTOVÁ AND Š. STAROSTA, *Languages invariant under more symmetries: overlapping factors versus palindromic richness*, Discrete Math., 313 (2013), pp. 2432–2445.

[36] ——, *Palindromic richness for languages invariant under more symmetries*, Theor. Comput. Sci, 518 (2014), pp. 42–63.

[37] ——, *Constructions of words rich in palindromes and pseudopalindromes*, Discrete Math. Theoret. Comput. Sci., 18 (2016).

[38] J. F. POWER, *Thue's 1914 paper: a translation*, available at http://arxiv.org/abs/1308.5858.

[39] E. PROUHET, *Mémoire sur quelques relations entre les puissances des nombres*, C. R. Acad. Sci. Paris, 33 (1851), p. 225.

[40] A. RESTIVO AND G. ROSONE, *Balancing and clustering of words in the Burrows-Wheeler transform*, Theoret. Comput. Sci., 412 (2011), pp. 3019–3032.

[41] M. RIGO, *Formal Languages, Automata and Numeration Systems 1*, Wiley, 2014.

[42] ——, *Formal Languages, Automata and Numeration Systems 2*, Wiley, 2014.

[43] G. ROZENBERG AND A. SALOMAA, *The book of L*, Springer, 1986.

[44] Š. STAROSTA, *Generalized Thue-Morse words and palindromic richness*, Kybernetika, 48 (2012), pp. 361–370.

[45] ——, *Morphic images of episturmian words having finite palindromic defect*, Eur. J. Combin., 51 (2016), pp. 359–371.

[46] M. STEINBY AND W. THOMAS, *Trees and term rewriting in 1910 : on a paper by Axel Thue*, EATCS Bull., 72 (2000), pp. 256–269.

[47] B. TAN, *Mirror substitutions and palindromic sequences*, Theoret. Comput. Sci., 389 (2007), pp. 118–124.

[48] J. VESTI, *Extensions of rich words*, Theoret. Comput. Sci., 548 (2014), pp. 14–24.

Ing. Štěpán Starosta, Ph.D.

Contact

Email:	`stepan.starosta@fit.cvut.cz`
Telephone:	+420 224 359 886
Web:	`users.fit.cvut.cz/~staroste`
Adress:	ČVUT v Praze, FIT Thákurova 9 160 00 Praha 6 Czech Republic

Personal information

Born 3 June 1983.
Language skills: English, French (advanced), Spanish (basics), Czech (first language).

Education

2008 - 2012	CTU in Prague, FNSPE, Department of Mathematics - PhD studies specialization: Mathematical modelling dissertation thesis: Generalized Palindromes in Infinite Words advisor: prof. Ing. Edita Pelantová CSc.
2002 - 2008	CTU in Prague, FNSPE, Department of Mathematics - Master studies specialization: Mathematical modelling diploma thesis: Application of adaptive filtering in recognition of spoken Czech advisor: Ing. Pavel Bolek

Work experience

2011 – now	*researcher* and *assistant professor* at FIT CTU in Prague
2004 – 2013	*programmer and software consultant*

Research

Domains of interest

Combinatorics on words, Symbolic dynamical systems, Theoretical computer science, Number theory, `SageMath` - open source computer algebra system

Publications published in peer reviewed journals

Author or co-author of 21 articles published in peer-reviewed journals.
Selected publications:

1. Š. Starosta, *On Theta-palindromic Richness*, Theoret. Comp. Sci. **412** (2011), 1111–1121, DOI: 10.1016/j.tcs.2010.12.011
2. P. Arnoux and Š. Starosta, *Rauzy gasket*, in: J. Barral and S. Seuret (Eds.), Further Developments in Fractals and Related Fields, Trends in Mathematics, Springer Science+Business Media New York 2013, DOI 10.1007/978-0-8176-8400-6_1
3. E. Pelantová and Š. Starosta: *Languages invariant under more symmetries: overlapping factors versus palindromic richness*, Discrete Math. 313 (2013), 2432–2445, DOI: 10.1016/j.disc.2013.07.002
4. E. Pelantová and Š. Starosta: *Palindromic richness for languages invariant under more symmetries*, Theoret. Comput. Sci. 518 (2014), 42–63, DOI: 10.1016/j.tcs.2013.07.021
5. M. Kupsa, Š. Starosta: *On the partitions with Sturmian-like refinements and an application to factor mappings from Sturmian subshifts*, Discrete and Continuous Dynamical Systems - Series A, Volume 35, Issue 8, August 2015, 3483–3501, DOI: 10.3934/dcds.2015.35.3483
6. K. Klouda, Š. Starosta: *An Algorithm Enumerating All Infinite Repetitions in a D0L-System*, J. Discrete Algorithms 33 (2015), 130—138, DOI: 10.1016/j.jda.2015.03.006
7. Š. Starosta: *Morphic images of episturmian words having finite palindromic defect*, Eur. J. Combin. 51 (2016), 359–371, DOI: 10.1016/j.ejc.2015.07.001
8. E. Pelantová, Š. Starosta, and M. Znojil: *Markov Constant and Quantum Instabilities*, Journal of Physics A: Mathematical and Theoretical 49 (2016), Number 15, DOI: 10.1088/1751-8113/49/15/155201
9. S. Labbé, E. Pelantová and Š. Starosta: *On the Zero Defect Conjecture*, Eur. J. Combin. 62 (2017), 132–146, DOI: 10.1016/j.ejc.2016.12.006

Grants

2013 – 2015	head investigator of 3-year postdoctoral research project of Czech Science Foundation (GA ČR) entitled *Algoritmy pro cirkulární morfismy a jejich pevné body (Algorithms for circular morphisms and their fixed points)*
2010, 2011	head investigator of student's grant SGS CTU in Prague entitled *Kombinatorické a algebraické aspekty nestandardních číselných systémů (Combinatorial and algebraic aspects of nonstandard numeration systems)*

Honors and Awards

2013	received award *Cena rektora za vynikající doktorskou práci za rok 2012 - I. stupeň (Rector's award for an excellent doctoral thesis in 2012 - I. class)*
2012	received award *Cena Josefa Hlávky pro nejlepší studenty a absolventy (Josef Hlávka's Award for best students and graduates)*

Teaching activities

Supervised 3 master students and 6 bachelor students.
Currently participating in teaching and preparation of courses *Matematika pro znalostní inženýrství (Mathematics for Knowledge Engineering)*, *Matematika pro kryptologii (Mathematics for Cryptology)*, *Matematika pro informatiku (Mathematics for Informatics)* at FIT CTU in Prague.