Morphisms, infinite words, and symmetries

HABILITATION THESIS

Prague, 2016

Štěpán Starosta

Abstract

This habilitation thesis is a collection of articles covering various topics, published or submitted from 2013 to 2016. Most of the articles are set in Combinatorics on Words and are dealing with languages generated by morphisms that have reversal symmetry: with each element of the language, the mirror image of this element is also in the language.

The first part of the articles investigates some of the recent conjectures in Combinatorics on Words. The first conjecture is the Brlek–Reutenaurer conjecture, connecting palindromic defect with factor and palindromic complexities. We solve this conjecture by giving an affirmative answer. The second conjecture is the Class P conjecture, stating that if a language generated by a morphism contains infinitely many palindromes, then the morphism belongs to a special class of morphism called class P. The third conjecture is the Zero defect conjecture, which states that if the generating morphism is primitive, then the palindromic defect of the language is zero or infinity. We give only partial answers to the last two conjectures: we deal with some specific subclasses of the morphisms in question. Namely, we give an affirmative answer for morphisms fixing 3 interval exchange transformation for Class P conjecture, and for binary and primitive marked morphisms for Zero defect conjecture.

The second part of the articles present many new constructions of words with finite palindromic defect, also in a generalized sense. We enlarge the family of known examples of such words by following the construction of Rote words, by doing letter-to-letter projections of episturmian words, and by investigating generalized Thue–Morse words.

The third part of the thesis deals with efficient algorithmic analysis of languages generated by morphisms and leads toward two efficient algorithms: the first algorithm enumerates all primitive factors that occur in the generated language in any power; the second algorithm tests whether a morphism is circular.

The last part is constituted from two various results: the study of the Rauzy gasket, a set representing letter frequencies of all ternary episturmian words; and the study of a generalization of Markov constant motivated by the study of spectrum of a a differential operator. This part serves as an illustration of connection of Combinatorics on Words to other research domains.

Acknowledgements

The works included in this thesis were supported by the following grants: Czech Science Foundation GAČR grants 13-35273P, 13-03538S, 201/09/0584; grants MSM6840770039 and LC06002 of the Ministry of Education, Youth, and Sports of the Czech Republic; and grant of the Grant Agency of the Czech Technical University in Prague grant No. SGS11/162/OHK4/3T/14.

The experimental results were done using the open-source mathematical software `SageMath` [77], including PARI/GP [1] and FLINT [38].

The author also wishes to thank his co-authors for their joyful cooperation.

List of included articles

This thesis is a collection of the following articles. They are included as chapters 2–11 and are accompanied by an introduction (Chapter 1) where they are referred to by Roman numerals.

[I] L. Balková, E. Pelantová, and Š. Starosta, *Proof of the Brlek-Reutenauer conjecture*, Theor. Comput. Sci., 475 (2013), pp. 120–125.

[II] Z. Masáková, E. Pelantová, and Š. Starosta, *Exchange of three intervals: itineraries, substitutions and palindromicity*, submitted to Eur. J. Combin., 2015.

[III] S. Labbé, E. Pelantová, and Š. Starosta, *Zero defect conjecture*, submitted to Eur. J. Combin., 2016.

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

[V] E. Pelantová and Š. Starosta, *Constructions of words rich in palindromes and pseudopalindromes*, accepted in Discrete Math. Theor. Comput. Sci., (2015).

[VI] T. Jajcayová, E. Pelantová, and Š. Starosta, *Palindromic closures using multiple antimorphisms*, Theor. Comput. Sci., 533 (2014), pp. 37–45.

[VII] K. Klouda and Š. Starosta, *An algorithm enumerating all infinite repetitions in a D0L-system*, J. Discrete Algorithms, 33 (2015), pp. 130–138.

[VIII] K. Klouda and Š. Starosta, *Characterization of circular D0L-systems*, submitted to Discrete Math., 2015.

[IX] P. Arnoux and Š. Starosta, *The Rauzy gasket*, in Further Developments in Fractals and Related Fields, J. Barral and S. Seuret, eds., Trends in Mathematics, Springer Science+Business Media New York, 2013, pp. 1–23.

[X] E. Pelantová, Š. Starosta, and M. Znojil, *Markov constant and quantum instabilities*, J. Phys. A: Math. Theor., 49 (2016), p. 155201.

Contents

Chapter 1

Introduction

This introduction serves as an overview of the included results accompanied with comments. A summary of used notions is part of this overview, including a brief general overview of the research domain.

1.1 On Combinatorics on Words

The research domain of most of the included articles is Combinatorics on Words which is catalogued in the Mathematics Subject Classification database under 68R15. We start by giving a brief overview of this domain.

1.1.1 Brief history and connection to other domains

The beginning of Combinatorics on Words is mostly attributed to Axel Thue and his article [78] from 1906 and 3 following articles until 1914. The reason is that he 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 [9, 74, 62] for translations of Thue's papers and comments on his results.

In 1921, Marston Morse published an article studying geodesics [54]. The article contained an overlap with Axel Thue's study. This overlap gave rise to a name of a famous infinite word: the **Thue–Morse word**. The Thue–Morse word, denoted $\mathbf{t}$, is an infinite word of elements from the alphabet $\{0, 1\}$, i.e., an infinite sequence of 0s and 1s. It can be constructed by building prefixes of $\mathbf{t}$, i.e., finite sequences that form the beginning of $\mathbf{t}$. We start by setting the first prefix to $p_1 = 0$, the finite word of length 1 consisting of the letter 0. We apply the following rewriting rule to all the elements of p_1: we replace 0 by

01 and 1 by 10. The obtained word is $p_2 = 01$ and it is the next constructed prefix of $\mathbf{t}$. Repeating the procedure, we obtain

$$\begin{aligned}
p_1 &= 0,\\
p_2 &= 01,\\
p_3 &= 0110,\\
p_4 &= 01101001,\\
p_5 &= 0110100110010110,\\
p_6 &= 01101001100101101001011001101001.
\end{aligned}$$

We can observe that p_i is also a prefix of p_{i+1} and the length of p_{i+1} is twice the length of p_i. Therefore, there is a unique infinite word having each p_i as a prefix, i.e., the procedure to construct $\mathbf{t}$ is unambiguous.

The constructions of the Thue–Morse word of Thue and Morse differ as they appear in different context. As already mentioned, Morse was studying geodesics. Thue was solving the following problem: does there exist an infinite word over a 2-letter alphabet which does not contain a cube? A *cube* is a contiguous subsequence that can be written as a repetition of 3 words. For instance, the word 011011011 is a cube since 011 is repeated 3 times. The Thue–Morse word $\mathbf{t}$ has such a property: there are no cubes in the Thue–Morse word.

Let us note that the word $\mathbf{t}$ is also sometimes called Prouhet–Thue–Morse since is appeared already in 1851 in [63] by Eugène Prouhet.

The next stepping stone in the history of Combinatorics on Words is the article [55] 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 $n + 1$.

Such finite contiguous subsequence is called a *factor*, thus the name *factor complexity* since it is one of the measures of chaos (or order) of an infinite word. The basic property of factor complexity is the following: 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 the notion of Sturmian words by defining the most famous word of this class: the Fibonacci word $\mathbf{f}$. The word $\mathbf{f}$ may be defined using the construction procedure of the Thue–Morse word $\mathbf{t}$ but using a distinct rewriting rule. 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 have $q_2 = 01$. We repeat the procedure using the new rewriting rule and obtain

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

As in the case of the word $\mathbf{t}$, 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 $\mathbf{f}$. Its name comes from the fact that we can recover the Fibonacci sequence by looking at the lengths of the prefixes q_i, denoted by $|q_i|$. We have

$$|q_{i+1}| = |q_i| + |q_{i-1}| \qquad \text{for all } i > 2.$$

In other words, the length of a prefix q_{i+1} equals the sum of the sum of the lengths of the previous two prefixes. As the initial conditions of this recurrence are $|q_1| = 1$ and $|q_2| = 2$, we retrieve the Fibonacci sequence.

After the mentioned works, the field of Combinatorics on Words has been growing steadily. The reader may refer to [10] 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 [49]. Two more books by M. Lothaire were published later, *Algebraic Combinatorics on Words* in 2002 [50] and *Applied Combinatorics on Words* in 2005 [51].

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 [33] is a basic reference for the connection of Combinatorics of Words and Symbolic Dynamics. The publication *Combinatorics, Automata, and Number Theory* of 2010 [14] 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 [40, 52, 67, 68].

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 [41] for an overview of some mentioned applications.

Let us mention, as a brief illustration of one of the connections, an attractive application in Biology, namely in DNA sequencing. One of the basic questions is the following: let w be a finite word of length n over an alphabet $\mathcal{A}$. Determine the least number k such that the word w can be reconstructed from the knowledge of k distinct subwords of w, where subword is a subsequence of w. The question is already stated using the terminology of Combinatorics on Words, which serves as a theoretical basis to study problems such as DNA sequencing. In the terms of DNA sequencing, the alphabet is fixed to be the four-elements set of nucleotides, the word w is the part of the DNA sequence to be recovered, and the subwords are the results of the experiments when analysing w. The reader may refer to [48] for more details on the links between Combinatorics on Words and DNA sequencing.

1.1.2 Current challenges

In this section we manifest the current state of the Combinatorics on Words by listing some of the actual challenges recognized by the international community. While listing below some of the current challenges, we omit giving specific details since the list is quite extensive and may be extended even more. In Section 1.3, we give more details for topics that are directly related to the results contained in this thesis.

We start the list by giving the reference to *Ten Conferences WORDS: Open Problems and Conjectures* [57] published in 2016 by Jean Néraud. As the title indicates, the author enumerates some of the questions that came up during WORDS conferences in the last ten years. The mentioned conference may be considered to be one of the most important conference series for researchers working in Combinatorics on Words.

The challenges mentioned in [57] are grouped as follows.

- *Pattern avoidance*: questions related to existence and properties of words avoiding certain patterns such as cubes in the case of the Thue–Morse word are being investigated.

- *Complexity studies*: besides factor complexity, open questions related to palindromic complexity, Abelian complexity, arithmetical complexity and other complexity measures are in the focus of researchers.

- *Factorization of words and equations*: questions concerning specific factorizations of words or equations such as $xy = yx$ with x, y being finite words are within the scope of ongoing research topics.

Inspired by the overview done by Jean Néraud, we may investigate the most frequent topics in some other important conference series and regular scientific meetings focusing on Combinatorics on Words: *Mons Theoretical Computer Science Days*, *RuFiDiM - Russian Finnish Symposium on Discrete Mathematics*, *Development in Language Theory*, and *International school and conference on Combinatorics, Automata and Number Theory.* We list some other frequent topics:

- balance properties of words,
- Rauzy fractals and tilings generated by words,
- morphisms and morphic words,
- automatic sequences,
- properties of words in link with other domains such as:
  - diophantine approximation and number theory in general,
  - numeration,
  - formal language theory,
  - automata theory.

We finish this overview by mentioning the growing software support for Combinatorics on Words in the open-source computer algebra system `SageMath` [77]. The library is being developed by some of the researchers in the domain and its possibilities range from basic support for finite or infinite words and morphisms to more advanced methods and algorithms. Many researchers use this library and advance faster in their discoveries thanks to it. As the system `SageMath` is built to represent mathematical knowledge, not just implement various algorithms, we can observe the connections of Combinatorics on Words by looking at how the library is used by other parts of the software[1].

Before giving more specific overview of the results included in this thesis, we summarize the used notation and terminology.

[1] The source code of the library can be explored at `https://git.sagemath.org/sage.git/tree/src/sage/combinat/words` (October 2016).

1.2 Notation and terminology

We mostly follow the usual terminology of Combinatorics on Words, see for instance [49]. 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 $\mathcal{A}$. Examples of infinite words over the alphabet $\{0, 1\}$ are the Thue–Morse word $\mathbf{t}$ and the Fibonacci word $\mathbf{f}$ defined above. The set of all (right-)infinite words over $\mathcal{A}$ is denoted $\mathcal{A}^{\mathbb{N}}$.

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 $\mathcal{L}_n(\mathbf{u})$.

A factor $w \in \mathcal{L}(\mathbf{u})$ is *right special* if there exist two distinct letters a and b such that $wa, wb \in \mathcal{L}(\mathbf{u})$. Analogously, it is *left special* if $aw, bw \in \mathcal{L}(\mathbf{u})$. A factor is *bispecial* if it is both left and right special. For a factor w we define its *bilateral multiplicity* (or *bilateral order*), see [23], $\mathrm{m}(w)$ as follows:

$$\begin{aligned}\mathrm{m}(w) = {} & \#\{awb \in \mathcal{L}(\mathbf{u}) \colon a, b \in \mathcal{A}\} \\ & - \#\{wb \in \mathcal{L}(\mathbf{u}) \colon b \in \mathcal{A}\} - \#\{aw \in \mathcal{L}(\mathbf{u}) \colon a \in \mathcal{A}\} \\ & + 1.\end{aligned}$$

An *occurrence* of $w = w_0 w_1 \cdots w_{n-1} \in \mathcal{A}^n$ in a word $v = v_0 v_1 v_2 \ldots$ 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.

We say that an infinite word $\mathbf{u}$ is *recurrent* if every its factor has infinitely many occurrences in $\mathbf{u}$. The word $\mathbf{u}$ is *uniformly recurrent* if it is recurrent and every its factor has a finite number of complete return words in $\mathbf{u}$. In other words, the gaps between successive occurrences of a factor are bounded.

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

$$R(w) = w_{n-1} w_{n-2} \cdots w_1 w_0 \quad \text{where } w = w_0 w_1 \cdots w_{n-1} \in \mathcal{A}^n.$$

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

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

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

With $\mathrm{Pal}(\mathbf{u})$ being 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

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

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

1.2.1 Palindromic defect

One property of words which is in the research focus is related to palindromic factors of a finite or infinite word. In [27], 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 $\#\mathrm{Pal}(w)$:

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

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 [17]. It is denoted $D(w)$. We have

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

A basic property of the palindromic defect is that $D(w) \ge 0$ for any w and $D(v) \le D(w)$ 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.

There exist words that have palindromic defect 0. The famous Sturmian words are an example. Such words are also called *rich* or *full*, and are being investigated as they possess some notable properties. The first interesting property is the existence of the following theorem listing many known characterizations of words with zero palindromic defect (provided the language in question is closed under reversal). For a palindrome $w \in \mathcal{L}(\mathbf{u})$, we set $\mathrm{Pext}(w) = \{awa \in \mathcal{L}(\mathbf{u}) \colon a \in \mathcal{A}\}$.

Theorem 1. *For an infinite word $\mathbf{u}$ with language closed under reversal the following statements are equivalent:*

1. $D(\mathbf{u}) = 0$ *(i.e.,* $\mathbf{u}$ *is rich) ([27]);*

2. *any complete return word of any palindromic factor of* $\mathbf{u}$ *is a palindrome ([34]);*

3. *for any factor* w *of* $\mathbf{u}$*, every factor of* $\mathbf{u}$ *that contains* w *only as its prefix and* $R(w)$ *only as its suffix is a palindrome ([34]);*

4. *the longest palindromic suffix of any factor* $w \in \mathcal{L}(\mathbf{u})$ *is unioccurrent in* w *([27, 34]);*

5. *for each* n *the following equality holds*

$$\mathcal{C}(n+1) - \mathcal{C}(n) + 2 = \mathcal{P}(n) + \mathcal{P}(n+1)$$

   ([20]);

6. *any bispecial factor* $w \in \mathcal{L}(\mathbf{u})$ *satisfies*

$$\mathrm{m}(w) = \begin{cases} 0 & \text{if } w \neq R(w); \\ \#\mathrm{Pext}(w) - 1 & \text{otherwise.} \end{cases}$$

   ([6]).

1.2.2 Infinite words with finite palindromic defect

A significant part of this thesis is connected to infinite words having finite palindromic defect. We give a list of their characterizations, which can be seen as a generalization of Theorem 1.

Theorem 2. *For an infinite word* $\mathbf{u}$ *with language closed under reversal the following statements are equivalent:*

1. $D(\mathbf{u})$ *is finite ([17]);*

2. *there exists an integer* P *such that any prefix of* $\mathbf{u}$ *longer than* P *has a unioccurrent longest palindromic suffix ([27, 7]);*

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 ([7, 60]);*

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* $R(w)$ *only as its suffix is a palindrome ([7, 60]);*

5. *there exists an integer N such that for each $n \geq N$ the following equality holds*
$$\mathcal{C}(n+1) - \mathcal{C}(n) + 2 = \mathcal{P}(n) + \mathcal{P}(n+1)$$
([I]);

6. *there exists an integer N such that any bispecial factor $w \in \mathcal{L}(\mathbf{u})$ of length at least N satisfies*
$$\mathrm{m}(w) = \begin{cases} 0 & \text{if } w \neq R(w); \\ \#\mathrm{Pext}(w) - 1 & \text{otherwise;} \end{cases}$$
and
$$\mathcal{C}(N+1) - \mathcal{C}(N) + 2 = \mathcal{P}(N) + \mathcal{P}(N+1)$$
([61]).

Note that a generalization of property 4 of Theorem 1 is not included in the last theorem. It may be included only if we add the assumption of uniform recurrence. In other words, if $\mathbf{u}$ is uniformly recurrent and has finite palindromic defect, then there exists an integer T such that any factor of $\mathbf{u}$ longer than T has a unioccurrent longest palindromic suffix. See [61, Theorem 35] for a proof[2]. The converse follows directly from Theorem 2, item 2.

To see that uniform recurrence is indeed required, consider the following example. Let $p = 1231321$. We have $D(p) = 1$. Set $w_0 = p$ and $w_i = w_{i-1}0^i w_{i-1}$ for all $i > 0$ where 0^i is the word consisting of the letter 0 repeated i times. Let $\mathbf{w}$ be the infinite word having w_i as its prefix for all i. We have

$$\mathbf{w} = \underbrace{\overbrace{\underbrace{p0p}_{w_1}\, 00p0p}^{w_2}\, 000p0p00p0p}_{w_3}\, 0000p\ldots$$

It can be verified that $D(\mathbf{w})$ is finite:

Proposition 3. *The word $\mathbf{w}$ satisfies $D(\mathbf{w}) = D(p) = 1$.*

Proof. We obtain $D(p) = 1$ by direct calculation.

By [I, Corollary 3], the palindromic defect of $\mathbf{w}$ is equal to the number of its prefixes such that their longest palindromic suffix is not unioccurrent. We show for all i that each prefix of w_i longer than $|p|$ has a unioccurrent longest

[2] The proof in the mentioned reference is given in a more general context of more possible symmetries than just the reversal symmetry.

palindromic suffix. We proceed by induction on i. First note that for all i, the word w_i is a palindrome and it contains exactly two occurrences of w_{i-1}.

The word $w_0 = p$ has no prefix longer than $|p|$ thus the claim is true.

Let $i > 0$. Assume that for w_{i-1}, each its prefix longer than $|p|$ has a unioccurrent longest palindromic suffix.

The prefixes of w_i of length less than $|w_{i-1}| + 1$ satisfy the claim by the induction hypothesis. The prefix of w_i of the form $w_{i-1}0^k$ with $0 < k < i$ has its longest palindromic suffix equal to $0^k w_{i-2} 0^k$. As there are exactly two occurrences of w_{i-2} in w_{i-1} and the other occurrence is as a prefix, this longest palindromic suffix is unioccurrent. The prefix of w_i of the form $w_{i-1}0^i s$ where s is a prefix of w_{i-1} has its longest palindromic suffix equal to $R(s)0^i s$ since w_{i-1} is a palindrome. As 0^i has exactly one occurrence in w_i, the longest palindromic suffix $R(s)0^i s$ is unioccurrent.

As w_i is a prefix of $\mathbf{w}$ for all i, we have $D(\mathbf{w}) = D(p) = 1$. □

To see that the word $\mathbf{w}$ is the counterexample we are looking for, consider for each integer k the word $0^k 1231$ which is a factor of $\mathbf{w}$. Its longest palindromic suffix is 1 and it is not unioccurrent.

1.2.3 Morphisms and languages they generate

Morphisms are an important tool to generate infinite words and their languages. A *morphism* φ is a mapping $\mathcal{A}^* \to \mathcal{B}^*$ where $\mathcal{A}$ and $\mathcal{B}$ are alphabets such that $\forall v, w$ we have $\varphi(vw) = \varphi(v)\varphi(w)$ (it is a homomorphism of 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 each $a, b \in \mathcal{A}$ there exists an integer k such that b occurs in $\varphi^k(a)$. It is well known that fixed points of primitive morphisms are uniformly recurrent.

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 4. 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

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

As φ is primitive, the word $\mathbf{u}$ is uniformly recurrent. 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 5. 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 [39] in the context of discrete Schrödinger operators.

1.3 Overview of results and comments

We give an overview of the results of the research articles which are part of this thesis. The overview is divided into the following groups:

1. Three conjectures related to palindromes,
2. Construction of words with finite palindromic defect,
3. D0L-systems and algorithms,
4. Rauzy gasket and generalized Markov constant.

1.3.1 Three conjectures related to palindromes

This section contains results on recent conjectures from Combinatorics on Words dealing with general properties of words related to palindromes and palindromic defect.

Brlek–Reutenauer conjecture ([I])

The first conjecture relates the palindromic defect of an infinite word with its factor and palindromic complexities.

Conjecture 1 (Brlek–Reutenauer conjecture [18]). *Let* $\mathbf{u}$ *be an infinite word with language closed under reversal. We have*

$$2D(\mathbf{u}) = \sum_{n=0}^{+\infty} \left(\mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)\right).$$

The authors of the conjecture proved already in [18] that the conjecture is true for periodic infinite words. The article [20] gives a positive answer for words having palindromic defect zero (it is a consequence of already mentioned Theorem 1, item 5). The article [I] completes the study by giving an affirmative answer to Conjecture 1. The proof is done by showing the two following theorems.

Theorem 6. *If* $\mathbf{u}$ *is an infinite word with language closed under reversal such that both* $D(\mathbf{u})$ *and* $\sum_{n=0}^{+\infty} \left(\mathcal{C}_{\mathbf{u}}(n+1) + \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)\right)$ *are finite, then*

$$2D(\mathbf{u}) = \sum_{n=0}^{+\infty} \left(\mathcal{C}_{\mathbf{u}}(n+1) + \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)\right).$$

Theorem 7. *If* $\mathbf{u}$ *is an infinite word with language closed under reversal, then* $D(\mathbf{u}) < +\infty$ *if and only if*

$$\sum_{n=0}^{+\infty} \left(\mathcal{C}_{\mathbf{u}}(n+1) + \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)\right) < +\infty.$$

The relation of palindromic defect and factor and palindromic complexity given by the statement of Conjecture 1 is not very practical to compute the palindromic defect of an infinite word. However, it gives more insight into words with finite palindromic defect. The methods used to prove the conjecture resulted in giving the characterization 5 of Theorem 2 and subsequently also the characterization 6. The article [III] exploits these results in the proof of its main theorem (see below).

Class P conjecture ([II])

The second conjecture is the following.

Conjecture 2 (Class P conjecture [39])**.** *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 [39] is ambiguous and allows for more interpretations, see also [46] or [37]. The above given statement of Conjecture 2 follows from two results. First, for binary alphabet the question is solved in [76]: 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 [45], 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 [37] note that the language of the word $\mathbf{w}$ may indeed be generated by a morphism of class P.

An affirmative answer to Conjecture 2 would provide a strong tool to investigate fixed points of primitive morphisms that contain infinitely many palindromic factors. However, at this moment, only partial answers are known: as already mentioned, the binary case is solved ([76]); for larger alphabets an affirmative answer is provided only for some special classes of morphisms ([46] and [II]).

The affirmative answer in [II] is for the case of φ fixing a word coding a non-degenerate exchange of 3 intervals. Moreover, the result states that φ itself or φ^2 is in class P'.

Words coding exchange of k intervals form a well-known class of words over ternary alphabet. Such words may be defined as follows. Let J be a left-closed right-open interval. Consider a partition $J = J_0 \cup \cdots \cup J_{k-1}$ of J into a disjoint union of left-closed right-open subintervals such that $\forall x \in J_i, \forall y \in J_{i+1}, x < y$. A bijection $T : J \to J$ is an *exchange of* k *intervals with permutation* π if there exist numbers $c_0, \ldots, c_{k-1}$ such that for $0 \leq i < k$ one has

$$T(x) = x + c_i \text{ for } x \in J_i,$$

where π is a permutation of $\{0, 1, \ldots, k-1\}$ determining the order of intervals $T(J_i)$, i.e., such that $\pi(i) < \pi(j)$ implies $\forall x \in T(J_i), \forall y \in T(J_j), x < y$. If π is the permutation $i \mapsto k-i-1$, then T is called a *symmetric* interval exchange transformation. The orbit of a given point $\rho \in J$ is the infinite sequence ρ, $T(\rho)$, $T^2(\rho)$, $T^3(\rho)$, ... It can be coded by an infinite word $\mathbf{u}_\rho = u_0u_1u_2\ldots$ over the alphabet $\{0, 1, \ldots, k-1\}$ as follows:

$$u_n = X \quad \text{if } T^n(\rho) \in J_X \qquad \text{for } X \in \{0, 1, \ldots, k-1\}.$$

If for every $n \in \mathbb{N}$ we have $\mathcal{C}_{\mathbf{u}_\rho}(n) = (k-1)n + 1$, then the transformation T and the word $\mathbf{u}_\rho$ are said to be *non-degenerate*.

The class of words coding symmetric interval exchange transformations are well-studied (see [31, 32, 30]). In particular, they have zero palindromic defect and their language is closed under reversal. It follows that they contain infinitely many palindromic factors. In [II], besides giving an affirmative answer to the Class P conjecture in the case of 3 intervals, we enlarge the knowledge of languages of words coding symmetric 3 interval exchange transformation by giving a detailed description of the return times to a subinterval and the corresponding itineraries. This description is then refined with respect to the reversal symmetry present in the language. The partial solution to the Class P conjecture is based on these results and known results on a relation with Sturmian words. The exploited relation with Sturmian words is also the reason why this result is only for the ternary case, not for a general coding of k-interval exchange transformation.

Zero defect conjecture ([III])

The last conjecture is the following.

Conjecture 3 (Zero defect conjecture [15])**.** *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 5) having their language closed under reversal. We have $D(\mathbf{f}) = 0$ (in fact, it is true for all Sturmian words, as already mentioned). On the other hand, $D(\mathbf{t}) = +\infty$.

In 2012, Bucci and Vaslet [22] found a counterexample to this conjecture on a ternary alphabet. They showed that the fixed point of the primitive morphism determined by

$$a \mapsto aabcacba, b \mapsto aa, c \mapsto a$$

has finite positive palindromic defect and is not periodic. In [8], the author also gives a counterexample. 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. For instance, the mentioned counterexample of [22] is not injective. Indeed, in [III] 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 fact, any non-trivial morphism on binary alphabet is marked.

In [III] we show the following theorem:

Theorem 8. *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 there exists a conjugate of φ distinct from φ itself, then $D(\mathbf{u}) = 0$.*

The proof is based on the characterization 6 of Theorem 2. The characterization is expressed using the notion of an *extension graph* of a factor, which is used to describe possible extensions of factors in a language. This notion is also used in [12] to study languages having some specific properties of extension graphs of its elements. In [11], this notion is further studied and the authors make a connection to words with finite palindromic defect.

Let us comment also on the assumptions of the last theorem. 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. The proof in [III] is not applicable to this case and probably it is not easily extendable to it.

In [III], the special case of binary alphabet is also solved. In this case, the requirements on the morphism may be dropped:

Theorem 9. *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.*

1.3.2 Construction of words with finite palindromic defect ([VI, V, IV])

Many examples and properties are known for words with zero finite palindromic defect, i.e., rich or full words. Besides the articles mentioned in connection with Theorem 1 ([27, 34, 20, 6]), let us mention some other relevant results:

- In [21], the authors give another characterization of words with zero palindromic defect.
- In [19], 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 [26].
- The number of all words with zero palindromic defect of a given length and other properties are investigated in [80, 35].
- As already mentioned, words coding symmetric interval exchange transformations have zero palindromic defect by [5] and Theorem 1 item 5.
- In [16], the authors show that words coding rotation on the unit circle with respect to partition consisting of two intervals have zero palindromic defect.
- In [66], the author show a connection of words having zero palindromic defect with Burrows–Wheeler transform.

The mentioned results give rise to many examples of words with zero palindromic defect. On the other hand, there are not many specific classes of words with finite but non-zero palindromic defect. In [7], the following relation of words with finite palindromic defect to words with zero palindromic is given:

Theorem 10. *If* $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ *is a uniformly recurrent infinite word such that* $D(\mathbf{u}) < +\infty$*, 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*

$$\mathbf{u} = \varphi(\mathbf{v}) \quad \textit{and} \quad D(\mathbf{v}) = 0.$$

This theorem gives insight on the relation of words with finite and zero palindromic defect, but it cannot be used directly to construct words of finite palindromic defect. Although the morphism φ is of a very special class, so-called class P_{ret} introduced in [7], the idea cannot be reversed as there exists a word with zero palindromic defect and a morphism of class P_{ret} such that the image of this word by the morphism has infinite palindromic defect [7, Proposition 5.7].

The general goal of the 3 articles of this section is to enlarge the family of known examples of words with finite palindromic defect and investigate their properties, and also to broaden the family of known words having finite generalized palindromic defect introduced in [60] and studied in [61] (see below for a definition).

In [IV], we focus on episturmian words, see [27]. A word is *episturmian* if its language is closed under reversal and for each n there is at most one right special factor of length n. Aperiodic binary episturmian words are exactly the Sturmian words. Episturmian words have all zero palindromic defect.

The first results of [IV] is that that the image of an episturmian word by a morphism of class P_{ret} has always finite palindromic defect. This result is used to show the second result:

Theorem 11. *Let* $\mathbf{u}$ *be an episturmian word over a ternary alphabet* $\mathcal{A}$*. Let* $\mathcal{B}$ *be an alphabet and* $\pi : \mathcal{A}^* \to \mathcal{B}^*$ *be a letter-to-letter morphism. We have* $D(\pi(\mathbf{u})) = 0$*.*

In other words, any letter-to-letter image of a ternary episturmian word has zero palindromic defect.

To illustrate this theorem, let us consider the so-called Tribonacci word which is the fixed point of the morphism $\varphi : 0 \mapsto 01, 1 \mapsto 02, 2 \mapsto 0$. For $\mathbf{u} = \varphi(\mathbf{u})$ we have

$$\mathbf{u} = 0102010010201010201001\ldots$$

Let $\pi : \{0, 1, 2\} \to \{0, 1\}$ be a letter-to-letter morphism determined by $\pi(0) = 0, \pi(1) = 1$ and $\pi(2) = 1$. The last theorem then implies that the word

$$\pi(\mathbf{u}) = 0101010010101010101001\ldots$$

has zero palindromic defect.

Both main mentioned results of [IV] exploit many properties of episturmian words and reveal more of the ingenious structure they possess. Theorem 11 is shown only for ternary episturmian words; however, computational evidence suggests that we may drop the requirement on the size of the alphabet. Unfortunately, the small size of the alphabet is crucial in the presented proof in [IV] and it does not allow a simple extension to even 4-letter alphabet.

The article [IV] investigates also the generalized palindromic defect, as do the articles [VI, V]. We introduce here the notion for the case of binary alphabet $\{0,1\}$. Let $E : \{0,1\}^* \to \{0,1\}^*$ be an antimorphism, i.e., $\forall v, w \in \{0,1\}^*$ we have $E(vw) = E(w)E(v)$, that is given by $E(0) = 1$ and $E(1) = 0$. The mapping E exchanges the two letters and reverses the order of letters in a word as the mapping R does. For instance, we have $E(011) = E(1)E(1)E(0) = 001$. This mapping can be viewed as a generalization of the concept of the mirror mapping R. Its fixed points are called *pseudopalindromes*, *antipalindromes*, or *E-palindromes* (studied for instance in [15, 35]). The E-palindromic defect measuring the number of missing E-palindromes can be defined analogously to the classical (R-)palindromic defect, see [72].

A further generalization of these notions is given in [60] where the so-called *H-palindromic defect* is introduced with $H = \{E, R, ER, \mathrm{Id}\}$. This value considers classical palindromes and E-palindromes at the same time and measures the difference of their actual count and their maximal count. To count the palindromes and E-palindromes at the same time we consider, given a word f, the set $[f] = \{\mu(f) \colon \mu \in H\}$ and for a factor w we count the following number: $\tau(w) = \#\{[f] \colon f \text{ factor of } w, f = R(f) \text{ or } f = E(f)\}$. In other word, the number $\tau(w)$ counts the number of sets $[f]$ where f is a palindrome or an E-palindrome, thus counting both at the same time. The maximal count, for the case of H, is given by $\tau(w) \leq |w| + 1$. Finally, the H-palindromic defect is given by the difference of this upper bound $|w| + 1$ and $\tau(w)$. The basic property of H-palindromic defect is also its nonnegativity and its interpretation is analogous to classical palindromic defect: it measures how many palindromes and E-palindromes, counted at the same time using the classes $[f]$, are missing to attain the maximum. It is also extended to infinite words in a similar way.

The Thue–Morse word contains infinitely many palindromes and also infinitely many E-palindromes. Moreover, it has infinite palindromic defect and E-palindromic defect, and its H-palindromic defect is zero. To demonstrate it, let us take the prefix $w = 0110100110$ of the Thue–Morse word. The count of palindromes and E-palindromes is

$$\tau(w) = \#\Big\{[\varepsilon], [0], [11], [01], [101], [0110], [1010], [0011], [110100], [100110], [01101001]\Big\} = 11 = |w| + 1.$$

In [V] we show that the mapping S defined on $\{0,1\}^*$ by $S(u_0u_1u_2\ldots) = v_1v_2v_3\ldots$ with $v_i = u_{i-1} + u_i \bmod 2$ can be used to construct new classes of words with zero or finite palindromic defect and words with zero or finite H-palindromic defect.s The article is inspired by the results of [69] stating that a word $\mathbf{u}$ is a complementary-symmetric Rote word if and only if $\mathcal{S}(\mathbf{u})$ is a Sturmian word, and by the results of [16] claiming that every complementary-symmetric Rote word has zero palindromic defect. A complementary-symmetric Rote word may be defined as a coding of an irrational rotation of the unit circle with respect to the partition into two intervals of same length.

Namely, we show the following theorems.

Theorem 12. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ be an infinite word having its language closed under all elements of $H = \{\mathrm{Id}, E, R, ER\}$. The word $\mathbf{u}$ has zero H-palindromic defect (resp. finite H-palindromic defect) if and only if $\mathcal{S}(\mathbf{u})$ has zero palindromic defect (resp. finite palindromic defect).*

A corollary of the last theorem is that complementary-symmetric Rote words have zero H-palindromic defect since their languages are indeed closed under all elements of H. As already mentioned, they have also zero palindromic defect. Thus, they have finite H-palindromic and palindromic defect at the same time.

Theorem 13. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ be a uniformly recurrent word. If $\mathbf{u}$ has finite palindromic defect, then the word $\mathcal{S}(\mathbf{u})$ has finite palindromic defect.*

The last theorem gives a procedure constructing possibly many words with finite palindromic defect. For example, if $\mathbf{u}$ is a Sturmian word, then $\mathcal{S}^k(\mathbf{u})$ has finite palindromic defect for all $k > 0$. In this case, computer experiments suggest that $\mathcal{S}^k(\mathbf{u})$ is not Sturmian for all k and its palindromic defect is not zero.

In [IV], we combine Theorems 11 and 12 in order to exhibit another new class of words with zero H-palindromic defect. Namely, we show that applying an operation inverse to $\mathcal{S}$ to a non-trivial letter-to-letter projection of a ternary Arnoux–Rauzy word, we obtain a new class of words with zero H-palindromic defect.

In [VI] we investigate infinite words which are constructed using iterated palindromic closure operators with respect to R and E. Such words are in general called *generalized pseudostandard words*. The operator constructs successively palindromic and E-palindromic prefixes of an infinite word. We focus on so-called generalized Thue–Morse words which are a generalization of the Thue–Morse word $\mathbf{t}$ for larger alphabets: given two integers b and m such that $b > 1$ and $m > 1$, the generalized Thue–Morse word $\mathbf{t}_{b,m}$ is defined over $\{0, \ldots, m-1\}$ by $\mathbf{t}_{b,m} = \big(s_b(n) \bmod m\big)_{n=0}^{\infty}$, where the number $s_b(n)$ denotes the digit sum of the expansion of the number n in the base b. The classical Thue–Morse word $\mathbf{t}$ equals $\mathbf{t}_{2,2}$. These words are studied for instance in [79, 73, 2].

The article [VI] is mainly motivated by [25] where the authors show that the Thue–Morse word is a generalized pseudostandard word. Its construction using this method is governed by two directive sequences 01^ω and $(RE)^\omega$, where the superscript ω denotes infinite repetition of the word, i.e., $01^\omega = 01111\ldots$. The first prefix is constructed by taking the first element of the first sequence, i.e, 0, and creating its R-palindromic closure, i.e., finding the shortest R-palindrome such that 0 is its prefix. The R-palindromic closure of a word w is denoted w^R. We have the first prefix to be $f_1 = 0^R = 0$. To obtain the next prefix, we take the second element of the first directive sequence, append it to f_1 and find its E-palindromic closure since E is the second element of the second directive sequence. Thus we have $f_2 = (f_1 1)^E = (01)^E = 01$. We continue this process to obtain more prefixes of the Thue–Morse word, taking again at step i the ith elements of both directive sequences:

$$\begin{aligned}
f_3 &= (f_2 1)^R = (011)^R = 0110,\\
f_4 &= (f_3 1)^E = (01101)^E = 01101001,\\
f_5 &= (f_4 1)^R = (011010011)^R = 0110100110010110.
\end{aligned}$$

In [VI], we investigate the generalized Thue–Morse words and find a characterization of those that can be constructed using the same procedure.

Theorem 14. *The generalized Thue–Morse word $\mathbf{t}_{b,m}$ is a generalized pseudostandard word if and only if* $b \le m$ *or* $b - 1 = 0 \pmod m$.

In [73] is is shown that all the generalized Thue–Morse words have zero $I_2(m)$-palindromic defect for some group $I_2(m)$ in the sense of the mentioned H-palindromic richness (see [60, 61] for an exact definition of this notion).

The group $I_2(m)$ is isomorphic to the dihedral group of order $2m$. A section of [V] is inspired by this fact and investigates the generalized palindromic defect of the words $\mathcal{S}^k(\mathbf{t}_{b,m})$ where the mapping $\mathcal{S}$ is generalized to the alphabet $\{0, 1, \ldots, m-1\}$ as follows: for every word $w = w_0 \cdots w_n$ with $w_i \in \{0, 1, \ldots, m-1\}$ we set

$$\mathcal{S}(w_0 w_1 \cdots w_n) = v_1 \cdots v_n,$$

where $v_i = (w_{i-1} + w_i) \bmod m$ for every $i \in \{1, \ldots, n\}$.

The main result of the last part of [V] is the following theorem.

Theorem 15. *Let $m, b \in \mathbb{Z}$ such that $m \geq 3$ and $b \geq 3$. There exists a group $I_2'(m)$ with $I_2'(m) = I_2(m)$ if m is odd, and $I_2'(m)$ being isomorphic to $I_2(\frac{m}{2})$ if m is even, such that*

1. *the word $\mathcal{S}(\mathbf{t}_{b,m})$ has finite $I_2'(m)$-palindromic defect;*
2. *if m or b is odd, the word $\mathcal{S}(\mathbf{t}_{b,m})$ has zero $I_2'(m)$-palindromic defect.*

1.3.3 D0L-systems and algorithms ([VII, VIII])

The name "D0L-system" stands for deterministic context-independent Lindenmayer system. It is defined as a triple $G = (\mathcal{A}, \varphi, w)$ with $\mathcal{A}$ an alphabet, φ an endomorphism of $\mathcal{A}^*$, and $w \in \mathcal{A}^*$. The word w is the *axiom* of G. The language of G is the set $\{\varphi^i(w) : i \leq 0\}$. The set of all factors of the elements of the language of G is denoted $\mathcal{F}(G)$. See more in [70, 71] on D0L-systems and more general concept of L-systems which were introduced by Aristid Lindenmayer to model the plant growth.

Our motivation to study this notion stems from the fact that it generalizes the notion of the language of a fixed point of a morphism. Indeed, by taking the axiom to be the first letter of a fixed point of morphism we can recover the language of the fixed point by considering the set $\mathcal{F}(G)$ of the corresponding D0L-system G.

For instance, consider the morphism φ_{TM} fixing the Thue–Morse word $\mathbf{t}$ (see Example 5) and set $G_w = (\{0,1\}, \varphi_{TM}, w)$ for $w \in \{0,1\}^*$. We have $\mathcal{F}(G_0) = \mathcal{L}(\mathbf{t})$. As φ_{TM} is primitive, all its fixed points have the same language, thus $\mathcal{F}(G_1) = \mathcal{L}(\mathbf{t})$. In general, if $w \in \mathcal{L}(\mathbf{t})$, we have $\mathcal{F}(G_w) = \mathcal{L}(\mathbf{t})$.

Further motivation for research in this domain comes from many open questions concerning efficient[3] analysis of a given D0L-system. Let us list 3 of these open questions concerning a given D0L-system G:

[3]By efficient analysis we mean an algorithm that in general does not need to enumerate factors of some length as this is feasible only for small lengths even if we use for instance bounds on linear recurrence constant for primitive morphisms.

> **Question 1**: Find an efficient algorithm to generate all elements of $\mathcal{F}(G)$ of given length (and thus calculate the factor complexity of the generated language).
>
> **Question 2**: Detect symmetries in $\mathcal{F}(G)$ (by a symmetry we mean for instance the closedness of $\mathcal{F}(G)$ under some involutive antimorphism).
>
> **Question 3**: Test efficiently whether the morphism of G is injective on $\mathcal{F}(G)$, which is the following condition: $\forall w, v \in \mathcal{F}(G)$ we have $w \neq v \Rightarrow \varphi(w) \neq \varphi(v)$.

To complete the overview, let us also give an example of an answered question. The problem whether $\mathcal{F}(G)$ is periodic and ultimately periodic is treated and solved in [36, 59, 47] and [75] the binary case is fully described. The question of possible order of magnitude of factor complexities of $\mathcal{F}(G)$ is also solved, see [58].

It seems that a helpful property of a D0L-system G that would be needed to answer at least Question 1 is the *synchronizing delay*. Simply put, knowing the synchronizing delay of a system, one can find preimages of elements that are longer than certain bound. Moreover, these preimages are unambiguous except for some prefix and suffix of bounded length of the word in question. Knowing the unique preimages gives much insight into the language. D0L-systems admitting a finite synchronizing delay are called *circular*. The formal definition follows.

Definition 16. Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system, $\varphi(a) \neq \varepsilon$ for all $a \in \mathcal{A}$ and $u \in \mathcal{F}(G)$.

A triplet (p, v, s) where $p, s \in \mathcal{A}^*$ and $v = v_1 \cdots v_n \in \mathcal{F}(G)$ with $n > 0$ is an *interpretation of the word* u if $\varphi(v) = pus$.

Let (p, v, s) and (p', v', s') be two interpretations of a non-empty word $u \in \mathcal{F}(G)$ with $v = v_1 \cdots v_n \in \mathcal{A}^n$, $v' = v'_1 \cdots v'_m \in \mathcal{A}^m$ and $u = u_1 \cdots u_\ell \in \mathcal{A}^\ell$.

We say that G is *circular* with *synchronization delay* $D > 0$ if whenever we have

$$|\varphi(v_1 \cdots v_i)| - |p| > D \quad \text{and} \quad |\varphi(v_{i+1} \cdots v_n)| - |s| > D$$

for some i such that $1 \leq i \leq n$, then there exists j such that $1 \leq j \leq m$ and

$$|\varphi(v_1 \cdots v_{i-1})| - |p| = |\varphi(v'_1 \cdots v'_{j-1})| - |p'|$$

and $v_i = v'_j$.

See Section 3 of [VIII] for an overview concerning the notion of synchronizing delay and circularity. This notion is tightly connected to *recognizability*, see [56], where the author also shows that a primitive morphism is circular.

If we restrict ourselves to circular morphisms, then using [42] we may design an efficient algorithm to answer Question 1 of enumerating all factors. However, to do this, the synchronizing delay is still required.

The task to efficiently determine the synchronizing delay is also an unsolved question. It is known for binary alphabets and uniform morphisms, see [43] and the references therein. For larger alphabets and primitive morphisms, the recent result [28] follows the work [56] and unveils an explicit upped bound. However, this upper bound is far from optimal and it cannot be used in a practical computation.

Our works [VII, VIII] aim towards this partial goal in the analysis of $\mathcal{F}(G)$ for a given D0L-system G.

In [VII] we show the following theorem. A non-empty word w is *primitive* if it is not a non-trivial (integer) power of another word, i.e., $w = v^k$ implies $k = 1$.

Theorem 17. *Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. There exists a constant M, a finite set of primitive words P such that $v^k \in \mathcal{F}(G)$ for every $k \geq M$ and $v \in P$.*

Moreover, using the known results on periodicity of D0L-systems of [47], we design a simple algorithm finding all primitive words v such that $v^k \in \mathcal{F}(G)$ for all k. The algorithm may be outlined as follows. A letter $a \in \mathcal{A}$ is *unbounded* or *growing* if $|\varphi^k(a)| \to +\infty$ as $k \to +\infty$.

1. Construct an injective simplification of φ (injective simplification is an injective morphism that may be used to generate the same language, see [29]).

2. Find all unbounded letters (done by inspecting the incidence matrix of the morphism).

3. Construct two finite directed labeled graphs describe relation between unbounded letters (done by investigating occurrences of unbounded letters in letter images of unbounded letters).

4. Loop over all cycles in the constructed graphs and look for specific labels that indicate one kind of primitive word we are looking for.

5. Test all growing letters if they are first letters of purely periodic fixed points of the morphism (using [47]). If a purely periodic fixed point

is found, then the second kind of primitive word we are looking for is found.

6. If the original morphism was simplified, an inverse operation needs to be applied to the list of discovered primitive words.

7. Output the list of all found primitive words and their conjugates.

In [VIII], we give the following characterization of circularity of a D0L-system. We say that a system $G = (\mathcal{A}, \varphi, w)$ is *unboundedly repetitive* if there exists $v \in \mathcal{F}(G)$ such that $v^k \in \mathcal{F}(G)$ for all k and v contains at least one unbounded letter.

Theorem 18. *An D0L-system $G = (\mathcal{A}, \varphi, w)$ injective on $\mathcal{F}(G)$ is not circular if and only if it is unboundedly repetitive.*

To obtain this result we prove that every non-circular D0L-system contains arbitrarily long words of the form v^k, using the results of [VII]. The last theorem is important as it may be used to design an efficient algorithm to test circularity, provided that the tested system is injective on its language. An efficient test of injectivity on a language is still an open question (Question 3), thus for the moment, the test may be done using a stronger condition of general injectivity that is easy to test.

Both algorithms presented in this section are efficient and present a step forward in solving Question 1 which, at least by algorithms enumerating factors, may then be used to solve the other two questions and other possible similar challenges.

1.3.4 Rauzy gasket and generalized Markov constant ([IX, X])

This section contains a short overview of two articles included in this thesis. Each article deals with a problem with no direct relation to the above problems. They may be perceived as witnesses of the mentioned strong connection of Combinatorics on Words to other research domains.

In [IX] we deal with ternary episturmian words (see the definition in Section 1.3.2). We study the set of all triples of letters frequencies of all episturmian words. Given an infinite word $\mathbf{u}$, the frequency of a letter a is given by the limit

$$\lim_{n\to+\infty} \frac{\text{number of occurrences of the letter } a \text{ in the prefix of } \mathbf{u} \text{ of length } n}{n},$$

if it exists. The letter frequencies of episturmian words exist.

Let us give an example, the Tribonacci word already mentioned above is a ternary episturmian word. It is the fixed point of the morphism $\varphi : 0 \mapsto 01, 1 \mapsto 02, 2 \mapsto 0$. For $\mathbf{u} = \varphi(\mathbf{u})$ we have

$$\mathbf{u} = 0102010010201010201001$$

and the frequencies of 0, 1, and 2 are ζ, ζ^2, and ζ^3 respectively with $\zeta + \zeta^2 + \zeta^3 = 1$, see [64].

The question to study the set of all letter frequencies of all ternary episturmian words is motivated by the study of systems generating such words in [3, 65]. The set in question is called "the Rauzy gasket" to honour Gérard Rauzy and due to its nature (namely its relation to the Sierpínski and Appollonian gaskets). Thus, a point of the Rauzy gasket is the point $(\zeta, \zeta^2, \zeta^3)$ of the letter frequencies of the Tribonacci word. Figure 1 shows an approximate shape of the Rauzy gasket.

The main result of [IX] is a proof of the fact that the Rauzy gasket is of Lebesgue measure 0 (in $\mathbb{R}^2$). The presented proof is done by finding a relation to known results concerning the so-called Fully subtractive algorithm, see [53, 44, 13].

In [4], the authors continue the study of the Rauzy gasket by showing that its Hausdorff dimension is less than 2.

In [X] we investigate the set

$$\mathcal{S}(\alpha) = \text{set of all accumulation points of } \left\{ m^2 \left(\frac{k}{m} - \alpha \right) : k, m \in \mathbb{Z} \right\}$$

where $\alpha \in \mathbb{R}$. The motivation stems from spectral properties of a differential operator connected to properties of metamaterials. From another point of view, the value $\inf |\mathcal{S}(\alpha)|$ is well-known in the literature and sometimes it is called *the Markov constant of* α. See for instance [24].

The notion of Markov constant is connected to a famous theorem of Hurwitz: for every irrational number ξ there exist infinitely many relatively prime integers m and n such that

$$\left| \frac{m}{n} - \xi \right| < \frac{1}{\sqrt{5} n^2}$$

and the bound $\frac{1}{\sqrt{5}}$ is optimal. The optimality is for ξ equal to the golden ratio (and other numbers having the tail of their continued fraction expansion equals to an infinite sequence of 1s). Thus, the Markov constant of the golden ratio $\frac{1+\sqrt{5}}{2}$ is $\frac{1}{\sqrt{5}}$ which is the maximum value of a Markov constant of any number.

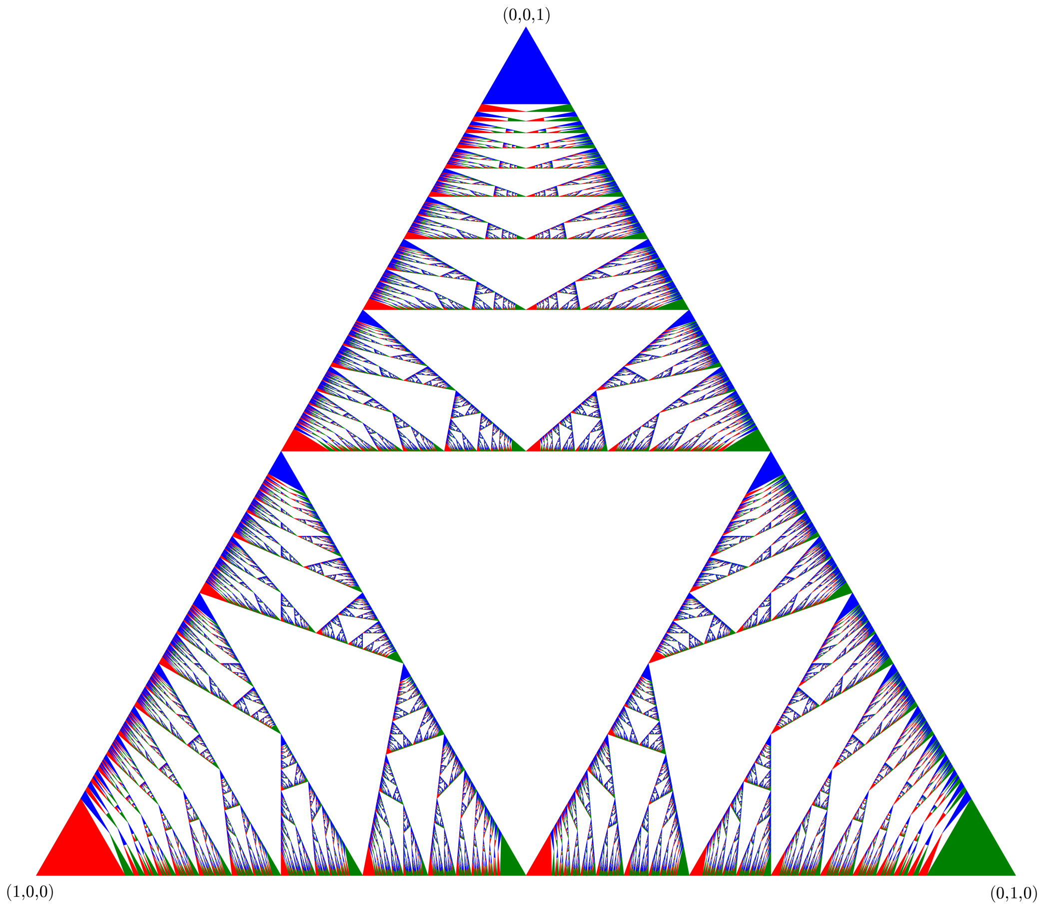

Figure 1: An approximation of the Rauzy gasket (a superset of the Rauzy gasket). The Rauzy gasket is a set contained in the convex set $\{(x,y,z) \in \mathbb{R}^3 \colon x+y+z=1, x,y,z \geq 0\}$. The plane containing this set is depicted.

In [X] we use methods and results from Number Theory and Combinatorics on Words in order to describe the set $\mathcal{S}(\alpha)$. Besides some basic properties of the set $\mathcal{S}(\alpha)$ such as the closedness under multiplication by z^2 for every $z \in \mathbb{Z}$ and its connection to known results on Markov constants and continued fractions, we exhibit different behaviour of the set for different parameter α. We show that there exist irrational numbers α_1, α_2 and α_3 such that

1. $\mathcal{S}(\alpha_1) = \mathbb{R}$;
2. $\mathcal{S}(\alpha_2) = (-\infty, -\varepsilon] \cup [\varepsilon, +\infty)$, where $\varepsilon = \frac{\sqrt{2}}{8} \approx 0.18$;
3. the Hausdorff dimension of $\mathcal{S}(\alpha_3) \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$ is positive but less than 1.

1.4 Contribution of the author

Since this thesis contains also co-authored articles, we briefly comment on the author's contribution. For all the co-authored articles, it is difficult to determine exactly the author's total contribution. The creation process of these articles consisted of a collective work towards the solution of the given question, either in person or by the means of electronic communication, followed by a collective creation of the article itself. The following specific contributions may be separated:

- In [I, II, V, VI], the author is responsible for doing all the supporting computer experiments.
- In [X], the author's contribution is in Sections 2–5.

References

[1] *PARI/GP, version 2.8.0*, Bordeaux, 2016. `http://pari.math.u-bordeaux.fr`.

[2] J.-P. Allouche and J. Shallit, *Sums of digits, overlaps, and palindromes*, Discrete Math. Theoret. Comput. Sci., 4 (2000), pp. 1–10.

[3] P. Arnoux and G. Rauzy, *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. Math. France, 119 (1991), pp. 199–215.

[4] A. Avila, P. Hubert, and A. Skripchenko, *On the Hausdorff dimension of the Rauzy gasket*, Bull. Soc. Math. France, 144 (2016), pp. 539–568.

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

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

[7] ——, *Infinite words with finite defect*, Adv. in Appl. Math., 47 (2011), pp. 562–574.

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

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

[10] J. Berstel and D. Perrin, *The origins of combinatorics on words*, Eur. J. Combin., 28 (2007), pp. 996–1022.

[11] V. Berthé, C. De Felice, V. Delecroix, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer, and G. Rindone, *Specular sets*, in Combinatorics on Words: 10th International Conference, WORDS 2015, Kiel, Germany, September 14–17, 2015, Proceedings, F. Manea and D. Nowotka, eds., Springer International Publishing, 2015, pp. 210–222.

[12] V. Berthé, C. De Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer, and G. Rindone, *Acyclic, connected and tree sets*, Monatsh. Math., 176 (2015), pp. 521–550.

[13] V. Berthé, Domenjoud, D. Jamet, and X. Provençal, *Fully subtractive algorithm, tribonacci numeration and connectedness of discrete planes*, RIMS Kokyuroku Bessatsu, 46 (2014), pp. 159–174.

[14] V. Berthé and M. Rigo, eds., *Combinatorics, automata, and number theory.*, vol. 135 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2010.

[15] 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.

[16] A. Blondin Massé, S. Brlek, S. Labbé, and L. Vuillon, *Palindromic complexity of codings of rotations*, Theoret. Comput. Sci., 412 (2011), pp. 6455–6463.

[17] 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.

[18] S. Brlek and C. Reutenauer, *Complexity and palindromic defect of infinite words*, Theoret. Comput. Sci., 412 (2011), pp. 493–497.

[19] 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.

[20] 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 (2009), pp. 60–74.

[21] ——, *A new characteristic property of rich words*, Theoret. Comput. Sci., 410 (2009), pp. 2860–2863.

[22] M. Bucci and E. Vaslet, *Palindromic defect of pure morphic aperiodic words*, in Proceedings of the 14th Mons Days of Theoretical Computer Science, 2012.

[23] J. Cassaigne, *Complexity and special factors*, Bull. Belg. Math. Soc. Simon Stevin 4, 1 (1997), pp. 67–88.

[24] T. Cusick and M. Flahive, *The Markoff and Lagrange Spectra*, Mathematical surveys and monographs, American Mathematical Society, 1989.

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

[26] A. de Luca, A. Glen, and L. Q. Zamboni, *Rich, sturmian, and trapezoidal words*, Theoret. Comput. Sci., 407 (2008), pp. 569–573.

[27] 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.

[28] F. Durand and J. Leroy, *The constant of recognizability is computable for primitive morphisms*, arXiv:1610.05577 [cs, math], (2016).

[29] A. Ehrenfeucht and G. Rozenberg, *Elementary homomorphisms and a solution of the d0l sequence equivalence problem*, Theoret. Comput. Sci., 7 (1978), pp. 169–183.

[30] S. Ferenczi, C. Holton, and L. Zamboni, *Structure of three-interval exchange transformations II: a combinatorial description of the tranjectories*, J. Anal. Math., 89 (2003), pp. 239–276.

[31] S. Ferenczi and L. Q. Zamboni, *Languages of k-interval exchange transformations*, Bull. Lond. Math. Soc., 40 (2008), pp. 705–714.

[32] S. Ferenczi and L. Q. Zamboni, *Structure of k-interval exchange transformations: Induction, trajectories, and distance theorems*, J. Anal. Math., 112 (2010), pp. 289–328.

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

[34] A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni, *Palindromic richness*, Eur. J. Combin., 30 (2009), pp. 510–531.

[35] C. Guo, J. Shallit, and A. M. Shur, *Palindromic rich words and run-length encodings*, Inform. Process. Lett., 116 (2016), pp. 735–738.

[36] T. Harju and M. Linna, *On the periodicity of morphisms on free monoids*, RAIRO-Theoret. Inf. Appl., 20 (1986), pp. 47–54.

[37] 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.

[38] W. Hart, F. Johansson, and S. Pancratz, *FLINT: Fast Library for Number Theory*, 2015. Version 2.5, `http://flintlib.org`.

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

[40] C. R. Jean Berstel, Dominique Perrin, *Codes and Automata*, Encyclopedia of Mathematics and its Applications 129, Cambridge University Press, 1 ed., 2009.

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

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

[43] K. Klouda and K. Medková, *Synchronizing delay for binary uniform morphisms*, Theoret. Comput. Sci., 615 (2016), pp. 327–334.

[44] C. Kraaikamp and R. Meester, *Ergodic properties of a dynamical system arising from percolation theory*, Ergod. Theor. Dyn. Syst., 15 (1995), pp. 653–661.

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

[46] S. Labbé and E. Pelantová, *Palindromic sequences generated from marked morphisms*, Eur. J. Combin., 51 (2016), pp. 200–214.

[47] B. Lando, *Periodicity and ultimate periodicity of D0L systems*, Theoret. Comput. Sci., 82 (1991), pp. 19–33.

[48] P. Ligeti, *Combinatorics on words and its applications*, PhD thesis, Eötvös Loránd University, 2010.

[49] 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.

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

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

[52] 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.

[53] R. W. J. Meester and T. Nowicki, *Infinite clusters and critical values in two dimensional circle percolation*, Isr. J. Math., 68 (1989), pp. 63–81.

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

[55] M. Morse and G. A. Hedlund, *Symbolic dynamics II - Sturmian trajectories*, Amer. J. Math., 62 (1940), pp. 1–42.

[56] B. Mossé., *Reconnaissabilité des substitutions et complexité des suites automatiques*, Bull. Soc. Math. Fr., 124 (1996), pp. 329–346.

[57] J. Néraud, *Ten Conferences WORDS and their contribution to the field of Combinatorics on Words*, EATCS Bull., 118 (2016), pp. 145–155.

[58] J.-J. Pansiot, *Complexité des facteurs des mots infinis engendrés par morphismes itérés*, in 11th ICALP, Antwerpen, J. Paredaens, ed., vol. 172 of LNCS, Springer, Jul 1984, pp. 380–389.

[59] ——, *Decidability of periodicity for infinite words*, RAIRO-Theoret. Inf. Appl., 20 (1986), pp. 43–46.

[60] E. Pelantová and Š. Starosta, *Languages invariant under more symmetries: overlapping factors versus palindromic richness*, Discrete Math., 313 (2013), pp. 2432–2445.

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

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

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

[64] G. Rauzy, *Nombres algébriques et substitutions*, Bull. Soc. Math. France, 110 (1982), pp. 147–178.

[65] ——, *Suites à termes dans un alphabet fini*, Séminaire de Théorie des Nombres de Bordeaux, Anné 1982–1983 (1983).

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

[67] M. Rigo, *Formal Languages, Automata and Numeration Systems 1*, Wiley, 2014.

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

[69] G. Rote, *Sequences with subword complexity* $2n$, J. Number Th., 46 (1993), pp. 196–213.

[70] G. Rozenberg and A. Salomaa, *The mathematical theory of L systems*, Academic Press, 1980.

[71] G. Rozenberg and A. Salomaa, *The book of L*, Springer, 1986.

[72] Š. Starosta, *On theta-palindromic richness*, Theoret. Comput. Sci., 412 (2011), pp. 1111–1121.

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

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

[75] P. Séébold, *An effective solution to the DOL periodicity problem in the binary case*, EATCS Bull., 36 (1988), pp. 137—151.

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

[77] The Sage Developers, *SageMath, the Sage Mathematics Software System*, 2016. http://www.sagemath.org.

[78] A. Thue, *Über unendliche Zeichenreihen*, vol. 7 of Norske Vid. Selsk. Skr., I. Mat. Nat. Kl., Christiana, 1906.

[79] J. Tromp and J. Shallit, *Subword complexity of a generalized Thue-Morse word*, Inform. Process. Lett., (1995), pp. 313–316.

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

Chapter 2

Proof of the Brlek–Reutenauer conjecture

Authors:	L. Balková, E. Pelantová, and Š. Starosta
Publication reference:	Theor. Comput. Sci., 475 (2013), pp. 120–125
Cited as:	[I]

Theoretical Computer Science 475 (2013) 120–125

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Note

Proof of the Brlek–Reutenauer conjecture

L'. Balková [a,*], E. Pelantová [a], Š. Starosta [b]

[a] Department of Mathematics, FNSPE, Czech Technical University in Prague, Trojanova 13, 120 00 Praha 2, Czech Republic

[b] Department of Applied Mathematics, FIT, Czech Technical University in Prague, Thákurova 9, 160 00 Praha 6, Czech Republic

ARTICLE INFO

Article history:
Received 13 August 2012
Accepted 26 December 2012
Communicated by D. Perrin

ABSTRACT

Brlek and Reutenauer conjectured that any infinite word $\mathbf{u}$ with language closed under reversal satisfies the equality $2D(\mathbf{u}) = \sum_{n=0}^{+\infty} T_{\mathbf{u}}(n)$ in which $D(\mathbf{u})$ denotes the defect of $\mathbf{u}$ and $T_{\mathbf{u}}(n)$ denotes $\mathcal{C}_{\mathbf{u}}(n+1)-\mathcal{C}_{\mathbf{u}}(n)+2-\mathcal{P}_{\mathbf{u}}(n+1)-\mathcal{P}_{\mathbf{u}}(n)$, where $\mathcal{C}_{\mathbf{u}}$ and $\mathcal{P}_{\mathbf{u}}$ are the factor and palindromic complexity of $\mathbf{u}$, respectively. This conjecture was verified for periodic words by Brlek and Reutenauer themselves. Using their results for periodic words, we have recently proved the conjecture for uniformly recurrent words. In the present article we prove the conjecture in its general version by a new method without exploiting the result for periodic words.

1. Introduction

Brlek and Reutenauer conjectured in [6] a nice equality which combines together the factor complexity $\mathcal{C}_{\mathbf{u}}$, the palindromic complexity $\mathcal{P}_{\mathbf{u}}$, and the palindromic defect $D(\mathbf{u})$ of an infinite word $\mathbf{u}$. It sounds as follows.

Brlek–Reutenauer Conjecture. *If $\mathbf{u}$ is an infinite word with language closed under reversal, then*

$$2D(\mathbf{u}) = \sum_{n=0}^{+\infty} T_{\mathbf{u}}(n),$$

where $T_{\mathbf{u}}(n) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n+1) - \mathcal{P}_{\mathbf{u}}(n)$.

Brlek and Reutenauer proved ibidem that their conjecture holds for periodic infinite words. It is known from [7] that the Brlek–Reutenauer conjecture holds for words with zero defect. In [3], we proved the conjecture for uniformly recurrent words. In our proof, we constructed for any uniformly recurrent word $\mathbf{u}$ whose language is closed under reversal a periodic word $\mathbf{v}$ with language closed under reversal such that $D(\mathbf{u}) = D(\mathbf{v})$ and $T_{\mathbf{u}}(n) = T_{\mathbf{v}}(n)$ for any n. Then we used validity of the conjecture for periodic words.

In this paper, we will prove that the Brlek–Reutenauer conjecture holds in full generality without exploiting the result for periodic words. Since both sides of the equality in the Brlek–Reutenauer conjecture are non-negative, validity of the conjecture will be shown if we prove the following two theorems.

Theorem 1. *If $\mathbf{u}$ is an infinite word with language closed under reversal such that both $D(\mathbf{u})$ and $\sum_{n=0}^{+\infty} T_{\mathbf{u}}(n)$ are finite, then*

$$2D(\mathbf{u}) = \sum_{n=0}^{+\infty} T_{\mathbf{u}}(n). \tag{1}$$

* Corresponding author. Tel.: +420 224358544.
E-mail address: lubomira.balkova@gmail.com (L. Balková).

0304-3975/$ – see front matter
doi:10.1016/j.tcs.2012.12.024

Theorem 2. *If* $\mathbf{u}$ *is an infinite word with language closed under reversal, then*

$$D(\mathbf{u}) < +\infty \quad \text{if and only if} \quad \sum_{n=0}^{+\infty} T_{\mathbf{u}}(n) < +\infty.$$

In [3], which is devoted mainly to the uniformly recurrent words, we already stated in the section Open problems one part of Theorem 2, namely that $D(\mathbf{u}) < +\infty$ implies $\sum_{n=0}^{+\infty} T_{\mathbf{u}}(n) < +\infty$. As pointed out in [4], there is a gap in our proof, and its corrected version can be found in [2]. In order to make the present paper self-sustained so that the reader understands and checks all steps of the proof without having all previous papers at hand, we recall necessary notations and statements together with the proofs of the essential ones.

2. Preliminaries

By $\mathcal{A}$ we denote a finite set of symbols called *letters*; the set $\mathcal{A}$ is therefore called an *alphabet*. A finite string $w = w_0w_1\ldots w_{n-1}$ of letters from $\mathcal{A}$ is said to be a *finite word*, its length is denoted by $|w| = n$. Finite words over $\mathcal{A}$ together with the operation of concatenation and the empty word ϵ as the neutral element form a free monoid $\mathcal{A}^*$. The map

$$w = w_0w_1\ldots w_{n-1} \quad \mapsto \quad \overline{w} = w_{n-1}w_{n-2}\ldots w_0$$

is a bijection on $\mathcal{A}^*$, the word $\overline{w}$ is called the *reversal* or the *mirror image* of w. A word w which coincides with its mirror image is a *palindrome*.

Under an *infinite word* we understand an infinite string $\mathbf{u} = u_0u_1u_2\ldots$ of letters from $\mathcal{A}$. A finite word w is a *factor* of a word v (finite or infinite) if there exist words p and s such that $v = pws$. If $p = \epsilon$, then w is said to be a *prefix* of v, if $s = \epsilon$, then w is a *suffix* of v.

The *language* $\mathcal{L}(v)$ of a finite or an infinite word v is the set of all its factors. Factors of v of length n form the set denoted by $\mathcal{L}_n(v)$. We say that the language of an infinite word $\mathbf{u}$ is *closed under reversal* if $\mathcal{L}(\mathbf{u})$ contains with every factor w also its reversal $\overline{w}$.

For any factor $w \in \mathcal{L}(\mathbf{u})$, there exists an index i such that w is a prefix of the infinite word $u_iu_{i+1}u_{i+2}\ldots$. Such an index is called an *occurrence* of w in $\mathbf{u}$. 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 reversal, then $\mathbf{u}$ is recurrent (a proof can be found in [9]). 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.

If any factor $w \in \mathcal{L}(\mathbf{u})$ has only finitely many complete return words, then the infinite word $\mathbf{u}$ is called *uniformly recurrent*.

The *factor complexity* of an infinite word $\mathbf{u}$ is the map $\mathcal{C}_{\mathbf{u}} : \mathbb{N} \mapsto \mathbb{N}$ defined by the prescription $\mathcal{C}_{\mathbf{u}}(n) := \#\mathcal{L}_n(\mathbf{u})$. To determine the first difference of the factor complexity, one has to count the possible extensions of factors of length n. A *right extension* of $w \in \mathcal{L}(\mathbf{u})$ is a letter $a \in \mathcal{A}$ such that $wa \in \mathcal{L}(\mathbf{u})$. Of course, any factor of $\mathbf{u}$ has at least one right extension. A factor w is called *right special* if w has at least two right extensions. Similarly, one can define a *left extension* and a *left special* factor. We will deal mainly with recurrent infinite words $\mathbf{u}$. In such a case, any factor of $\mathbf{u}$ has at least one left extension.

In [8] it is shown that any finite word w contains at most $|w| + 1$ distinct palindromes (including the empty word). The *defect* $D(w)$ of a finite word w is the difference between the utmost number of palindromes $|w| + 1$ and the actual number of palindromes contained in w.

In accordance with the terminology introduced in [8], the factor with a unique occurrence in another factor is called *unioccurrent*.

The following corollary gives an insight into the birth of defects.

Corollary 3 *([8]). The defect* $D(w)$ *of a finite word* w *is equal to the number of prefixes* w' *of* w *for which the longest palindromic suffix of* w' *is not unioccurrent in* w'*. In other words, if* b *is a letter and* w *a finite word, then* $D(wb) = D(w) + \delta$*, where* $\delta = 0$ *if the longest palindromic suffix of* wb *occurs exactly once in* wb *and* $\delta = 1$ *otherwise.*

Corollary 3 implies that $D(v) \geq D(w)$ whenever w is a factor of v. It enables to give a reasonable definition of the defect of an infinite word (see [5]).

Definition 4. The defect of an infinite word $\mathbf{u}$ is the number (finite or infinite)

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

Let us point out two facts.

(1) If we consider all factors of a finite or an infinite word $\mathbf{u}$, we obtain the same defect, i.e.,

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

(2) Any infinite word with finite defect contains infinitely many palindromes.

Using Corollary 3 and Definition 4, we obtain immediately the following corollary.

Corollary 5. *Let* $\mathbf{u}$ *be an infinite word with language closed under reversal. The following statements are equivalent.*

(1) *The defect of* $\mathbf{u}$ *is finite.*
(2) *There exists an integer* H *such that the longest palindromic suffix of any prefix* w *of length* $|w| \geq H$ *occurs in* w *exactly once.*

For *the longest palindromic suffix* of a word w we will sometimes use the notation $lps(w)$.

The number of palindromes of a fixed length occurring in an infinite word is measured by the so called *palindromic complexity* $\mathcal{P}_{\mathbf{u}}$, the map which assigns to any non-negative integer n the number

$$\mathcal{P}_{\mathbf{u}}(n) := \#\{w \in \mathcal{L}_n(u) \colon w \text{ is a palindrome}\}.$$

Denote

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

The following proposition is proven in [1] for uniformly recurrent words; however, as also noted in [6], the uniform recurrence is not needed in the proof and it holds for any infinite word with language closed under reversal.

Proposition 6 (*[1]*). *If* $\mathbf{u}$ *is an infinite word with language closed under reversal, then*

$$T_{\mathbf{u}}(n) \geq 0 \quad \text{for all } n \in \mathbb{N}. \tag{2}$$

Let $\mathbf{u}$ be an infinite word with language closed under reversal. Using the proof of Proposition 6, those $n \in \mathbb{N}$ for which $T_{\mathbf{u}}(n) = 0$ can be characterized in the graph language. Before doing that we need to introduce some more notions.

An *n-simple path* e is a factor of $\mathbf{u}$ of length at least $n+1$ such that the only special (right or left) 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 starts in w and ends in v. We will denote by $G_n(\mathbf{u})$ an undirected graph whose set of vertices is formed by unordered pairs $(w, \overline{w})$ such that $w \in \mathcal{L}_n(\mathbf{u})$ is right or left special. We connect two vertices $(w, \overline{w})$ and $(v, \overline{v})$ by an unordered pair $(e, \overline{e})$ if e or $\overline{e}$ is an n-simple path starting in w or $\overline{w}$ and ending in v or $\overline{v}$. Note that the graph $G_n(\mathbf{u})$ may have multiple edges and loops.

Lemma 7. *If* $\mathbf{u}$ *is an infinite word with language closed under reversal and* $n \in \mathbb{N}$, *then* $T_{\mathbf{u}}(n) = 0$ *if and only if both of the following conditions are met.*

(1) *The graph obtained from* $G_n(\mathbf{u})$ *by removing loops is a tree.*
(2) *Any* n*-simple path forming a loop in the graph* $G_n(\mathbf{u})$ *is a palindrome.*

Proof. It is a direct consequence of the proof of Theorem 1.2 in [1] (recalled in this paper as Proposition 6). □

3. Proof of Theorem 1

The aim of this section is to prove Theorem 1, i.e., to prove the Brlek–Reutenauer conjecture under the additional assumption that the defect $D(\mathbf{u})$ of an infinite word $\mathbf{u}$ and the sum $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n)$ are finite. As observed in [6], it is easy to prove the "finite analogy" of the conjecture, which deals only with finite words. We will also make use of this result.

Theorem 8 (*[6]*). *For every finite word* w *we have*

$$2D(w) = \sum_{n=0}^{|w|} T_w(n),$$

where $T_w(n) = \mathcal{C}_w(n+1) - \mathcal{C}_w(n) + 2 - \mathcal{P}_w(n+1) - \mathcal{P}_w(n)$ *and the index* w *means that we consider only factors of* w.

It may seem that the Brlek–Reutenauer conjecture for an infinite word $\mathbf{u}$ can be obtained from Theorem 8 by a "limit transition". However, this transition would be far from being kosher. The following lemmas enable us to avoid the incorrectness.

Lemma 9. *Let* $\mathbf{u}$ *be an infinite word with language closed under reversal and finite defect. If* q *is its prefix satisfying* $D(\mathbf{u}) = D(q)$, *then for* $H = |q| + 1$ *one has*

$$\mathcal{C}_{\mathbf{u}}(H) - \mathcal{P}_{\mathbf{u}}(H) = 2\#\{x \in \mathcal{L}(\mathbf{u}) \colon x \text{ is a palindrome shorter than } H \text{ which is not contained in } q\}.$$

Proof. Let us define a mapping $f : S \to T$, where

$$S = \{x \in \mathcal{L}(\mathbf{u}) \colon x \notin \mathcal{L}(q), |x| < H, x = \overline{x}\}$$

and

$$T = \left\{\{w, \overline{w}\} \colon w \in \mathcal{L}_H(\mathbf{u}), w \neq \overline{w}\right\}.$$

Let x be a palindrome from S and i be the first occurrence of x in $\mathbf{u}$. Put $w = u_{i+|x|-H} \cdots u_{i+|x|-1}$. It means that w is a factor of $\mathbf{u}$ of length H and x is a suffix of w. Since $H > |x|$, the factor w is not a palindrome — otherwise it contradicts the fact that i is the first occurrence of the palindrome x. We put $f(x) = \{w, \overline{w}\}$.

To show that f is surjective, we consider $w \in \mathcal{L}_H(\mathbf{u})$ such that $w \neq \overline{w}$. Let p be the prefix of $\mathbf{u}$ which ends in the first occurrence of w or $\overline{w}$ in $\mathbf{u}$. Since $|p| \geq H = |w| > |q|$, we have according to Corollary 3 that $D(q) = D(p)$ and consequently, $lps(p)$ is unioccurrent in p, which implies that $lps(p)$ is not a factor of q. Moreover, $lps(p)$ is shorter than H — otherwise it contradicts the choice of the prefix p. We found $x = lps(p) \in S$ such that $f(x) = \{w, \overline{w}\}$, i.e., f is surjective.

To show that f is injective, we consider two palindromes $y, z \in S$ and we denote $f(y) = \{w_y, \overline{w_y}\}$ and $f(z) = \{w_z, \overline{w_z}\}$. From the definition of w_x we know that the palindrome x occurs as a factor of w_x exactly once, namely as its suffix. It means that x equals $lps(w_x)$. Let us suppose that $f(y) = f(z)$. We have to discuss two cases.

(1) Case $w_y = w_z$. It gives $lps(w_y) = lps(w_z)$ and thus $y = z$.
(2) Case $w_y = \overline{w_z}$. It implies that y is a prefix of w_z and z is a prefix of w_y. The fact that y is a prefix of w_z forces the first occurrence of w_y to be strictly smaller than the first occurrence of w_z. Simultaneously, since z is a prefix of w_y, the first occurrence of w_z is strictly smaller than the first occurrence of w_y - a contradiction.

Consequently, the assumption $f(y) = f(z)$ implies $z = y$ and the mapping f is injective as well.

Existence of the bijection f between the finite sets T and S means $\#T = \#S$. Since from the definition of T it follows that $\mathcal{C}_{\mathbf{u}}(H) - \mathcal{P}_{\mathbf{u}}(H) = 2\#T$, the equality stated in the lemma is proven. □

Remark 10. As it was pointed out by Bojan Bašić, Lemma 9 may be stated in a more general form for $H > |q|$, then the equality changes to

$$\mathcal{C}_{\mathbf{u}}(H) - \mathcal{P}_{\mathbf{u}}(H) = 2\#\{x \in \mathcal{L}(\mathbf{u}) : x \notin \mathcal{L}(q),\ |x| < H,\ x = \overline{x}\} - 2(H - |q| - 1).$$

Thanks to him, we added the assumption $H = |q| + 1$ in Lemma 9 necessary for the validity of the statement.

Lemma 11. *Let $\mathbf{u}$ be an infinite word with language closed under reversal and finite defect. If q is its prefix satisfying $D(\mathbf{u}) = D(q)$, then for any prefix p of $\mathbf{u}$ such that $|p| > |q|$ the number*

$$\#\{x \in \mathcal{L}(p) : x \text{ is a palindrome of length at most } |q| \text{ which is not contained in } q\} + \sum_{n=|q|+1}^{|p|} \mathcal{P}_p(n)$$

equals $|p| - |q|$.

Proof. At first we will show the equality

$$|p| - |q| = \#\{x \in \mathcal{L}(p) \setminus \mathcal{L}(q) : x = \overline{x}\}. \tag{3}$$

Let us denote by $u^{(i)}$ the prefix of $\mathbf{u}$ of length i. For any palindrome $x \in \mathcal{L}(p) \setminus \mathcal{L}(q)$ we find the minimal index i such that x occurs in $u^{(i)}$. Since $x \in \mathcal{L}(p) \setminus \mathcal{L}(q)$, we have $|q| < i \leq |p|$. Thus we map any element of $\{x \in \mathcal{L}(p) \setminus \mathcal{L}(q) : x = \overline{x}\}$ to an index $i \in \{|q| + 1, |q| + 2, \ldots, |p|\}$.

Let us look at the details of this mapping. The minimality of i guarantees that x is unioccurrent in $u^{(i)}$. Palindromicity of x gives that $x = lps(u^{(i)})$. It implies that no two different palindromes are mapped to the same index i, i.e., the mapping is injective.

Since $D(q) = D(\mathbf{u})$, according to Corollary 3, $lps(u^{(i)})$ is unioccurrent in $u^{(i)}$ and thus $lps(u^{(i)}) \notin \mathcal{L}(q)$. Thus any index i such that $|q| < i \leq |p|$ has its preimage $x = lps(u^{(i)})$. Therefore the mapping is a bijection and its domain and range have the same cardinality as stated in (3).

To finish the proof, we split elements of $\{x \in \mathcal{L}(p) \setminus \mathcal{L}(q) : x = \overline{x}\}$ into two disjoint parts: elements of length smaller than or equal to $|q|$ and elements of length greater than $|q|$. Since

$$\#\{x \in \mathcal{L}(p) \setminus \mathcal{L}(q) : x = \overline{x},\ |x| > |q|\} = \#\{x \in \mathcal{L}(p) : x = \overline{x},\ |x| > |q|\} = \sum_{n=|q|+1}^{|p|} \mathcal{P}_p(n),$$

the statement of Lemma 11 is proven. □

Now we can complete the proof of Theorem 1.

Proof of Theorem 1. Finiteness of defect means that there exists a constant $L \in \mathbb{N}$ such that $D(\mathbf{u}) = D(q)$ for any prefix q of $\mathbf{u}$ which is longer than or of length equal to L. On the other hand, finiteness of the sum $\sum_{n=0}^{+\infty} T_{\mathbf{u}}(n)$ together with the fact $0 \leq T_{\mathbf{u}}(n) \in \mathbb{Z}$ for any $n \in \mathbb{N}$ implies that there exists a constant $M \in \mathbb{N}$ such that $T_{\mathbf{u}}(n) = 0$ for any $n > M$. Let us fix an integer $H > \max\{L, M\}$ and denote by q the prefix of $\mathbf{u}$ of length $|q| = H - 1$. Consequently,

$$T_{\mathbf{u}}(n) = 0 \quad \text{for any } n \geq H \quad \text{and} \quad D(\mathbf{u}) = D(q).$$

In order to show the equality (1), it thus remains to show $2D(q) = \sum_{n=0}^{H-1} T_{\mathbf{u}}(n)$.

Let us consider a prefix p of $\mathbf{u}$ containing all factors of length H. In this case p is longer than q, thus it holds by Corollary 3 that $D(q) = D(p)$. Using Theorem 8, we have

$$2D(p) = \sum_{n=0}^{|p|} T_p(n) = \sum_{n=0}^{H-1} T_p(n) + \sum_{n=H}^{|p|} T_p(n) = \sum_{n=0}^{H-1} T_{\mathbf{u}}(n) + \sum_{n=H}^{|p|} T_p(n),$$

where the last equality is due to the fact that p contains all factors of length H. It remains to prove that $\sum_{n=H}^{|p|} T_p(n) = 0$. Let us rewrite the sum by definition.

$$\begin{aligned}\sum_{n=H}^{|p|} T_p(n) &= \sum_{n=H}^{|p|} \big(\mathcal{C}_p(n+1) - \mathcal{C}_p(n) + 2 - \mathcal{P}_p(n+1) - \mathcal{P}_p(n)\big) \\ &= -\mathcal{C}_p(H) + 2(|p| - H + 1) - 2\sum_{n=H}^{|p|} \mathcal{P}_p(n) + \mathcal{P}_p(H) \\ &= -\mathcal{C}_{\mathbf{u}}(H) + 2(|p| - H + 1) - 2\sum_{n=H}^{|p|} \mathcal{P}_p(n) + \mathcal{P}_{\mathbf{u}}(H), \end{aligned} \tag{4}$$

where in the last equality we again used the fact that p contains all factors of length H. This fact also allows us to rewrite the set $\{x \in \mathcal{L}(p) : x \notin \mathcal{L}(q), x = \overline{x}, |x| \le |q|\}$ from Lemma 11 as $\{x \in \mathcal{L}(\mathbf{u}) : x \notin \mathcal{L}(q), x = \overline{x}, |x| < H\}$. Denote the cardinality of this set by B.

In this notation, Lemmas 9 and 11 say

$$\mathcal{C}_{\mathbf{u}}(H) - \mathcal{P}_{\mathbf{u}}(H) = 2B \quad \text{and} \quad B + \sum_{n=H}^{|p|} \mathcal{P}_p(n) = |p| - H + 1.$$

This implies that the last expression in (4) is zero as desired. □

4. Proof of Theorem 2

If an infinite word $\mathbf{u}$ is periodic with language closed under reversal, then $D(\mathbf{u}) < +\infty$ and $\sum_{n=0}^{+\infty} T_{\mathbf{u}}(n) < +\infty$, as shown in [6]. Consequently, we will limit our considerations in the sequel to aperiodic words.

Proposition 12. *If $\mathbf{u}$ is an aperiodic infinite word with language closed under reversal and N is an integer, then $T_{\mathbf{u}}(n) = 0$ for all $n \ge N$ if and only if for any factor w such that $|w| \ge N$, any factor longer than w beginning in w or $\overline{w}$ and ending in w or $\overline{w}$, with no other occurrences of w or $\overline{w}$, is a palindrome.*

Proof. ($\Leftarrow$): Let us show for any $n \ge N$ that the assumptions of Lemma 7 are satisfied. We have to show two properties of $G_n(\mathbf{u})$ for any $n \ge N$.

(1) Any loop in $G_n(\mathbf{u})$ is a palindrome.
Since any loop e in $G_n(\mathbf{u})$ at a vertex $(w, \overline{w})$ is a word longer than w beginning in a special factor w or $\overline{w}$ and ending in w or $\overline{w}$, with no other occurrences of w or $\overline{w}$, the loop e is a palindrome by the assumption.
(2) The graph obtained from $G_n(\mathbf{u})$ by removing loops is a tree.
Or equivalently, we have to show that in $G_n(\mathbf{u})$ there exists a unique path between any two different vertices $(w', \overline{w'})$ and $(w'', \overline{w''})$. Let p be a factor of $\mathbf{u}$ such that w' or $\overline{w'}$ is its prefix, w'' or $\overline{w''}$ is its suffix and p has no other occurrences of w', $\overline{w'}$, w'', $\overline{w''}$. Let v be a factor starting in p, ending in w' or $\overline{w'}$ and containing no other occurrences of w' or $\overline{w'}$. By the assumption the factor v is a palindrome, thus $\overline{p}$ is a suffix of v. It is then a direct consequence of the construction of v that the next factor with the same properties as p, i.e., representing a path in the undirected graph $G_n(\mathbf{u})$ between w' and w'', which occurs in $\mathbf{u}$ after p, is $\overline{p}$. This shows that there is only one such path.

Consequently, Lemma 7 implies that $T_n(\mathbf{u}) = 0$ for any $n \ge N$.
($\Rightarrow$): First we prove an auxiliary claim.

Claim. *If $\mathbf{u}$ is an aperiodic infinite word with language closed under reversal and N is an integer such that $T_{\mathbf{u}}(n) = 0$ for all $n \ge N$, then for any w such that $|w| \ge N$ and any factor v longer than w beginning in w and ending in w or $\overline{w}$, with no other occurrences of w or $\overline{w}$, there exists a letter $a \in \mathcal{A}$ such that v has prefix wa and suffix $a\overline{w}$.*

It is clear that repeated application of the previous claim to factors w of length gradually increased by one gives the proof of implication ($\Rightarrow$) of Proposition 12.

We split the proof of the auxiliary claim into two cases.

- Case 1: Assume that w is a special factor.
  If v does not contain any other special factor of length $n = |w|$ except for w and $\overline{w}$, then v is a loop in the graph $G_n(\mathbf{u})$ and according to Lemma 7, the factor v is a palindrome. Necessarily, v begins in wa for some letter a and ends in $a\overline{w}$.
  Suppose now that $v = v_0 v_1 \cdots v_m$ contains a special factor $z \ne w, \overline{w}$ of length n at the position i, i.e., $z = v_i v_{i+1} \cdots v_{n+i-1}$. Without loss of generality, we consider the smallest index i with this property. The pair $(z, \overline{z})$ is a vertex in the graph $G_n(\mathbf{u})$ and a prefix of v, say e, corresponds to an edge in $G_n(\mathbf{u})$ starting in $(w, \overline{w})$ and ending in $(z, \overline{z})$. Since the graph $G_n(\mathbf{u})$ is a tree, the word v which corresponds to a walk from $(w, \overline{w})$ to the same vertex $(w, \overline{w})$ has a suffix f representing an edge in $G_n(\mathbf{u})$ connecting again vertices $(z, \overline{z})$ and $(w, \overline{w})$. It means that the suffix f starts in z or $\overline{z}$ and ends in w or $\overline{w}$. Since $G_n(\mathbf{u})$ has no multiple edges connecting distinct vertices, necessarily $f = \overline{e}$, which already gives the claim.

- Case 2: Assume w is not a special factor.

It means that there exists a unique letter a such that wa belongs to the language of $\mathbf{u}$. As the language is closed under reversal, the factor $\overline{w}$ has a unique left extension, namely a. If v starts in w and ends in $\overline{w}$, then the claim is proven.

It remains to exclude that v begins and ends in a non-palindromic factor w. Suppose this situation happens. In this case, there exists a unique q such that wq is a right special factor and it is the shortest right special factor having the prefix w. The factor wq has only one occurrence of the factor w — otherwise we can find a shorter prolongation of w which is right special. Since w is a suffix of v, we deduce that $|wq| < |v|$. Because wq is the shortest right special factor with prefix w, the factor vq belongs to the language and its prefix and suffix wq is a special factor. According to already proven Case 1, we have $wq = \overline{wq} = \overline{q}\,\overline{w}$. It means together with the inequality $|wq| < |v|$ that $\overline{w}$ is contained in v as well — a contradiction. □

The proof of the implication ($\Rightarrow$) of Proposition 12 is taken from [10], where we showed a more general statement for an infinite word whose language is closed under a larger group of symmetries.

Corollary 13. *Let $\mathbf{u}$ be an aperiodic infinite word with language closed under reversal and let N be an integer. If $T_{\mathbf{u}}(n) = 0$ for all $n \geq N$, then the occurrences of w and $\overline{w}$ in $\mathbf{u}$ alternate for any factor w of $\mathbf{u}$ of length at least N.*

The following lemma builds a bridge between Corollary 5 and Proposition 12.

Lemma 14. *Let $\mathbf{u}$ be an aperiodic infinite word with language closed under reversal. There exists $H \in \mathbb{N}$ such that the longest palindromic suffix of any prefix w of $\mathbf{u}$ of length $|w| \geq H$ occurs in w exactly once if and only if there exists $N \in \mathbb{N}$ such that for any factor w with $|w| \geq N$, any factor longer than w beginning in w or $\overline{w}$ and ending in w or $\overline{w}$, with no other occurrences of w or $\overline{w}$, is a palindrome.*

Proof. ($\Rightarrow$): We will show that N may be set equal to H. Let us proceed by contradiction. Suppose there exists a factor $w \in \mathcal{L}(\mathbf{u})$ such that $|w| \geq H$ and there exists a non-palindromic factor of $\mathbf{u}$ longer than w beginning in w or $\overline{w}$ and ending in w or $\overline{w}$, with no other occurrences of w or $\overline{w}$. Let us find the first non-palindromic factor of the above form in $\mathbf{u}$ and let us denote it as r. Let p be the prefix of $\mathbf{u}$ ending in the first occurrence of r in $\mathbf{u}$, i.e., $p = tr$ for some word t and r is unioccurrent in p. Denote by s the longest palindromic suffix of p. By the assumption, s is unioccurrent in p. No matter how long the suffix s is, we will obtain a contradiction.

(1) If $|s| \leq |w|$, then we have a contradiction to the unioccurrence of s.
(2) If $|r| > |s| > |w|$, then we can find at least 3 occurrences of w or $\overline{w}$ in r which is a contradiction to the form of r.
(3) The equality $|r| = |s|$ contradicts the fact that we supposed r to be non-palindromic.
(4) Finally, if $|r| < |s|$, then there is an occurrence of the mirror image of r which is a non-palindromic factor having the same properties as r which occurs before r and contradicts the choice of p.

($\Leftarrow$): Take a prefix containing all factors of length N. Set H equal to its length. Let us show that any prefix p of length greater than or equal to H has $lps(p)$ of length greater than or equal to N. Consider a suffix of p of length N, say w. Either w is a palindrome, then $lps(p)$ is of length greater than or equal to N. Or w is not a palindrome, then we find a suffix of p beginning in $\overline{w}$ and containing exactly two occurrences of w or $\overline{w}$. Such a suffix exists since all factors of length N are contained in p. By assumptions, such a suffix is a palindrome, hence $lps(p)$ is longer than N.

Any prefix p of $\mathbf{u}$ of length greater than or equal to H has $lps(p)$ unioccurrent. Assume there are more occurrences of $lps(p)$ in p and consider its suffix v starting in the last-but-one occurrence of $lps(p)$. Since the length of $lps(p)$ is greater than or equal to N, the factor v is a palindrome by assumptions, which contradicts the choice of $lps(p)$. □

Proof of Theorem 2. For periodic words, the statement was shown in [6]. If $\mathbf{u}$ is aperiodic, then the statement is a direct consequence of Lemma 14, Corollary 5, and Proposition 12. □

Acknowledgments

This work was supported by the Czech Science Foundation grant GAČR 201/09/0584 and by the grant MSM6840770039 of the Ministry of Education, Youth, and Sports of the Czech Republic.

We would like to thank Bojan Bašić for his careful reading, useful remarks and in particular for his correction of a lemma.

References

[1] P. Baláži, Z. Masáková, E. Pelantová, Factor versus palindromic complexity of uniformly recurrent infinite words, Theoret. Comput. Sci. 380 (2007) 266–275.
[2] Ľ. Balková, E. Pelantová, Š. Starosta, Corrigendum: "On Brlek–Reutenauer conjecture", Theoret. Comput. Sci. 465 (2012) 73–74.
[3] Ľ. Balková, E. Pelantová, Š. Starosta, On Brlek-Reutenauer conjecture, Theoret. Comput. Sci. 412 (2011) 5649–5655.
[4] B. Bašić, A note on the paper On Brlek-Reutenauer conjecture, Theoret. Comput. Sci. 448 (2012) 94–96.
[5] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer, On the palindromic complexity of infinite words, Internat. J. Found. Comput. 15 (2004) 293–306.
[6] S. Brlek, C. Reutenauer, Complexity and palindromic defect of infinite words, Theoret. Comput. Sci. 412 (2011) 493–497.
[7] M. Bucci, A. De Luca, A. Glen, L.Q. Zamboni, A connection between palindromic and factor complexity using return words, Adv. in Appl. Math. 42 (2009) 60–74.
[8] X. Droubay, J. Justin, G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (2001) 539–553.
[9] A. Glen, J. Justin, S. Widmer, L.Q. Zamboni, Palindromic richness, European J. Combin. 30 (2009) 510–531.
[10] E. Pelantová, Š Starosta, Palindromic richness and Coxeter groups, preprint available at http://arxiv.org/abs/1108.3042.

Chapter 3

Exchange of three intervals: itineraries, substitutions and palindromicity

Authors:	Z. Masáková, E. Pelantová, and Š. Starosta
Publication reference:	submitted to Eur. J. of Combin., 2015
Cited as:	[II]

Exchange of three intervals: itineraries, substitutions and palindromicity

Zuzana Masáková Edita Pelantová Štěpán Starosta

Abstract

Given a symmetric exchange of three intervals, we provide a detailed description of the return times to a subinterval and the corresponding itineraries. We apply our results to morphisms fixing words coding non-degenerate three interval exchange transformation. This allows us to prove that the conjecture stated by Hof, Knill and Simon is valid for such infinite words.

1 Introduction

Interval exchange transformations have been extensively studied since the works on their ergodic aspects by Sinai [27], Keane [14], Veech [31], Rauzy [24], and others. For an overview, see [30] and references therein. Among general dynamical systems, interval exchanges have the interesting property that the Poincaré first return map is again a mapping of the same type, i.e. an exchange of (possibly different number of) intervals. Rauzy [23] used this fact to present a generalization of the classical continued fractions expansion.

It is commonly known that interval exchange transformations provide a very useful framework for the study of infinite words arising by coding of rotations, in particular Sturmian words. These are usually defined as aperiodic infinite words with lowest factor complexity. Equivalently, one obtains Sturmian words by binary coding of the trajectory under exchange T of two intervals $[0, \alpha)$, $[\alpha, 1)$ with α irrational. Given a subinterval $I \subset [0, 1)$, the first return map T_I to I is an exchange of at most three intervals, although the return itineraries of points can take up to four values. The set of these itineraries can be used to describe certain characteristics of Sturmian words, namely the return words, see [32], or abelian return words [25], and invariance under morphisms [34].

Infinite words coding exchange of k-intervals, $k \geq 3$, are also in focus for several decades [10, 11] and, here too, one finds a close relation between their combinatorial features and the properties of the induced map, see for example [33] for a result on return words or [1] about substitutivity of interval exchange words. A generalized version of the Poincaré first return map was used in [9] for description of palindromic complexity in codings of rotations. These words are in intimate relation with three interval exchange words.

In this paper we focus on codings of a non-degenerate symmetric exchange $T : J \to J$ of three intervals. First we describe the return times to a general interval $I \subset J$ and provide an insight on the structure of the set of I-itineraries. These results are given as Theorem 5.1 and then interpreted as analogues of the well known three gap and three distance theorems.

A particular attention is paid to the special cases when the set of I-itineraries has only three elements. These cases belong to the most interesting from the combinatorial point of view, since they provide information about return words to factors, and about the morphisms preserving three interval exchange words. For mutually conjugated morphisms, we describe in Theorem 9.10

the relation between intercepts of their fixed points, as was done for Sturmian morphisms in [21]. We also show that morphisms conjugated only to themselves do not have a non-degenerated fixed point.

The most important application of our results is a contribution to the solution of the question stated by Hof, Knill and Simon [13] for palindromic words. We refer to it as the HKS conjecture and adopt its reformulation by Tan [29] who showed its validity for binary words. Labbé [16] presented a counterexample for the conjecture on ternary alphabet; the ternary word not satisfying the hypothesis turns out to be a degenerate three interval exchange word. In fact, degenerate three interval exchange words are just morphic images of binary, in fact Sturmian, words. In this paper we show that for non-degenerated words coding exchange of three intervals the HKS conjecture holds. Let us mention that the latter result has been announced at the DLT conference [18]. Here we provide a full proof.

This paper is organized as follows. Section 2 contains the necessary notions from combinatorics on words. Symmetric k-interval exchange transformations and their properties with respect to the first return map are treated in Section 3. Section 4 focuses on specific properties when $k = 3$. The main theorem about return times in three interval exchanges is given in Section 5. In Section 6 we put our results into context of three gaps and three distance theorems.

The specific case when the set of I-itineraries has only three elements is studied in Section 7. This allows us to describe the return words to palindromic bispecial factors. In Section 8 we focus on substitution invariance of words coding interval exchange transformations. The key lemma for the demonstration of our Theorem 10.3 on HKS conjecture requires some knowledge about the relation of substitutions fixing words coding three interval exchange and Sturmian morphisms. This topic is treated in Section 9. The proof of Theorem 10.3 is then provided in Section 10 together with some other comments.

2 Preliminaries

Let us recall necessary notions and notation from combinatorics on words. For a basic overview we refer to [17]. An *alphabet* is a finite set of symbols, called *letters.* A *finite word* w over an alphabet $\mathcal{A}$ of length $|w| = n$ is a concatenation $w = w_0 \cdots w_{n-1}$ of letters $w_i \in \mathcal{A}$. The set of all finite words over $\mathcal{A}$ equipped with the operation of concatenation and the empty word ϵ is a monoid denoted by $\mathcal{A}^*$. For a fixed letter $a \in \mathcal{A}$, the number of occurrences of a in w, i.e., the number of indices i such that $w_i = a$, is denoted by $|w|_a$. The *reversal* or *mirror image* of the word w is the word $\overline{w} = w_{n-1} \cdots w_0$. A word w for which $w = \overline{w}$ is called a *palindrome.* An *infinite word* $\mathbf{u}$ is an infinite concatenation $\mathbf{u} = u_0u_1u_2 \ldots \in \mathcal{A}^{\mathbb{N}}$. An infinite word $\mathbf{u} = wvvv\ldots$ with $w, v \in \mathcal{A}^*$ is said to be *eventually periodic*; it is said to be *aperiodic* if it is not of such form. We say that $w \in \mathcal{A}^*$ is a *factor* of $v \in \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$ if $v = w'ww''$ for some $w' \in \mathcal{A}^*$ and $w'' \in \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$. If $w' = \epsilon$ or $w'' = \epsilon$, then w is a *prefix* or *suffix* of v, respectively. If $v = wu$, then we write $u = w^{-1}v$ and $w = vu^{-1}$.

The set $\mathcal{L}(\mathbf{u})$ of all finite factors of an infinite word $\mathbf{u}$ is called the *language of* $\mathbf{u}$. If for any factor $w \in \mathcal{L}(\mathbf{u})$ there exist at least two indices i such that w is a prefix of the infinite word $u_iu_{i+1}u_{i+2}\cdots$, the word $\mathbf{u}$ is *recurrent.* Given a factor $w \in \mathcal{L}(\mathbf{u})$, a finite word v such that vw belongs to $\mathcal{L}(\mathbf{u})$ and the word w occurs in vw exactly twice, once as a prefix and once as a suffix of vw, is called a *return word of* w. If any factor w of an infinite recurrent word $\mathbf{u}$ has only finitely many return words, the word $\mathbf{u}$ is called *uniformly recurrent.*

The *factor complexity* $\mathcal{C}_{\mathbf{u}}$ is the function $\mathbb{N} \to \mathbb{N}$ counting the number of factors of $\mathbf{u}$ of length n. It is known that the factor complexity of an aperiodic infinite word $\mathbf{u}$ satisfies $\mathcal{C}_{\mathbf{u}}(n) \geq n+1$ for all n. Aperiodic infinite words having the minimal complexity $\mathcal{C}_{\mathbf{u}}(n) = n+1$ for all n are called *Sturmian words*. Since $\mathcal{C}_{\mathbf{u}}(1) = 2$, they are binary words. Sturmian words can be equivalently defined in many different frameworks, one of them is coding of an exchange of two intervals.

Let $\mathcal{A}$ and $\mathcal{B}$ be alphabets. Let $\varphi : \mathcal{A}^* \to \mathcal{B}^*$ be a *morphism*, i.e., $\varphi(wv) = \varphi(w)\varphi(v)$ for all $w, v \in \mathcal{A}^*$. We say that φ is *non-erasing* if $\varphi(b) \neq \epsilon$ for every $b \in \mathcal{A}$. The action of φ can be naturally extended to infinite words $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ by setting $\varphi(\mathbf{u}) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots$. If $\mathcal{A} = \mathcal{B}$ and $\varphi(\mathbf{u}) = \mathbf{u}$, then $\mathbf{u}$ is said to be a *fixed point* of φ. A non-erasing morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ such that there is a letter $a \in \mathcal{A}$ satisfying $\varphi(a) = aw$ for some non-empty word w is called a *substitution*. Obviously, a substitution has always a fixed point, namely $\lim_{n\to\infty} \varphi^n(a)$ where the limit is taken over the product topology. An infinite word which is a fixed point of a substitution is called a *pure morphic* word. Let $\mathcal{A} = \{a_1, \ldots, a_k\}$ and $\mathcal{B} = \{b_1, \ldots, b_\ell\}$. One associates to every morphism $\varphi : \mathcal{A} \to \mathcal{B}$ its *incidence matrix* $M_\varphi \in \mathbb{N}^{k\times \ell}$ defined by

$$(M_\varphi)_{ij} = |\varphi(a_i)|_{b_j}\,, \qquad \text{for } 1 \le i \le k,\ 1 \le j \le \ell.$$

A morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ is said to be *primitive* if all elements of some power of its incidence matrix $M_\varphi \in \mathbb{N}^{k\times k}$ are positive. A specific class of morphisms is formed by the so-called *Sturmian morphisms* which are defined over the binary alphabet $\{0, 1\}$ and for which there exists a Sturmian word $\mathbf{u}$ such that $\varphi(\mathbf{u})$ is also Sturmian. For an overview about properties of Sturmian morphisms see [17].

3 Itineraries in symmetric exchange of intervals

For disjoint intervals K and K' we write $K < K'$ if for $x \in K$ and $x' \in K'$ we have $x < x'$. Let J be a semi-closed interval. Consider a partition $J = J_0 \cup \cdots \cup J_{k-1}$ of J into a disjoint union of semi-closed subintervals $J_0 < J_1 < \cdots < J_{k-1}$. A bijection $T : J \to J$ is called an *exchange of k intervals with permutation π* if there exist numbers $c_0, \ldots, c_{k-1}$ such that for $0 \le i < k$ one has

$$T(x) = x + c_i \text{ for } x \in J_i\,, \tag{1}$$

where π is a permutation of $\{0, 1, \ldots, k-1\}$ such that $T(J_i) < T(J_j)$ for $\pi(i) < \pi(j)$. In other words, the permutation π determines the order of intervals $T(J_i)$. If π is the permutation $i \mapsto k - i - 1$, then T is called a *symmetric* interval exchange.

The orbit of a given point ρ is the infinite sequence ρ, $T(\rho)$, $T^2(\rho)$, $T^3(\rho)$, ... It can be coded by an infinite word $\mathbf{u}_\rho = u_0u_1u_2\ldots$ over the alphabet $\{0, 1, \ldots, k-1\}$ given by

$$u_n = X \quad \text{if } T^n(\rho) \in J_X \qquad \text{for } X \in \{0, 1, \ldots, k-1\}.$$

The point ρ is called the *intercept* of $\mathbf{u}_\rho$. An exchange of intervals satisfies the *minimality condition* if the orbit of any given $\rho \in [0, 1)$ is dense in J. In this case, the word $\mathbf{u}_\rho$ is aperiodic, uniformly recurrent, and the language of $\mathbf{u}_\rho$ depends only on the parameters of the transformation T and not on the intercept ρ itself. The complexity of an infinite word $\mathbf{u}_\rho$ is known to satisfy $\mathcal{C}_{\mathbf{u}_\rho}(n) \le (k-1)n + 1$ (see [10]). If for every $n \in \mathbb{N}$ we have $\mathcal{C}_{\mathbf{u}_\rho}(n) = (k-1)n + 1$, then the transformation T and the word $\mathbf{u}_\rho$ are said to be *non-degenerate*. A sufficient and necessary condition on T to be non-degenerate is that the orbits of the discontinuity points of T are infinite and disjoint. This condition is known under the abbreviation i.d.o.c.

Definition 3.1. Let T be an exchange of k intervals satisfying the minimality condition. Given a subinterval $I \subset J$, we define the mapping $r_I : I \to \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ by

$$r_I(x) = \min\{n \in \mathbb{Z}^+ : T^n(x) \in I\}\,,$$

the so-called *return time* to I. The prefix of length $r_I(x)$ of the word $\mathbf{u}_x$ coding the orbit of a point $x \in I$ is called the I*-itinerary* of x and denoted $R_I(x)$. The set of all I-itineraries is denoted by $\mathrm{It}_I = \{R_I(x) \colon x \in I\}$. The mapping $T_I : I \to I$ defined by

$$T_I(x) = T^{r_I(x)}(x)$$

is said to be the *first return map of* T *to* I, or *induced map of* T *on* I.

Throughout the paper, when it is clear from the context, we sometimes omit the index I in r_I or R_I. It is known from Keane [14] that if T is an exchange of k intervals and $I \subset J$, then It_I has at most $k+2$ elements, and, consequently, T_I is an exchange of at most $k+2$ intervals.

Remark 3.2. Let $X \in \{0, 1, \ldots, k-1\}$. If $I \subset J_X$, then $T(I)$ is an interval and we have

$$R \text{ is an } I\text{-itinerary} \quad \Leftrightarrow \quad X^{-1}RX \text{ is a } T(I)\text{-itinerary.}$$

Similarly, if $I \subset T(J_X)$, then $T^{-1}(I)$ is an interval and we have

$$R \text{ is an } I\text{-itinerary} \quad \Leftrightarrow \quad XRX^{-1} \text{ is a } T^{-1}(I)\text{-itinerary.}$$

We will use another fact about itineraries of an interval exchange. Without loss of generality, we consider $J = [0, 1)$. The intervals J_X are left-closed right-open for all $X \in \{0, 1 \ldots, k-1\}$. Such interval exchange T is right-continuous. Therefore, if $I = [\gamma, \delta)$, then every word $w \in \mathrm{It}_I = \{R(x) \colon x \in I\}$ is an I-itinerary $R(x)$ for infinitely many $x \in I$, which form an interval, again left-closed right-open.

Proposition 3.3. *Let T be a k-interval exchange satisfying the minimality condition and let $I = [\gamma, \delta) \subset [0, 1)$. There exist neighbourhoods H_γ and H_δ of γ and δ, respectively, such that for every $\tilde\gamma \in H_\gamma$ and $\tilde\delta \in H_\delta$ with $0 \le \tilde\gamma < \tilde\delta \le 1$ one has*

$$It_{\tilde I} \supseteq It_I\,, \quad \text{where } \tilde I = [\tilde\gamma, \tilde\delta).$$

In particular, if $\# It_I = k+2$, then $It_{\tilde I} = It_I$.

Proof. Let $It_I = \{R_1, \ldots, R_m\}$, $m \in \mathbb{N}$, and $I_i = \{x \in I \colon R_I(x) = R_i\}$ for $1 \le i \le m$. As already mentioned, $m \le k+2$. For every $i = 1, \ldots, m$, consider arbitrary x_i in the interior of I_i. Such x_i satisfies that $T^j(x_i) \notin \{\gamma, \delta\}$ for $0 \le j \le r_I(x_i) = |R_I(x_i)|$. The reason is simple: if $T^j(x_i)$ were equal to γ (or δ), then a point x in I_i with $x < x_i$ (or $x > x_i$) would have a longer return time than x_i itself, which is a contradiction. Denote $M = \{T^j(x_i) \colon 0 \le j \le r_I(x_i),\ i = 1, \ldots, m\}$ and

$$\varepsilon := \min\big\{|y - z| \colon y \in M, z \in \{\gamma, \delta\}\big\}\,.$$

Let $H_\gamma = (\gamma - \varepsilon, \gamma + \varepsilon)$ and $H_\delta = (\delta - \varepsilon, \delta + \varepsilon)$ be neighbourhoods of γ and δ, respectively. If $\tilde\gamma \in H_\gamma$ and $\tilde\delta \in H_\delta$ with $0 \le \tilde\gamma < \tilde\delta \le 1$ and $\tilde I = [\tilde\gamma, \tilde\delta)$, then clearly for every $i = 1, \ldots, m$ we have $x_i \in \tilde I$ and

$$T^j(x_i) \in I \quad \Leftrightarrow \quad T^j(x_i) \in \tilde I \qquad \text{for } 0 \le j \le r_I(x_i)\,.$$

Therefore the point x_i has the same return time with respect to I as to $\tilde I$. Consequently, the $\tilde I$-itinerary of x_i coincides with the I-itinerary of x_i. Thus $It_I \subseteq It_{\tilde I}$. ☐

For the rest of the section, we consider only symmetric interval exchange. In order to state a property of such interval exchanges, for an interval $K = [c, d) \subset [0, 1)$ we set $\overline{K} = [1-d, 1-c)$. In this notation

$$T(J_X) = \overline{J_X} \quad \text{for any letter } X \in \{0, 1, \ldots, k-1\}\,. \tag{2}$$

4

Proposition 3.4. *Let $T : [0,1) \to [0,1)$ be a symmetric exchange of k intervals satisfying the minimality condition. Let $I \subset [0,1)$ and let $R_1, \ldots, R_m$ be the I-itineraries. The $\overline{I}$-itineraries are the mirror images of the I-itineraries, namely $\overline{R_1}, \ldots, \overline{R_m}$. Moreover, if*

$$[\gamma_j, \delta_j) := \{x \in I \colon R_I(x) = R_j\} \quad \textit{and} \quad [\gamma_j', \delta_j') := T_I[\gamma_j, \delta_j)\,,$$

for $j = 1, \ldots, m$, then

$$\{x \in \overline{I} \colon R_{\overline{I}}(x) = \overline{R_j}\} = [1 - \delta_j', 1 - \gamma_j')\,.$$

Proof. Consider the restriction of the transformation T to the set

$$S = [0,1) \setminus \{T^j(\alpha) \colon j \in \mathbb{Z},\ \alpha \text{ is a discontinuity of } T\}\,.$$

Such a restriction is a bijection $S \to S$. We will show by induction that for any $i \geq 1$ and $y \in S$

$$T^{-i}(y) = 1 - T^i(1-y)\,. \tag{3}$$

Let $y \in S$ and $j \in \{0, \ldots, k-1\}$ such that $y \in I_j$. Since T is symmetric, we have

$$1 - y \in I_j \;\Leftrightarrow\; y \in T(I_j).$$

The last equivalence and the definition of T imply

$$T(1-y) = 1 - y + c_j \qquad \text{and}$$
$$T^{-1}(y) = y - c_j.$$

Summing the last two equalities we obtain

$$T^{-1}(y) = 1 - T(1-y).$$

Then, using the induction hypothesis, we have for $y \in S$ that

$$T^{-(i+1)}(y) = T^{-1}\big(T^{-i}(y)\big) = 1 - T\big(1 - T^{-i}(y)\big) = 1 - T\big(T^i(1-y)\big) = 1 - T^{i+1}(1-y)\,,$$

which proves (3).

Using (2) we can write for $y \in S$

$$T^{-1}(y) \in J_X \;\Leftrightarrow\; y \in T(J_X) \;\Leftrightarrow\; 1 - y \in J_X\,. \tag{4}$$

More generally,

$$T^{-i}(y) = T^{-1}\big(T^{-(i-1)}(y)\big) \in J_X \;\Leftrightarrow\; 1 - T^{-(i-1)}(y) = T^{i-1}(1-y) \in J_X\,, \tag{5}$$

where we have first used (4) and then (3).

Now we show that if R_j is an I-itinerary, then its mirror image $\overline{R_j}$ is an $\overline{I}$-itinerary. Consider $\rho \in (\gamma_j, \delta_j) \cap S$ and let $R_I(\rho) = a_0 a_1 \cdots a_{n-1}$ be its I-itinerary, i.e., $a_i = X$ if and only if $T^i(\rho) \in J_X$. Moreover, $T^i(\rho) \notin I$ for $1 \leq i < n$, and $T^n(\rho) \in I$. Let

$$\rho' := 1 - T^n(\rho) = 1 - T_I(\rho) \in (1 - \delta_j', 1 - \gamma_j') \cap S \subset \overline{I}\,. \tag{6}$$

By (3), we have $\rho' = T^{-n}(1-\rho)$, and therefore again by (3), $T^i(\rho') = T^{-(n-i)}(1-\rho) = 1 - T^{n-i}(\rho) \notin \overline{I}$ for $0 < i < n$. On the other hand, $T^n(\rho') = 1 - \rho \in \overline{I}$. By (5), we have for $i = 0, 1, \ldots, n-1$ that

$$J_X \ni T^i(\rho') = T^{-(n-i)}(1-\rho) \;\Leftrightarrow\; T^{n-i-1}(\rho) \in J_X\,,$$

which implies that the $\overline{I}$-itinerary of ρ' is $R_{\overline{I}}(\rho') = a_{n-1} a_{n-2} \cdots a_0$, as we wanted to show.

By right continuity of T, all points from $[1 - \delta_j', 1 - \gamma_j')$ have the same $\overline{I}$-itinerary as $\rho' \in (1 - \delta_j', 1 - \gamma_j') \cap S$. □

The above auxiliary statements will be used in Section 5 for the description of I-itineraries in exchanges of three intervals. Analogous result for exchange of two intervals was given in [19]. The claim of the last proposition also partially follows from the work done in [10].

4 Exchange of three intervals

We will be particularly interested in exchange of three intervals. For reasons that will appear later, we prefer to use for its coding the ternary alphabet $\{A, B, C\}$ instead of $\{0, 1, 2\}$. Without loss of generality let $0 < \alpha < \beta < 1$. Let $T : [0, 1) \to [0, 1)$ be given by

$$T(x) = \begin{cases} x + 1 - \alpha & \text{if } x \in [0, \alpha) =: J_A\,, \\ x + 1 - \alpha - \beta & \text{if } x \in [\alpha, \beta) =: J_B\,, \\ x - \beta & \text{if } x \in [\beta, 1) =: J_C\,. \end{cases} \tag{7}$$

The transformation T is an exchange of three intervals with the permutation (321). It is often called a *3iet* for short. The infinite word $\mathbf{u}_\rho$ coding the orbit of a point $\rho \in [0, 1)$ under a 3iet is called a *3iet word*.

We require that $1 - \alpha$ and β be linearly independent over $\mathbb{Q}$, which is known to be a necessary and sufficient condition for minimality of the 3iet T. Non-degeneracy of T is equivalent to the condition of minimality together with

$$1 \notin (1 - \alpha)\mathbb{Z} + \beta\mathbb{Z}\,, \tag{8}$$

see [10]. This means that the 3iet word $\mathbf{u}$ has complexity $\mathcal{C}_{\mathbf{u}}(n) = 2n + 1$ if and only if the parameters α and β of the corresponding 3iet T satisfy (8).

For a given subinterval $I \subset [0, 1)$ there exist at most five I-itineraries under a 3iet T. In particular, from the paper of Keane [14], one can deduce what are the intervals of points with the same itinerary. We summarize it as the following lemma.

Lemma 4.1. *Let T be a 3iet defined by* (7) *and let $I = [\gamma, \delta) \subset [0, 1)$ such that $\delta < 1$. Denote*

$$\begin{aligned} k_\alpha &:= \min\left\{k \in \mathbb{Z},\, k \geq 0 \colon T^{-k}(\alpha) \in (\gamma, \delta)\right\}, \\ k_\beta &:= \min\left\{k \in \mathbb{Z},\, k \geq 0 \colon T^{-k}(\beta) \in (\gamma, \delta)\right\}, \\ k_\gamma &:= \min\left\{k \in \mathbb{Z},\, k \geq 1 \colon T^{-k}(\gamma) \in (\gamma, \delta)\right\}, \\ k_\delta &:= \min\left\{k \in \mathbb{Z},\, k \geq 1 \colon T^{-k}(\delta) \in (\gamma, \delta)\right\}, \end{aligned}$$

and further

$$\mathfrak{A} := T^{-k_\alpha}(\alpha),\ \mathfrak{B} := T^{-k_\beta}(\beta),\ \mathfrak{C} := T^{-k_\gamma}(\gamma),\ \mathfrak{D} := T^{-k_\delta}(\delta).$$

For $x \in I$, let K_x be a maximal interval such that for every $y \in K_x$, we have $R(y) = R(x)$. Then K_x is of the form $[c, d)$ with $c, d \in \{\gamma, \delta, \mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}\}$. Consequently, $\#\mathrm{It}_I \leq 5$.

For a 3iet T, Lemma 4.1 implies the already mentioned result that T_I is an exchange of at most 5 intervals. Consequently, the transformation T_I has at most four discontinuity points. In fact, the following result of [6] says that independently of the number of I-itineraries, the induced map T_I has always at most two discontinuity points.

Proposition 4.2 ([6])**.** *Let $T : J \to J$ be a 3iet with the permutation* (321) *and satisfying the minimality condition, and let $I \subset J$ be an interval. The first return map T_I is either a 3iet with permutation* (321) *or a 2iet with permutation* (21)*. In particular, in the notation of Lemma 4.1, we have $\mathfrak{D} \leq \mathfrak{C}$.*

Convention: For the rest of the paper, let T be a non-degenerate exchange of three intervals with permutation (321) given by (7).

5 Return time in a 3iet

The aim of this section is to describe the possible return times of a non-degenerate 3iet T to a general subinterval $I \subset [0,1)$. Our aim is to prove the following theorem.

Theorem 5.1. *Let T be a non-degenerate 3iet and let $I \subset [0,1)$. There exist positive integers r_1, r_2 such that the return time of any $x \in I$ takes value in the set $\{r_1, r_1+1, r_2, r_1+r_2, r_1+r_2+1\}$ or $\{r_1, r_1+1, r_2, r_2+1, r_1+r_2+1\}$.*

First, we will formulate an important lemma, which needs the following notation. Given letters $X, Y, Z \in \{A, B, C\}$ and a finite word $w \in \{A, B, C\}^*$, Let $\omega_{XY \to Z}(w)$ be the set of words obtained from w replacing one factor XY by the letter Z, i.e.

$$\omega_{XY\to Z}(w) = \{w_1 Z w_2 \colon w = w_1 XY w_2\}\,.$$

Similarly,

$$\omega_{Z\to XY}(w) = \{w_1 XY w_2 \colon w = w_1 Z w_2\}\,.$$

Clearly,

$$v \in \omega_{XY\to Z}(w) \;\Leftrightarrow\; w \in \omega_{Z\to XY}(v). \tag{9}$$

By abuse of notation, we write $v = \omega_{XY\to Z}(w)$ instead of $v \in \omega_{XY\to Z}(w)$.

Lemma 5.2. *Assume that the orbits of points α, β, γ and δ are mutually disjoint. For sufficiently small $\varepsilon > 0$, we have the following relations between I-itineraries of points in I*

(a) $R(\mathfrak{A} - \varepsilon) = \omega_{B\to AC}\big(R(\mathfrak{A} + \varepsilon)\big)$,

(b) $R(\mathfrak{A} + \varepsilon) = \omega_{AC\to B}\big(R(\mathfrak{A} - \varepsilon)\big)$,

(c) $R(\mathfrak{B} - \varepsilon) = \omega_{CA\to B}\big(R(\mathfrak{B} + \varepsilon)\big)$,

(d) $R(\mathfrak{B} + \varepsilon) = \omega_{B\to CA}\big(R(\mathfrak{B} - \varepsilon)\big)$,

(e) $R(\mathfrak{D} + \varepsilon) = R(\mathfrak{D} - \varepsilon)R(\delta - \varepsilon)$,

(f) $R(\mathfrak{C} - \varepsilon) = R(\mathfrak{C} + \varepsilon)R(\gamma + \varepsilon)$,

where $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ are given in Lemma 4.1.

Proof. We will first demonstrate the proof of the case a. Let $K = [\mathfrak{A} - \varepsilon, \mathfrak{A} + \varepsilon]$ with ε chosen such that $K \subset I$ and

$$\alpha, \beta, \gamma, \delta \notin T^i(K) \text{ for all } 0 \le i \le k_\alpha \text{ with the only exception of } T^{k_\alpha}(\mathfrak{A}) = \alpha. \tag{10}$$

For simplicity, denote $t = \max\{r_I(x) \colon x \in K\}$ the maximal return time. The existence of such ε follows trivially from the definition of the interval exchange transformation and the assumptions of the lemma.

Let $K_- = [\mathfrak{A} - \varepsilon, \mathfrak{A})$ and $K_+ = [\mathfrak{A}, \mathfrak{A} + \varepsilon]$. It follows from the definition of $\mathfrak{A}$ and condition (10) that for all i such that $0 < i \le k_\alpha$ we have $T^i(K) \cap I = \emptyset$. Moreover, condition (10) implies that all such $T^i(K)$ are intervals. It implies that for any $x, y \in K$, the prefixes of $R(x)$ and $R(y)$ of length $k_\alpha + 1$ are the same. Denote this prefix by w.

The definition of k_α implies that $\alpha \in T^{k_\alpha}(K)$. Since $T^{k_\alpha}(K_+) = [\alpha, \alpha + \varepsilon] \subset J_B$, we obtain

$$T^{k_\alpha+1}(K_+) = \big[T(\alpha), T(\alpha) + \varepsilon\big)\,.$$

Furthermore, since $T^{k_\alpha}(K_-) = [\alpha - \varepsilon, \alpha) \subset J_A$, we obtain

$$T^{k_\alpha+1}(K_-) = [1-\varepsilon, 1) \subset J_C\,,$$

and thus

$$T^{k_\alpha+2}(K_-) = \big[T(\alpha) - \varepsilon, T(\alpha)\big)\,.$$

This implies that the set $K' = T^{k_\alpha+2}(K_-) \cup T^{k_\alpha+1}(K_+) = [T(\alpha)-\varepsilon, T(\alpha)+\varepsilon]$ is an interval. As above, condition (10) implies that the set $T^i(K')$ is an interval for all i such that $0 \le i \le t - k_\alpha - 1$. It follows that $\min\{i\colon T^i(K') \cap K \neq \emptyset\} = t - k_\alpha - 2$ and condition (10) moreover implies that $T^{t-k_\alpha-2}(K') \subset K$. Thus, the iterations $x, T(x), \ldots, T^{t-k_\alpha-2}(x)$ of every $x \in K'$ are coded be the same word, say v.

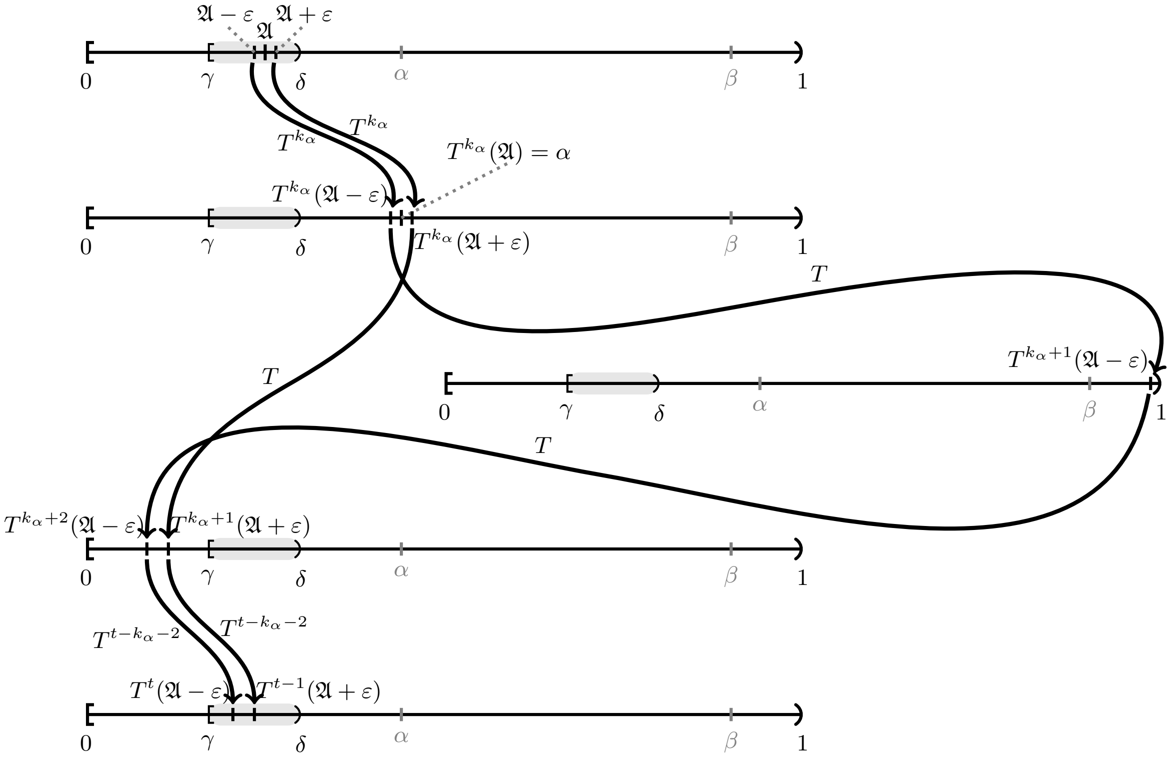

Figure 1: Situation in the proof of Lemma 5.2, case a.

.

The whole situation is depicted in Figure 1. From what is said above, we can write down the I-itineraries of points from K,

$$R(x) = \begin{cases} wACv & \text{if } x \in K_-,\\ wBv & \text{if } x \in K_+. \end{cases}$$

This finishes the proof of (a).

The claim in item (c) is analogous to (a). Cases (b) and (d) are derived from (a) and (c) by the use of equivalence (9).

Let us now demonstrate the proof of the case (e). Denote $s = \min\{n \in \mathbb{Z}^+ \colon T^n(\delta) \in I\}$. Let $K = [\mathfrak{D} - \varepsilon, \mathfrak{D} + \varepsilon]$ with ε chosen such that $K \subset I$ and

$$\alpha, \beta, \gamma, \delta \notin T^i(K) \text{ for all } 0 \le i \le k_\delta + s \text{ with the only exception of } T^{k_\delta}(\mathfrak{D}) = \delta\,. \tag{11}$$

The existence of such ε follows trivially from the definition of the interval exchange transformation and the assumptions of the lemma.

Condition (11) implies that $T^i(K)$ is an interval for all i such that $0 < i \le k_\delta + s$. Moreover, $T^i(K) \cap I = \emptyset$ for all i such that $0 < i < k_\delta$. We obtain $T^{k_\delta}(K) \cap I = [\delta - \varepsilon, \delta)$. In other words, the I-itineraries of all points of K start with a prefix of length k_δ which is equal to $R(\mathfrak{D} - \varepsilon)$. Condition (11) and the definition of s implies that for all i such that $k_\delta < i < s + k_\delta$ we have $T^i(K) \subset J_X$ for some $X \in \{A, B, C\}$ and $T^i(K) \cap I = \emptyset$. Moreover, $T^i(K) \subset I$ for $i = k_\delta + s$. Thus, the iterations of points of $T^{k_\delta}(K) = [\delta - \varepsilon, \delta) \cup T^{k_\delta}[\mathfrak{D}, \mathfrak{D} + \varepsilon]$ are coded by the same word of length s, namely $R(\delta - \varepsilon)$. Altogether, we can conclude that the I-itinerary of points in the interval $[\mathfrak{D}, \mathfrak{D} + \varepsilon]$ is equal to $R(\mathfrak{D} - \varepsilon)R(\delta - \varepsilon)$. The situation is depicted in Figure 2.

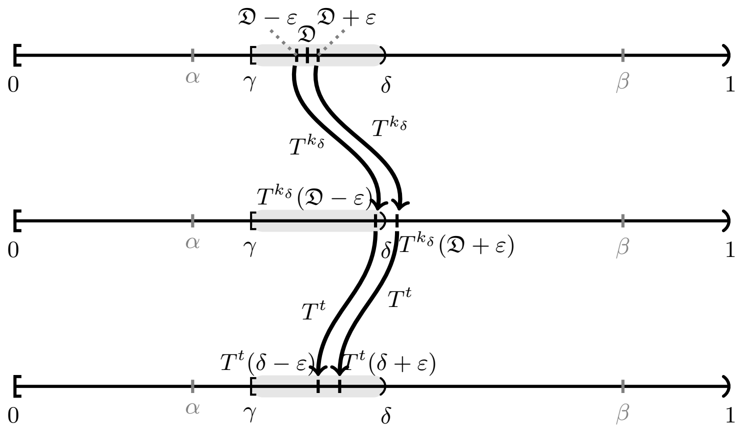

Figure 2: Situation in the proof of Lemma 5.2, case d.

.

Case (f) can be treated in a way analogous to case (e). □

Now we are in the state to prove the main theorem describing the return times in 3iet. In the proof, it is sufficient to focus on the case when $\#\mathrm{It}_I = 5$, since, as we have seen from Proposition 3.3, the set of I-itineraries, and thus also their return times, for the other cases is only a subset of $\mathrm{It}_{\tilde{I}}$ for some "close enough" generic subinterval $\tilde{I} \subset [0, 1)$. So throughout the rest of this section, suppose that $\#\mathrm{It}_I = 5$. This means by Lemma 4.1 that points $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ lie in the interior of the interval $I = [\gamma, \delta)$ and are mutually distinct, moreover, by Proposition 4.2, we have $\mathfrak{D} < \mathfrak{C}$. Such conditions imply 12 possible orderings of $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ which give rise to 12 cases in the study of return times. We will describe them in the proof of Theorem 5.1 as cases (i)–(xii) and then show in Example 5.5 that all 12 cases may occur.

Remark 5.3. Note that if $\gamma = 0$, i.e. we induce on an interval $I = [0, \delta)$, we have $T^{-1}(\gamma) = \beta$ and therefore necessarily $\mathfrak{B} = \mathfrak{C}$. Thus there are at most four I-itineraries. Due to Proposition 3.4, similar situation happens if $\delta = 1$.

Proof of Theorem 5.1. We will discuss the 12 possibilities of ordering of points $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}$ in the interior of the interval $[\gamma, \delta)$ with the condition $\mathfrak{D} < \mathfrak{C}$. The structure of the set of I-itineraries will be best shown in terms of I-itineraries of points in the left neighbourhood of the point $\mathfrak{D}$ and right neighbourhood of the point $\mathfrak{C}$. For simplicity, we thus denote for sufficiently small positive ε

$$R_1 = R(\mathfrak{D} - \varepsilon),\ R_2 = R(\mathfrak{C} + \varepsilon) \quad \text{and} \quad |R_1| = t_1, |R_2| = t_2.$$

In order to be allowed to use Lemma 5.2, we will assume that the orbits of points α, β, γ and δ are mutually disjoint. Otherwise, we use Proposition 3.3 to find a modified interval $\tilde{I}$ where this is satisfied and $\mathrm{It}_{\tilde{I}} = \mathrm{It}_I$.

(i) Let $\mathfrak{A} < \mathfrak{B} < \mathfrak{D} < \mathfrak{C}$. We know that $R(x)$ is constant on the intervals $[\gamma, \mathfrak{A})$, $[\mathfrak{A}, \mathfrak{B})$, $[\mathfrak{B}, \mathfrak{D})$, $[\mathfrak{D}, \mathfrak{C})$, and $[\mathfrak{C}, \delta)$. By definition $R(x) = R_2$ for $x \in [\mathfrak{C}, \delta)$ and $R(x) = R_1$ for $x \in [\mathfrak{B}, \mathfrak{D})$. We can derive from rule (e) of Lemma 5.2 that if $x \in [\mathfrak{D}, \mathfrak{C})$, then $R(x) = R_1 R_2$. Further, we use rule (c) to show that $R(x) = \omega_{CA\to B}(R_1)$ for $x \in [\mathfrak{A}, \mathfrak{B})$ and further by applying rule (a), we obtain that $R(x) = \omega_{B\to AC}\big(\omega_{CA\to B}(R_1)\big)$ for $x \in [\gamma, \mathfrak{A})$. Summarized,

$$R(x) = \begin{cases} \omega_{B\to AC}\big(\omega_{CA\to B}(R_1)\big) & \text{for } x \in [\gamma, \mathfrak{A}) \\ \omega_{CA\to B}(R_1) & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{D}) \\ R_1 R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta). \end{cases}$$

It is easy to show that the lengths of the above I-itineraries are t_1, $t_1 - 1$, t_1, $t_1 + t_2$, t_2, respectively. Setting $r_1 = t_1 - 1$ and $r_2 = t_2$, we obtain the desired return times.

The proofs of the other cases are analogous, we state the results in terms of R_1 and R_2.

(ii) Let $\mathfrak{D} < \mathfrak{C} < \mathfrak{A} < \mathfrak{B}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_2 R_1 & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{A}) \\ \omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ \omega_{B\to CA}\big(\omega_{AC\to B}(R_2)\big) & \text{for } x \in [\mathfrak{B}, \delta), \end{cases}$$

with lengths t_1, $t_1 + t_2$, t_2, $t_2 - 1$, t_2, respectively. We set $r_1 = t_1$ and $r_2 = t_2 - 1$.

(iii) Let $\mathfrak{B} < \mathfrak{A} < \mathfrak{D} < \mathfrak{C}$. A discussion as above leads to

$$R(x) = \begin{cases} \omega_{CA\to B}\big(\omega_{B\to AC}(R_1)\big) & \text{for } x \in [\gamma, \mathfrak{B}) \\ \omega_{B\to AC}(R_1) & \text{for } x \in [\mathfrak{B}, \mathfrak{A}) \\ R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{D}) \\ R_1 R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta), \end{cases}$$

with the corresponding lengths t_1, $t_1 + 1$, t_1, $t_1 + t_2$, t_2, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

(iv) Let $\mathfrak{D} < \mathfrak{C} < \mathfrak{B} < \mathfrak{A}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_2R_1 & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{B}) \\ \omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{B}, \mathfrak{A}) \\ \omega_{AC\to B}\big(\omega_{B\to CA}(R_2)\big) & \text{for } x \in [\mathfrak{A}, \delta), \end{cases}$$

with lengths t_1, t_1+t_2, t_2, t_2+1, t_2, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

(v) Let $\mathfrak{A} < \mathfrak{D} < \mathfrak{B} < \mathfrak{C}$. We obtain

$$R(x) = \begin{cases} \omega_{B\to AC}(R_1) & \text{for } x \in [\gamma, \mathfrak{A}) \\ R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{D}) \\ R_1R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{B}) \\ R_2\omega_{B\to AC}(R_1) & \text{for } x \in [\mathfrak{B}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta), \end{cases}$$

with lengths t_1+1, t_1, t_1+t_2, t_1+t_2+1, t_2, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

(vi) Let $\mathfrak{D} < \mathfrak{A} < \mathfrak{C} < \mathfrak{B}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_1\omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{D}, \mathfrak{A}) \\ R_2R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{B}) \\ \omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{B}, \delta), \end{cases}$$

with lengths t_1, t_1+t_2+1, t_1+t_2, t_2, t_2+1, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

(vii) Let $\mathfrak{B} < \mathfrak{D} < \mathfrak{A} < \mathfrak{C}$. We obtain

$$R(x) = \begin{cases} \omega_{CA\to B}(R_1) & \text{for } x \in [\gamma, \mathfrak{B}) \\ R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{D}) \\ R_1R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{A}) \\ R_2\omega_{CA\to B}(R_1) & \text{for } x \in [\mathfrak{A}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta), \end{cases}$$

with lengths t_1-1, t_1, t_1+t_2, t_1+t_2-1, t_2, respectively. We set $r_1 = t_1-1$ and $r_2 = t_2$.

(viii) Let $\mathfrak{D} < \mathfrak{B} < \mathfrak{C} < \mathfrak{A}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_1\omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{D}, \mathfrak{B}) \\ R_2R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{A}) \\ \omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{A}, \delta), \end{cases}$$

with lengths t_1, t_1+t_2-1, t_1+t_2, t_2, t_2-1, respectively. We set $r_1 = t_1$ and $r_2 = t_2-1$.

11

(ix) Let $\mathfrak{A} < \mathfrak{D} < \mathfrak{C} < \mathfrak{B}$. We obtain

$$R(x) = \begin{cases} \omega_{B\to AC}(R_1) & \text{for } x \in [\gamma, \mathfrak{A}) \\ R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{D}) \\ R_1\omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{B}) \\ \omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{B}, \delta), \end{cases}$$

with lengths t_1+1, t_1, t_1+t_2+1, t_2, t_2+1, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

(x) Let $\mathfrak{B} < \mathfrak{D} < \mathfrak{C} < \mathfrak{A}$. We obtain

$$R(x) = \begin{cases} \omega_{CA\to B}(R_1) & \text{for } x \in [\gamma, \mathfrak{B}) \\ R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{D}) \\ R_1\omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \mathfrak{A}) \\ \omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{A}, \delta), \end{cases}$$

with lengths t_1-1, t_1, t_1+t_2-1, t_2, t_2-1, respectively. We set $r_1 = t_1-1$ and $r_2 = t_2-1$.

(xi) Let $\mathfrak{D} < \mathfrak{A} < \mathfrak{B} < \mathfrak{C}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_1R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{A}) \\ \omega_{CA\to B}(R_2R_1) & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ R_2R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta), \end{cases}$$

with lengths t_1, t_1+t_2, t_1+t_2-1, t_1+t_2, t_2, respectively. We set $r_1 = t_1-1$ and $r_2 = t_2$.

(xii) Let $\mathfrak{D} < \mathfrak{B} < \mathfrak{A} < \mathfrak{C}$. We obtain

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{D}) \\ R_1R_2 & \text{for } x \in [\mathfrak{D}, \mathfrak{B}) \\ \omega_{B\to AC}(R_2R_1) & \text{for } x \in [\mathfrak{B}, \mathfrak{A}) \\ R_2R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{C}) \\ R_2 & \text{for } x \in [\mathfrak{C}, \delta), \end{cases}$$

with lengths t_1, t_1+t_2, t_1+t_2+1, t_1+t_2, t_2, respectively. We set $r_1 = t_1$ and $r_2 = t_2$.

□

Remark 5.4. When describing the I-itineraries using the words R_1, R_2, we could apply the rules of Lemma 5.2 in a different order. By doing so, we would obtain the itineraries expressed differently, which yields interesting relations between words R_1, R_2. For example, in the case (ix), we derive that the I-itinerary of $x \in [\mathfrak{D}, \mathfrak{C})$ is $R(x) = R_1\omega_{B\to CA}(R_2) = R_2\omega_{B\to AC}(R_1)$.

Note also the symmetries between the cases (i) and (ii), (iii) and (iv), (v) and (vi), (vii) and (viii), in consequence of Proposition 3.4. Indeed, if we exchange pair of points $\mathfrak{D} \leftrightarrow \mathfrak{C}$, $\mathfrak{B} \leftrightarrow \mathfrak{A}$, letters $A \leftrightarrow C$, and finally the inequalities $<$ and $>$, we obtain a symmetric situation in the list of cases we discussed in the proof. In this sense, each of cases (ix) up to (xii) is symmetric to itself.

γ	δ	type	lengths
$\frac{6}{25}$	$\frac{99}{100}$	$\mathfrak{A} < \mathfrak{B} < \mathfrak{D} < \mathfrak{C}$	$[2,1,2,3,1]$
$\frac{29}{100}$	$\frac{71}{100}$	$\mathfrak{D} < \mathfrak{C} < \mathfrak{A} < \mathfrak{B}$	$[1,15,14,13,14]$
$\frac{77}{100}$	$\frac{4}{5}$	$\mathfrak{B} < \mathfrak{A} < \mathfrak{D} < \mathfrak{C}$	$[88,89,88,109,21]$
$\frac{7}{25}$	$\frac{3}{4}$	$\mathfrak{D} < \mathfrak{C} < \mathfrak{B} < \mathfrak{A}$	$[1,13,12,13,12]$
$\frac{1}{100}$	$\frac{3}{4}$	$\mathfrak{A} < \mathfrak{D} < \mathfrak{B} < \mathfrak{C}$	$[2,1,2,3,1]$
$\frac{1}{100}$	$\frac{29}{100}$	$\mathfrak{D} < \mathfrak{A} < \mathfrak{C} < \mathfrak{B}$	$[2,14,13,11,12]$
$\frac{1}{4}$	$\frac{99}{100}$	$\mathfrak{B} < \mathfrak{D} < \mathfrak{A} < \mathfrak{C}$	$[1,2,3,2,1]$
$\frac{71}{100}$	$\frac{99}{100}$	$\mathfrak{D} < \mathfrak{B} < \mathfrak{C} < \mathfrak{A}$	$[2,13,14,12,11]$
$\frac{1}{25}$	$\frac{37}{50}$	$\mathfrak{A} < \mathfrak{D} < \mathfrak{C} < \mathfrak{B}$	$[2,1,4,2,3]$
$\frac{29}{100}$	$\frac{99}{100}$	$\mathfrak{B} < \mathfrak{D} < \mathfrak{C} < \mathfrak{A}$	$[1,2,4,3,2]$
$\frac{1}{100}$	$\frac{99}{100}$	$\mathfrak{D} < \mathfrak{A} < \mathfrak{B} < \mathfrak{C}$	$[1,2,1,2,1]$
$\frac{1}{4}$	$\frac{3}{4}$	$\mathfrak{D} < \mathfrak{B} < \mathfrak{A} < \mathfrak{C}$	$[1,12,13,12,11]$

Table 1: The cases (i)–(xii) from the proof of Theorem 5.1 for $\alpha = \frac{1}{5}\sqrt{5} - \frac{1}{5}$, $\beta = -\frac{1}{6}\sqrt{5} + \frac{2}{3}$ as in Example 5.5. The endpoints of the interval $I = [\gamma, \delta)$ are in the first and second column. The last column contains a list of lengths of I-itineraries of all 5 subintervals of I starting from the leftmost one.

Example 5.5. Set $\alpha = \frac{1}{5}\sqrt{5} - \frac{1}{5}$ and $\beta = -\frac{1}{6}\sqrt{5} + \frac{2}{3}$. Table 1 shows 12 choices of $I = [\gamma, \delta)$ which produce 12 distinct orders of the points $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ and $\mathfrak{D}$, shown in the third column. The last column contains the respective lengths of the 5 distinct I-itineraries.

Let us describe in detail one of the cases, namely the case $\mathfrak{B} < \mathfrak{D} < \mathfrak{C} < \mathfrak{A}$. The induced interval is determined by setting $\gamma = \frac{29}{100}$ and $\delta = \frac{99}{100}$. One can verify that

$$\begin{aligned}
\mathfrak{B} = T^{-0}(\beta) &= \frac{1}{6}\sqrt{5} + \frac{1}{3} \approx 0.706011329583298;\\
\mathfrak{D} = T^{-2}(\delta) &\frac{11}{30}\sqrt{5} + \frac{37}{300} \approx 0.943224925083256;\\
\mathfrak{C} = T^{-3}(\gamma) &= \frac{8}{15}\sqrt{5} - \frac{73}{300} \approx 0.949236254666554;\\
\mathfrak{A} = T^{-1}(\alpha) &= \frac{11}{30}\sqrt{5} + \frac{2}{15} \approx 0.953224925083256.
\end{aligned}$$

It corresponds to the case (x) in the proof of Theorem 5.1 with $R_1 = CA$ and $R_2 = CAC$. The I-itinerary of a point $x \in I = [\gamma, \delta)$ is

$$R(x) = \begin{cases} B & \text{for } x \in [\gamma, \mathfrak{B}) \\ CA & \text{for } x \in [\mathfrak{B}, \mathfrak{D}) \\ CACB & \text{for } x \in [\mathfrak{D}, \mathfrak{C}) \\ CAC & \text{for } x \in [\mathfrak{C}, \mathfrak{A}) \\ CB & \text{for } x \in [\mathfrak{A}, \delta) \end{cases}$$

6 Gaps and distance theorems

Let us reinterpret the statement of Theorem 5.1 in point of view of three gap and three distance theorems which are narrowly connected with exchange of two intervals. Under the name three gap theorem one usually refers to the description of gaps between neighbouring elements of the set

$$\mathcal{G}(\alpha,\delta) := \big\{n \in \mathbb{N} \colon \{n\alpha\} < \delta\big\} \subset \mathbb{N}\,,$$

where $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $\delta \in (0,1)$ and $\{x\} = x - \lfloor x \rfloor$ stands for the fractional part of x, see [28]. Sometimes one uses a more general formulation, namely the set

$$\mathcal{G}(\alpha,\rho,\gamma,\delta) := \big\{n \in \mathbb{N} : \gamma \le \{n\alpha + \rho\} < \delta\big\} \subset \mathbb{N}\,,$$

where moreover $\rho \in \mathbb{N}$, $0 \le \gamma < \delta < 1$. The three gap theorem states that there exist integers r_1, r_2 such that gaps between neighbours in $\mathcal{G}(\alpha,\rho,\gamma,\delta)$ take at most three values, namely in the set $\{r_1, r_2, r_1 + r_2\}$.

Let us interpret the three gap theorem in the frame of exchange of two intervals $J_0 = [0, 1-\alpha)$, $J_1 = [1-\alpha, 1)$. The transformation $T : [0,1) \to [0,1)$ is of the form

$$T(x) = \begin{cases} x + \alpha & \text{for } x \in [0, 1-\alpha)\,, \\ x + \alpha - 1 & \text{for } x \in [1-\alpha, 1)\,, \end{cases} \qquad \text{i.e.} \qquad T(x) = \{x + \alpha\}\,.$$

Therefore we can write

$$\mathcal{G}(\alpha,\rho,\gamma,\delta) := \big\{n \in \mathbb{N} \colon T^n(\rho) \in [\gamma,\delta)\big\}\,, \tag{12}$$

and the gaps in this set correspond to return times to the interval $[\gamma,\delta)$ under the transformation T.

Our Theorem 5.1 is an analogue of the three gap theorem in the form (12) generalized for the case when the transformation T is a non-degenerate 3iet. We see that there are 5 gaps, but still expressed using two basic values r_1, r_2.

The so-called three distance theorem focuses on distances between neighbours of the set

$$\mathcal{D}(\alpha,\rho,N) := \big\{\{\alpha n + \rho\} \colon n \in \mathbb{N},\ n < N\big\} \subset [0,1)\,.$$

The three distance theorem ensures existence of $\Delta_1, \Delta_2 > 0$ such that distances between neighbours in $\mathcal{D}(\alpha,\rho,N)$ take at most three values, namely in $\{\Delta_1, \Delta_2, \Delta_1 + \Delta_2\}$.

In the framework of 2iet T, we can write for the distances

$$\mathcal{D}(\alpha,\rho,N) := \big\{T^n(\rho) \colon n \in \mathbb{N},\ n < N\big\} \subset [0,1)\,. \tag{13}$$

We could try to study the analogue of the three distance theorem in the form (13) for exchanges of three intervals. In fact, it can be derived from the results of [12] that if T is a 3iet with discontinuity points α, β, then

$$\mathcal{D}(\alpha,\beta,\rho,N) := \big\{T^n(\rho) \colon n \in \mathbb{N},\ n < N\big\}$$

has again at most three distances Δ_1, Δ_2, and $\Delta_1 + \Delta_2$ for some positive Δ_1, Δ_2.

The three distance theorem can also be used to derive that the frequencies of factors of length n in a Sturmian word take at most three values. Recall that the frequency of a factor w in the infinite word $\mathbf{u} = u_0 u_1 u_2 \ldots$ is given by

$$\operatorname{freq}(w) := \lim_{N\to\infty} \frac{1}{N}\big(\#\{0 \le i < N \colon w \text{ is a prefix of } u_i u_{i+1} \ldots\}\big)\,,$$

if the limit exists.

It is a well known fact that the frequencies of factors of length n in a coding of an exchange of intervals are given by the lengths of cylinders corresponding to the factors. The boundary points of these cylinders are $T^{-j}(1-\alpha)$, for $j = 0, \ldots, n-1$. Consequently, the distances in the set $\mathcal{D}(\alpha, 1-\alpha, N)$ are precisely the frequencies of factors, and the three distance theorem implies the well known fact that Sturmian words have for each n only three values of frequencies of factors of length n, namely $\varrho_1, \varrho_2, \varrho_1 + \varrho_2$.

The frequencies of factors of length n in 3iet words are given by distances between neighbours of the set

$$\left\{T^{-n}(\alpha) \colon n \in \mathbb{N},\ n < N\right\} \cup \left\{T^{-n}(\beta) \colon n \in \mathbb{N},\ n < N\right\}.$$

In [7] it is shown, based on the study of Rauzy graphs, that the number of distinct values of frequencies in infinite words with reversal closed language satisfies

$$\#\{\mathrm{freq}(w) \colon w \in \mathcal{L}(\mathbf{u}), |w| = n\} \leq 2\Big(\mathcal{C}_{\mathbf{u}}(n) - \mathcal{C}_{\mathbf{u}}(n-1)\Big) + 1\,,$$

which in case of 3iet words reduces to ≤ 5. Paper [11] shows, that the set of integers n for which this bound is achieved, is of density 1 in $\mathbb{N}$.

7 Description of the case of three I-itineraries

The cases (i) – (xii) in the proof of Theorem 5.1 correspond to the generic instances of a subinterval I in a non-degenerate 3iet which lead to 5 different I-itineraries. Let us focus on the cases where, on the contrary, the set of I-itineraries has only 3 elements. First we recall two reasons why such cases are interesting.

For a factor w from the language of a non-degenerate 3iet transformation T, denote

$$[w] = \{\rho \in [0,1) \colon w \text{ is a prefix of } \mathbf{u}_\rho\}\,.$$

It is easy to see that $[w]$ – usually called the cylinder of w – is a semi-closed interval and its boundaries belong to the set $\{T^{-i}(z) \colon 0 \leq i < n, z \in \{\alpha, \beta\}\}$. Clearly, a factor v is a return word to the factor w if and only if v is a $[w]$-itinerary. It is well known [33] that any factor of an infinite word coding a non-degenerate 3iet has exactly three return words and thus the set $It_{[w]}$ has three elements.

The second reason why to study intervals I yielding three I-itineraries is that any morphism fixing a non-degenerate 3iet word corresponds to such an interval I. Details of this correspondence will be explained in Section 8.

Proposition 7.1. *Let T be a non-degenerate 3iet and let $I = [\gamma, \delta) \subset [0,1)$ be such that $\#\mathrm{It}_I = 3$. One of the following cases occurs:*

(i) $\mathfrak{B} = \mathfrak{D} < \mathfrak{A} = \mathfrak{C}$ *and*

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{B}) \\ \omega_{B\to CA}(R_1R_2) = \omega_{B\to AC}(R_2R_1) & \text{for } x \in [\mathfrak{B}, \mathfrak{A}) \\ R_2 & \text{for } x \in [\mathfrak{A}, \delta) \end{cases}$$

(ii) $\mathfrak{A} = \mathfrak{D} < \mathfrak{B} = \mathfrak{C}$ *and*

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{A}) \\ \omega_{AC\to B}(R_1R_2) = \omega_{CA\to B}(R_2R_1) & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ R_2 & \text{for } x \in [\mathfrak{B}, \delta) \end{cases}$$

(iii) $\mathfrak{B} < \mathfrak{A} = \mathfrak{C} = \mathfrak{D}$ *and*

$$R(x) = \begin{cases} \omega_{CA\to B}(R_1) & \text{for } x \in [\gamma, \mathfrak{B}) \\ R_1 & \text{for } x \in [\mathfrak{B}, \mathfrak{A}) \\ R_2 & \text{for } x \in [\mathfrak{A}, \delta) \end{cases}$$

(iv) $\mathfrak{B} = \mathfrak{C} = \mathfrak{D} < \mathfrak{A}$ *and*

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{A}) \\ R_2 & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ \omega_{AC\to B}(R_2) & \text{for } x \in [\mathfrak{B}, \delta) \end{cases}$$

(v) $\mathfrak{A} = \mathfrak{C} = \mathfrak{D} < \mathfrak{B}$ *and*

$$R(x) = \begin{cases} R_1 & \text{for } x \in [\gamma, \mathfrak{A}) \\ R_2 & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ \omega_{B\to CA}(R_2) & \text{for } x \in [\mathfrak{B}, \delta) \end{cases}$$

(vi) $\mathfrak{A} < \mathfrak{B} = \mathfrak{C} = \mathfrak{D}$ *and*

$$R(x) = \begin{cases} \omega_{B\to AC}(R_1) & \text{for } x \in [\gamma, \mathfrak{A}) \\ R_1 & \text{for } x \in [\mathfrak{A}, \mathfrak{B}) \\ R_2 & \text{for } x \in [\mathfrak{B}, \delta) \end{cases}$$

Sketch of a proof. Since by Lemma 4.1 the subintervals of I corresponding to the same itinerary are delimited by the points $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$ and $\mathfrak{D}$, we may have $\#\mathrm{It}_I = 3$ only if some of these points coincide, more precisely if $\#\{\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \mathfrak{D}\} = 2$. The non-degeneracy of the considered 3iet implies that always $\mathfrak{A} \neq \mathfrak{B}$, which further limits the discussion.

The six cases listed in the statement are the possibilities of how this may happen, respecting the condition $\mathfrak{D} < \mathfrak{C}$ or $\mathfrak{D} = \mathfrak{C}$. In order to describe the itineraries, denote again

$$R_1 = R(\mathfrak{D} - \varepsilon) \quad \text{and} \quad R_2 = R(\mathfrak{C} + \varepsilon)$$

for $\varepsilon > 0$ sufficiently small. One can then follow the ideas of the proof of Lemma 5.2. $\square$

Let us apply Proposition 7.1 in order to provide the description of return words to factors of a non-degenerate 3iet word. If a factor w has a unique right prolongation in the language $\mathcal{L}(T)$, i.e. there exists only one letter $a \in \mathcal{A}$ such that $wa \in \mathcal{L}(T)$, then the set of return words to w and the set of return words to wa coincide. And (almost) analogously, if a factor w has a unique left prolongation in the language $\mathcal{L}(T)$, say aw for some $a \in \mathcal{A}$, then a word v is a return word to w if and only if ava^{-1} is a return word to aw. Consequently, to describe the structure of return words to a given factor w, we can restrict to factors which have at least two right and at least two left prolongations. Such factors are called *bispecial.* It is readily seen that the language of an aperiodic recurrent infinite word $\mathbf{u}$ contains infinitely many bispecial factors. Before giving the description of return words to bispecial factors, we state the following lemma.

Lemma 7.2. *Let w belong to language of a non-degenerate 3iet T. Denote $n = |w|$. For the cylinder of its reversal $\overline{w}$, one has*

$$[\overline{w}] = \overline{T^n([w])}\,.$$

16

Proof. According to definition of $[w]$, for each $[w]$-itinerary r, the word rw belongs to the language and w occurs in rw exactly twice, as a prefix and as a suffix. In other words r is a return word to w. Moreover, $[w]$ is the maximal (with respect to inclusion) interval with this property. According to Remark 3.2, if r is an $[w]$-itinerary, then the word $w^{-1}rw$ is an $T^n([w])$-itinerary. Applying Proposition 3.4 to the interval $T^n([w])$ we obtain that $s := \overline{w^{-1}rw}$ is an $\overline{T^n([w])}$-itinerary. Since the word $s\overline{w} = \overline{rw}$ has a prefix $\overline{w}$ and a suffix $\overline{w}$, with no other occurrences of $\overline{w}$, the word s is a return word to $\overline{w}$ and thus by definition of the cylinder, $s = \overline{w^{-1}rw}$ belongs to $[\overline{w}]$-itinerary for any $\overline{T^n([w])}$-itinerary s. From the maximality of the cylinder we have $\overline{T^n([w])} \subset [\overline{w}]$. Since lengths of the intervals $[w]$ and $T^n([w])$ coincide we have, in particular, that the length of interval $[w]$ is less or equal to the length of the interval $[\overline{w}]$. But from the symmetry of the role w and $\overline{w}$, their length must be equal and thus $\overline{T^n([w])} = [\overline{w}]$. □

The language of T contains two types of bispecial factors: palindromic and non-palindromic. In [10], Ferenczi, Holton and Zamboni studied the structure of return words to non-palindromic bispecial factors. The following proposition completes this description.

Proposition 7.3. *Let w be a bispecial factor. If w is a palindrome, then its return words are described by the cases (i) and (ii) of Proposition 7.1. If w is not a palindrome, then its return words are described by the cases (iii) – (iv) of Proposition 7.1.*

Proof. Let w be a bispecial factor. If w is not a palindrome, the claim follows from Theorem 4.6 of [10].

Assume w is a palindrome and let $[w] = [T^{-\ell}(L), T^{-r}(R))$ with $L, R \in \{\alpha, \beta\}$ and $0 \le \ell, r < |w|$. By Lemma 7.2 we have $[\overline{w}] = \overline{T^{|w|}([w])}$. Since $w = \overline{w}$, we have $I_w = I_{\overline{w}}$, and thus

$$I_w = [T^{-\ell}(L), T^{-r}(R)) = [1 - T^{n-r}(R), 1 - T^{n-\ell}(L)) = I_{\overline{w}}.$$

Since the considered 3iet is non-degenerate, the parameters α, β satisfy (8). Consequently, the equation $T^{-\ell}(L) = 1 - T^{n-r}(R)$ has a solution if and only if $R \neq L$. Thus, we have neither $\mathfrak{A} = \mathfrak{C} = \mathfrak{D}$ nor $\mathfrak{B} = \mathfrak{C} = \mathfrak{D}$ and we are in the case (i) or (ii) of Proposition 7.1. □

8 Substitution invariance and conjugation of substitutions

Let us recall the relation of induction to a subinterval I to substitution invariance of 3iet words. Let I be an interval $I \subset [0, 1)$ such that the set It_I of I-itineraries has three elements, say R_A, R_B and R_C. For every $\rho \in I$, the infinite word $\mathbf{u}_\rho$ can be written as a concatenation of words R_A, R_B and R_C. For a letter $Y \in \{A, B, C\}$ denote $I_Y = \{x \in I : R(x) = R_Y\}$. Obviously, $I = I_A \cup I_B \cup I_C$, and the induced mapping T_I is an exchange of these three intervals. The order of the words R_A, R_B and R_C in the concatenation is determined by the iterations of $T_I(\rho)$.

Suppose that T_I is homothetic to T. Recall that mappings $f : I_f \to I_f$ and $g : I_g \to I_g$ are homothetic if there exists an affine bijection $\Phi : I_f \to I_g$ with $\Phi(x) = \lambda x + \mu$ such that

$$\Phi f(x) = g\Phi(x) \quad \text{for all } x \in I_f\,. \tag{14}$$

This means that f and g behave in the same way, up to a scaling factor λ and a shift μ of the domains I_f and I_g. In other words, the graphs of the mappings f and g are the same, up to their scale and placing. The homothety of T and T_I implies that $\Phi(J_Y) = I_Y$ for all $Y \in \{A, B, C\}$. From (14), we derive for every $k \in \mathbb{N}$ that $\Phi T^k(x) = T_I^k \Phi(x)$ for $x \in [0, 1)$, and thus $\Phi T^k(\rho) = T_I^k(\rho)$ whenever

$$\Phi(\rho) = \rho\,, \tag{15}$$

i.e., ρ is the homothety center. From the relation $\Phi(J_Y) = I_Y$ it follows that the k-th element in the concatenation of itineraries R_A, R_B and R_C is equal to R_Y if and only if the k-th letter in the infinite word $\mathbf{u}_\rho$ is equal to Y. This is equivalent to saying that the infinite word $\mathbf{u}_\rho$ is invariant under the substitution η given by

$$\eta(A) = R_A,\ \eta(B) = R_B,\ \eta(C) = R_C\,. \tag{16}$$

We conclude that the existence of an interval I with three itineraries and T_I homothetic to T leads to a substitution fixing a 3iet word whose intercept is the homothety center ρ. In fact, the converse holds, too, as shown in [5]. We summarize both statements as follows.

Theorem 8.1 ([5]). *Let ξ be a primitive substitution over $\{A, B, C\}$ with incidence matrix M and let T be a non-degenerate 3iet. The substitution ξ fixes the word $\mathbf{u}_\rho$ coding the orbit of a point $\rho \in [0, 1)$ under T if and only if there exists an interval $I \subset [0, 1)$ with I-itineraries $It_I = \{R_A, R_B, R_C\}$ such that T_I is homothetic to T, ρ is the homothety center, and the substitution η given by*

$$\eta = \begin{cases} \xi & \textit{if no eigenvalue of } M \textit{ belongs to } (-1,0), \\ \xi^2 & \textit{otherwise,} \end{cases}$$

satisfies $\eta(A) = R_A$, $\eta(B) = R_B$ and $\eta(C) = R_C$.

Let us mention that the scaling factor $\lambda \in (0, 1)$ in the homothety mapping $\Phi(x) = \lambda x + \mu$ is equal to the length of the interval $I = [\gamma, \delta)$, i.e., $\lambda = \delta - \gamma$, and the shift μ is equal to the left end-point of the interval I, namely γ. Moreover, it is related to the intercept ρ of an infinite word $\mathbf{u}_\rho$ in the following way: one has $\mu = \gamma = \rho(1 - \lambda)$, as follows from (15). In fact, λ is an eigenvalue of the incidence matrix of η. It follows from [5] that if ξ has such an eigenvalue, then the choice $\eta = \xi$ is sufficient. Otherwise, the incidence matrix of ξ^2 has such an eigenvalue.

By Theorem 8.1, if $\mathbf{u}_\rho$ is invariant under a substitution, we find an interval I such that T_I is homothetic to T. If $I' = T(I)$ is again an interval, then $T_{I'}$ is also homothetic to T, and the I'-itineraries change with respect to the I-itineraries, as described in Remark 3.2. To show the relation of the corresponding substitutions, we need the following definition.

Definition 8.2. Let φ and ψ be morphisms over $\mathcal{A}^*$ and let $w \in \mathcal{A}^*$ be a word such that $w\varphi(a) = \psi(a)w$ for every letter $a \in \mathcal{A}$. The morphism φ is said to be a *left conjugate* of ψ and ψ is said to be a *right conjugate* of φ. We write $\varphi \lhd \psi$. If φ is a left or right conjugate of ψ, then we say φ is *conjugate* to ψ. If the only left conjugate of φ is φ itself, then φ is called the *leftmost conjugate* of ψ and we write $\varphi = \psi_L$. If the only right conjugate of ψ is ψ itself, then ψ is called the *rightmost conjugate* of φ and we write $\psi = \varphi_R$.

Note that given a substitution ξ, its leftmost and rightmost conjugates ξ_L and ξ_R may not exist. If this happens, it can be shown that its fixed point is a periodic word. All the substitutions considered here thus possess their leftmost and rightmost conjugates.

Proposition 8.3. *Let $\mathbf{u}_\rho$ be a 3iet word coding the orbit of the point $\rho \in [0, 1)$ under a non-degenerate 3iet T. Moreover, assume that $\mathbf{u}_\rho$ is a fixed point of a primitive substitution η such that the corresponding interval I of Theorem 8.1 is of length λ. Let η' be a left conjugate of η, i.e., $\eta(a)w = w\eta'(a)$ for some word $w \in \mathcal{A}^*$. The morphism η' fixes the infinite word $\mathbf{u}_{\rho'}$ with ρ' satisfying*

$$(1 - \lambda)\rho' = T^n\big((1 - \lambda)\rho\big)\,, \qquad \textit{where } n = |w|. \tag{17}$$

Moreover, the interval I' corresponding to η' by Theorem 8.1 satisfies $I' = T^n(I)$.

Proof. Suppose that w is a letter, i.e., $w = X \in \mathcal{A}$. Necessarily, the words $\eta(a)$ start with the letter X for all $a \in \mathcal{A}$. This means for the interval I that $I \subset J_X$. According to Remark 3.2, the interval $I' = T(I)$ has three I'-itineraries. Moreover, the induced mapping $T_{I'}$ is also homothetic to T. Denote $I = [\gamma, \delta)$. The homothety between the transformations T and T_I is achieved by the map $\Phi(x) = \lambda x + \gamma$. The homothety between T and $T_{I'}$ is the map $\Phi'(x) = \lambda x + T(\gamma)$. Since the intercepts ρ and ρ' are by (15) fixed by the homotheties Φ and Φ', respectively, we have

$$\Phi(\rho) = \lambda\rho + \gamma = \rho \qquad \text{and} \qquad \Phi'(\rho') = \lambda\rho' + T(\gamma) = \rho' . \tag{18}$$

Eliminating γ, we obtain

$$(1-\lambda)\rho' = T(\gamma) = T\big((1-\lambda)\rho\big) .$$

Since conjugation by any word w can be performed letter by letter, the proof is finished. $\square$

Remark 8.4. Note that the first of equalities in (18) implies for the left boundary point γ of the interval I that $\gamma = (1-\lambda)\rho$. If w is as in Proposition 8.3, then for for $0 \le k < n = |w|$, the iterations $T^k(I)$ are all intervals, and hence the coding $\mathbf{u}_x$ of every point $x \in I$ starts with the same prefix w.

In the following, we will also need to see the relation of the substitution η corresponding to the interval $I = [\gamma, \delta)$ with the substitution corresponding to the interval $\overline{I} = [1-\delta, 1-\gamma)$. It turns out that it is the mirror substitution of η, defined in general as follows. For a morphism $\xi : \mathcal{A} \to \mathcal{A}$, we define the morphism $\overline{\xi} : \mathcal{A} \to \mathcal{A}$ by $\overline{\xi}(a) = \overline{\xi(a)}$ for $a \in \mathcal{A}$.

Proposition 8.5. *Let $I \subset [0,1)$ be a left-closed right-open interval such that $\#It_I = 3$ and T_I is an exchange of three intervals with the permutation (321). The interval $\overline{I}$ satisfies $\#It_{\overline{I}} = 3$ and the induced map $T_{\overline{I}}$ is homothetic to T_I. If, moreover, T_I is homothetic to T and the substitution η corresponding to I fixes the infinite word $\mathbf{u}_\rho$, then the substitution corresponding to $\overline{I}$ is $\overline{\eta}$ and fixes the infinite word $\mathbf{u}_{\overline{\rho}}$, where $\overline{\rho} = 1-\rho$.*

Proof. Denote $It_I = \{R_1, R_2, R_3\}$ and $I_j = \{x \in I \colon R_I(x) = R_j\} = [\gamma_j, \delta_j)$ for $j = 1,2,3$ so that $I_1 < I_2 < I_3$. By Proposition 3.4, the $\overline{I}$-itineraries are $\overline{R_1}, \overline{R_2}$ and $\overline{R_3}$, where

$$I'_j = \{x \in \overline{I} \colon R_{\overline{I}}(x) = \overline{R_j}\} = [1-\delta'_j, 1-\gamma'_j) ,$$

where $[\gamma'_j, \delta'_j) = T_I[\gamma_j, \delta_j)$. Since T_I is an exchange of three intervals with the permutation (321) we have

$$T_I[\gamma_1, \delta_1) > T_I[\gamma_2, \delta_2) > T_I[\gamma_3, \delta_3) ,$$

and therefore $I'_1 < I'_2 < I'_3$. The induced map $T_{\overline{I}}$ is therefore an exchange of three intervals I'_1, I'_2 and I'_3 with the permutation (321) and since $|I'_j| = |I_j|$ for $j = 1,2,3$, the transformation $T_{\overline{I}}$ is homothetic to T_I.

Suppose that T_I is homothetic to the original 3iet T. By Theorem 8.1, there is a substitution η corresponding to the interval I and satisfying $\eta(A) = R_1$, $\eta(B) = R_2$ and $\eta(C) = R_3$. The mapping $T_{\overline{I}}$ is homothetic to T_I and thus also to T, the corresponding substitution η' satisfies $\eta'(A) = \overline{R_1}$, $\eta'(B) = \overline{R_2}$ and $\eta'(C) = \overline{R_3}$. We can see that $\eta' = \overline{\eta}$.

Let ρ be the intercept of the infinite word which is fixed by the substitution η. It is the center of homothety between T_I and T, i.e., it is the fixed point of the mapping $\Phi(x) = (\delta-\gamma)x + \gamma$. We have $\rho = (\delta-\gamma)\rho + \gamma$, which implies

$$\rho = \frac{\gamma}{1-\delta+\gamma} .$$

Similarly, the intercept $\overline{\rho}$ of the substitution $\overline{\eta}$ satisfies $\overline{\rho} = (\delta - \gamma)\overline{\rho} + 1 - \delta$, whence

$$\overline{\rho} = \frac{1-\delta}{1-\delta+\gamma} = 1 - \rho\,. \qquad \square$$

For a finite word w, $\mathrm{Fst}(w)$ and $\mathrm{Lst}(w)$ denote the first and last letters of w, respectively.

Remark 8.6. Let η be a primitive substitution given by Theorem 8.1 fixing a 3iet word. Necessarily, the first and the last letters of $\eta(A)$, $\eta(B)$ and $\eta(C)$ satisfy

$$\mathrm{Fst}\big(\eta(A)\big) \leq \mathrm{Fst}\big(\eta(B)\big) \leq \mathrm{Fst}\big(\eta(C)\big) \text{ and } \mathrm{Lst}\big(\eta(A)\big) \leq \mathrm{Lst}\big(\eta(B)\big) \leq \mathrm{Lst}\big(\eta(C)\big),$$

where we consider the order $A < B < C$. The inequalities for the first letters follow from the definition of an exchange of intervals, namely from the fact that the words $\eta(A)$, $\eta(B)$ and $\eta(C)$ are given as I-itineraries. By Proposition 8.5, the last letters of the words $\eta(A)$, $\eta(B)$ and $\eta(C)$ are the first letters of the words $\overline{\eta(A)}$, $\overline{\eta(B)}$ and $\overline{\eta(C)}$ which proves the second set of inequalities.

9 Ternarization

A characterization of 3iet words over the alphabet $\{A, B, C\}$ by morphic images of Sturmian words over $\{0, 1\}$ is derived in [5]. Let σ_{01} and σ_{10} be morphisms $\{A, B, C\} \to \{0, 1\}$ defined by

$$\begin{aligned}
\sigma_{01}(A) &= \sigma_{10}(A) = 0\,,\\
\sigma_{01}(B) &= 01,\ \sigma_{10}(B) = 10\,,\\
\sigma_{01}(C) &= \sigma_{10}(C) = 1\,.
\end{aligned}$$

Theorem 9.1 ([5]). *An infinite word $\mathbf{u} \in \{A, B, C\}^{\mathbb{N}}$ is a 3iet word if and only if $\sigma_{01}(\mathbf{u})$ and $\sigma_{10}(\mathbf{u})$ are Sturmian words.*

This theorem was an important tool in [4] for the description of substitutions η from Theorem 8.1 fixing a 3iet word. Since this result is important for our further considerations, we cite it as Theorem 9.5, but first, we need some definitions.

Definition 9.2. Let u and v be finite or infinite words over the alphabet $\{0, 1\}$. We say that u is *amicable* to v, and denote it by $u \propto v$, if there exists a ternary word w over $\{A, B, C\}$ such that $u = \sigma_{01}(w)$ and $v = \sigma_{10}(w)$. In such case, we set $w := \mathrm{ter}(u, v)$ and say that w is the *ternarization* of u and v.

Definition 9.3. Let $\varphi, \psi : \{0, 1\}^* \to \{0, 1\}^*$ be two morphisms. We say that φ is *amicable* to ψ, and denote it by $\varphi \propto \psi$, if the three following relations hold

$$\begin{aligned}
\varphi(0) &\propto \psi(0)\,,\\
\varphi(1) &\propto \psi(1)\,,\\
\varphi(01) &\propto \psi(10)\,.
\end{aligned} \tag{19}$$

The morphism $\eta : \{A, B, C\}^* \to \{A, B, C\}^*$ given by

$$\begin{aligned}
\eta(A) &:= \mathrm{ter}(\varphi(0), \psi(0))\,,\\
\eta(B) &:= \mathrm{ter}(\varphi(01), \psi(10))\,,\\
\eta(C) &:= \mathrm{ter}(\varphi(1), \psi(1))\,,
\end{aligned}$$

is called the ternarization of φ and ψ and denoted by $\eta := \mathrm{ter}(\varphi, \psi)$.

Remark 9.4. If u and v is a pair of amicable words over $\{0,1\}$, then $|u|_0 = |v|_0$ and $|u|_1 = |v|_1$. Consequently, if φ and ψ are two amicable morphisms, then they have the same incidence matrix.

Theorem 9.5 ([4])**.** *Let η be a primitive substitution from Theorem 8.1 fixing a non-degenerate 3iet word $\mathbf{u}$. There exist Sturmian morphisms φ and ψ having fixed points such that $\varphi \propto \psi$ and $\eta = \mathrm{ter}(\varphi,\psi)$. On the other hand, if φ and ψ are Sturmian morphisms with fixed points such that $\varphi \propto \psi$, then the morphism $\eta = \mathrm{ter}(\varphi,\psi)$ has a 3iet fixed point.*

Example 9.6. Consider the following Sturmian morphisms $\varphi, \psi : \{0,1\}^* \to \{0,1\}^*$,

$$\begin{aligned}\varphi(0) &= 0110101\\ \varphi(1) &= 01101\end{aligned}\,, \qquad \begin{aligned}\psi(0) &= 1010101\\ \psi(1) &= 10101\end{aligned}\,.$$

We verify the condition given in (19) and in the same time construct the ternarization $\eta = \mathrm{ter}(\varphi,\psi)$. We check that $\varphi(0) \propto \psi(0)$,

$$\begin{array}{r|c|c|c|c|c|c|} \varphi(0) = & 0\,1 & 1 & 0 & 1 & 0 & 1 \\ \psi(0) = & 1\,0 & 1 & 0 & 1 & 0 & 1 \\ \hline \eta(A) = & B & C & A & C & A & C \end{array}\,,$$

and $\varphi(1) \propto \psi(1)$

$$\begin{array}{r|c|c|c|c|} \varphi(1) = & 0\,1 & 1 & 0 & 1 \\ \psi(1) = & 1\,0 & 1 & 0 & 1 \\ \hline \eta(C) = & B & C & A & C \end{array}\,,$$

and lastly that $\varphi(01) \propto \psi(10)$

$$\begin{array}{r|c|c|c|c|c|c|c|c|c|} \varphi(01) = & 0\,1 & 1 & 0 & 1 & 0\,1 & 0\,1 & 1 & 0 & 1 \\ \psi(10) = & 1\,0 & 1 & 0 & 1 & 1\,0 & 1\,0 & 1 & 0 & 1 \\ \hline \eta(B) = & B & C & A & C & B & B & C & A & C \end{array}\,.$$

We obtained a ternarization of a pair of amicable Sturmian morphisms. By Theorem 9.5, η fixes a 3iet word.

Theorem 9.5 expresses the relation between substitutions fixing 3iet words and Sturmian morphisms. Recall that by a result of [26], all Sturmian morphisms with the same incidence matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ can be ordered by the relation $\lhd$ of conjugation into a chain

$$\xi_1 \lhd \xi_2 \lhd \cdots \lhd \xi_N\,, \qquad \text{where } N = a+b+c+d-1\,. \tag{20}$$

This implies that for every i and j such that $1 \le i < j \le N$, there exists a word $u \in \{0,1\}^*$ of length $j-i$ such that $u\xi_i(a) = \xi_j(a)u$ for $a \in \{0,1\}$.

Lemma 9.7. *Let $\eta = \mathrm{ter}(\varphi,\psi)$, $\eta' = \mathrm{ter}(\varphi',\psi')$, where φ,ψ and φ',ψ' are pairs of amicable Sturmian morphisms over the alphabet $\{0,1\}$. If $\eta \lhd \eta'$, then $\varphi \lhd \varphi'$ and $\psi \lhd \psi'$, and, moreover, $\varphi = \xi_i$, $\psi = \xi_j$, $\varphi' = \xi_{i'}$, $\psi' = \xi_{j'}$ where $j-i = j'-i'$.*

Proof. Since $\eta \lhd \eta'$, there exists a word $w \in \{A,B,C\}^*$ such that $w\eta_1(X) = \eta_2(X)w$ for every $X \in \{A,B,C\}$. We will show that then there exists an amicable pair of words $u,v \in \{0,1\}^*$ with $|u| = |v|$ such that $w = \operatorname{ter}(u,v)$ and

$$\begin{aligned} u\varphi(b) &= \varphi'(b)u \\ v\psi(b) &= \psi'(b)v \end{aligned} \quad \text{for } b \in \{0,1\}. \tag{21}$$

It suffices to prove this statement for w of length 1, i.e., $\operatorname{Fst}(\eta'(X)) = w$ for $X \in \{A,B,C\}$. If $w = A$, then necessarily $\operatorname{Fst}\big(\varphi'(b)\big) = \operatorname{Fst}\big(\psi'(b)\big) = \operatorname{Lst}\big(\varphi(b)\big) = \operatorname{Lst}\big(\psi(b)\big) = 0$ for $b \in \{0,1\}$. Therefore $u = v = 0$. If $w = C$, then similarly, $\operatorname{Fst}\big(\varphi'(b)\big) = \operatorname{Fst}\big(\psi'(b)\big) = \operatorname{Lst}\big(\varphi(b)\big) = \operatorname{Lst}\big(\psi(b)\big) = 1$ for $b \in \{0,1\}$, and thus $u = v = 1$. If $w = B$, then $\varphi'(0)$ and $\varphi'(1)$ have prefix 01, and $\psi'(0)$ and $\psi'(1)$ have prefix 10. Thus $u = 01$, $v = 10$ and clearly, $w = \operatorname{ter}(u,v)$.

Now φ and ψ are amicable Sturmian morphisms with the same incidence matrix M, and since φ', ψ' are their conjugates, they also have the same incidence matrix, thus $\varphi = \xi_i$, $\psi = \xi_j$, $\varphi' = \xi_{i'}$, $\psi' = \xi_{j'}$ for some $1 \le i,j,i',j' \le N$. The relation $|u| = j - i = j' - i'$ follows from (21). □

Lemma 9.8. *Let η be a primitive substitution given by Theorem 8.1 fixing a 3iet word. We have*

$$\begin{aligned} &\Big(\operatorname{Fst}\big(\eta_L(A)\big), \operatorname{Fst}\big(\eta_L(B)\big), \operatorname{Fst}\big(\eta_L(C)\big)\Big) = \\ &\Big(\operatorname{Lst}\big(\eta_R(A)\big), \operatorname{Lst}\big(\eta_R(B)\big), \operatorname{Lst}\big(\eta_R(C)\big)\Big) = (A,B,B)\,, \end{aligned}$$

or

$$\begin{aligned} &\Big(\operatorname{Fst}\big(\eta_L(A)\big), \operatorname{Fst}\big(\eta_L(B)\big), \operatorname{Fst}\big(\eta_L(C)\big)\Big) = \\ &\Big(\operatorname{Lst}\big(\eta_R(A)\big), \operatorname{Lst}\big(\eta_R(B)\big), \operatorname{Lst}\big(\eta_R(C)\big)\Big) = (B,B,C)\,. \end{aligned}$$

Proof. Since the words $\eta(A)$, $\eta(B)$ and $\eta(C)$ are I-itineraries for some interval I, the first letters of $\eta(A)$, $\eta(B)$ and $\eta(C)$ cannot all be distinct. On the contrary, suppose that the discontinuity points α and β of the transformation T belong to the interval I. It implies that these points coincide with the discontinuity points $\mathfrak{D}$ and $\mathfrak{C}$ of the induced map T_I. But this means that T_I is not homothetic to T, which is a contradiction.

By Remark 8.6, the only possibilities for the triple of letters

$$\Big(\operatorname{Fst}\big(\eta(A)\big), \operatorname{Fst}\big(\eta(B)\big), \operatorname{Fst}\big(\eta(C)\big)\Big) \quad \text{and} \quad \Big(\operatorname{Lst}\big(\eta(A)\big), \operatorname{Lst}\big(\eta(B)\big), \operatorname{Lst}\big(\eta(C)\big)\Big)$$

are (A,A,B), (A,B,B), (B,B,C), and (B,C,C).

We will prove the following claim: Let φ, ψ be the pair of amicable Sturmian morphisms over the alphabet $\{0,1\}$ such that $\eta = \operatorname{ter}(\varphi,\psi)$.

(i) If $\eta = \eta_L$, i.e., η is the leftmost conjugate of itself, then either

$$\begin{aligned} \psi = \psi_L \quad &\text{and} \quad \Big(\operatorname{Fst}\big(\eta_L(A)\big), \operatorname{Fst}\big(\eta_L(B)\big), \operatorname{Fst}\big(\eta_L(C)\big)\Big) = (A,B,B)\,, \quad \text{or} \\ \varphi = \varphi_L \quad &\text{and} \quad \Big(\operatorname{Fst}\big(\eta_L(A)\big), \operatorname{Fst}\big(\eta_L(B)\big), \operatorname{Fst}\big(\eta_L(C)\big)\Big) = (B,B,C)\,. \end{aligned}$$

(ii) If $\eta = \eta_R$, i.e., η is the rightmost conjugate of itself, then either

$$\begin{aligned} \varphi = \varphi_R \quad &\text{and} \quad \Big(\operatorname{Lst}\big(\eta_L(A)\big), \operatorname{Lst}\big(\eta_L(B)\big), \operatorname{Lst}\big(\eta_L(C)\big)\Big) = (A,B,B)\,, \quad \text{or} \\ \psi = \psi_R \quad &\text{and} \quad \Big(\operatorname{Lst}\big(\eta_L(A)\big), \operatorname{Lst}\big(\eta_L(B)\big), \operatorname{Lst}\big(\eta_L(C)\big)\Big) = (B,B,C)\,. \end{aligned}$$

In order to prove (i), let us discuss the case $\eta = \eta_L$ and $\mathrm{Fst}\big(\eta(A)\big) = A$, $\mathrm{Fst}\big(\eta(C)\big) = B$. Since $\eta(A) = \mathrm{ter}\big(\varphi(0), \psi(0)\big)$, necessarily $\mathrm{Fst}\big(\varphi(0)\big) = \mathrm{Fst}\big(\psi(0)\big) = 0$. As, $\eta(C) = \mathrm{ter}\big(\varphi(1), \psi(1)\big)$, necessarily $\mathrm{Fst}\big(\varphi(1)\big) = 0$ and $\mathrm{Fst}\big(\psi(1)\big) = 1$. Thus, the first letter of $\eta(B) = \mathrm{ter}\big(\varphi(01), \psi(10)\big)$ is B. Therefore the triple (A, A, B) is excluded. Moreover, we see that $\psi = \psi_L$. By the same reasoning, we proceed in the case that $\mathrm{Fst}\big(\eta(A)\big) = B$, $\mathrm{Fst}\big(\eta(C)\big) = C$ to exclude the triple (B, C, C) and prove $\varphi = \varphi_L$. The proof of (ii), i.e., the case $\eta = \eta_R$ is analogous.

Consider $\big(\mathrm{Fst}\big(\eta_L(A)\big), \mathrm{Fst}\big(\eta_L(B)\big), \mathrm{Fst}\big(\eta_L(C)\big)\big) = (A, B, B)$ and $\psi = \psi_L$. If $\xi_1 \triangleleft \cdots \triangleleft \xi_N$ are Sturmian morphisms of (20) with the same incidence matrix, then we have $\psi = \xi_1$, and $\varphi = \xi_j$ for some $1 < j \leq N$. Consider now the substitution η_R and denote φ', ψ' the amicable Sturmian morphisms such that $\eta_R = \mathrm{ter}\big(\varphi', \psi'\big)$. By item (ii), either φ' or ψ' is equal to ξ_N. Due to Lemma 9.7, we know that $\varphi' = \xi_N$, whence by item (ii), the substitution η_R satisfies $\big(\mathrm{Lst}\big(\eta_R(A)\big), \mathrm{Lst}\big(\eta_R(B)\big), \mathrm{Lst}\big(\eta_R(C)\big)\big) = (A, B, B)$.

The case $\big(\mathrm{Fst}\big(\eta_L(A)\big), \mathrm{Fst}\big(\eta_L(B)\big), \mathrm{Fst}\big(\eta_L(C)\big)\big) = (B, B, C)$ is treated similarly. $\square$

Corollary 9.9. *Let η be a primitive substitution given by Theorem 8.1 fixing a 3iet word $\mathbf{u}_\rho$. If η satisfies $\big(\mathrm{Fst}\big(\eta(A)\big), \mathrm{Fst}\big(\eta(B)\big), \mathrm{Fst}\big(\eta(C)\big)\big) = (A, B, B)$, then $\rho = \alpha$, and if it satisfies $\big(\mathrm{Fst}\big(\eta(A)\big), \mathrm{Fst}\big(\eta(B)\big), \mathrm{Fst}\big(\eta(C)\big)\big) = (B, B, C)$, then $\rho = \beta$.*

Proof. Let I be the interval corresponding to η such that T_I is homothetic to T. Denote $I_X = \{x \in I \colon R_I(x) = X\}$. If $\big(\mathrm{Fst}\big(\eta(A)\big), \mathrm{Fst}\big(\eta(B)\big), \mathrm{Fst}\big(\eta(C)\big)\big) = (A, B, B)$, then the boundary between intervals I_A and I_B, i.e., the discontinuity point of T_I, is equal to the point α. Since T_I is homothetic to T, the homothety map Φ maps the discontinuity points of T to the discontinuity points of T_I, i.e., $\Phi(\alpha) = \alpha$. Since the fixed point of the homothety is equal to the intercept of the infinite word coded by η, we have $\rho = \alpha$. The second implication is analogous. $\square$

As a byproduct of our results, it is possible, for a given substitution ξ admitting a non-degenerate 3iet word $\mathbf{u}$ as a fixed point, to give a formula allowing to determine the parameters of $\mathbf{u}$, i.e., the parameters α, β of the transformation T and the intercept ρ such that $\mathbf{u} = \mathbf{u}_\rho$ is a coding of ρ under T. Similar formula for Sturmian morphisms, i.e., those having some word coding an exchange of two intervals as a fixed point, has been given in [21].

The identification of the parameters α and β of the 3iet T is a straightforward task: The values α, $\beta - \alpha$, and $1 - \beta$ are frequencies of the letters A, B and C, respectively, in any infinite word coding some orbit of T. Moreover, the frequencies of letters of a fixed point of a primitive substitution form can be easily determined from the eigenvector corresponding to the dominant eigenvalue of the incidence matrix of the substitution, see [22].

Therefore, the only nontrivial task is to determine the intercept ρ. For this purpose we use the substitution η assigned to the substitution ξ by Theorem 8.1 and its leftmost conjugate η_L. The substitution η has exactly one eigenvalue belonging to the interval $(0, 1)$, see comments below Theorem 8.1. Let this eigenvalue be denoted by λ. Let w be the word of conjugacy between η and η_L, i.e., $\eta(a)w = w\eta_L(a)$ for any $a \in \{A, B, C\}$. Recall that symbols $|w|_A$, $|w|_B$, and $|w|_C$ stand for the number of the letters A, B, and C occuring in w.

Theorem 9.10. *Let $\xi : \{A, B, C\}^* \to \{A, B, C\}^*$ be a primitive substitution such that it has a fixed point $\mathbf{u}$. Suppose that $\mathbf{u}$ is a coding of an orbit of a point, say ρ, under a non-degenerate 3iet T with parameters α and β. Let λ, η, η_L and w be as above. We have*

$$\rho = \rho_L + \frac{1}{1-\lambda}(1-\alpha, 1-\alpha-\beta, -\beta)\begin{pmatrix} |w|_A \\ |w|_B \\ |w|_C \end{pmatrix},$$

where $\rho_L = \alpha$ if $\eta_L(A)$ starts with A and $\rho_L = \beta$ if $\eta_L(A)$ starts with B.

23

Proof. According to Lemma 9.8, $\eta_L(A)$ starts in A or B. Denote by ρ_L the intercept of the 3iet word fixed by η_L. By Corollary 9.9, ρ_L is equal to α if $\eta_L(A)$ starts in A and it is equal to β if $\eta_L(A)$ starts in B. We can use Proposition 8.3 to derive

$$(1-\lambda)\rho_L = T^n\big((1-\lambda)\rho\big)\,, \qquad \text{where } n = |w|.$$

The definition of the transformation T implies the following observation: If w of lenght n is a prefix of $\mathbf{u}_x$, then $\mathbf{u}_x = w\mathbf{u}_{T^n(x)}$ and

$$T^n(x) = x + (1-\alpha, 1-\alpha-\beta, -\beta)\begin{pmatrix} |w|_A \\ |w|_B \\ |w|_C \end{pmatrix}.$$

Combining these two facts with Remark 8.4, we get

$$(1-\lambda)\rho = (1-\lambda)\rho_L + (1-\alpha, 1-\alpha-\beta, -\beta)\begin{pmatrix} |w|_A \\ |w|_B \\ |w|_C \end{pmatrix}.$$

The statement follows. □

10 Applications

10.1 Class P conjecture for non-degenerate 3iet

This subsection is devoted to a question coming from another field, namely mathematical physics, where notions from combinatorics on words appear naturally in the study of the spectra of Schrödinger operators associated to infinite sequences. The question is stated in an article of Hof, Knill and Simon [13] and concerns infinite sequences generated by a substitution over a finite alphabet. The authors show in their paper that if a sequence contains infinitely many palindromic factors (such sequences are called *palindromic*), then the associated operator has a purely singular continuous spectrum. In the same paper, the following class of substitutions is defined.

Definition 10.1. Let φ be a substitution over an alphabet $\mathcal{A}$. We say that φ belongs to the *class* P if there exists a palindrome p such that for every $a \in \mathcal{A}$ one has $\varphi(a) = pp_a$ where p_a is a palindrome. We say that φ is of *class* P' if it is conjugate to some morphism in class P.

Hof, Knill and Simon ask the following question: "Are there (minimal) sequences containing arbitrarily long palindromes that arise from substitutions none of which belongs to class P?" A discussion on how to transform this question into a mathematical formalism can be found in [15].

The first result concerning class P was given by Tan in [29]. The author extended class P by morphisms conjugated to the elements of class P, since it is well-known that fixed points of conjugated morphisms have the same set of factors. This extended class is denoted by P'.

The conjecture, stemming from the question of Hof, Knill and Simon, states that every pure morphic (uniformly recurrent) palindromic sequence is a fixed point of a morphism of class P'. It is referred to as 'class P conjecture'.

In [29], it is shown that if a fixed point of a primitive substitution φ over a binary alphabet is palindromic, then the substitution φ or φ^2 belongs to class P'. In [16], Labbé shows that the assumption of a binary alphabet in the theorem of Tan is essential. He shows that the fixed point of the substitution

$$A \mapsto ABA,\ B \mapsto C,\ C \mapsto BAC,$$

is palindromic. The substitution clearly does not belong to class P'. Moreover, no other substitution fixing the same infinite word belongs to class P'. Is is easy to see that Labbé's substitution fixes a degenerate 3iet word, namely 3iet word coding the orbit of $\rho = \frac{2-\sqrt{2}}{4}$ under the 3iet with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{3-\sqrt{2}}{2}$.

We show that a ternary analogue of the theorem of Tan holds in the context of codings of non-degenerate 3iet with the permutation (321). The following lemma is a generalization of a result obtained for binary alphabets by Tan [29], also shown in [15]. We provide a different proof.

Proposition 10.2. *Let $\varphi : \mathcal{A} \to \mathcal{A}$ be a non-erasing morphism. The morphism φ is conjugate to $\overline{\varphi}$ if and only if φ is of class P'.*

Proof. ($\Leftarrow$): Since φ is of class P', there exists a morphism φ' of class P which is conjugate to φ, i.e., there exists a word w such that $w\varphi(a) = \varphi'(a)w$ or $\varphi(a)w = w\varphi'(a)$ for every letter a.

We can suppose that $w\varphi(a) = \varphi'(a)w$ for every letter a as the other case is analogous. It implies $\varphi(a) = w^{-1}pp_aw$ for some palindromes p_a and p. Thus, $\overline{\varphi(a)} = \overline{w}p_ap(\overline{w})^{-1}$ for every letter a. In other words, the morphism $\overline{\varphi}$ is conjugate to $\overline{\varphi'}$. Since $\overline{\varphi'}$ is clearly conjugate to φ, we conclude that φ is conjugate to $\overline{\varphi}$.

($\Rightarrow$): Since φ is conjugate to $\overline{\varphi}$, there exists a word $w \in \mathcal{B}^*$ such that for every $a \in \mathcal{A}$, we have

$$\varphi(a)w = w\overline{\varphi(a)} \quad \text{or} \quad w\varphi(a) = \overline{\varphi(a)}w\,.$$

Suppose first that $\varphi(a)w = w\overline{\varphi(a)}$ holds. By Lemma 1 in [8], this implies that w is a palindrome. Let $u \in \mathcal{A}^*$ and $c \in \{\varepsilon\} \cup \mathcal{A}$ be such that $w = uc\overline{u}$. We can thus write

$$\varphi(a)uc\overline{u} = uc\overline{u}\overline{\varphi(a)}\,.$$

By applying $(uc)^{-1}$ from the left and $(c\overline{u})^{-1}$ from the right, we obtain for any $a \in \mathcal{A}$

$$c^{-1}u^{-1}\varphi(a)u = \overline{u}\overline{\varphi(a)}\overline{u}^{-1}c^{-1} = \overline{c^{-1}u^{-1}\varphi(a)u}\,.$$

This means that the word $p_a := c^{-1}u^{-1}\varphi(a)u$ is a palindrome. Set $p := c$. Denote by φ' the morphism defined for all $a \in \mathcal{A}$ by $\varphi'(a) = pp_a = u^{-1}\varphi(a)u$. Obviously, φ is conjugate to φ' which is of class P. Therefore $\varphi \in P'$.

The case $w\varphi(a) = \overline{\varphi(a)}w$ is analogous. □

We are now in position to complete the proof the main theorem.

Theorem 10.3. *If ξ is a primitive substitution fixing a non-degenerate 3iet word, then ξ or ξ^2 belongs to class P'.*

Proof. Denote by $\eta \in \{\xi, \xi^2\}$ the substitution from Theorem 8.1. There exist intervals I_L and $I_R \subset [0,1)$ such that $\eta_L(A), \eta_L(B)$ and $\eta_L(C)$ are the I_L-itineraries, $\eta_R(A), \eta_R(B)$ and $\eta_R(C)$ are the I_R-itineraries, and such that T_{I_L} and T_{I_R} are 3iets homothetic to T.

Lemma 9.8 implies that

$$\Big(\mathrm{Fst}\big(\eta_L(A)\big), \mathrm{Fst}\big(\eta_L(B)\big), \mathrm{Fst}\big(\eta_L(C)\big)\Big) = \Big(\mathrm{Lst}\big(\eta_R(A)\big), \mathrm{Lst}\big(\eta_R(B)\big), \mathrm{Lst}\big(\eta_R(C)\big)\Big)$$

and this triple of letters equals (A, B, B) or (B, B, C). Suppose it is equal to (A, B, B). Note that by Corollary 9.9, η_L fixes the infinite word $\mathbf{u}_\alpha$.

According to Proposition 8.5, the induced transformation $T_{\overline{I_R}}$ is again homothetic to T and the corresponding substitution is $\overline{\eta_R}$. Since it is the mirror substitution to η_R, we have $\Big(\mathrm{Fst}\big(\overline{\eta_R}(A)\big), \mathrm{Fst}\big(\overline{\eta_R}(B)\big), \mathrm{Fst}\big(\overline{\eta_R}(C)\big)\Big) = (A, B, B)$. By Corollary 9.9, the substitution $\overline{\eta_R}$ also fixes the infinite word $\mathbf{u}_\alpha$. Since the intervals I_L and $\overline{I_R}$ are of the same length and are homothetic to the interval $[0,1)$ with the same homothety center α, necessarily $I_L = \overline{I_R}$ and thus $\overline{\eta_R} = \eta_L$. Consequently, η_R is conjugate to its mirror image. We apply Proposition 10.2 to finish the proof.

In case that $\big(\mathrm{Fst}\big(\eta_L(A)\big), \mathrm{Fst}\big(\eta_L(B)\big), \mathrm{Fst}\big(\eta_L(C)\big)\big) = (B,B,C)$, we proceed in a similar way. In this case, the center of the homothety of the intervals $I_L = \overline{I_R}$ and $[0,1)$ is β. □

Let us mention that another analogue of Tan's result is already known for marked morphisms. Recall that a substitution ξ over an alphabet $\mathcal{A}$ is called *marked* if its leftmost conjugate ξ_L and its rightmost conjugate ξ_R satisfy

$$\mathrm{Fst}\big(\xi_L(a)\big) \neq \mathrm{Fst}\big(\xi_L(b)\big) \quad \text{and} \quad \mathrm{Lst}\big(\xi_R(a)\big) \neq \mathrm{Lst}\big(\xi_R(b)\big)$$

for distinct $a, b \in \mathcal{A}$. It can be shown that if ξ is marked, then all its powers are marked. In [15], it is shown that for a marked morphism ξ with fixed point $\mathbf{u}$ having infinitely many palindromes, some power ξ^k belongs to class P'.

Our Lemma 9.8 shows that a substitution fixing a non-degenerated 3iet word cannot be marked. Theorem 10.3 thus provides a new family of substitutions satisfying class P conjecture.

10.2 Properties of 3iet preserving substitutions

A morphism which maps a 3iet word to a 3iet word is called 3iet preserving. Morphisms which preserve 2iet words, i.e. Sturmian words, are called Sturmian and they have been extensively studied for many years. Sturmian morphisms form a monoid which is generated by three morphisms only [20]. In contrary to Sturmian morphisms, the class of 3iet preserving morphisms is not completely described. Only partial results are known. For example, the monoid of 3iet preserving morphisms is not finitely generated [2] and contains the ternarizations we defined in Section 9, see [4]. Our previous considerations lead to some comments on properties of 3iet preserving substitution.

- Our results of Sections 7 and 8 allow us to say more about the structure of substitutions fixing 3iet words.

  Corollary 10.4. *Let η be a primitive substitution of Theorem 8.1 fixing a non-degenerate 3iet word over the alphabet $\{A,B,C\}$. We have*

  $$\eta(B) = \omega_{AC\to B}\big(\eta(AC)\big) = \omega_{CA\to B}\big(\eta(CA)\big) \quad \text{or} \quad \eta(B) = \omega_{B\to CA}\big(\eta(AC)\big) = \omega_{B\to AC}\big(\eta(CA)\big)\,.$$

  Proof. By Theorem 8.1, η corresponds to an interval I such that T_I is homothetic to T. Since T is non-degenerate, also T_I is non-degenerate, and therefore its discontinuity points $\mathfrak{C}, \mathfrak{D}$ are distinct. By Proposition 7.1, the three I-itineraries are of the form given by cases (i) or (ii). □

  Example 10.5. We can illustrate the above corollary on substitutions from Example 9.6. We have

  $$\begin{matrix} \varphi(0) = 0110101 \\ \varphi(1) = 01101 \end{matrix}\,, \qquad \begin{matrix} \psi(0) = 1010101 \\ \psi(1) = 10101 \end{matrix}\,,$$

  and

  $$\eta(A) = BCACAC, \quad \eta(B) = BCACBBCAC, \quad \eta(C) = BCAC.$$

  We can check that η satisfies the property given in Corollary 10.4, namely that

  $$\begin{aligned} \eta(B) = BCACBBCAC &= \omega_{AC\to B}\big(\eta(AC)\big) = \omega_{AC\to B}\big(BCACACBCAC\big) = \\ &= \omega_{CA\to B}\big(\eta(CA)\big) = \omega_{CA\to B}\big(BCACBCACAC\big)\,. \end{aligned}$$

The above corollary implies a relation of numbers of occurrences of letters in letter images of η which may be used to get the following relation:

$$M_\eta \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} |\eta(A)|_A \\ |\eta(A)|_B \\ |\eta(A)|_C \end{pmatrix} - \begin{pmatrix} |\eta(B)|_A \\ |\eta(B)|_B \\ |\eta(B)|_C \end{pmatrix} + \begin{pmatrix} |\eta(C)|_A \\ |\eta(C)|_B \\ |\eta(C)|_C \end{pmatrix} = \pm \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Thus, $(1,-1,1)^\top$ is an eigenvector of the incidence matrix of η corresponding to the eigenvalue 1 and -1, respectively. This fact has been already derived in [3] by other methods.

- One can ask whether a substitution η fixing a 3iet word can be in the same time the leftmost and rightmost conjugate of itself, i.e. $\eta = \eta_L = \eta_R$. It can be easily seen that such a situation never occurs for non-degenerate 3iets. Indeed, the proof of Theorem 10.3 implies that for any primitive substitution η fixing a non-degenerate 3iet, η_L and $\overline{\eta_R}$ fix the same infinite word $\mathbf{u}_\rho$, where $\rho \in \{\alpha, \beta\}$. If, moreover, $\eta_R = \overline{\eta_R}$, then by Proposition 8.5 we have $1-\rho = \rho$. This implies that $\rho = \frac{1}{2} \in \{\alpha, \beta\}$. However, this cannot happen for a non-degenerate 3iet T.

- It can be observed from Lemma 9.8 that given a substitution η fixing a 3iet word, its leftmost conjugate η_L has always two fixed points, namely either $\lim_{n\to\infty} \eta_L^n(A)$ and $\lim_{n\to\infty} \eta_L^n(B)$, or $\lim_{n\to\infty} \eta_L^n(B)$ and $\lim_{n\to\infty} \eta_L^n(C)$. One can show that one of these fixed points is a coding of the point $\rho = \alpha$ or β, respectively, under a 3iet T. The other fixed point is a coding of the same point, but under an exchange of three intervals defined over $(0,1]$, where all the intervals are of the form $(\cdot,\cdot]$.

  Example 10.6. Consider the substitution η from Example 9.6. We have

$$\eta_L(A) = ACBCAC, \quad \eta_L(B) = BBCACBCAC, \quad \eta_L(C) = BCAC.$$

  This substitution has two fixed points, namely

$$\lim_{n\to\infty} \eta_L^n(A) = ACBCACBCACBBCACBCACBCACACBCACBCAC\cdots,$$
$$\lim_{n\to\infty} \eta_L^n(B) = BBCACBCACBBCACBCACBCACACBCACBCAC\cdots.$$

  It can be verified that the two infinite words differ only by the prefix AC vs. B. The infinite word $\mathbf{u}_\alpha$, coding α under the 3iet T is equal to the fixed point $\lim_{n\to\infty} \eta_L^n(B)$.

Acknowledgements

Z.M. and E.P. acknowledge financial support by the Czech Science Foundation grant GAČR 13-03538S, Š.S. acknowledges financial support by the Czech Science Foundation grant GAČR 13-35273P.

References

[1] B. Adamczewski, *Codages de rotations et phénomènes d'autosimilarité*, J. Théor. Nombres Bordeaux, 14 (2002), pp. 351–386.

[2] P. Ambrož, L. Háková, and E. Pelantová, *Properties of 3iet preserving morphisms and their matrices*, in Proceedings WORDS 2007, 2007, pp. 18–24.

[3] P. Ambrož, Z. Masáková, and E. Pelantová, *Matrices of 3-iet preserving morphisms*, Theor. Comput. Sci., 400 (2008), pp. 113–136.

[4] P. Ambrož, Z. Masáková, and E. Pelantová, *Morphisms fixing words associated with exchange of three intervals*, RAIRO-Theor. Inf. Appl., 44 (2010), pp. 3–17.

[5] P. Arnoux, V. Berthé, Z. Masáková, and E. Pelantová, *Sturm numbers and substitution invariance of 3iet words*, Integers, 8 (2008). Article A14.

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

[7] L. Balková and E. Pelantová, *A note on symmetries in the rauzy graph and factor frequencies*, Theor. Comput. Sci., 410 (2009), pp. 2779–2783.

[8] A. Blondin Massé, S. Brlek, and S. Labbé, *Palindromic lacunas of the Thue-Morse word*, in Proc. GASCom 2008, 2008, pp. 53–67.

[9] A. Blondin Massé, S. Brlek, S. Labbé, and L. Vuillon, *Palindromic complexity of codings of rotations*, Theor. Comput. Sci., 412 (2011), pp. 6455–6463.

[10] S. Ferenczi, C. Holton, and L. Q. Zamboni, *Structure of three-interval exchange transformations II: a combinatorial description of the tranjectories*, J. Anal. Math., 89 (2003), pp. 239–276.

[11] S. Ferenczi and L. Q. Zamboni, *Languages of k-interval exchange transformations*, Bull. Lond. Math. Soc., 40 (2008), pp. 705–741.

[12] L. S. Guimond, Z. Masáková, and E. Pelantová, *Combinatorial properties of infinite words associated with cut-and-project sequences*, J. Théor. Nombres Bordeaux, (2003), pp. 697–725.

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

[14] M. Keane, *Interval exchange transformations*, Math. Z., 141 (1975), pp. 25–31.

[15] S. Labbé and E. Pelantová, *Palindromic sequences generated from marked morphisms*, Eur. J. Comb., 51 (2016), pp. 200–214.

[16] S. Labbé, *A counterexample to a question of Hof, Knill and Simon*, Electron. J. Comb., 21 (2014).

[17] M. Lothaire, *Algebraic combinatorics on words*, no. 90 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2002.

[18] Z. Masáková, E. Pelantová, and Š. Starosta, *Interval exchange words and the question of Hof, Knill, and Simon*, in Proceedings of DLT 2015, Liverpool, vol. 9168 of LNCS, 2015, pp. 377–388.

[19] Z. Masáková and E. Pelantová, *Itineraries induced by exchange of two intervals*, Acta Polytechnica, 53 (2013), pp. 444–449.

[20] F. Mignosi and P. Séébold, *Morphismes sturmiens et règles de Rauzy*, J. Théor. Nombres Bordeaux, 5 (1993), pp. 221–233.

[21] L. Peng and B. Tan, *Sturmian sequences and invertible substitutions*, Discr. Math. Theor. Comput. Sci., 13 (2011), pp. 63–68.

[22] M. Queffélec, *Substitution dynamical systems—spectral analysis*, vol. 1294 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, second ed., 2010.

[23] G. Rauzy, *Une généralisation du développement en fraction continue*, in Séminaire Delange-Pisot-Poitou, 18e année: 1976/77, Théorie des nombres, Fasc. 1, Secrétariat Math., Paris, 1977, pp. Exp. No. 15, 16.

[24] ——, *Échanges d'intervalles et transformations induites*, Acta Arith., 34 (1979), pp. 315–328.

[25] M. Rigo, P. Salimov, and E. Vandomme, *Some properties of abelian return words*, J. Integer Seq., 16 (2013), pp. Article 13.2.5, 21.

[26] P. Séébold, *On the conjugation of standard morphisms*, Theor. Comput. Sci., 195 (1998), pp. 91–109.

[27] Y. G. Sinai, *Introduction to ergodic theory (Russian)*, Erevan State University, 1973. English translation: Mathematical Notes, Princeton Univ. Press, 1977.

[28] N. B. Slater, *Gaps and steps for the sequence $n\theta$ mod 1*, Math. Proc. Cambridge Phil. Soc., 63 (1967), pp. 1115–1123.

[29] B. Tan, *Mirror substitutions and palindromic sequences*, Theor. Comput. Sci., 389 (2007), pp. 118–124.

[30] W. A. Veech, *Topological dynamics*, Bull. Amer. Math. Soc., 83 (1977), pp. 775–830.

[31] ——, *Interval exchange transformations*, J. Analyse Math., 33 (1978), pp. 222–272.

[32] L. Vuillon, *A characterization of Sturmian words by return words*, Eur. J. Combin., 22 (2001), pp. 263–275.

[33] ——, *On the number of return words in infinite words constructed by interval exchange transformations*, Pure Math. Appl. (PU.M.A.), 18 (2007), pp. 345–355.

[34] S.-I. Yasutomi, *On Sturmian sequences which are invariant under some substitutions*, in Number theory and its applications (Kyoto, 1997), vol. 2 of Dev. Math., Kluwer Acad. Publ., Dordrecht, 1999, pp. 347–373.

Chapter 4

On the Zero Defect Conjecture

Authors:	S. Labbé, E. Pelantová, and Š. Starosta
Publication reference:	submitted to Eur. J. of Combin., 2016
Cited as:	[III]

On the Zero Defect Conjecture

Sébastien Labbé*[1], Edita Pelantová[2], and Štěpán Starosta†[3]

[1] Université de Liège, Bât. B37 Institut de Mathématiques, Grande Traverse 12, 4000 Liège, Belgium
[2] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Czech Republic
[3] Faculty of Information Technology, Czech Technical University in Prague, Czech Republic

Abstract

Brlek et al. conjectured in 2008 that any fixed point of a primitive morphism with finite palindromic defect is either periodic or its palindromic defect is zero. Bucci and Vaslet disproved this conjecture in 2012 by a counterexample over ternary alphabet. We prove that the conjecture is valid on binary alphabet. We also describe a class of morphisms over multiliteral alphabet for which the conjecture still holds. The proof is based on properties of extension graphs.

Keywords: palindromic defect, marked morphism, special factor, extension graph
2000 MSC: 68R15, 37B10

1 Introduction

Palindromes — words read the same from the left as from the right — are a favorite pun in various languages. For instance, the words ressasser, ťahať, and šílíš are palindromic words in the first languages of the authors of this paper. The reason for a study of palindromes in formal languages is not only to deepen the theory, but it has also applications.

The theoretical reasons include the fact that a Sturmian word, i.e., an infinite aperiodic word with the least factor complexity, can be characterized using the number of palindromic factors of given length that occur in a word, see [10]. The application motives include the study of the spectra of discrete Schrödinger operators, see [12, 13].

In [9], the authors provide an elementary observation that a finite word of length n cannot contain more than $n+1$ (distinct) palindromic factors, including the empty word as a palindromic factor. We illustrate this on the following 2 examples of words of length 9:

$$w^{(1)} = 010010100 \quad \text{and} \quad w^{(2)} = 011010011.$$

The word $w^{(1)}$ is a prefix of the famous Fibonacci word and w_2 is a prefix of (also famous) Thue–Morse word. There are 10 palindromic factors of $w^{(1)}$: 0, 1, 00, 010, 101, 1001, 01010, 010010, 0010100, and the empty word. The word $w^{(2)}$ contains only 9 palindromes: 0, 1, 11, 0110, 101, 010, 00, 1001, and the empty word.

*Electronic address: `slabbe@ulg.ac.be`
†Electronic address: `stepan.starosta@fit.cvut.cz`

The existence of the upper bound on the number of distinct palindromic factors lead to the definition of *palindromic defect* (or simply *defect*) of a finite word w, see [5], as the value

$$D(w) = n + 1 - \text{ the number of palindromic factors of } w$$

with n being the length of w. Our examples satisfy $D(w^{(1)}) = 0$, i.e., the upper bound is attained, and $D(w^{(2)}) = 1$. The notion of palindromic defect naturally extends to infinite words. For an infinite word $\mathbf{u}$ we set

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

In this paper, we deal with infinite words that are generated by a primitive morphism of a free monoid $\mathcal{A}^*$ with $\mathcal{A}$ being a finite alphabet. A morphism φ is completely determined by the images of all letters $a \in \mathcal{A}$: $a \mapsto \varphi(a) \in \mathcal{A}^*$. A morphism is *primitive* if there exists a power k such that any letter $b \in \mathcal{A}$ appears in the word $\varphi^k(a)$ for any letter $a \in \mathcal{A}$.

The two mentioned infinite words can be generated using a primitive morphism. Consider the morphism φ_F over $\{0, 1\}^*$ determined by $0 \mapsto 01$ and $1 \mapsto 0$. By repeated application of φ_F, starting from 0, we obtain

$$0 \mapsto 01 \mapsto 010 \mapsto 01001 \mapsto 01001010\ldots$$

Since $\varphi_F^n(0)$ is a prefix of $\varphi_F^{n+1}(0)$ for all $n \in \mathbb{N}$, there exists an infinite word $\mathbf{u}_F$, called the Fibonacci word, such that $\varphi_F^n(0)$ is its prefix for all n. Consider a natural extension of φ_F to infinite words, we obtain that $\mathbf{u}_F$ is a fixed point of φ_F since

$$\mathbf{u}_F = \varphi_F(\mathbf{u}_F) = \varphi_F(u_0 u_1 u_2 \ldots) = \varphi_F(u_0)\varphi_F(u_1)\varphi_F(u_2)\ldots$$

where $u_i \in \{0, 1\}$.

Similarly, let φ_{TM} be a morphism determined by $0 \mapsto 01$ and $1 \mapsto 10$. By repeated application of φ_{TM}, starting again from 0, we obtain

$$0 \mapsto 01 \mapsto 0110 \mapsto 01101001 \mapsto 0110100110010110\ldots$$

The infinite word having $\varphi_{TM}^n(0)$ as a prefix for each n is the Thue–Morse word, sometimes also called Prouhet–Thue–Morse word.

The present article focuses on palindromic defect of infinite words which are fixed points of primitive morphisms. In order for the palindromic defect of such an infinite word to be finite, the word must contain an infinite number of palindromic factors. This property is satisfied by the two mentioned words $\mathbf{u}_F$ and $\mathbf{u}_{TM}$. However, for their palindromic defect, we have $D(\mathbf{u}_F) = 0$, whilst $D(\mathbf{u}_{TM}) = +\infty$.

There exist fixed points $\mathbf{u}$ of primitive morphisms with $0 < D(\mathbf{u}) < +\infty$, but on a two-letter alphabet, only ultimately periodic words are known. In [5], examples of such words are given by Brlek, Hamel, Nivat and Reutenauer as follows: for any $k \in \mathbb{Z}, k \geq 2$ denote by z the finite word

$$z = 01^k01^{k-1}001^{k-1}01^k0\,.$$

Then the infinite periodic word z^ω has palindromic defect k. Of course, the periodic word z^ω is fixed by the primitive morphism $0 \mapsto z, 1 \mapsto z$. In [4], the authors stated the following conjecture:

Conjecture (Zero Defect Conjecture)**.** *If* $\mathbf{u}$ *is a fixed point of a primitive morphism such that* $D(\mathbf{u}) < +\infty$*, then* $\mathbf{u}$ *is periodic or* $D(\mathbf{u}) = 0$.

In 2012, Bucci and Vaslet [7] found a counterexample to this conjecture on a ternary alphabet. They showed that the fixed point of the primitive morphism determined by

$$a \mapsto aabcacba, b \mapsto aa, c \mapsto a$$

has finite positive palindromic defect and is not periodic.

In this article, we show that the conjecture is valid on a binary alphabet. Then we generalize the method used for morphisms on a binary alphabet to marked morphisms on a multiliteral alphabet. The main result of the article is the following theorem.

Theorem 1. *Let φ be a primitive marked morphism and let $\mathbf{u}$ be its fixed point with finite defect. If all complete return words of all letters in $\mathbf{u}$ are palindromes or there exists a conjugate of φ distinct from φ itself, then $D(\mathbf{u}) = 0$.*

Observe that in the case of primitive marked morphisms, as it was noted in [14, Cor. 30, Cor. 32], there is no ultimately periodic infinite word $\mathbf{u}$ fixed point of a primitive marked morphisms such that $0 < D(\mathbf{u}) < \infty$.

The main ingredients for the presented proofs of Theorem 1 and Theorem 24 are the following:

1. description of bilateral multiplicities of factors in a word with finite palindromic defect ([1]),
2. description of the form of marked morphisms with their fixed points containing infinitely many palindromic factors ([14]).
3. observation that non-zero palindromic defect of a word implies an existence of a factor with a specific property, see Lemma 23 for the binary case and Theorem 26 for the multiliteral case.

The paper is organized as follows: First we recall notions from combinatorics on words and we list known results that we use in the sequel. In Section 3, the properties of marked morphisms are studied. In Section 4, we introduce the notion of a graph of a factor and we interpret bilateral multiplicity of factors in the language of graph theory. Section 5 is focused on properties of a graph of a factor in the case of language having finite palindromic defect. The validity of the Zero Defect Conjecture on binary alphabet is demonstrated in Section 6 (Theorem 24). Section 7 contains the proof of Theorem 1. Comments on counterexamples to two conjectures concerning palindromes form the last Section 8.

2 Preliminaries

An *alphabet* $\mathcal{A}$ is a finite set of symbols called *letters*. A *finite word* $w = w_0w_1 \cdots w_{n-1}$ is a finite sequence over $\mathcal{A}$, i.e., $w_i \in \mathcal{A}$. The *length* of w is n and is denoted by $|w|$. An *infinite word* is an infinite sequence over $\mathcal{A}$. Given words p, f, s with p and f being finite such that $w = pfs$, we say that p is a *prefix* of w, f is a *factor* of w, and s is a *suffix* of w.

2.1 Language of an infinite word

Consider an infinite word $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ over the alphabet $\mathcal{A}$. An index $i \in \mathbb{N}$ is an occurrence of a factor $w = w_0w_1 \cdots w_{n-1}$ of $\mathbf{u}$ if $u_iu_{i+1} \cdots u_{i+n-1} = w$, in other words w is prefix of the infinite word $u_iu_{i+1}u_{i+2} \cdots$. The set of all factors of $\mathbf{u}$ is referred to as the *language* of $\mathbf{u}$ and denoted

$\mathcal{L}(\mathbf{u})$. The mapping $\mathcal{C}(n)$, defined by $\mathcal{C}(n) = \#\mathcal{L}(\mathbf{u}) \cap \mathcal{A}^n$, is the *factor complexity* of $\mathbf{u}$. A word $\mathbf{u}$ is called recurrent if any factor $w \in \mathcal{L}(\mathbf{u})$ has infinitely many occurrences. If $i < j$ are two consecutive occurrences of the factors w, then the factor $u_i u_{i+i} \cdots u_j u_{j+1} \cdots u_{j+n-1}$ is the *complete return word* to w in $\mathbf{u}$. If any factor of a recurrent word $\mathbf{u}$ has only finitely many complete return words, then $\mathbf{u}$ is called *uniformly recurrent*.

Reversal of a finite word $w = w_0 w_1 \cdots, w_{n-1}$ is the word $\widetilde{w} = w_{n-1} w_{n-2} \cdots w_0$. A word w is a *palindrome* if $w = \widetilde{w}$. The language of $\mathbf{u}$ is said to be *closed under reversal* if $w \in \mathcal{L}(\mathbf{u})$ implies $\widetilde{w} \in \mathcal{L}(\mathbf{u})$; $\mathbf{u}$ is said to be *palindromic* if $\mathcal{L}(\mathbf{u})$ contains infinitely many palindromes. If a uniformly recurrent word $\mathbf{u}$ is palindromic, then its language is closed under reversal. The mapping counting the palindromes of length n in $\mathcal{L}(\mathbf{u})$ is the *palindromic complexity* and is denoted by $\mathcal{P}(n)$, i.e., we have $\mathcal{P}(n) = \#\{w \in \mathcal{L}(\mathbf{u}) \colon |w| = n \text{ and } w = \widetilde{w}\}$.

A letter $b \in \mathcal{A}$ is called *right* (resp. *left*) *extension* of w in $\mathcal{L}(\mathbf{u})$ if wb (resp. bw) belongs to $\mathcal{L}(\mathbf{u})$. In a recurrent word $\mathbf{u}$ any factor has at least one right and at least one left extension. A factor w is *right special* (resp. *left special*) if it has more than one right (resp. left) extension. A factor w which is simultaneously left and right special is *bispecial*. To describe one-sided and both-sided extensions of a factor w we put

$$E^+(w) = \{b \in \mathcal{A} \colon wb \in \mathcal{L}(\mathbf{u})\}, \quad E^-(w) = \{a \in \mathcal{A} \colon aw \in \mathcal{L}(\mathbf{u})\},$$

$$\text{and} \quad E(w) = \{(a, b) \in \mathcal{A}^2 \colon awb \in \mathcal{L}(\mathbf{u})\}.$$

The *bilateral multiplicity* $m(w)$ of a factor $w \in \mathcal{L}(\mathbf{u})$ is defined as

$$m(w) = \#E(w) - \#E^+(w) - \#E^-(w) + 1.$$

Under the assumption of recurrent language, the second difference of the factor complexity may be expressed using bilateral multiplicities as follows:

$$\Delta^2\mathcal{C}(n) = \mathcal{C}(n+2) - 2\mathcal{C}(n+1) + \mathcal{C}(n) = \sum_{\substack{w \in \mathcal{L}(\mathbf{u}) \\ |w| = n \\ w \text{ is bispecial}}} m(w) = \sum_{\substack{w \in \mathcal{L}(\mathbf{u}) \\ |w| = n}} m(w). \tag{1}$$

(See [8], Section 4.5.2 for the equation (1) and Section 4 for a recent reference on factor complexity in general.) Note that the last equality in (1) follows from the fact that $m(w)$ is nonzero only for bispecial factors in the case of a recurrent language.

A bispecial factor $w \in \mathcal{L}(\mathbf{u})$ is said to be *strong* if $m(w) > 0$, *weak* if $m(w) < 0$ and *neutral* if $m(w) = 0$.

2.2 Palindromic defect

As shown in [9] finite words with zero defect can be characterized using palindromic suffixes. More specifically, a word $w = w_0 w_1 \cdots w_{n-1}$ has defect 0 if and only if for any $i = 0, 1, \ldots, n-1$ the longest palindromic suffix of $w_0 w_1 \cdots w_i$ is unioccurrent in w. To illustrate this important property, consider the words

$$w^{(1)} = 010010100 \quad \text{and} \quad w^{(2)} = 011010011.$$

mentioned in Introduction. The longest palindromic suffix of $w^{(1)}$ is 0010100 and it is unioccurrent in $w^{(1)}$, whereas the longest palindromic suffix of $w^{(2)}$ is 11 and occurs in $w^{(2)}$ twice. The index i for which the longest palindromic suffix is not unioccurrent is called a *lacuna* and the number of lacunas equals the palindromic defect of w.

Since the number of palindromes in w and in its reversal $\widetilde{w}$ is the same, we have $D(w) = D(\widetilde{w})$. Therefore, instead of the longest palindromic suffix one can consider the longest palindromic prefix as well.

The complete return words were applied in [11] to characterize infinite words with zero defect.

Theorem 2 ([11])**.** *$D(\mathbf{u}) = 0$ if and only if for all palindromes $w \in \mathcal{L}(\mathbf{u})$ all complete return words to w in $\mathbf{u}$ are palindromes.*

Before stating a generalization of the previous result we need a new notion.

Definition 3. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ and $w \in \mathcal{L}(\mathbf{u})$. A word $c = c_1c_2\cdots c_n \in \mathcal{L}(\mathbf{u})$ is a* complete mirror return *to w in $\mathbf{u}$ if*

1. *neither w nor $\widetilde{w}$ is a factor of $c_2\cdots c_{n-1}$, and*
2. *either w is a prefix of c and $\widetilde{w}$ is suffix of c, or $\widetilde{w}$ is a prefix of c and w is a suffix of c.*

Note that c is a complete mirror return to w if and only if it is a complete mirror return to $\widetilde{w}$.

Example 4. We illustrate the notion of complete mirror return word on the Fibonacci word $\mathbf{u}_F$. The factors r_1, r_2 and r_3 are complete mirror returns to $w_1 = 0101$, $w_2 = 001$ and $w_3 = 00$ respectively.

$$\mathbf{u}_F = 010\,\underbrace{01010}_{r_1}\,0100101\,\underbrace{0010100}_{r_2}\,1\,\underbrace{0010100}_{r_3}\,10\cdots$$

Note that if $w = \widetilde{w}$, then the complete mirror return words of w and $\widetilde{w}$ coincide with complete return words of w.

Using the notion of complete mirror return word we can reformulate Proposition 2.3 from [6].

Proposition 5 ([6])**.** *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$. We have $D(\mathbf{u}) = 0$ if and only if for each factor $w \in \mathcal{L}(\mathbf{u})$ any complete mirror return word to w in $\mathbf{u}$ is a palindrome.*

A generalization of the previous statement to finite defect follows from [2, Cor. 5 and Lemma 14].

Theorem 6 ([2])**.** *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ be aperiodic and have its language closed under reversal. $D(\mathbf{u}) < +\infty$ if and only if there exists a positive integer K such that for every factor w of length at least K the occurrences of w and $\widetilde{w}$ alternate and every complete mirror return to w in $\mathbf{u}$ is a palindrome.*

2.3 Morphisms

In this section we concentrate on primitive morphisms. For a morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ consider the maps $\textsc{Fst}(\varphi), \textsc{Lst}(\varphi) : \mathcal{A} \to \mathcal{A}$ defined by the formula

$$\textsc{Fst}(\varphi)(a) = \text{the first letter of } \varphi(a) \qquad \text{and} \qquad \textsc{Lst}(\varphi)(a) = \text{the last letter of } \varphi(a)$$

for all $a \in \mathcal{A}$. A morphism φ may have more fixed points, see for example the Thue–Morse morphism. The number of fixed points of a primitive morphism φ is the number of letters for which $\textsc{Fst}(\varphi)(a) = a$. It is easy to see that the languages of all fixed points of a primitive morphism coincide and therefore all its fixed points have the same defect.

Recall from Lothaire [17] (Section 2.3.4) that a morphism ψ is a *left conjugate* of φ, or that φ is a *right conjugate* of ψ, denoted $\psi \triangleright \varphi$, if there exists $w \in \mathcal{A}^*$ such that

$$\varphi(x)w = w\psi(x), \quad \text{for all words } x \in \mathcal{A}^*, \tag{2}$$

or equivalently that $\varphi(a)w = w\psi(a)$, for all letters $a \in \mathcal{A}$. We say that the word w is the *conjugate word of the relation* $\psi \triangleright \varphi$. If, moreover, the map $\mathrm{FST}(\psi)$ is not constant, then ψ is the *leftmost conjugate* of φ. Analogously, if $\mathrm{LST}(\varphi)$ is not constant, then φ is the *rightmost conjugate* of ψ.

Example 7. Let

$$\varphi : \begin{array}{ll} a & \mapsto abab \\ b & \mapsto abb \end{array} \quad \text{and} \quad \psi : \begin{array}{ll} a & \mapsto baba \\ b & \mapsto bba \end{array}.$$

We have $\psi \triangleright \varphi$ and the conjugate word of the relation is $w = a$. The leftmost conjugate of φ (and of ψ) is the morphism

$$a \mapsto abab \quad \text{and} \quad b \mapsto bab.$$

If φ is a primitive morphism, then any of its (left or right) conjugate is primitive as well and the languages of their fixed points coincide.

A morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ is *cyclic* [16] if there exists a word $w \in \mathcal{A}^*$ such that $\varphi(a) \in w^*$ for all $a \in \mathcal{A}$. Otherwise, we say that φ is *acyclic*. If φ is cyclic, then the fixed point of φ is $wwww\dots$ and is periodic. Remark that the converse does not hold. For example, the morphism determined by $a \mapsto aba$ and $b \mapsto bab$ is acyclic but its fixed point is periodic. Obviously, a morphism is cyclic if and only if it is conjugate to itself with a nonempty conjugate word. If a morphism is acyclic, it has a leftmost and a rightmost conjugate. See [14] for a proof of these statements on cyclic morphisms.

3 Marked morphisms

A morphism φ over binary alphabet has a trivial but important property: the leftmost conjugate of φ maps both letters to words with a distinct starting letter and analogously for the rightmost conjugate. The notion of *marked morphism* extends this important property to any alphabet.

Definition 8. *Let φ be an acyclic morphism. We say that φ is* marked *if*

$$\mathrm{FST}(\varphi_L) \textit{ and } \mathrm{LST}(\varphi_R) \textit{ are injective}$$

and that φ is well-marked *if*

$$\textit{it is marked and if } \mathrm{FST}(\varphi_L) = \mathrm{LST}(\varphi_R)$$

where φ_L (resp. φ_R) is the leftmost (resp. rightmost) conjugate of φ.

Remark 9. Any injective mapping f on a finite set is a permutation and thus there exists a power k such that f^k is the identity map. It implies that for any marked morphism φ there exists a power k such that φ^k is well-marked and moreover $\mathrm{FST}(\varphi_L^k) = \mathrm{LST}(\varphi_R^k) = \mathrm{Id}$.

Theorem 10 ([14]). *Let φ be a primitive well-marked morphism and* $\mathbf{u}$ *be its palindromic fixed point. The conjugacy word w of $\varphi_L \triangleright \varphi_R$ is a palindrome and*

$$\widetilde{\varphi_R(a)} = \varphi_L(a) \quad \textit{for all} \quad a \in \mathcal{A}\,.$$

We are interested in the defect of fixed points of primitive marked morphisms. We can consider, instead of the marked morphism φ, a suitable power of φ. Thus, without loss of generality we assume that φ is well marked and that $\mathrm{FST}(\varphi_L) = \mathrm{LST}(\varphi_R) = \mathrm{Id}$. For such φ with the conjugacy word w of $\varphi_L \triangleright \varphi_R$ we define the mapping $\Phi : \mathcal{A}^* \to \mathcal{A}^*$ by

$$\Phi(u) = \varphi_R(u)w \quad \text{for all} \quad u \in \mathcal{A}^* .$$

As φ is primitive, each of its powers and also each of its conjugates have the same language. Moreover, if we assume that $\mathbf{u}$ is palindromic, we can deduce using [14, Lemma 15, Lemma 27, Prop. 28] remarkable properties of the mapping Φ.

Lemma 11 ([14])**.** *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ and $u \in \mathcal{A}^*$. If φ satisfies assumptions of Theorem 10, we have*

(I) *If $u \in \mathcal{L}(\mathbf{u})$, then $\Phi(u) \in \mathcal{L}(\mathbf{u})$.*

(II) $\widetilde{\Phi(u)} = \Phi(\widetilde{u})$.

(III) *The word u is a palindrome if and only if $\Phi(u)$ is a palindrome.*

(IV) *For any $a, b \in \mathcal{A}$, $aub \in \mathcal{L}(\mathbf{u})$ implies $a\Phi(u)b \in \mathcal{L}(\mathbf{u})$.*

(V) *If u is a palindromic (respectively non-palindromic) bispecial factor, then $\Phi(u)$ is a palindromic (respectively non-palindromic) bispecial factor.*

Proof. (I) Let us find v such that $uv \in \mathcal{L}(\mathbf{u})$ with $|\varphi_L(v)| \geq w$. We have

$$\varphi_R(uv)w = \varphi_R(u)w\varphi_L(v).$$

Since $\varphi_R(uv) \in \mathcal{L}(\mathbf{u})$, by erasing a suffix of length greater than or equal to $|w|$ from $\varphi_R(u)w\varphi_L(v)$ we obtain a factor of $\mathcal{L}(\mathbf{u})$, in particular $\varphi_R(u)w \in \mathcal{L}(\mathbf{u})$.

(II) Let $u = u_1 u_2 \cdots u_n$ with $u_i \in \mathcal{A}$. We obtain

$$\Phi(u) = w\varphi_L(u_1) \cdots \varphi_L(u_n) = \varphi_R(u_1) \cdots \varphi_R(u_n)w.$$

Using Theorem 10 we obtain

$$\widetilde{\Phi(u)} = \widetilde{w}\widetilde{\varphi_R(u_n)} \cdots \widetilde{\varphi_R(u_1)} = w\varphi_L(u_n) \cdots \varphi_L(u_1) = \Phi(\widetilde{u}).$$

(III) Let us note that any marked morphism is injective and thus Φ is injective as well. If u is a palindrome, then $\widetilde{\Phi(u)} = \Phi(\widetilde{u}) = \Phi(u)$ from Item (II), therefore $\Phi(u)$ is a palindrome. Conversely, if $\Phi(u)$ is a palindrome, then $\Phi(u) = \widetilde{\Phi(u)} = \Phi(\widetilde{u})$. As φ_L is injective, Φ is injective and the claim follows.

(IV) Let $aub \in \mathcal{L}(\mathbf{u})$. We have $\Phi(aub) \in \mathcal{L}(\mathbf{u})$ and

$$\Phi(aub) = \varphi_R(a)\varphi_R(u)w_\varphi\varphi_L(b) = \varphi_R(a)\Phi(u)\varphi_L(b).$$

By our assumption, $\mathrm{LST}(\varphi_R)(c) = \mathrm{FST}(\varphi_L)(c) = \mathrm{Id}(c) = c$ for any $c \in \mathcal{A}$. Thus, $a\Phi(u)b$ is a factor $\Phi(aub) \in \mathcal{L}(\mathbf{u})$.

(V) The statement follows from the previous properties. □

4 Extension graphs of a factor

To study the Zero Defect Conjecture on a multiliteral alphabet, we assign graphs to palindromic and non-palindromic bispecial factors. These graphs were used already in [1] where only words with zero defect are considered. These graphs enable to represent extensions of a bispecial factor and to determine factor complexity, see [8, p.234–235]. They also appear in a more general context in [3]. We use these graphs to demonstrate that the definition of bilateral multiplicity of bispecial factors is related to basic notions of graph theory which we use later in the proofs.

Definition 12 ($\mathbf{\Gamma(w)}$)**.** *Let* $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$*. We assign to a factor* $w \in \mathcal{L}(\mathbf{u})$ *the bipartite* extension graph $\Gamma(w) = (V, U)$ *whose vertices* V *consist of the disjoint union of* $E^-(w)$ *and* $E^+(w)$

$$V = \big(E^-(w) \times \{-1\}\big) \cup \big(E^+(w) \times \{+1\}\big)$$

and whose edges U *are essentially the elements of* $E(w)$*:*

$$U = \{\{(a,-1),(b,+1)\} : (a,b) \in E(w)\}.$$

Lemma 13. *If* $\Gamma(w)$ *is connected, then* $m(w) \geq 0$ *and*

- $m(w) > 0$ *if and only if* $\Gamma(w)$ *contains a cycle,*
- $m(w) = 0$ *if and only if* $\Gamma(w)$ *is a tree.*

Proof. Let $G = (V, U)$ be a graph with vertices V and edges U. If G is connected then $\#U - \#V + 1 \geq 0$. A connected graph $G = (V, U)$ is a tree if and only if $\#U - \#V\# + 1 = 0$ and it contains a cycle if and only if $\#U - \#V + 1 > 0$. In the case of the graph $\Gamma(w)$, it turns out that

$$\#U - \#V + 1 = \#E(w) - \#E^-(w) - \#E^+(w) + 1 = m(w). \qquad \square$$

Another graph will be useful in the case when $w = \widetilde{w}$ and when the language $\mathcal{L}(\mathbf{u})$ is closed under reversal. These two hypotheses imply that $E^-(w) = E^+(w)$ and that $E(w)$ is symmetric, i.e. $(a,b) \in E(w)$ if and only if $(b,a) \in E(w)$.

Definition 14 ($\mathbf{\Theta(w)}$)**.** *Let* $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ *having its language closed under reversal. To a palindromic factor* $w \in \mathcal{L}(\mathbf{u})$ *we assign a graph* $\Theta(w) = (V, U)$ *whose vertices* $V = E^-(w) = E^+(w)$ *are exactly the right (or left) extensions of* w *and whose edges* U *are unordered pairs of distinct elements of* $E(w)$*:*

$$U = \{\{a,b\} : (a,b) \in E(w), a \neq b\}.$$

In particular, $\Theta(w)$ does not contain loops.

The next lemma uses the both-sided symmetric extensions of a factor w which are denoted by

$$E^=(w) = \{a \in \mathcal{A} \colon awa \in \mathcal{L}(\mathbf{u})\}.$$

Lemma 15. *Suppose that the language* $\mathcal{L}(\mathbf{u})$ *is closed under reversal and* $w = \widetilde{w}$*. If* $\Theta(w)$ *is connected, then* $m(w) \geq \#E^=(w) - 1$ *and*

- $m(w) > \#E^=(w) - 1$ *if and only if* $\Theta(w)$ *contains a cycle,*
- $m(w) = \#E^=(w) - 1$ *if and only if* $\Theta(w)$ *is a tree.*

Proof. Using the same argument as for the previous lemma, we compute that

$$\#U = \frac{1}{2}\left(\#E(w) - \#E^{=}(w)\right) \quad \text{and} \quad \#V = \#E^{-}(w) = \#E^{+}(w).$$

Therefore,

$$\begin{aligned}\#U - \#V + 1 &= \frac{1}{2}\big(\#E(w) - \#E^{=}(w) - \#E^{-}(w) - \#E^{+}(w)\big) + 1\\ &= \frac{1}{2}\left(m(w) - \#E^{=}(w) + 1\right)\end{aligned}$$

□

Example 16. Let $\mathbf{u}$ be the fixed point of the substitution $\eta : a \mapsto aabcacba, b \mapsto aa, c \mapsto a$ used by Bucci and Vaslet. The list of all factors of length 2 is:

$$aa, ab, ac, ba, ca, bc, cb.$$

The list of all factors of length 3 is:

$$aaa, aab, abc, acb, baa, bca, cac, cba.$$

This allows to construct the graphs $\Theta(w)$ and $\Gamma(w)$ for $w \in \{\varepsilon, a, b, c\}$ (see Fig. 1) and the following table of values for the bilateral multiplicity:

w	ε	a	b	c
$m(w)$	2	-1	-1	-1
$\#E^{=}(w) - 1$	0	0	-1	-1

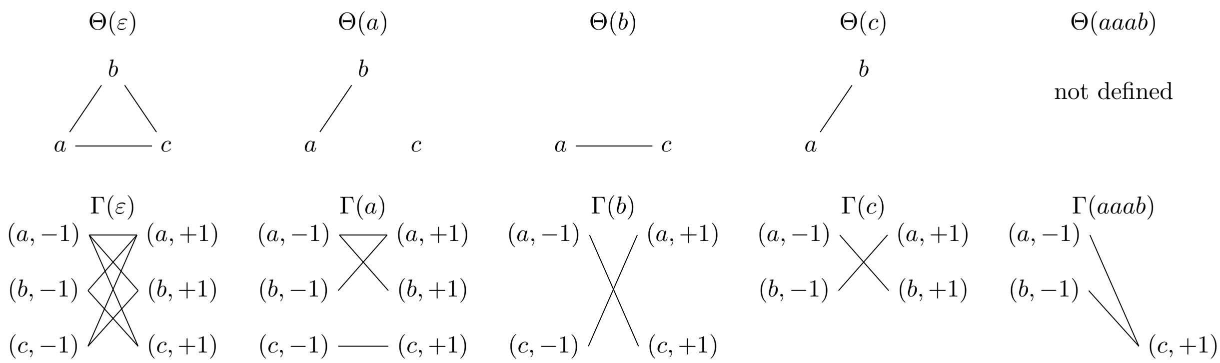

Figure 1: Example of graphs $\Theta(w)$ and $\Gamma(w)$ for $w \in \{\varepsilon, a, b, c, aaab\}$ in the language of the fixed point of the morphism $a \mapsto aabcacba, b \mapsto aa, c \mapsto a$.

1. The graph $\Theta(\varepsilon)$ has vertices $V = \{a, b, c\}$ and edges $U = \big\{\{a,b\}, \{a,c\}, \{b,c\}\big\}$. The graph $\Theta(\varepsilon)$ contains a cycle. The bilateral multiplicity equals $m(\varepsilon) = 2 > 0 = \#E^{=}(\varepsilon) - 1$.
2. The graph $\Theta(a)$ has vertices $V = \{a, b, c\}$ and edges $U = \big\{\{a,b\}\big\}$. The graph $\Theta(a)$ is not connected. The bilateral multiplicity equals $m(a) = -1 < 0 = \#E^{=}(a) - 1$.
3. The graph $\Theta(b)$ has vertices $V = \{a, c\}$ and edges $U = \big\{\{a,c\}\big\}$. The graph $\Theta(b)$ is a tree. The bilateral multiplicity equals $m(b) = -1 = \#E^{=}(b) - 1$.

It is easy to see that the graph $\Theta(c)$ is isomorphic to $\Theta(b)$. The construction of the graphs $\Gamma(w)$ is analogous. From the extension set $E(aaab) = \{(a,c), (b,c)\}$ of the non-palindromic left special word $w = aaab$, the graph $\Gamma(aaab)$ can be constructed (see Fig. 1). Notice that it is a tree.

5 Words with finite palindromic defect

The graphs introduced in the previous section allow to interpret the palindromic defect in terms of graph theory. In this section we focus on properties of graphs of a factor under the assumption of finite palindromic defect (Theorem 21 and Corollary 22). In Section 7, we study properties of a graph of a factor under the assumption of positive palindromic defect (Theorem 26).

The proof of the main results of this section, namely Theorem 21, can be excerpted from [1, proof of Theorem 3.10]. However, the mentioned theorem has a stronger assumption (the palindromic defect of $\mathbf{u}$ is zero) and it is not stated in terms of graphs as done below in Corollary 22. Therefore, Theorem 21 is accompanied here with an independent proof. The proof requires the next two lemmas, which explain the link between complete mirror return word to a factor w and the connectedness of its associated graphs, and a proposition on the relation of factor and palindromic complexity in a word having finite palindromic defect.

Lemma 17. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. Suppose that v is a palindromic complete mirror return word to $w \in \mathcal{L}(\mathbf{u})$ such that $b\widetilde{w}$ is a suffix of v and $av \in \mathcal{L}(\mathbf{u})$ for some letters $a, b \in \mathcal{A}$. Then $\{(a,-1),(b,+1)\}$ is an edge of the graph $\Gamma(w)$. If w is a palindrome and $a \neq b$, then $\{a,b\}$ is an edge of the graph $\Theta(w)$.*

Proof. Let $s \in \mathcal{A}^*$ such that $v = sb\widetilde{w}$. Since v is a palindrome, we get $v = wb\widetilde{s}$. Therefore, $awb \in \mathcal{L}(\mathbf{u})$ being a prefix of av and $(a,b) \in E(w)$. We conclude that $\{(a,-1),(b,+1)\}$ is an edge of the graph $\Gamma(w)$. Also if $w = \widetilde{w}$ and $a \neq b$, we conclude that $\{a,b\}$ is an edge of the graph $\Theta(w)$. □

Lemma 18. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal, $w \in \mathcal{L}(\mathbf{u})$ and suppose that occurrences of w and $\widetilde{w}$ alternate in $\mathbf{u}$. Suppose that all complete mirror return words to w are palindromes. Then $\Gamma(w)$ is connected. If w is a palindrome, then $\Theta(w)$ is connected.*

Proof. It suffices to show that there is a path from any vertex $(a,-1)$ to any vertex $(b,+1)$ in $\Gamma(w)$. Let $(a,-1)$ and $(b,+1)$ be two vertices of $\Gamma(w)$. Then $aw, wb \in \mathcal{L}(\mathbf{u})$. Let $v \in \mathcal{L}(\mathbf{u})$ be such that aw is a prefix of av and $b\widetilde{w}$ is a suffix of av. If there are no other occurrences of factors of $\mathcal{A}w \cup \mathcal{A}\widetilde{w}$ in v, then $\{(a,-1),(b,+1)\}$ is an edge of the graph $\Gamma(w)$ from Lemma 17. Suppose that

$$a_1w,\ b_1\widetilde{w},\ a_2w,\ b_2\widetilde{w},\ \ldots,a_nw,\ b_n\widetilde{w}$$

are consecutive occurrences of factors of $\mathcal{A}w \cup \mathcal{A}\widetilde{w}$ in v where $a = a_1$, $b = b_n$ and $n \geq 2$. From Lemma 17, $\{(a_i,-1),(b_i,+1)\}$ is an edge of the graph $\Gamma(w)$ for all i with $1 \leq i \leq n$. Also, $\{(a_{i+1},-1),(b_i,+1)\}$ is an edge of the graph $\Gamma(w)$ for all i with $1 \leq i \leq n-1$. Therefore, we conclude that there exists a path from $(a,-1)$ to $(b,+1)$.

Assume $w = \widetilde{w}$. Let $a, b \in E^-(w) = E^+(w)$ be two distinct vertices of $\Theta(w)$. Then $aw, bw \in \mathcal{L}(\mathbf{u})$. We want to show that there exists a path from a to b in $\Theta(w)$. Among the occurrences of factors in $\mathcal{A}w$, if there exist two consecutive occurrences of aw and bw, then $\{a,b\}$ is an edge of $\Theta(w)$ from Lemma 17. Otherwise, we conclude that there exists a path from a to b by transitivity. □

Corollary 19. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. If $D(\mathbf{u}) < +\infty$, then there exists an integer K such that for each bispecial factor $w \in \mathcal{L}(\mathbf{u})$ with $|w| \geq K$ the graph $\Gamma(w)$ is connected. If w is moreover a palindrome, then also the graph $\Theta(w)$ is connected.*

Proof. If $\mathbf{u}$ is not aperiodic, then the claim is trivially satisfied as there is only finite number of bispecial factors.

If $\mathbf{u}$ is aperiodic, Theorem 6 implies that there exists a positive integer K such that for every factor $w \in \mathcal{L}(\mathbf{u})$ longer than K, the occurrences of w and $\widetilde{w}$ alternate and every complete mirror return to w in $\mathbf{u}$ is a palindrome. We conclude from Lemma 18 that the graph $\Gamma(w)$ is connected. Also if w is a palindrome, then $\Theta(w)$ is connected. $\square$

The following claim may be deduced from [2].

Proposition 20 ([2, Th. 2, Prop. 6])**.** *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. If $D(\mathbf{u}) < +\infty$, then there exists an integer M such that for all $n \geq M$ we have*

$$\Delta^2\mathcal{C}(n) = \mathcal{P}(n+2) - \mathcal{P}(n).$$

Proof. Since $\mathcal{L}(\mathbf{u})$ is closed under reversal, Proposition 6 from [2] says that there exists an integer M such that for all $n \geq M$ we have

$$\Delta\mathcal{C}(n) + 2 \geq \mathcal{P}(n+1) + \mathcal{P}(n).$$

Since $D(\mathbf{u}) < +\infty$, Theorem 2 from [2] together with the above inequality implies that there exists an integer M such that for all $n \geq M$ we have

$$\Delta\mathcal{C}(n) + 2 = \mathcal{P}(n+1) + \mathcal{P}(n).$$

From this we conclude:

$$\Delta^2\mathcal{C}(n) = \Delta\mathcal{C}(n+1) - \Delta\mathcal{C}(n) = \mathcal{P}(n+2) + \mathcal{P}(n+1) - \mathcal{P}(n+1) - \mathcal{P}(n) = \mathcal{P}(n+2) - \mathcal{P}(n). \quad \square$$

Theorem 21. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. If $D(\mathbf{u}) < +\infty$, then there exists an integer K such that each bispecial factor $w \in \mathcal{L}(\mathbf{u})$ with $|w| \geq K$ satisfies*

$$m(w) = \begin{cases} 0 & \text{if } w \neq \widetilde{w}, \\ \#E^{=}(w) - 1 & \text{if } w = \widetilde{w}. \end{cases}$$

Proof. Let K_1 be the constant given by Corollary 19. If w is a bispecial factor with $|w| > K_1$, we conclude from Lemma 18 that the graph $\Gamma(w)$ is connected. Also if w is a palindrome, then $\Theta(w)$ is connected. It follows from Lemma 13 that $m(w) \geq 0$. If w is a palindrome, Lemma 15 implies $m(w) \geq \#E^{=}(w) - 1$.

If w is not a bispecial factor, then $m(w) = 0$ and, moreover, if w is not a bispecial factor and $w = \widetilde{w}$, then by closedness under reversal we have $\#E^{=}(w) = 1$, and thus $m(w) = 0 = \#E^{=}(w) - 1$.

Suppose by contradiction that for every integer N there exists a non-palindromic factor v of length $|v| > N$ such that $m(v) > 0$ or there exists a palindromic factor v of length $|v| > N$ such that $m(v) > \#E^{=}(v) - 1$. As closedness under reversal implies recurrence, using (1) we obtain that for every integer N there exists $n = |v| > N$ such that

$$\Delta^2\mathcal{C}(n) = \sum_{\substack{w \in \mathcal{L}(\mathbf{u}) \\ |w| = n \\ w \neq \widetilde{w}}} m(w) + \sum_{\substack{w \in \mathcal{L}(\mathbf{u}) \\ |w| = n \\ w = \widetilde{w}}} m(w) > 0 + \sum_{\substack{w \in \mathcal{L}(\mathbf{u}) \\ |w| = n \\ w = \widetilde{w}}} (\#E^{=}(w) - 1) = \mathcal{P}(n+2) - \mathcal{P}(n). \tag{3}$$

This contradicts Proposition 20 and ends the proof of the theorem with $K = \max\{K_1, M\}$. $\square$

The following result is a direct consequence of Lemma 13 and Lemma 15. It allows to interpret the previous theorem in terms of graph theory.

Corollary 22. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ be an infinite word with its language closed under reversal and $D(\mathbf{u}) < +\infty$. There exists a positive integer K such that for every $w \in \mathcal{L}(\mathbf{u})$ of length at least K*

- *if w is not a palindrome, then the graph $\Gamma(w)$ is a tree,*
- *if w is a palindrome, then the graph $\Theta(w)$ is a tree.*

Proof. Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. From Corollary 19, there exists an integer K_1 such that for each bispecial factor $w \in \mathcal{L}(\mathbf{u})$ with $|w| \geq K_1$ the graph $\Gamma(w)$ is connected. If w is moreover a palindrome, then also the graph $\Theta(w)$ is connected.

From Theorem 21, it follows that there exists a constant K_2 such that every factor w longer than K_2 satisfies

$$m(w) = \begin{cases} 0 & \text{if } w \neq \widetilde{w}, \\ \#E^{=}(w) - 1 & \text{if } w = \widetilde{w}. \end{cases}$$

Let $K = \max\{K_1, K_2\}$ and w be a factor of $\mathcal{L}(\mathbf{u})$ such that $|w| > K$. If $w \neq \widetilde{w}$, Lemma 13 implies that $\Gamma(w)$ is a tree. If $w = \widetilde{w}$, Lemma 15 implies that $\Theta(w)$ is a tree. $\square$

6 Proof of Zero Defect Conjecture for binary alphabet

The binary alphabet offers less variability for the construction of a strange phenomenon. The recent counterexamples to two conjectures concerning palindromes in fixed points of primitive morphisms — namely the Bucci-Vaslet counterexample to the Zero Defect Conjecture and the Labbé counterexample to the Hof-Knill-Simon (HKS) conjecture — use ternary alphabet. That conjecture [12] asks whether all palindromic fixed points of primitive substitutions are fixed by some conjugate of a morphism of the form $\alpha \mapsto p_\alpha p$ where p_α and p are palindromes. On a binary alphabet, Tan demonstrated the validity of the HKS conjecture, see [20]. Here we prove the Zero Defect Conjecture on a binary alphabet.

Lemma 23. *Let $\mathcal{A} = \{0,1\}$ and $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$. If $\mathcal{L}(\mathbf{u})$ is closed under reversal and $D(\mathbf{u}) > 0$, then there exists a non-palindromic factor $q \in \mathcal{L}(\mathbf{u})$ such that $0q0, 0q1, 1q0, 1q1 \in \mathcal{L}(\mathbf{u})$.*

Proof. By Proposition 5, as $D(\mathbf{u}) > 0$, there exist factors v and w in $\mathcal{L}(\mathbf{u})$ such that v is a complete mirror return word to w and v is not a palindrome. Let us consider the shortest v with this property. For this fixed v we find the longest w such that v is a complete mirror return word to w. It means that v has a prefix wa and a suffix $b\widetilde{w}$ where $a, b \in \mathcal{A}$ and $a \neq b$. Since on a binary alphabet every complete mirror return word to a letter is always a palindrome, we have $|w| > 1$. Without loss of generality we can write $w = 0q$ with $q \neq \varepsilon$. Consequently $v = 0u0$. Clearly u has a prefix q, the word u has a suffix $\widetilde{q}$ and u is not a palindrome. Our choice of v (to be the shortest non-palindromic mirror return to a factor) implies that u is not a complete mirror return word to q and thus q or $\widetilde{q}$ has another occurrence inside u. Since v is a complete mirror return word to $w = 0q$,

$$0q \text{ and } \widetilde{q}0 \text{ do not occur in } u. \tag{4}$$

Let us suppose that $q = \widetilde{q}$. Consider the shortest prefix of u which has exactly two occurrences of q. It is palindrome. Since v has a prefix $wa = 0qa$ the second occurrence of q is extended to the left as aq. Analogously, consider the shortest suffix of u which contains exactly two occurrences of q. It is a palindrome and thus the penultimate occurrence of q is extended to the right as qb. This contradicts (4) as $a \neq b$. We conclude that q is not a palindrome.

Now we show that occurrences of q and $\widetilde{q}$ in u alternate. Assume that there exists a factor of u, denoted by u', such that q is a prefix and a suffix of u' and u' does not contain $\widetilde{q}$. It follows that the longest palindromic suffix of u' is not unioccurrent in u'. Therefore $D(u') \geq 1$ (see Section 2.2), which contradicts the minimality of $|v|$.

The minimality of $|v|$ implies that all mirror return words to q in u are palindromes. Therefore, the leftmost occurrence of $\widetilde{q}$ in u is extended to the left as $a\widetilde{q}$ and the rightmost occurrence of q in u is extended to the right as qb. From (4) we deduce that $0qa$, $a\widetilde{q}1$, $1qb$, and $b\widetilde{q}0$ belong to $\mathcal{L}(\mathbf{u})$. The assumption that $\mathcal{L}(\mathbf{u})$ is closed under reversal and the fact that $a \neq b$ finish the proof. $\square$

Theorem 24. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ be a fixed point of a primitive morphism φ over a binary alphabet $\mathcal{A}$. If $D(\mathbf{u}) < +\infty$, then $D(\mathbf{u}) = 0$ or $\mathbf{u}$ is periodic.*

Proof. Assume the contrary, i.e., $\mathbf{u}$ is not periodic and $D(\mathbf{u}) > 0$ and let $\mathcal{A} = \{0, 1\}$.

Since $D(\mathbf{u})$ is finite, $\mathbf{u}$ is palindromic. As φ is primitive, $\mathcal{L}(\mathbf{u})$ is uniformly recurrent. Any uniformly recurrent word which is palindromic has its language closed under reversal. Due to Lemma 23 there exists a strong bispecial non-palindromic factor q with $m(q) = 1$.

Since $\mathbf{u}$ is not periodic, the morphism φ is acyclic. On the binary alphabet, it means that φ is well-marked. Applying repeatedly Lemma 11 (IV) and (V), we can construct an infinite sequence of strong bispecial factors $q, \Phi(q), \Phi^2(q), \Phi^3(q), \ldots$, each with bilateral multiplicity 1. By Lemma 11 (III), all these bispecial factors are non-palindromic. This contradicts Theorem 21. $\square$

7 Proof of Zero Defect Conjecture for marked morphisms

At first we have to stress that unlike the binary version, the statement of Theorem 1 does not speak about periodic fixed points. The following result from [14] allows to deduce that on a larger alphabet there is no ultimately periodic infinite word $\mathbf{u}$ fixed point of a primitive marked morphism such that $0 < D(\mathbf{u}) < \infty$.

Proposition 25. [14, Cor. 30, Cor. 32] *Let $\mathbf{u}$ be an eventually periodic fixed point of a primitive marked morphism φ over an alphabet $\mathcal{A}$. If $\mathbf{u}$ is palindromic, then $\mathcal{A} = \{0, 1\}$ is a binary alphabet and $\mathbf{u}$ equals $(01)^\omega$ or $(10)^\omega$.*

Due to the previous proposition, a fixed point of a marked morphisms on binary alphabet is either not eventually periodic or equal to $(01)^\omega$ or $(10)^\omega$. Since both words $(01)^\omega$ and $(10)^\omega$ have defect zero and the Zero Defect Conjecture for binary alphabet is proven by Theorem 24, we may restrict ourselves to alphabets with cardinality at least three.

First, we prove a multiliteral analogue of Lemma 23 for words with its language closed under reversal and with positive palindromic defect.

Theorem 26. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ have its language closed under reversal. If $D(\mathbf{u}) > 0$, then either*

1. *there exists a non-palindrome $q \in \mathcal{L}(\mathbf{u})$ such that $\Gamma(q)$ contains a cycle or*

2. *there exists a palindrome $q \in \mathcal{L}(\mathbf{u})$ such that $\Theta(q)$ contains a cycle.*

Moreover, if the empty word is the unique factor q with the above property, then there exists a letter with a non-palindromic complete return word.

Proof. Since $D(\mathbf{u}) > 0$, there exists a word $v = v_0 v_1 \cdots v_n$ such that w is a prefix of v, $\widetilde{w}$ is a suffix of v, v does not contain other occurrences of w or $\widetilde{w}$, v is not a palindrome and $|w| \geq 1$. Suppose that v is a word of minimal length with this property and suppose that w is the longest prefix of v such that $\widetilde{w}$ is a suffix of v. Then there exist letters $\alpha \neq \beta$ such that $w\alpha$ is a prefix and $\beta\widetilde{w}$ is a suffix of v. Let us define $t \in \mathcal{A}$ and $q \in \mathcal{A}^*$ to satisfy $w = tq$ (see Figure 2).

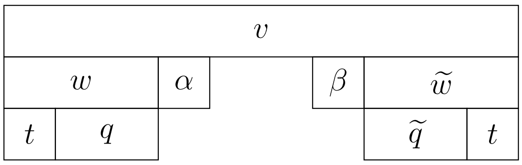

Figure 2: The complete mirror return word v to the factor w.

We discuss three cases:

1. Let us suppose $q = \widetilde{q} \neq \varepsilon$. Due to the minimality of $v = v_0 v_1 \dots v_n = tq\alpha \cdots \beta qt$, the non-palindromic factor $v_1 v_2 \dots v_{n-1} = q\alpha \cdots \beta q$ cannot be a complete return word to q and thus contains at least 3 occurrences of q. Let k be the number of occurrences q in v. For $i = 1, 2, \dots, k$, denote by γ_i the letter which precedes the i^{th} occurrence of q and by δ_i the letter which succeeds the i^{th} occurrence of q.

   • Obviously, $\gamma_1 = t$, $\delta_1 = \alpha$, and $\gamma_k = \beta$ and $\delta_k = t$.

   • Since v is a complete mirror return word to the factor $w = tq$, necessarily $t \neq \gamma_i$ for $i = 2, \dots, k$ and $t \neq \delta_i$ for $i = 1, \dots, k-1$. In particular, $\alpha \neq t$ and $\beta \neq t$.

   • Since each complete return word to q in v is a palindrome, $\delta_i = \gamma_{i+1}$ for $i = 1, 2, \dots, k-1$. We artificially put $\gamma_{k+1} = \delta_k = t$.

   According to the definition of $\Theta(q)$, if $\gamma_i \neq \gamma_{i+1} = \delta_i$, then the pair $\{\gamma_i, \gamma_{i+i}\}$ forms an edge. We want to find a cycle in $\Theta(q)$. For this purpose we modify the sequence of letters $\gamma_1, \gamma_2, \dots, \gamma_k, \gamma_{k+1}$ as follows: If $\gamma_{j+1} = \gamma_j$ for some index $j = 1, \dots, k$, then we erase from the sequence the $(j+1)^{th}$ entry γ_{j+1}. Then the modified sequence is a path in $\Theta(q)$ which starts and ends at t. The second vertex on the path is α, the penultimate vertex is β. As $\alpha \neq \beta$, the graph $\Theta(q)$ contains a cycle.

2. Let us suppose that $q = \varepsilon$. Now $v = v_0 v_1 \dots v_n = t\alpha v_2 v_3 \cdots \beta t$. It means that v is a complete return to the letter t which is non-palindromic. If $v_i \neq v_{i+1}$, the pair of consecutive letters $\{v_i, v_{i+1}\}$ is an edge in the graph $\Theta(\varepsilon)$ connecting vertices v_i and v_{i+1}. If we erase from the sequence $v_0, v_1, \dots, v_n$ each vertex v_{j+1} which coincides with its predecessor v_j, we get a path starting and ending in the vertex t. The first edge on this path is $\{t, \alpha\}$, the last one is $\{t, \beta\}$. As $\alpha \neq \beta$, the graph $\Theta(\varepsilon)$ contains a cycle.

3. Now we assume that $q \neq \widetilde{q}$. Note that occurrences of q and $\widetilde{q}$ alternate inside v. Indeed, suppose the contrary, that is there exists a complete return word z of q that has no occurrences of $\widetilde{q}$ and z is a factor of v. The longest palindrome suffix of z must be shorter than q. Therefore the longest palindromic suffix of z is not unioccurrent in z. This contradicts the minimality of v. Note also that v must contain other occurrences of q or $\widetilde{q}$ inside or otherwise we get

a contradiction on minimality of v. Let us denote k the number of occurrences of q in v. Clearly k equals to the number of occurrences of $\widetilde{q}$ as well.

Again we denote by γ_i the letter which precedes the i^{th} occurrence of q and by δ_i the letter which succeeds the i^{th} occurrence of q. In particular, $\gamma_1 = t$ and $\delta_1 = \alpha$. Analogously, we denote by $\tilde{\gamma_i}$ the letter which precedes the i^{th} occurrence of $\widetilde{q}$ and by $\tilde{\delta_i}$ the letter which succeeds the i^{th} occurrence of $\widetilde{q}$. In particular, $\tilde{\gamma_k} = \beta$ and $\tilde{\delta_k} = t$. Point out three important facts:

- $\gamma_i q \delta_i \in \mathcal{L}(\mathbf{u})$ implies $\{(\gamma_i, -1), (\delta_i, +1)\}$ is an edge in $\Gamma(q)$ for $i = 1, 2, \ldots, k$.
- As the language $\mathcal{L}(\mathbf{u})$ is closed under reversal, $\tilde{\gamma_i}\widetilde{q}\tilde{\delta_i} \in \mathcal{L}(\mathbf{u})$ implies $\{(\tilde{\delta_i}, -1), (\tilde{\gamma_i}, +1)\}$ is an edge in $\Gamma(q)$ for $i = 1, 2, \ldots, k$.
- Due to minimality of v, any mirror return to q in v is a palindrome. Thus $\delta_i = \tilde{\gamma_i}$ for $i = 1, 2, \ldots, k$ and $\tilde{\delta_i} = \gamma_{i+1}$ for $i = 1, 2, \ldots, k-1$.

Therefore, $\{(\gamma_i, -1), (\tilde{\gamma_i}, +1)\}$ is an edge in $\Gamma(q)$ for $i = 1, 2, \ldots, k$, $\{(\tilde{\gamma_i}, +1), (\gamma_{i+1}, -1)\}$ is an edge in $\Gamma(q)$ for $i = 1, 2, \ldots, k-1$ and $\{(\tilde{\gamma_k}, +1), (\tilde{\delta_k}, -1)\}$ is an edge in $\Gamma(q)$. We can summarize that the sequence of vertices

$$(\gamma_1, -1), (\tilde{\gamma_1}, +1), (\gamma_2, -1), (\tilde{\gamma_2}, +1), \ldots, (\gamma_k, -1), (\tilde{\gamma_k}, +1), (\tilde{\delta_k}, -1)$$

forms a path in the bipartite graph $\Gamma(q)$ with $\gamma_1 = \tilde{\delta_k} = t$ and $\tilde{\gamma_1} = \alpha \neq \beta = \tilde{\gamma_k} = t$. In this path the first and the last vertices coincide and the second and the penultimate vertices are distinct. Thus the graph $\Gamma(q)$ contains a cycle. □

As we have seen in Example 16 for the fixed point $\mathbf{u}$ of the morphism $\eta : a \mapsto aabcacba, b \mapsto aa, c \mapsto a$ for which the defect is known to be positive, the graph $\Theta(\varepsilon)$ contains a cycle. Since the defect of $\mathbf{u}$ is finite, Corollary 22 also applies. Thus there are no arbitrarily large palindromic factors w containing a cycle in their graph $\Theta(w)$ nor non-palindromic factors w containing a cycle in their graph $\Gamma(w)$. This is readily seen on the conjugacy word of $\eta_L \triangleright \eta_R$ which is aaa (see Fig. 3).

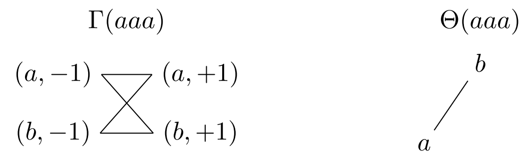

Figure 3: $\Gamma(aaa)$ contains a cycle but $\Theta(aaa)$ is a tree in the language of the fixed point of the morphism $a \mapsto aabcacba, b \mapsto aa, c \mapsto a$.

We are now ready to finish the proof for the multiliteral case.

Proof of Theorem 1. As the languages of the fixed points of φ and φ^k coincide, we may assume without loss of generality that the marked morphism φ has already the property $\text{LST}(\varphi_R) = \text{FST}(\varphi_L) = \text{Id}$.

Proving that the Zero Defect Conjecture holds in the case of marked morphisms amounts to prove that the defect is either zero or $+\infty$. Let us assume on the contrary that $0 < D(\mathbf{u}) < +\infty$. It follows that $\mathbf{u}$ is palindromic. The primitivity of φ implies that $\mathcal{L}(\mathbf{u})$ is closed under reversal.

Theorem 26 implies that there exists a factor q such that if $q \neq \widetilde{q}$ the graph $\Gamma(q)$ contains a cycle, or if $q = \widetilde{q}$, the graph $\Theta(q)$ contains a cycle. Lemma 11, property (IV), implies that for all n, there is a cycle in the graph of $\Phi^n(q)$.

If $q \neq \varepsilon$, then the primitivity of φ implies that $\lim_{n\to+\infty} |\Phi^n(q)| = +\infty$. If $q = \varepsilon$, then, again by Theorem 26, there exists a letter having non-palindromic complete return word. By the assumption of the theorem, there must exist a conjugate of φ distinct from φ itself. It implies that the conjugacy word of $\varphi_L \triangleright \varphi_R$ is nonempty, i.e., $\Phi(\varepsilon) \neq \varepsilon$. Moreover, $\lim_{n\to+\infty} |\Phi^n(q)| = +\infty$.

To conclude, we have that $\lim_{n\to+\infty} |\Phi^n(q)| = +\infty$ and there is a cycle in the graph of $\Phi^n(q)$ for all n. This is a contradiction with Corollary 22. $\square$

8 Comments

Let us comment two conjectures concerning palindromes in languages of fixed points of primitive morphisms.

- The counterexample to the Zero Defect Conjecture in full generality was already mentioned in the Introduction. It is taken from [7]. The fixed point of

  $$\varphi : a \mapsto aabcacba, b \mapsto aa, c \mapsto a$$

  has finite positive palindromic defect and is not periodic. There is a remarkable property of the fixed point $\mathbf{u} = \varphi(\mathbf{u})$.

  Let $\mu : a \mapsto ap, p \mapsto apaaaapaaaap$ be a morphism over the binary alphabet $\{a, p\}$. Let us denote $\mathbf{v}$ the fixed point of μ. Then one can easily verify that $\mathbf{u} = \pi(\mathbf{v})$, where $\pi : a \mapsto a, p \mapsto abcacba$. Moreover, $\mathbf{v}$ has zero defect.

  In other words, the counterexample word is just an image under π of a purely morphic binary word with zero defect.

- The counterexample to the question of Hof, Knill and Simon (recalled in Section 6) given in [15] by the first author is

  $$\psi : a \mapsto aca, b \mapsto cab, c \mapsto b.$$

  As mentioned in [18], the fixed point $\mathbf{u} = \psi(\mathbf{u})$ is again an image of a Sturmian word $\mathbf{v}$ under a morphism $\pi : \{0, 1\} \mapsto \{a, b, c\}$ and the Sturmian word $\mathbf{v}$ itself is a fixed point of a morphism over binary alphabet $\{0, 1\}$. Since $\mathbf{v}$ is Sturmian, its defect is zero.

Both counterexamples are in some sense degenerate. Both words are on ternary alphabet, but the binary alphabet is hidden in their structure. For further research in this area, it would be instructive to find another kind of counterexamples to both mentioned conjectures. In this context we mention that the second and third authors showed in [19] that any uniformly recurrent infinite word $\mathbf{u}$ with a finite defect is a morphic image of a word $\mathbf{v}$ with defect 0.

Acknowledgements

We are grateful to the anonymous referees for their many valuable comments one of them leading to a reorganisation of the structure of the article including a proof of Theorem 21 which generalizes Theorem 3.10 proved for rich words in [1].

The first author is supported by a postdoctoral Marie Curie fellowship (BeIPD-COFUND) cofunded by the European Commission. The second and the third authors acknowledge support of GAČR 13-03538S (Czech Republic).

References

[1] L. Balková, E. Pelantová, and Š. Starosta, *Sturmian jungle (or garden?) on multi-literal alphabets*, RAIRO-Theor. Inf. Appl., 44 (2010), pp. 443–470.

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

[3] V. Berthé, C. Felice, F. Dolce, J. Leroy, D. Perrin, C. Reutenauer, and G. Rindone, *Acyclic, connected and tree sets*, Monatsh. Math., 176 (2014), pp. 521–550.

[4] 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.

[5] 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.

[6] 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 (2009), pp. 60–74.

[7] M. Bucci and E. Vaslet, *Palindromic defect of pure morphic aperiodic words*, in Proceedings of the 14th Mons Days of Theoretical Computer Science, 2012.

[8] J. Cassaigne and F. Nicolas, *Factor complexity*, in Combinatorics, automata and number theory, vol. 135 of Encyclopedia Math. Appl., Cambridge Univ. Press, Cambridge, 2010, pp. 163–247.

[9] X. Droubay, J. Justin, and G. Pirillo, *Episturmian words and some constructions of de Luca and Rauzy*, Theor. Comput. Sci., 255 (2001), pp. 539–553.

[10] X. Droubay and G. Pirillo, *Palindromes and Sturmian words*, Theor. Comput. Sci., 223 (1999), pp. 73–85.

[11] A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni, *Palindromic richness*, Eur. J. Combin., 30 (2009), pp. 510–531.

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

[13] S. Jitomirskaya and B. Simon, *Operators with singular continuous spectrum: III. Almost periodic Schrödinger operators*, Comm. Math. Phys., 165 (1994), pp. 201–205.

[14] S. Labbé and E. Pelantová, *Palindromic sequences generated from marked morphisms*, Eur. J. Comb., 51 (2016), pp. 200–214.

[15] S. Labbé, *A counterexample to a question of Hof, Knill and Simon*, Electron. J. Comb., 21 (2014).

[16] 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, 1997.

[17] ——, *Algebraic combinatorics on words*, no. 90 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2002.

[18] Z. Masáková, E. Pelantová, and Š. Starosta, *Exchange of three intervals: itineraries, substitutions and palindromicity*, preprint available at `http://arxiv.org/abs/1503.03376`, (2015).

[19] E. Pelantová and Š. Starosta, *Almost rich words as morphic images of rich words*, Int. J. Found. Comput. Sci., 23 (2012), pp. 1067–1083.

[20] B. Tan, *Mirror substitutions and palindromic sequences*, Theor. Comput. Sci., 389 (2007), pp. 118–124.

Chapter 5

Morphic images of episturmian words having finite palindromic defect

Authors:	Š. Starosta
Publication reference:	Eur. J. Combin., 51 (2016), pp. 359–371
Cited as:	[IV]

European Journal of Combinatorics 51 (2016) 359–371

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

Morphic images of episturmian words having finite palindromic defect

Štěpán Starosta

Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00, Prague 6, Czech Republic

ARTICLE INFO

Article history:
Received 25 November 2014
Accepted 1 July 2015
Available online 17 July 2015

ABSTRACT

We study morphisms from certain classes and their action on episturmian words. The first class is P_{ret}. In general, a morphism of class P_{ret} can map an infinite word having zero palindromic defect to a word having infinite palindromic defect. We show that the image of an episturmian word, which has zero palindromic defect, under a morphism of class P_{ret} has always its palindromic defect finite. We also focus on letter-to-letter morphisms to binary alphabet: we show that images of ternary episturmian words under such morphisms have zero palindromic defect. These results contribute to the study of an unsolved question of characterization of morphisms that preserve finite, or even zero, palindromic defect. They also enable us to construct new examples of binary words having zero or finite H-palindromic defect, where $H = \{\mathrm{Id}, R, E, RE\}$ is the group generated by both involutory antimorphisms on a binary alphabet.

1. Introduction

In combinatorics on words, the most famous class of words probably is the class of Sturmian words: aperiodic words having minimal factor complexity possible (see [24]). Sturmian words are profoundly studied and many generalizations are known, see for instance [4]. One such generalization of Sturmian words are episturmian words. Episturmian words were inspired by Arnoux–Rauzy words (see [28,1]). An infinite word over a k-letter alphabet is *episturmian* if it is closed under reversal and has at most one left special factor of each length. Refer for instance to [17,22,19] for more results on this class.

E-mail address: stepan.starosta@fit.cvut.cz.

http://dx.doi.org/10.1016/j.ejc.2015.07.001
0195-6698/

A notion related to the study of episturmian words is a *palindrome*—a word equal to its reversal. Episturmian words are rich in palindromes: they contain the maximum number of distinct palindromic factors possible. Precisely, we say that a finite word w is *rich* if it contains exactly $|w| + 1$ distinct palindromic factors, which is the upper bound for the number of distinct palindromic factors in a finite word of length $|w|$ (see [17]). The notion is extended to infinite words: an infinite word is *rich* if every its factor is rich.

In the context of this upper bound on the number of palindromic factors, a measure of the count of missing palindromic factors was introduced in [9]: the *palindromic defect* $D(w)$ of a finite word w is

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

where $\mathrm{Pal}(w)$ is the set of all palindromic factors of w. The palindromic defect of an infinite word $\mathbf{u}$ is defined by $D(\mathbf{u}) = \sup\{D(w) : w \text{ is a factor of } \mathbf{u}\}$. If $D(\mathbf{u})$ is finite, we say that $\mathbf{u}$ is *almost rich*. (If it is zero, then $\mathbf{u}$ is rich as already mentioned.)

Besides episturmian words, examples of rich words include some well-explored word classes such as words coding symmetric interval exchange transformation (see [2]) and words coding rotation on two intervals (see [8]). Properties and characterizations of rich words are studied for instance in [20,4,13,11]. General properties and characterizations of almost rich words are studied in [20,5,6].

In this paper, we study richness and almost richness of images of episturmian words by a morphism from a specific class. Our first result states that we obtain an almost rich word while using a morphism of class P_{ret} introduced in [5] (see Section 2.2 later for the definition).

Theorem 1. *Let $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ be an episturmian word and $\pi : \mathcal{A}^* \to \mathcal{B}^*$ be a morphism of class P_{ret}. The word $\pi(\mathbf{u})$ is almost rich.*

The second main result involves a letter-to-letter projection of a ternary episturmian word to a binary alphabet. We use the following definition for such a projection.

Definition 2. Let $\mathcal{A}$ be an alphabet and $\mathcal{A}'$ its proper subset. A morphism $\zeta : \mathcal{A} \to \{A, B\}$ defined by

$$\zeta : a \mapsto \begin{cases} A \text{ if } a \in \mathcal{A}', \\ B \text{ otherwise} \end{cases}$$

is called a *binary projection* from $\mathcal{A}$.

The second main result states that we obtain a rich word by projecting a ternary episturmian word to a binary alphabet.

Theorem 3. *Let $\mathbf{u}$ be an episturmian word over a ternary alphabet $\mathcal{A}$ and ζ be a binary projection from $\mathcal{A}$. The word $\zeta(\mathbf{u})$ is rich.*

Our motivation for these results is to find new binary words which are rich in a generalized sense—with respect to both symmetries given by the involutive antimorphisms on a binary alphabet: the reversal R and the exchange of letters E. We give a definition in Section 4, see also [25,26] for more information on this generalization. To construct new binary words rich in this generalized sense, we use the recent results of [27] which provides theorems that relate the classical richness and the generalized richness on a binary alphabet.

Our computer experiments suggest that we can improve Theorem 3: we can drop the requirement on the size of the alphabet $\mathcal{A}$. We state this hypothesis in the last section along with some comments.

The paper is organized as follows. The next section contains some necessary definitions and basic results. Section 3 contains overview of results on episturmian words and proofs of the main results. Finally, Section 4 contains an application of the main results: a construction of binary words which are rich and almost rich in the generalized sense. The last section states some comments and open questions.

2. Preliminaries

2.1. Notions of combinatorics on words

Let $\mathcal{A}$ be an *alphabet*—a finite set of *letters*. A finite sequence $w = w_0w_1 \cdots w_{n-1}$ with $w_i \in \mathcal{A}$ for all i is a *finite word*. The *length* of the word w is denoted by $|w|$ and equals n. The unique word of length 0 is the *empty word*, it is denoted by ε. The set $\mathcal{A}^*$ is the set of all finite words over $\mathcal{A}$. The set $\mathcal{A}^*$ equipped with concatenation forms a free monoid with the neutral element ε. A word $v \in \mathcal{A}^*$ is a *factor* of a word $w \in \mathcal{A}^*$ if $w = uvz$ for some words $u, z \in \mathcal{A}^*$. If $u = \varepsilon$, then v is a *prefix* of w; if $z = \varepsilon$, then v is a *suffix* of w. If w is of the form $w = uz$ for some words u and z, we write $z = u^{-1}w$ and $u = wz^{-1}$.

An *infinite word* over $\mathcal{A}$ is a sequence $\mathbf{u} = (u_j)_{j\in\mathbb{N}} = u_0u_1u_2 \ldots$. The set of all infinite words over $\mathcal{A}$ is denoted by $\mathcal{A}^{\mathbb{N}}$. A finite word $w \in \mathcal{A}^*$ of length n is a *factor* of $\mathbf{u} = (u_j)_{j\in\mathbb{N}}$ if there exists an integer i such that $w = u_iu_{i+1} \cdots u_{i+n-1}$. The integer i is an *occurrence* of the factor w in $\mathbf{u}$. The *language* of $\mathbf{u}$ is the set of all its factors and is denoted by $\mathcal{L}(\mathbf{u})$. Given $a \in \mathcal{A}$ and $w \in \mathcal{A}^*$, if $wa \in \mathcal{L}(\mathbf{u})$, then wa is a *right extension* of the factor w in $\mathbf{u}$. The set of all right extensions of w is denoted by $\mathrm{Rext}(w)$. Any factor of $\mathbf{u}$ has at least one right extension. If w has at least two right extensions, it is *right special*. The notions of *left extension* of a factor, *left special* factor and $\mathrm{Lext}(w)$ are defined analogously. A factor w which is left and right special is *bispecial*.

An infinite word $\mathbf{u}$ is *recurrent* if any factor of $\mathbf{u}$ has infinitely many occurrences in $\mathbf{u}$. If for every factor the sequence of all its consecutive occurrences has its first difference bounded, then the word $\mathbf{u}$ is *uniformly recurrent*. Let rw be a factor of $\mathbf{u}$ such that rw has a prefix w and w occurs as a factor in rw exactly twice. Such a word r is a *return word* of w and the word rw is a *complete return word* of w. An infinite word $\mathbf{u}$ is uniformly recurrent if and only if every its factor has finitely many return words.

An infinite word $\mathbf{u}$ is *eventually periodic* if there exist words p and z such that $\mathbf{u} = pzzz \ldots = pz^{\omega}$. It is *periodic* if $p = \varepsilon$. If an infinite word is not eventually periodic, it is *aperiodic*.

Let $\mathcal{C}_{\mathbf{u}}$ be the mapping $\mathbb{N} \to \mathbb{N}$ which is determined by $\mathcal{C}_{\mathbf{u}}(n) = \#(\mathcal{L}(\mathbf{u}) \cap \mathcal{A}^n)$, i.e., it counts the factors of length n of the word $\mathbf{u}$. This mapping is the *factor complexity* of $\mathbf{u}$.

Given two alphabets $\mathcal{A}$ and $\mathcal{B}$, a mapping $\mu : \mathcal{A}^* \to \mathcal{B}^*$ is a *morphism* if $\mu(wv) = \mu(w)\mu(v)$ for all $w, v \in \mathcal{A}^*$. It is an *antimorphism* if $\mu(wv) = \mu(v)\mu(w)$ for all $w, v \in \mathcal{A}^*$. An infinite word $\mathbf{u}$ is *closed under the mapping* μ if for any factor $w \in \mathcal{L}(\mathbf{u})$ we have also $\mu(w) \in \mathcal{L}(\mathbf{u})$. A morphism $\nu : \mathcal{A}^* \to \mathcal{B}^*$ is a *conjugate morphism* to a morphism μ if there exists a word $w \in \mathcal{B}^*$ such that for any letter $a \in \mathcal{A}$ we have $w\mu(a) = \nu(a)w$ or $\mu(a)w = w\nu(a)$.

A mapping Ψ is *involutory* if $\Psi^2 = \mathrm{Id}$. The most frequent involutory antimorphism is the reversal mapping R which is given by

$$R(w_0w_1 \cdots w_{n-1}) = w_{n-1} \cdots w_0 \quad \text{with } w_i \in \mathcal{A}.$$

It can be easily seen that if a word $\mathbf{u}$ is closed under an involutory antimorphism, then $\mathbf{u}$ is recurrent.

As said above, if $w = R(w)$, then w is a palindrome. The set of all palindromes occurring as factors of a finite word w is denoted by $\mathrm{Pal}(w)$. The *palindromic complexity* of an infinite word $\mathbf{u}$ is the mapping $\mathcal{P}_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$ defined by $\mathcal{P}_{\mathbf{u}}(n) = \#\{p \in \mathcal{L}(\mathbf{u}) : p = R(p), |p| = n\}$, i.e., the number of palindromic factors of length n.

A palindrome w is *centered at* $x \in \mathcal{A} \cup \{\varepsilon\}$ if $w = vxR(v)$ for some word v. Obviously, a palindrome is centered at ε if and only if it is of even length.

Recall that a word is rich if its palindromic defect is 0 and it is almost rich if its palindromic defect is finite. To prove (almost) richness of a word we will use the characterization of rich words given in [3]. It exploits the notions of bilateral order $\mathrm{b}(w)$ of a factor w and palindromic extension of a palindrome. The *bilateral order* was introduced in [15] as

$$\mathrm{b}(w) = \#\{awb \in \mathcal{L}(\mathbf{u}) : a, b \in \mathcal{A}\} - \#\mathrm{Rext}(w) - \#\mathrm{Lext}(w) + 1. \tag{1}$$

If a and b are letters, we say that awb is a *both-sided extension of* w if $awb \in \mathcal{L}(\mathbf{u})$. Moreover, if w is a palindrome and $a = b$, then awa is its *palindromic extension*. The set of all palindromic extensions of a palindrome $w \in \mathcal{L}(\mathbf{u})$ is denoted by $\mathrm{Pext}(w)$: we have

$$\mathrm{Pext}(w) = \{awa : awa \in \mathcal{L}(\mathbf{u}), a \in \mathcal{A}\}.$$

Theorem 4 ([4]). *Let* $\mathbf{u}$ *be an infinite word that is closed under reversal. The word* $\mathbf{u}$ *is rich if and only if any bispecial factor* w *of* $\mathbf{u}$ *satisfies*

$$\mathrm{b}(w) = \begin{cases} \#\mathrm{Pext}(w) - 1 & \text{if } w \text{ is a palindrome;} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The next theorem is another characterization of rich words which will be useful.

Theorem 5 ([14]). *A word* $\mathbf{u}$ *is rich if and only if for every* $w \in \mathcal{L}(\mathbf{u})$*, any factor of* $\mathbf{u}$ *containing exactly two occurrences of* w *or* $R(w)$*, one as a prefix and one as a suffix, is a palindrome.*

2.2. Class P_{ret}

We are interested in classes of morphisms that may, under some conditions, preserve richness or almost richness. The following definition is from [5].

Definition 6. Let $\varphi : \mathcal{B}^* \to \mathcal{A}^*$ and $r \in \mathcal{A}^*$ be a palindrome. We say that φ is of class P_{ret} (with respect to r) if the following is true:

- $\varphi(b)r$ is a palindrome for any $b \in \mathcal{B}$,
- $\varphi(b)r$ contains exactly 2 distinct occurrences of r, one as a prefix and one as a suffix, for any $b \in \mathcal{B}$,
- $\varphi(b) \neq \varphi(c)$ for all $b, c \in \mathcal{B}$, $b \neq c$.

A direct consequence of the above definition is that any morphism $\varphi \in P_{ret}$ is injective and $\varphi(s)r$ is a palindrome if and only if $s \in \mathcal{B}^*$ is a palindrome. An important property of the class is also that it is closed under taking composition of morphisms, see [5].

In [5], this class is used to show relations between rich and almost rich words in general: every almost rich word is an image of a rich word by a morphism from P_{ret}. The class is also used in [10] to show that every episturmian word is an image of an Arnoux–Rauzy word.

The following example taken from [5] illustrates that there exists a morphism of class P_{ret} which maps a word having finite palindromic defect to a word with infinite palindromic defect. One of the main results of this article, Theorem 1, states this cannot happen if such a morphism acts on an episturmian word.

Example 7. Let $v_0 = \varepsilon$ and for all $i > 0$ set

$$v_i = v_{i-1}0v_{i-1}1v_{i-1}1v_{i-1}0v_{i-1}2v_{i-1}2v_{i-1}0v_{i-1}1v_{i-1}1v_{i-1}0v_{i-1}.$$

Let $\mathbf{v} \in \{0, 1, 2\}^{\mathbb{N}}$ be determined by the limit

$$\mathbf{v} = \lim_{i \to +\infty} v_i$$

and $\varphi : \{0, 1, 2\}^* \to \{0, 1\}^*$ be given as follows:

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

Proposition 5.7 in [5] states that $D(\mathbf{v}) = 0$ and $D(\varphi(\mathbf{v})) = +\infty$.

The key property of $\mathbf{v}$ in this case is that the palindromes v_i have two palindromic extensions $1v_i1$ and $2v_i2$ which produce (by application of φ) the same palindrome $1\varphi(v_i)0101$ but they also produce a non-palindromic complete return word to it (it is contained in $\varphi(1v_i1v_i0v_i2v_i2)$). Having infinitely many non-palindromic complete return words to palindromes implies that the palindromic defect is infinite (see [20]).

3. Morphic images of episturmian words versus richness

In this section, we first recall and deduce some properties of episturmian words that will be needed later. A proof of Theorem 1 is given in the second subsection. The last subsection contains a proof of Theorem 3.

3.1. Properties of episturmian words

For a thorough treatment of properties of episturmian words, we refer the reader to [17,22] or survey papers [19,7].

As already mentioned, an infinite word $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is *episturmian* if for any n there exists at most one left special factor of length n and $\mathbf{u}$ is closed under reversal. Let $k = \#\mathcal{A}$. If for any n there exists exactly one left special factor of length n with k left extensions, then $\mathbf{u}$ is a k-ary *Arnoux–Rauzy* word.

One can show from the definition that an episturmian word is uniformly recurrent. The most important examples of episturmian words can be constructed using the following construction. Let $(\delta_i)_{i=0}^{+\infty} \in \mathcal{A}^{\mathbb{N}}$ and $w_0 = \varepsilon$. Set $w_{i+1} = (w_i\delta_i)^R$ where $\delta_i \in \mathcal{A}$ and w^R is the shortest palindrome having w as a prefix (the so-called *palindromic closure* of w). The word $\mathbf{u}$ defined by

$$\mathbf{u} = \lim_{i\to+\infty} w_i$$

is an episturmian word over $\mathcal{A}$ and it is called *standard episturmian*. The sequence (δ_i) is the *directive sequence* of $\mathbf{u}$. A standard episturmian word is Arnoux–Rauzy if and only if each letter occurs infinitely many times in its directive sequence. In general, if $w_0 \in \mathcal{A}^*$, then such a word $\mathbf{u}$ is called *standard word with seed w_0 and directive sequence $(\delta_i)_{i=0}^{+\infty}$*; such words are investigated in [12], where it is shown that they are almost rich.

The importance of standard episturmian words follows from the fact that given any episturmian word $\mathbf{u}$, there exists a unique standard episturmian word with the same language. Since (almost) richness is a property of language, when studying it, we can restrict ourselves to the standard episturmian words which are determined by their directive sequence. A standard episturmian word can be recognized by looking at its prefixes: an episturmian word is standard if and only if each its prefix is a left special factor.

A basic property of any episturmian word $\mathbf{u}$ is that one letter is *separating*: it occurs in every factor of length 2. The separating letter of $\mathbf{u}$ is the first letter of the directive sequence $\Delta = \delta_0\delta_1\delta_2\ldots$ of the corresponding standard episturmian word. Denote ℓ the least integer such that δ_0^{ℓ} is not a prefix of Δ. Then the word $x\delta_0^k y$ with $x, y \neq \delta_0$ is a factor of the episturmian word $\mathbf{u}$ only if $k = \ell - 1$ or $k = \ell$. Moreover, the word δ_0^{ℓ} is a factor of $\mathbf{u}$ if and only if the letter δ_0 occurs at least once in the sequence $\delta_\ell\delta_{\ell+1}\delta_{\ell+2}\ldots$.

Any palindromic prefix of a standard episturmian word $\mathbf{u}$ is equal to w_n for some $n \in \mathbb{N}$. The complete return words of w_n are described by Theorem 4.4 of [23]. Denote by $\mathcal{F}$ the set $\{\delta_m : m \geq n\}$. The set of all complete return words of w_n consists of elements of the form

$$(w_n x)^R = r_x w_n$$

for all $x \in \mathcal{F}$. The word r_x is the corresponding return word. Thus, the palindromic prefix w_n has exactly $\#\mathcal{F}$ return words. Note that we have $\mathcal{F} = \mathcal{A}$ for all n if and only if $\mathbf{u}$ is an Arnoux–Rauzy word.

Bispecial factors play a crucial role in the study of the language of an infinite word. We give essential properties of bispecial factors of an episturmian word. Any bispecial factor w of an episturmian word $\mathbf{u}$ is a palindrome; moreover there exists $n \in \mathbb{N}$ such that $w = w_n$, where w_n is a palindromic prefix of the standard episturmian word having the same language as $\mathbf{u}$. For a standard Arnoux–Rauzy word $\mathbf{u}$, the sets of palindromic prefixes and bispecial factors coincide and the set of all both-sided extensions of a bispecial factor w_n satisfies

$$\{xw_ny \in \mathcal{L}(\mathbf{u}) : x, y \in \mathcal{A}\} = \{\delta_n w_n x \in \mathcal{L}(\mathbf{u}) : x \in \mathcal{A}\} \cup \{xw_n\delta_n \in \mathcal{L}(\mathbf{u}) : x \in \mathcal{A}\}.$$

It follows that $\delta_n w \delta_n$ is the only palindromic extension of w_n in $\mathcal{L}(\mathbf{u})$ (see [16]).

In the more general case when $\mathbf{u}$ is a standard episturmian word, bispecial factors are a subset of all palindromic prefixes of $\mathbf{u}$. They are a proper subset if and only if $\mathbf{u}$ is periodic. Let $\mathcal{F}' = \{\delta_m : m > n\}$. The set of all both-sided extensions of the bispecial factor w_n satisfies

$$\{xw_ny \in \mathcal{L}(\mathbf{u}) : x, y \in \mathcal{A}\} = \{\delta_n w_n x \in \mathcal{L}(\mathbf{u}) : x \in \mathcal{F}'\} \cup \{xw_n\delta_n \in \mathcal{L}(\mathbf{u}) : x \in \mathcal{F}'\}. \tag{3}$$

Since in what follows palindromic extensions of w_n require some attention, note that this time we may have $\delta_n w_n \delta_n \notin \mathcal{L}(\mathbf{u})$. Precisely, we have $\delta_n w_n \delta_n \in \mathcal{L}(\mathbf{u})$ if and only if $\delta_n \in \mathcal{F}'$. It follows that the letter δ_n is also the unique letter such that $\delta_n v \in \mathcal{L}(\mathbf{u})$ for all $v \in \mathrm{Rext}(w_n)$. If $\mathbf{u}$ is aperiodic, δ_n is the unique letter such that $\delta_n w_n$ is right special.

This last property will be used when dealing with a bispecial factor w of a general episturmian word $\mathbf{u}$, i.e., for such a bispecial factor w there is a unique letter $a \in \mathcal{A}$ such that $av \in \mathcal{L}(\mathbf{u})$ for all $v \in \mathrm{Rext}(w)$. The word awa is the only possible palindromic extension of w in $\mathcal{L}(\mathbf{u})$. Note that such a letter a is well defined for any factor w of $\mathbf{u}$: if w is not left special, then a is the unique letter extending it to the left.

The next lemmas give an insight on the structure of return words of an episturmian word.

Lemma 8. *Let $\mathbf{u}$ be an episturmian word over $\mathcal{A}$ and $w \in \mathcal{L}(\mathbf{u})$. Let $r_1, r_2, \ldots, r_s$ be the list of all distinct return words of w in $\mathbf{u}$. Define a morphism Ψ over $\mathcal{E} = \{1, 2, \ldots, s\}$ by the rule $k \mapsto r_k$ for all $k \in \mathcal{E}$. There exists an episturmian word $\mathbf{v}$ over $\mathcal{E}$ such that $\mathbf{u} = g\Psi(\mathbf{v})$ for some finite word g.*

Proof. First, let us assume that w is a bispecial factor of $\mathbf{u}$. If $\mathbf{u}$ is an Arnoux–Rauzy word, then $\#\mathcal{E} = \#\mathcal{A}$ and the claim follows directly from the proof of Theorem 1 in [10] and from Theorem 3 ibidem. In case $\#\mathcal{E} < \#\mathcal{A}$, the proof is analogous and is left to the reader.

Suppose w is not bispecial and it can be extended in a unique way to the shortest bispecial factor $b = uwv$ in which it occurs. A factor r is a return word of b if and only if $u^{-1}ru$ is a return word of w. Thus the morphism Ψ (defined using the return words of w) is a conjugate morphism of the morphism defined using the return words of b and we can use the validity of the statement for the bispecial factor b.

The last case, w is not bispecial and it cannot be extended to a bispecial factor, is trivial as $\mathbf{u}$ is periodic and w has only one return word and thus $\#\mathcal{E} = 1$. □

The word $\mathbf{v}$ from Lemma 8 records the structure of return words of the factor w in $\mathbf{u}$. Such a word was also studied in [18] and we keep the same terminology: we say that the word $\mathbf{v}$ from the previous lemma is a *derived word of* $\mathbf{u}$ *with respect to the factor* w and its corresponding morphism is Ψ. By comparing the definitions we obtain that if w is a palindrome, then $\Psi \in P_{ret}$.

Lemma 9. *Let $\mathbf{u}$ be a standard episturmian word $\mathbf{u}$ over the alphabet $\mathcal{A} = \{0, 1, 2\}$ with the directive sequence $\delta_0\delta_1\delta_2\ldots$, where $\delta_0 = 0$. Denote by ℓ the least integer such that 0^ℓ is not a prefix of the directive sequence.*

If $w \in \{0^\ell, 1, 2\} \cap \mathcal{L}(\mathbf{u})$, then no two return words of w have the same length.

Proof. If w has exactly one return word, then the claim is trivially satisfied.

Suppose w has more than 1 return word. Let $b = uwv$ be the shortest bispecial factor containing w. Any return word of w has the form $u^{-1}ru$ where r is a return word of b. Therefore, it is enough to prove the statement only for the bispecial factor b. It follows that the bispecial factor b contains at least two distinct letters. Indeed, for $w = 0^\ell$, it follows from the fact that $0^{\ell+1}$ is not a factor. For $w \neq 0^\ell$ it follows from the fact that 0 is the separating letter.

As already mentioned, Theorem 4.4 of [23] describes the return words of bispecial factors. In particular, any complete return word of b is equal to the palindromic closure of bx for some letter $x \in \mathcal{A}$. Moreover, since $\mathbf{u}$ is standard, each bispecial factor b of $\mathbf{u}$ is a palindromic prefix of $\mathbf{u}$. Since the prefixes of the word are constructed using the palindromic closure and the directive sequence, the palindromic closures $(bx)^R$ and $(by)^R$ for distinct letters x and y have the same length if and only if neither x nor y occurs in b (see [21]). As b contains at least two distinct letters, this cannot happen on the ternary alphabet. □

Remark 10. As one can see from the proof of the last lemma, for alphabets having cardinality greater than 3 the claim of the last lemma does not hold. Let us suppose that $\mathbf{u}$ is a standard episturmian word with its directive sequence $\Delta = 01023\ldots$. We have

$$\mathbf{u} = 01001020100103010010201001\ldots.$$

Clearly, $\ell = 2$. Since there exists j such that $j \geq \ell$ and $\delta_j = \delta_0 = 0$, we have $0^\ell \in \mathcal{L}(\mathbf{u})$. The factors 0010201 and 0010301 are return words of $w = 0^\ell = 00$ and they are of the same length.

Lemma 11. *Let* $\mathbf{u}$ *be an episturmian word over* $\mathcal{A}$*. Let* $\mathcal{A}' \subset \mathcal{A}$ *and* xpy *be a factor of* $\mathbf{u}$ *such that* $x, y \in \mathcal{A}'$ *and* p *does not contain any letter from* $\mathcal{A}'$*. The word* p *is a palindrome.*

Proof. We can suppose without loss of generality that the word $\mathbf{u}$ is standard episturmian. We order the elements of $\mathcal{A}'$ according to their first occurrence in the directive sequence of $\mathbf{u}$, i.e.,

$$\mathcal{A}' = \{x_1, x_2, \ldots, x_n\},$$

where x_i appears for the first time in the directive sequence of $\mathbf{u}$ before the first occurrence of x_j if and only if $i < j$.

Suppose $x = x_1$. Let p_1 be the longest prefix of $\mathbf{u}$ that does not contain x_1 (or equivalently, the longest prefix of $\mathbf{u}$ over $\mathcal{A} \setminus \mathcal{A}'$). Since $\mathbf{u}$ is a standard word constructed using the palindromic closure operators, the word p_1 is a palindrome (it may be empty). The shortest palindromic prefix containing x_1 is $p_1x_1p_1$. Using Theorem 4.4 of [23] we have that any complete return word of $p_1x_1p_1$ has the form $(p_1x_1p_1z)^R$ for some $z \in \mathcal{A}$. If $p_1x_1p_1$ is bispecial, it is the shortest bispecial factor containing x_1; if it is not bispecial, then $\mathbf{u}$ is periodic. In both cases, any word x_1py is a prefix of a complete return word of $p_1x_1p_1$ without the leading p_1. If $z \in \mathcal{A}'$, then $y = z$ and $xpy = x_1p_1z$, i.e., $p = p_1$ is a palindrome. If $z \notin \mathcal{A}'$, then $y = x_1$ and $xpy = p_1^{-1}(p_1x_1p_1z)^Rp_1^{-1}$, i.e., p is again a palindrome.

Take $j > 1$ and suppose $x = x_j$. Let p_j be the longest prefix of $\mathbf{u}$ that does not contain x_j. The shortest palindromic prefix of $\mathbf{u}$ containing x_j is the word $p_jx_jp_j$. If $p_jx_jp_j$ is bispecial, it is the shortest bispecial factor containing x_j; if it is not bispecial, then $\mathbf{u}$ is periodic. Both cases imply that any complete return word of x_j starts with x_jp_j. Since $x_jp_j = x_jp_1x_1s$ for some word s, there is just one word of the desired form, precisely $x_jp_1x_1$, i.e., $y = x_1$, and the proof is finished. □

The following notion of ancestor is useful when dealing with images by a morphism. Suppose that w is a factor of $\Psi(\mathbf{v})$ for some infinite word $\mathbf{v}$ and morphism Ψ. A factor $e_0e_1 \cdots e_n \in \mathcal{L}(\mathbf{v})$ is an *ancestor* by Ψ of w if $\Psi(e_0e_1 \cdots e_n)$ contains w but neither $\Psi(e_1 \cdots e_n)$ nor $\Psi(e_0e_1 \cdots e_{n-1})$ contains w.

The following lemma can be considered as a generalization of Theorem 5. The lemma uses an additional notion. Let us fix a factor w of an episturmian word $\mathbf{u}$ over the alphabet $\mathcal{A}$ and a subset $\mathcal{E} \subset \mathcal{A}$. A both-sided extension xwy of w in $\mathbf{u}$ is called $\mathcal{E}$*-extension* if x and y belong to $\mathcal{E}$.

Lemma 12. *Let* w *be a palindromic factor of an episturmian word* $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ *and* $a \in \mathcal{A}$ *be the unique letter satisfying* $av \in \mathcal{L}(\mathbf{u})$ *for all* $v \in \mathrm{Rext}(w)$*. Let* $\mathcal{E} \subset \mathcal{A}$ *such that* $a \in \mathcal{E}$*. If* xuy *is a factor of* $\mathbf{u}$ *containing exactly two* $\mathcal{E}$*-extensions of* w *– one as its prefix and one as its suffix – then* u *is a palindrome.*

Proof. If w is not a bispecial factor, then the only both-sided extension of w is awa. It means that $xuy = aua$ is a complete return word of the palindrome awa and thus, since $\mathbf{u}$ is rich, using Theorem 5, the word u is a palindrome.

For a bispecial factor w, let n be the number of occurrences of the factor w in u.

If $n = 2$, then u is a complete return word of w and using again Theorem 5 we find that u is a palindrome.

If $n > 2$, we first show that the factor xuy begins with xwa and ends with awy. Suppose that xuy begins with xwz where $z \in \mathcal{A}$ and $z \neq a$. There is only one complete return word of w beginning with wz, it ends with zw and it is always followed by the letter a, thus we have a contradiction to $n > 2$.

We deal with two subcases:

(1) $x = y$.

Since the word xuy has a prefix $p = xwa$, it has a suffix $R(p) = awx$. The factor $xuy = xux$ does not contain further occurrences of p and $R(p)$. According to Theorem 5, such a factor is a palindrome in any rich word, in particular in the episturmian word $\mathbf{u}$, as we want to show.

(2) $x \neq y$.

Let $\mathcal{B}$ be the set of letters such that for all $z \in \mathcal{B}$ the word wz is a right extension of w. For all $z \in \mathcal{B}$, denote r_z the return word of w which has a prefix wz. Define a morphism $\Psi : \mathcal{B}^* \to \mathcal{A}^*$ by setting $\Psi(z) = r_z$ for all $z \in \mathcal{B}$. Using Lemma 8, let $\mathbf{v} \in \mathcal{B}^{\mathbb{N}}$ be the episturmian word which is a derived word of $\mathbf{u}$ with respect to the factor w such that Ψ is the corresponding morphism. The ancestor by Ψ of the factor xuy is a factor $xasay \in \mathcal{L}(\mathbf{v})$ with s containing only letters from $\{z : z = a \text{ or } z \notin \mathcal{E}\}$ and moreover, in any pair of neighboring letters in asa exactly one of the letters is a, i.e., aa does not occur in asa.

If $x, y \neq a$, then we apply Lemma 11 with $\mathcal{A}' = \mathcal{E} \setminus \{a\}$. We get that asa is a palindrome and thus $u = \Psi(asa)w$ is a palindrome as well.

If $x = a$ and $y \neq a$, then since a is the separating letter in $\mathbf{v}$, we have that $xasay = aasay$ is an image of the morphism σ_a defined by $a \mapsto a$ and $b \mapsto ab$ for any $b \neq a$. Thus $aasay = \sigma_a(a\tilde{s}y)$ where $a\tilde{s}y$ is a factor of an episturmian word and $\tilde{s}$ is produced from s by erasing all letters a in it. Moreover, the word $\tilde{s}$ does not contain letters from $\mathcal{E}$. Applying Lemma 11 with $\mathcal{A}' = \mathcal{E}$ we get that $\tilde{s}$ is a palindrome. Since σ_a and Ψ belong to the class P_{ret}, their composition $\Psi\sigma_a$ is in P_{ret} with respect to $\Psi(a)w$. Consequently, $u = \Psi(\sigma_a(\tilde{s})a)w$ is a palindrome. □

Remark 13. The morphism σ_a is one of *episturmian morphisms*, see [22]. For instance, episturmian morphisms can be used to construct episturmian words.

3.2. Proof of Theorem 1

Besides the proof of Theorem 1 which concerns images of episturmian words under a morphism of class P_{ret}, we also provide in this section a description of the bilateral order of long bispecial factors in these infinite words. This result is needed also to prove Theorem 3 in the next section.

Lemma 14. *Let $\mathbf{u}$ be an episturmian word and let $\pi : \mathcal{A}^* \to \mathcal{B}^*$ be a morphism of class P_{ret} with respect to the palindrome r. If $w \in \mathcal{L}(\pi(\mathbf{u}))$ is a bispecial factor containing at least one occurrence of r, then its bilateral order $b(w)$ is -1 or 0, and satisfies (2).*

Proof. Let $w \in \mathcal{L}(\pi(\mathbf{u}))$ be a bispecial factor containing r. Let us denote sr and rp the prefix and the suffix of w containing the factor r exactly once. As π belongs to the class P_{ret}, there exists a unique factor $v \in \mathcal{L}(\mathbf{u})$ such that $w = s\pi(v)rp$.

Let $\{z_1, \dots, z_t\} \subset \mathcal{B}$ be the set of all letters such that wz_i is a right extension of w. Set $R_i = \{b \in \mathcal{A}\colon \pi(v)rpz_i$ is a prefix of $\pi(vb)r$ and $vb \in \mathcal{L}(\mathbf{u})\}$. This set is well defined due to the choice of p and v. Since π is of class P_{ret}, it follows that $R_i \cap R_j = \emptyset$ if $i \neq j$. As w is bispecial, we have $t > 1$ and thus v is right special.

Let $\{y_1, \dots, y_\ell\}$ be the set of all letters such that $y_i w$ is a left extension of w. Analogously to R_i, we define the sets L_i. We conclude that v is left special, thus it is bispecial.

Denote by a the unique letter such that $ag \in \mathcal{L}(\mathbf{u})$ for all $g \in \mathrm{Rext}(w)$. Next, we show that there exists an index i such that $a \in R_i$. Suppose the contrary: for all i and $b \in R_i$ we have that wz_i is a factor of $\pi(cvb)$ for some letter c (possibly depending on b). Using (3), since $b \neq a$, it implies that $c = a$, i.e., v is not special—a contradiction. Let without loss of generality $a \in R_1$. Analogously, we have $a \in L_1$.

Using (3) for v, we have that if $y_j wz_k \in \mathcal{L}(\pi(\mathbf{u}))$, then $k > 1$ implies $j = 1$ and symmetrically $j > 1$ implies $k = 1$. Thus we have

$$\{ywz \in \mathcal{L}(\pi(\mathbf{u}))\} \setminus \{y_1wz_1\} = \{y_1wz_k\colon k > 1\} \cup \{y_jwz_1\colon j > 1\}. \tag{4}$$

To calculate the bilateral order of w, it remains to see if y_1wz_1 is its both-sided extension or not: if $y_1wz_1 \in \mathcal{L}(\pi(\mathbf{u}))$, then $b(w) = 0$; otherwise $b(w) = -1$. We first show a few more claims.

Since v is a palindrome, the word w is a palindrome if and only if $s = R(p)$. As rp is the longest common prefix of all the words of the form $\pi(b)r$ with $b \in \bigcup R_k$, and s is the longest common suffix of all the words of the form $\pi(c)$ with $c \in \bigcup L_j$, using the fact that $\pi \in P_{ret}$ we conclude that $s = R(p)$ if and only if $\bigcup R_k = \bigcup L_j$. Moreover, for each k there exists j such that $R_k = L_j$ (and especially $R_1 = L_1$). It follows from (4) that y_jwz_k is a palindrome if and only if $j = k = 1$, i.e., y_1wz_1 is the only possible palindromic extension of w.

We now show that if $R_1 = L_1 = \{a\}$, then w is a palindrome. Suppose that $R_1 = L_1 = \{a\}$ and $s \neq R(p)$, i.e., w is not a palindrome. We start with the case $|s| < |R(p)|$. Let $b \in R_k$ with $k > 1$. As the longest common prefix of $\pi(b)r$ and $\pi(a)r$ is rp, the longest common suffix of $\pi(b)$ and $\pi(a)$ is $R(p)$. As $|s| < |R(p)|$, we obtain that s is a suffix of $R(p)$. Since $bva \in \mathcal{L}(\mathbf{u})$, we obtain $b \in L_1$ which is a contradiction since $b \neq a$ and $L_1 = \{a\}$. The other case is analogous.

If $R_1 \neq \{a\}$, then there exists $b \in R_1$ such that $b \neq a$. It implies that $avb \in \mathcal{L}(\mathbf{u})$, which implies that $y_1wz_1 \in \mathcal{L}(\pi(\mathbf{u}))$. Similarly, $L_1 \neq \{a\}$ implies $y_1wz_1 \in \mathcal{L}(\pi(\mathbf{u}))$. We conclude that $y_1wz_1 \in \mathcal{L}(\pi(\mathbf{u}))$ if and only if $ava \in \mathcal{L}(\mathbf{u})$ or $R_1 \neq \{a\}$ or $L_1 \neq \{a\}$.

Let us finish the evaluation of $b(w)$. If $y_1wz_1 \notin \mathcal{L}(\pi(\mathbf{u}))$, i.e., $b(w) = -1$, then $R_1 = L_1 = \{a\}$, which implies that w is a palindrome and w has no palindromic extension. Therefore, $b(w) = \#\mathrm{Pext}(w) - 1 = -1$ and (2) is satisfied. Suppose $y_1wz_1 \in \mathcal{L}(\pi(\mathbf{u}))$. If w is a palindrome, since y_1wz_1 is the only palindromic extension, (2) is again satisfied. If w is not a palindrome, we have $b(w) = 0$, and (2) is also satisfied. □

Note that if $\mathbf{u}$ is an Arnoux–Rauzy word, then the case $y_1wz_1 \notin \mathcal{L}(\pi(\mathbf{u}))$ of the last proof does not occur since we always have $ava \in \mathcal{L}(\mathbf{u})$.

Proof of Theorem 1. Let π be a morphism from P_{ret} with respect to r. Since $\mathbf{u}$ is a derived word of $\pi(\mathbf{u})$ with respect to r, then $\mathbf{u}$ is periodic if and only if $\pi(\mathbf{u})$ is periodic.

Suppose $\mathbf{u}$ is periodic. Thus, $\pi(\mathbf{u})$ is also periodic. In Theorem 4 of [9], it is shown that an infinite periodic word contains infinitely many palindromes if and only if it is equal to $(pq)^\omega$ for some palindromes p and q. Since $\mathbf{u}$ contains infinitely many palindromes, $\pi(\mathbf{u})$ contains infinitely many palindromes and we can write $\pi(\mathbf{u}) = (pq)^\omega$ for some palindromes p and q from $\mathcal{L}(\pi(\mathbf{u}))$. It can be easily seen that such a word has finite palindromic defect.

Suppose $\mathbf{u}$ is aperiodic. As shown in [5], a uniformly recurrent infinite word $\mathbf{v}$ containing infinitely many palindromes has finite palindromic defect if and only if there exists an integer M such that any complete return word in $\mathbf{v}$ of any palindrome w of length at least M is palindromic. As for each palindrome $p \in \mathcal{L}(\mathbf{u})$, the word $\pi(p)r$ is a palindrome, and moreover a factor of $\pi(\mathbf{u})$, we conclude that $\pi(\mathbf{u})$ contains infinitely many palindromes. Moreover, uniform recurrence of $\mathbf{u}$ implies uniform recurrence of $\pi(\mathbf{u})$. Thus, to show finiteness of the palindromic defect of $\pi(\mathbf{u})$, it remains to verify palindromicity of complete return words to palindromic factors of length N such that $N \geq M = 2\max\{|\pi(a)r| : a \in \mathcal{A}\}$. Let w be a palindromic factor of $\pi(\mathbf{u})$ such that $|w| \geq M$.

First, suppose that w is a bispecial factor. The word w contains r and according to the proof of Lemma 14 there exist a bispecial palindromic factor v, a letter a and a factor s such that we have

$$w = s\pi(v)rR(s), \quad av \in \mathcal{L}(\mathbf{u}) \text{ for all } v \in \mathrm{Rext}(w), \text{ and } s \text{ is a suffix of } \pi(a).$$

Let f be a complete return word of w. Clearly, $f = s\pi(u)rR(s)$ for some factor $u \in \mathcal{L}(\mathbf{u})$ and v is a prefix and a suffix of u. If s is the empty word ε, then u is a complete return word of v in $\mathbf{u}$. Since u is a palindrome, the factor $f = \pi(u)r$ is a palindrome as well.

If $s \neq \varepsilon$, then we apply Lemma 12 with $\mathcal{E} = \{b \in \mathcal{A} : s \text{ is a suffix of } \pi(b)\}$. We have $a \in \mathcal{E}$. According to our notation any ancestor of w has the form xvy, where $x, y \in \mathcal{E}$. Moreover, any ancestor of the complete return word $f = s\pi(u)rR(s)$ equals xuy with $x, y \in \mathcal{E}$ and the factor xuy has only two occurrences of $\mathcal{E}$-extensions of the factor v. According to Lemma 12, the factor u is a palindrome and thus $\pi(u)r$ is a palindrome as well. It implies that the complete return word $f = s\pi(u)rR(s)$ is a palindrome, too.

To complete the proof, we need to discuss the case when w is not bispecial. Since $\pi(\mathbf{u})$ is closed under reversal, such a palindrome w has a unique both-sided extension and this extension is palindromic. It implies that the shortest bispecial factor in which the palindrome w occurs is a bispecial palindrome, say $qwR(q)$. Moreover, u is a complete return word of w if and only if $quR(q)$ is a complete return word of u. Since $quR(q)$ is a palindrome, u is a palindrome as well. □

Remark 15. Given an episturmian word $\mathbf{u}$ and a morphism π of class P_{ret}, the palindromic defect of $\pi(\mathbf{u})$ can be nonzero. It suffices to choose π such that for some letter $a \in \mathcal{A}$ the palindromic defect of $\pi(a)$ is nonzero, which is possible in general.

For instance, let $\mathbf{u}$ be the Fibonacci word, i.e., the fixed point of the morphism given by $0 \mapsto 01$ and $1 \mapsto 0$. The Fibonacci word is Sturmian, thus also an episturmian word. Let $\pi \in \mathcal{P}_{ret}$ be determined by

$$0 \mapsto 1110100110010 \quad \text{and} \quad 1 \mapsto 1.$$

(The morphism π is of class P_{ret} with respect to the palindrome $r = 111$.) We have $D(\pi(0)) = 1$, thus $D(\pi(\mathbf{u})) \geq 1$. In fact, $D(\pi(\mathbf{u})) = 1$.

3.3. Proof of Theorem 3

Before proving Theorem 3, we need two lemmas dealing with images of episturmian words by morphisms of a special form.

Lemma 16. *Let* $\mathbf{u}$ *be an episturmian word over* $\mathcal{A}$ *and* $\pi : \mathcal{A} \to \{A, B\}$ *a morphism of class* P_{ret} *such that* $\pi(x) = AB^{k_x}$ *with* $k_x \geq 0$ *for all* x*. The word* $\pi(\mathbf{u})$ *is rich.*

Proof. Any morphism from class P_{ret} is injective, thus $k_x \neq k_y$ for any $x, y \in \mathcal{A}, x \neq y$. We first show that if w is a bispecial factor of $\pi(\mathbf{u})$, then its bilateral order $b(w)$ satisfies (2).

If the bispecial factor w contains the letter A, the claim follows from Lemma 14. Suppose w does not contain any occurrence of the letter A, i.e., $w = B^k$ with $k < \max_{x \in \mathcal{A}} k_x$.

The set of both-sided extensions of w clearly contains AB^{k+1} and $B^{k+1}A$. Thus the set of both-sided extensions of w has cardinality $2 + \#\mathrm{Pext}(w)$. Since $\#\mathrm{Rext}(w) = 2 = \#\mathrm{Lext}(w)$, according to Eq. (1) we have $\mathrm{b}(w) = \#\mathrm{Pext}(w) - 1$.

Since (2) is satisfied for any bispecial factor of $\pi(\mathbf{u})$, the richness of $\pi(\mathbf{u})$ follows from Theorem 4. □

Lemma 17. *Let* $\mathbf{u}$ *be a ternary episturmian word over* $\{0, 1, 2\}$ *with the separating letter* 0 *and let* $\pi : \mathcal{A} \to \{A, B\}$ *be a morphism such that* $\pi(0) = A$ *and* $\pi(x) = AB$ *for all* $x \neq 0$*. The word* $\pi(\mathbf{u})$ *is rich.*

Proof. Let ℓ be the minimal integer such that 0^ℓ is not a prefix of the directive sequence of the standard episturmian word having the same language as $\mathbf{u}$. It means that any two non-zero letters of $\mathbf{u}$ are separated by the block 0^ℓ or $0^{\ell-1}$.

If 0^ℓ does not belong to $\mathcal{L}(\mathbf{u})$, then $\pi(\mathbf{u}) = (A^\ell B)^\omega$, which is a rich word.

Let us suppose that 0^ℓ occurs in $\mathbf{u}$. Using Lemma 8, let $\mathbf{v} \in \mathcal{B}^{\mathbb{N}}$ be an episturmian word that is a derived word of $\mathbf{u}$ with respect to the factor 0^ℓ and let Ψ be the corresponding morphism. Thus, for any letter $b \in \mathcal{B}$ the image $\Psi(b)$ is a return word of 0^ℓ in $\mathbf{u}$. Using Lemma 9 one can see that the morphism $\pi\Psi$ is of class P_{ret} with respect to the palindrome $r = A^{\ell+1}$.

We will show that if w is a bispecial factor of $\pi(\mathbf{u})$, then the bilateral order $b(w)$ satisfies (2). If the bispecial factor w contains r, the claim follows from Lemma 14.

Suppose w does not contain any occurrence of the word r. Since any occurrence of the letter B in $\pi(\mathbf{u})$ is followed by the block A^ℓ or $A^{\ell+1}$ one can see that the longest factor (not necessarily bispecial) of $\pi(\Psi(\mathbf{v}))$ which does not contain $r = A^{\ell+1}$ is of the form $A^\ell(BA^\ell)^k$ for some integer k and that w is its factor. It implies that w is of the form $w = A^\ell(BA^\ell)^s$, with $s < k$. Since $A^\ell(BA^\ell)^k$ is a factor, we can see that AwB and BwA are factors of $\pi(\mathbf{u})$. Since w is a palindrome, we have $b(w) = \#\mathrm{Pext}(w) - 1$ and consequently w satisfies (2) as well.

Since (2) is satisfied for any bispecial factor of $\pi(\mathbf{u})$, Theorem 4 implies that $\pi(\mathbf{u})$ is rich. □

Proof of Theorem 3. Without loss of generality let us suppose that $\mathcal{A} = \{0, 1, 2\}$ and 0 is the separating letter of $\mathbf{u}$. Let $x \in \mathcal{A}$ be the letter such that $\zeta(x) \neq \zeta(y)$ for all $y \in \mathcal{A}$ such that $y \neq x$.

Suppose $x = 0$. Let $\mathbf{v}$ be a derived word of $\mathbf{u}$ with respect to the factor 0 and let Ψ be the corresponding morphism. Since the return words of $x = 0$ are 0 and $0y$ for all $y \neq 0$, we may choose Ψ such that $\Psi(0) = 0$. It implies that the morphism $\pi = \zeta\Psi$ and the word $\mathbf{v}$ satisfy the assumptions of Lemma 17 which implies that $\pi(\mathbf{v}) = \zeta(\mathbf{u})$ is rich.

Suppose $x \neq 0$, i.e., x is not the separating letter in $\mathbf{u}$. Let again $\mathbf{v}$ be a derived word of $\mathbf{u}$ with respect to the word x and let Ψ be the corresponding morphism. Using Lemma 9, one can see that $\pi = \zeta\Psi$ is of the form as in the assumptions of Lemma 16 which implies that $\pi(\mathbf{v}) = \zeta(\mathbf{u})$ is rich. □

4. Application: construction of H-rich words

A generalization of the notion of richness was introduced in [25]. As already mentioned, Theorems 1 and 3 and results of [27] may be used to produce examples of words that are rich in this generalized sense. In order to give these examples, and to give the definition of generalized richness on binary alphabets, we need a few more definitions.

Let E be the antimorphism over $\{0, 1\}$ which exchanges the two letters, i.e., for instance we have $E(011) = E(1)E(1)E(0) = 001$. Let $\Psi \in \{R, E\}$. If $p = \Psi(p)$, the word p is a Ψ-*palindrome*.

The Ψ*-palindromic complexity* of an infinite word $\mathbf{u}$ is the mapping $\mathcal{P}^{\Psi}_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$, defined by $\mathcal{P}^{\Psi}_{\mathbf{u}}(n) = \#\{p \in \mathcal{L}(\mathbf{u}) : p = \Psi(p), |p| = n\}$.

Let $H = \{R, E, RE, \mathrm{Id}\}$. For binary words closed under all elements of H, it is shown in [25] that

$$\mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 4 \geq \mathcal{P}^{R}_{\mathbf{u}}(n+1) + \mathcal{P}^{R}_{\mathbf{u}}(n) + \mathcal{P}^{E}_{\mathbf{u}}(n+1) + \mathcal{P}^{E}_{\mathbf{u}}(n) \quad \text{for every } n \geq 1. \tag{5}$$

When equality is reached for every $n \geq 1$, we say that the word is H*-rich*. When the equality holds except for finitely many n, then the word is *almost* H*-rich*. The notion of H-palindromic defect is also introduced in [26]: almost H-rich are words with finite H-palindromic defect and H-rich words are words with zero H-palindromic defect.

In [25], the notion of generalized richness is defined for an arbitrary alphabet and a finite group of morphisms and antimorphisms G containing at least one antimorphism. We also refer the reader to [26] where the generalized richness is investigated and various characterizations are shown. In this context, the "classical" richness defined in the previous sections is for $G = \{\mathrm{Id}, R\}$, and it coincides with (almost) $\{\mathrm{Id}, R\}$-richness, or shortly (almost) R-richness.

Known examples of H-rich words include binary generalized Thue–Morse words, see [30], and complementary symmetric Rote words, see [27]. A word from the first mentioned class can be constructed as follows. Let $b > 1$ and φ be determined by $0 \mapsto 012\cdots(b-2)(b-1)$ and $1 \mapsto 123\cdots(b-1)b$ where the letters are taken modulo 2. The fixed point of φ starting in 0 is a binary generalized Thue–Morse word. For $b = 2$, we retrieve the famous Thue–Morse word. The second class, complementary symmetric Rote words, consist of words having complexity $2n$ for all n and set of factors closed under E.

To construct H-rich and almost H-rich words, we shall use the operation S acting over words over the alphabet $\{0, 1\}$ and defined by $S(w_0w_1w_2\ldots) = v_1v_2v_3\ldots$, where $v_i = w_{i-1} + w_i \bmod 2$. Note that given $\mathbf{v} = (v_i)_{i=1}^{+\infty} \in \{0, 1\}^{\mathbb{N}}$, there exists $\mathbf{w} \in \{0, 1\}$ such that $\mathbf{v} = S(\mathbf{w})$: it suffices to choose $w_0 \in \{0, 1\}$ and then let $w_{i+1} = w_i + v_i \bmod 2$ for all $i > 0$. Thus, a preimage by S is unique up to permutation of letters.

We will use the two following results of [27].

Proposition 18 *([27]). Let $\mathbf{w} \in \{0, 1\}^{\mathbb{N}}$ and $\mathbf{v} = S(\mathbf{w}) \in \{0, 1\}^{\mathbb{N}}$ be uniformly recurrent. If $\mathcal{L}(\mathbf{v})$ contains infinitely many palindromes centered at the letter 1 and infinitely many palindromes not centered at the letter 1, then $\mathbf{w}$ is closed under all elements of H.*

Theorem 19 *([27]). Let $\mathbf{w} \in \{0, 1\}^{\mathbb{N}}$ and let $\mathbf{w}$ be closed under all elements of $H = \{\mathrm{Id}, E, R, ER\}$. Then $\mathbf{w}$ is H-rich (resp. almost H-rich) if and only if $S(\mathbf{w})$ is R-rich (resp. almost R-rich).*

We first use Proposition 18 and Theorem 19 to produce almost H-rich words.

Proposition 20. *Let $\mathbf{u}$ be an Arnoux–Rauzy word over $\mathcal{A}$ and $\pi : \mathcal{A}^* \to \{0, 1\}$ a morphism of class P_{ret} with respect to the palindrome r. Let $\mathbf{w}$ be a word such that $\pi(\mathbf{u}) = S(\mathbf{w})$. Assume that r or $\pi(a)r$ for some $a \in \mathcal{A}$ is centered at 1, and r or $\pi(b)r$ for some $b \in \mathcal{A}$ is not centered at 1. The word $\mathbf{w}$ is almost H-rich.*

Proof. Let $v \in \mathcal{A}^*$ be a palindrome. We have the following observation.

- If v is of even length, the palindrome $\pi(v)r$ is centered at the same letter as r.
- If v is of odd length, centered at $a \in \mathcal{A}$, then $\pi(v)r$ is centered at the same letter as $\pi(a)r$.

Since any Arnoux–Rauzy word over $\mathcal{A}$ has infinitely many palindromes centered at a for each $a \in \mathcal{A} \cup \{\varepsilon\}$, see [16], we conclude that $\pi(\mathbf{u})$ contains infinitely many palindromes centered at 1 and infinitely many palindromes not centered at 1.

Since $\mathbf{u}$ is uniformly recurrent, so is the word $\pi(\mathbf{u})$. We may now apply Proposition 18 and obtain that $\mathbf{w}$ is closed under all elements of H. By Theorem 1, the word $\pi(\mathbf{u}) = S(\mathbf{w})$ is almost R-rich. Theorem 19 implies that $\mathbf{w}$ is almost H-rich. □

Using Theorem 3, we can produce H-rich words.

Proposition 21. *Let $\mathbf{u}$ be a ternary Arnoux–Rauzy word over the alphabet $\mathcal{A} = \{0, 1, 2\}$ and ζ be a binary projection over $\mathcal{A}$. If $\mathbf{w} \in \{0, 1\}^{\mathbb{N}}$ is a preimage of $\zeta(\mathbf{u})$ by S, i.e., $S(\mathbf{w}) = \zeta(\mathbf{u})$, then $\mathbf{w}$ is H-rich.*

Proof. We use the two well-known properties of Arnoux–Rauzy words which were already mentioned: any Arnoux–Rauzy word is uniformly recurrent and any Arnoux–Rauzy word contains infinitely many palindromes centered at a with $a \in \mathcal{A} \cup \{\varepsilon\}$. Therefore, $\zeta(\mathbf{u})$ is uniformly recurrent and contains infinitely many palindromes centered at 1 and infinitely many palindromes centered at ε. Due to Proposition 18, the word $\mathbf{w}$ is closed under all elements of H. As Theorem 3 states that $\zeta(\mathbf{u}) = S(\mathbf{w})$ is R-rich, according to Theorem 19, the word $\mathbf{w}$ is H-rich. $\square$

Remark 22. Let us note that the last proposition indeed produces new H-rich words as claimed in the introduction.

First, it is known that the image by S of a complementary-symmetric Rote word is a Sturmian word, see [29]. As the preimage by S is unique up to permutation of letters, it remains to check that a binary projection of an Arnoux–Rauzy word $\zeta(\mathbf{u})$ is not a Sturmian word nor equal to $S(\mathbf{t}_b)$ where $\mathbf{t}_b$ is the binary generalized Thue–Morse word with parameter $b > 1$ as above.

Assume that $\mathbf{u}$ is standard and its directive sequence of $\mathbf{u}$ has prefix 01201. Let ζ be the binary projection given by $0 \mapsto 0$, $1 \mapsto 1$ and $2 \mapsto 0$. One can easily check that the set of all factors of length 5 of $\zeta(w_5)$ has cardinality 7. It implies that $C_{\zeta(\mathbf{u})}(5) \geq 7$. Therefore, $\zeta(\mathbf{u})$ is not a Sturmian word. Since neither 000 nor 111 is a factor of $\mathbf{t}_b$, we obtain that 00 is not a factor of $S(\mathbf{t}_b)$. As 00 is a factor of $\zeta(\mathbf{u})$, we conclude that $\zeta(\mathbf{u}) \neq S(\mathbf{t}_b)$ for all b.

Clearly, as infinitely many standard Arnoux–Rauzy sequences have the directive sequence with the prefix 01201, the class of newly obtained H-rich words is infinite.

5. Comments and open questions

As already mentioned in the introduction, our computer experiments suggest that we may conjecture a more general statement than Theorem 3:

Conjecture 23. *Let $\mathbf{u}$ be an episturmian word over $\mathcal{A}$ and ζ be a binary projection from $\mathcal{A}$. The word $\zeta(\mathbf{u})$ is rich.*

Let us indicate the difficulties that would arise when one would adapt the current proofs in this paper to prove the conjecture. As noted in Remark 10, the assumption of ternary alphabet is needed in Lemma 9. Thus, to adapt the proof one would first need to find a more general version of Lemma 9. Then, using the ideas of the proof of Theorem 3, one could conclude that $\zeta(\mathbf{u})$ is almost rich. However, the main difficulty here would arise when proving richness, which would require to deal with "short" factors in general (namely, when generalizing the results of Lemmas 16 and 17).

Besides this conjecture, one may inquire about the class of binary words that is obtained by applying a binary projection on an episturmian word. Study of special factors, factor complexity and structure of return words may provide insight into this class. It would also be of interest to see if there is a relation to some known class of binary words.

Acknowledgments

The author acknowledges financial support from the Czech Science Foundation grant GAČR 13-35273P. The computer experiments were performed using the open-source algebra system SAGE [31].

References

[1] P. Arnoux, G. Rauzy, Représentation géométrique de suites de complexité $2n + 1$, Bull. Soc. Math. France 119 (1991) 199–215.
[2] P. Baláži, Z. Masáková, E. Pelantová, Factor versus palindromic complexity of uniformly recurrent infinite words, Theoret. Comput. Sci. 380 (3) (2007) 266–275.
[3] L. Balková, E. Pelantová, Š. Starosta, Palindromes in infinite ternary words, RAIRO Theor. Inform. Appl. 43 (4) (2009) 687–702.
[4] L. Balková, E. Pelantová, Š. Starosta, Sturmian jungle (or garden?) on multiliteral alphabets, RAIRO Theor. Inform. Appl. 44 (2010) 443–470.
[5] L. Balková, E. Pelantová, Š. Starosta, Infinite words with finite defect, Adv. Appl. Math. 47 (3) (2011) 562–574.
[6] L. Balková, E. Pelantová, Š. Starosta, Proof of the Brlek-Reutenauer conjecture, Theoret. Comput. Sci. 475 (2013) 120–125.

[7] J. Berstel, Sturmian and episturmian words: a survey of some recent results, in: CAI'07: Proceedings of the 2nd International Conference on Algebraic Informatics, in: Lecture Notes in Computer Science, vol. 4728, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 23–47.
[8] A. Blondin Massé, S. Brlek, S. Labbé, L. Vuillon, Palindromic complexity of codings of rotations, Theoret. Comput. Sci. 412 (46) (2011) 6455–6463.
[9] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer, On the palindromic complexity of infinite words, Internat. J. Found Comput. Sci. 15 (2) (2004) 293–306.
[10] M. Bucci, 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, Springer-Verlag, 2009, pp. 259–266.
[11] M. Bucci, A. de Luca, A. De Luca, Rich and periodic-like words, in: Proceedings of the 13th International Conference on Developments in Language Theory, Springer-Verlag, Berlin, Heidelberg, 2009, pp. 145–155.
[12] M. Bucci, A. de Luca, A. De Luca, L.Q. Zamboni, On different generalizations of episturmian words, Theoret. Comput. Sci. 393 (1–3) (2008) 23–36.
[13] M. Bucci, A. De Luca, A. Glen, L.Q. Zamboni, A new characteristic property of rich words, Theoret. Comput. Sci. 410 (30–32) (2009) 2860–2863.
[14] M. Bucci, A. De Luca, A. Glen, L.Q. Zamboni, A connection between palindromic and factor complexity using return words, Adv. Appl. Math. 42 (1) (2009) 60–74.
[15] J. Cassaigne, Complexity and special factors, Bull. Belg. Math. Soc. Simon Stevin 4 (1) (1997) 67–88.
[16] D. Damanik, L.Q. Zamboni, Combinatorial properties of Arnoux-Rauzy subshifts and applications to Schrödinger operators, Rev. Math. Phys. 15 (07) (2003) 745–763.
[17] X. Droubay, J. Justin, G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (1–2) (2001) 539–553.
[18] F. Durand, A characterization of substitutive sequences using return words, Discrete Math. 179 (1–3) (1998) 89–101.
[19] A. Glen, J. Justin, Episturmian words: a survey, Theoret. Inf. Appl. 43 (3) (2009) 403–442.
[20] A. Glen, J. Justin, S. Widmer, L.Q. Zamboni, Palindromic richness, European J. Combin. 30 (2) (2009) 510–531.
[21] J. Justin, Episturmian morphisms and a Galois theorem on continued fractions, RAIRO. Theor. Inform. Appl. 39 (2005) 207–215.
[22] J. Justin, G. Pirillo, Episturmian words and episturmian morphisms, Theoret. Comput. Sci. 276 (1–2) (2002) 281–313.
[23] J. Justin, L. Vuillon, Return words in sturmian and episturmian words, Theoret. Inf. Appl. 34 (2000) 343–356.
[24] M. Morse, G.A. Hedlund, Symbolic dynamics II—Sturmian trajectories, Amer. J. Math. 62 (1940) 1–42.
[25] E. Pelantová, Š. Starosta, Languages invariant under more symmetries: overlapping factors versus palindromic richness, Discrete Math. 313 (2013) 2432–2445.
[26] E. Pelantová, Š. Starosta, Palindromic richness for languages invariant under more symmetries, Theoret. Comput. Sci. 518 (2014) 42–63.
[27] E. Pelantová, Š. Starosta, Constructions of words rich in palindromes and pseudopalindromes, preprint available at http://arxiv.org/abs/1409.2354.
[28] G. Rauzy, Suites à termes dans un alphabet fini, Séminaire de Théorie des Nombres de Bordeaux Anné 1982–1983 (exposé 25).
[29] G. Rote, Sequences with subword complexity $2n$, J. Number Theory 46 (1993) 196–213.
[30] Š. Starosta, Generalized Thue–Morse words and palindromic richness, Kybernetika 48 (3) (2012) 361–370.
[31] W. Stein, et al., Sage Mathematics Software (Version 6), The Sage Development Team, 2014. http://www.sagemath.org.

Chapter 6

Constructions of words rich in palindromes and pseudopalindromes

Authors: E. Pelantová and Š. Starosta
Publication reference: accepted in Discrete Math. Theor. Comput. Sci., (2015)
Cited as: [V]

Constructions of words rich in palindromes and pseudopalindromes

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

Štěpán Starosta
Faculty of Information Technology
Czech Technical University in Prague

Abstract

A narrow connection between infinite binary words rich in classical palindromes and infinite binary words rich simultaneously in palindromes and pseudopalindromes (the so-called H-rich words) is demonstrated. The correspondence between rich and H-rich words is based on the operation S acting over words over the alphabet $\{0, 1\}$ and defined by $S(u_0u_1u_2\ldots) = v_1v_2v_3\ldots$, where $v_i = u_{i-1} + u_i$ mod 2. The operation S enables us to construct a new class of rich words and a new class of H-rich words. Finally, the operation S is considered on the multiliteral alphabet $\mathbb{Z}_m$ as well and applied to the generalized Thue–Morse words. As a byproduct, new binary rich and H-rich words are obtained by application of S on the generalized Thue–Morse words over the alphabet $\mathbb{Z}_4$.

1 Introduction

In the present paper we concentrate on construction of infinite words which are filled with palindromes or pseudopalindromes to the highest possible level, the so-called *rich* words. Before we explain the expression "the highest possible level" we recall the basic notions we work with. We understand by an *infinite word* over a finite alphabet $\mathcal{A}$ a sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}} = u_0u_1u_2\ldots$, where $u_n \in \mathcal{A}$ for each $n \in \mathbb{N}$. A *factor* of $\mathbf{u}$ is a finite sequence $w = w_0w_1\cdots w_{n-1}$ of letters from $\mathcal{A}$ such that $w = u_iu_{i+1}\cdots u_{i+n-1}$ for some $i, n \in \mathbb{N}$. The set of all factors of $\mathbf{u}$ is the *language* of $\mathbf{u}$, usually denoted $\mathcal{L}(\mathbf{u})$. A finite word $w = w_0w_1\cdots w_{n-1}$ is called palindrome if w coincides with its reversal $R(w) = w_{n-1}w_{n-2}\cdots w_1w_0$.

Infinite words whose language contains infinitely many palindromes are being studied by many authors. Apart from the impulses from outside mathematics (such as [21] where these words are used in a model of solid materials with finite local complexity) the main reason of the interest of mathematicians is the variety of characterizations of rich words. To specify the expression "the highest possible level" one can adopt two distinct points of view: local and global.

From the local point of view, one looks at a finite piece of the infinite word, i.e., at a factor of $\mathbf{u}$, and counts the number of distinct palindromes occurring in this factor. A motivation for rich word definition was an inequality due to Droubay and Pirillo [19] which states that a finite word of length n contains at most $n + 1$ distinct palindromes (the empty word is counted as a palindrome). An infinite word is *rich*, or *full*, if every its factor of length n contains $n + 1$ distinct palindromes.

From the global point of view, one counts the palindromes of length n in the set of all factors of $\mathbf{u}$, i.e., in the language $\mathcal{L}(\mathbf{u})$. Let $\mathcal{C}_\mathbf{u}(n)$ and $\mathcal{P}_\mathbf{u}(n)$ denote the number of factors of length n and the number of palindromic factors of length n, respectively. As shown in [2], if $\mathcal{L}(\mathbf{u})$ is closed under reversal, then the number of palindromes in $\mathbf{u}$ is bounded from above by the relation

$$\mathcal{C}_\mathbf{u}(n+1) - \mathcal{C}_\mathbf{u}(n) + 2 \geq \mathcal{P}_\mathbf{u}(n+1) + \mathcal{P}_\mathbf{u}(n) \quad \text{for every } n \in \mathbb{N}. \tag{1}$$

In [15], Bucci, De Luca, Glen and Zamboni show that for infinite words with language closed under reversal the local and global points of view coincide. More precisely, $\mathbf{u}$ is rich if and only if the inequality in (1) can be written as an equality for every $n \in \mathbb{N}$.

Classic examples of rich words on binary alphabets include Sturmian words, i.e., infinite words over binary alphabet with the factor complexity $\mathcal{C}_{\mathbf{u}}(n) = n + 1$ for each $n \in \mathbb{N}$. Sturmian words can be generalized to multiliteral alphabets in many ways, see for example [6]. Two of these generalizations, namely k-ary Arnoux–Rauzy words and words coding k-interval exchange transformation with symmetric interval permutation, are rich as well. Both mentioned classes have their language closed under reversal.

Blondin Massé, Brlek, Garon and Labbé showed in [10] that rich words include complementary-symmetric Rote words. They can be defined as binary words with factor complexity $\mathcal{C}_{\mathbf{u}}(n) = 2n$ for every nonzero integer n and with language closed under the exchange of letters, see [26]. This implies that the language of a complementary-symmetric Rote word is closed under two mappings acting on the set $\{0,1\}^*$ of all finite binary words: the first is R and the second is E defined by $E(w_0 \cdots w_n) = E(w_n) \cdots E(w_0)$ for letters w_i and $E(0) = 1$ and $E(1) = 0$. Thus, the language of a complementary-symmetric Rote word is closed under all elements of a group $H = \{R, E, ER, \mathrm{Id}\}$. The same property has the language of the famous Thue–Morse word $\mathbf{t}$, nevertheless, it is well-known that $\mathbf{t}$ is not rich.

For binary words having language closed under all elements of H, we show in [24] that

$$\mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 4 \geq \mathcal{P}_{\mathbf{u}}(n+1) + \mathcal{P}_{\mathbf{u}}(n) + \mathcal{P}^{E}_{\mathbf{u}}(n+1) + \mathcal{P}^{E}_{\mathbf{u}}(n) \quad \text{for every } n \geq 1, \tag{2}$$

where $\mathcal{P}^{E}_{\mathbf{u}}$ is the function counting E-palindromes – words fixed by E – in the word $\mathbf{u}$. Analogously to the case of equality in (1), we say that an infinite word with language closed under all elements of H is H*-rich* if in (2) the equality holds for all $n \geq 1$. We also demonstrated that the Thue–Morse word $\mathbf{t}$ is H-rich. In [28] the second author proved that the binary generalization $\mathbf{t}_{b,2}$ of the Thue–Morse word is H-rich for all $b \geq 2$ (the definition of $\mathbf{t}_{b,2}$ is recalled in Preliminaries). In fact, the words $\mathbf{t}_{b,2}$ are the only H-rich words that have been found up to now.

One of the main aims of the present article is to describe a procedure which produces new H-rich words. We have found an inspiration in a connection between complementary-symmetric Rote words and Sturmian words due to Rote in [26]. Given an infinite word $\mathbf{u} = u_0 u_1 \ldots \in \{0,1\}^{\mathbb{N}}$, we set $S(\mathbf{u}) = v_1 v_2 \ldots \in \{0,1\}^{\mathbb{N}}$ with $v_i = (u_{i-1} + u_i) \bmod 2$ for all positive integer i. The operator S defines the mentioned relation: a word $\mathbf{u}$ is a complementary-symmetric Rote word if and only if $S(\mathbf{u})$ is a Sturmian word.

In Section 4, we investigate binary words which are simultaneously H-rich and also rich in the classical sense. In particular, we prove that every complementary-symmetric Rote word is H-rich, see Corollary 14. The main result concerning H-richness is presented in Theorem 24. On the one hand the theorem says that the operator S applied to an H-rich word produces a rich word. Using the examples of H-rich words mentioned earlier, we get a new class of rich words, namely the words $S(\mathbf{t}_{b,2})$ for all $b \geq 2$. On the other hand, the theorem transforms the task to discover new H-rich words to the task to discover a new class of rich words with special structures of palindromes. One such class is described in [29]. Section 5 is devoted to the notion of G-richness on a multiliteral alphabet. In particular, the operation S is defined over the alphabet $\mathbb{Z}_m$. Theorem 31 illustrates that even on a multiliteral alphabet the operation S connects G-richness and G'-richness for, in general, distinct groups G and G'. In this sense Theorem 31 is a weaker version of Theorem 24.

2 Preliminaries

The set $\mathcal{A}^*$ is the set of all finite words over the *alphabet* $\mathcal{A}$ which is a finite set of *letters*. The *length* of the word $w = w_0 w_1 \cdots w_{n-1} \in \mathcal{A}^*$ with $w_i \in \mathcal{A}$ for all i is denoted $|w|$ and equals n. The *empty word* – the unique word of length 0 – is denoted ε. The set $\mathcal{A}^*$ together with concatenation forms a free monoid with the neutral element ε. A word $v \in \mathcal{A}^*$ is a *factor* of $w \in \mathcal{A}^*$ if $w = uvz$ for some word $u, z \in \mathcal{A}^*$. If, moreover, $u = \varepsilon$, then we say that v is a *prefix* of w, if $z = \varepsilon$, the word v is a *suffix* of w. If w has the form $w = vz$, then z is denoted $z = v^{-1}w$ and the word $v^{-1}wv$ is a *conjugate* of the word w.

The *infinite word* over $\mathcal{A}$ is a sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}} = u_0u_1u_2\ldots$. The symbol $\mathcal{A}^{\mathbb{N}}$ denotes the set of all infinite words over $\mathcal{A}$. A finite word $w \in \mathcal{A}^*$ of length $n = |w|$ is a *factor* of $\mathbf{u}$ if there exists an index i such that $w = u_iu_{i+1}\cdots u_{i+n-1}$; the index i is an *occurrence* of the factor w. The symbol $\mathcal{L}_n(\mathbf{u})$ stands for the set of all factors of length n occurring in $\mathbf{u}$. The set of all factors of $\mathbf{u}$ is the *language* of $\mathbf{u}$ and is denoted by $\mathcal{L}(\mathbf{u})$.

An infinite word $\mathbf{u}$ is *recurrent* if any factor of $\mathbf{u}$ has at least two occurrences in $\mathbf{u}$. Equivalently, a word is recurrent if any factor has infinitely many occurrences. If moreover for any factor w the gaps between consecutive occurrences of w are bounded, then the word $\mathbf{u}$ is *uniformly recurrent*. Let w and vw be factors of $\mathcal{L}(\mathbf{u})$ such that vw has a prefix w and w occurs in vw exactly twice. The word v is a *return word* of w and vw is a *complete return word* of w. One can say equivalently: a recurrent word $\mathbf{u}$ is uniformly recurrent if any factor $w \in \mathcal{L}(\mathbf{u})$ has finite number of return words of w.

The *factor complexity* of $\mathbf{u}$ is the mapping $\mathcal{C}_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$, defined by $\mathcal{C}_{\mathbf{u}}(n) = \#\mathcal{L}_n(\mathbf{u})$. Given $a \in \mathcal{A}$ and $w \in \mathcal{A}^*$, a factor $wa \in \mathcal{L}(\mathbf{u})$ is a *right extension* of the factor w. Any factor of $\mathbf{u}$ has at least one right extension, the set of all right extensions of w is denoted $\mathrm{Rext}(w)$. If w has at least two right extensions we call it *right special*. Analogously one can define *left extension* and *left special* and $\mathrm{Lext}(w)$. In a recurrent word $\mathbf{u}$ any factor has at least one left extension. A factor w which is left and right special is *bispecial*. Special factors can be used to determine the factor complexity, in particular

$$\Delta\mathcal{C}_{\mathbf{u}}(n) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) = \sum_{w\in\mathcal{L}_n(\mathbf{u})} \big(\#\mathrm{Rext}(w) - 1\big).$$

If $\mathcal{A}$ is a binary alphabet, we get

$$\Delta\mathcal{C}_{\mathbf{u}}(n) = \#\{w \in \mathcal{L}_n(\mathbf{u}) \colon w \text{ is right special}\}. \tag{3}$$

A mapping $\mu : \mathcal{A}^* \to \mathcal{B}^*$ is a *morphism* if $\mu(wv) = \mu(w)\mu(v)$ for all $w, v \in \mathcal{A}^*$. It is an *antimorphism* if $\mu(wv) = \mu(v)\mu(w)$ for all $w, v \in \mathcal{A}^*$. An infinite word $\mathbf{u}$ is *closed under the mapping* μ if $\mu(w) \in \mathcal{L}(\mathbf{u})$ for any factor $w \in \mathcal{L}(\mathbf{u})$. Domain of a morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ can be naturally extended to $\mathcal{A}^{\mathbb{N}}$ by the prescription $\varphi(\mathbf{u}) = \varphi(u_0u_1u_2\ldots) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots$. An infinite word $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is called *fixed point* of a morphism φ if $\varphi(\mathbf{u}) = \mathbf{u}$.

An antimorphism Ψ is *involutory* if $\Psi^2 = \mathrm{Id}$. The most frequent involutory antimorphism is the reversal mapping R. If the word $\mathbf{u}$ is closed under an involutory antimorphism, then $\mathbf{u}$ is necessarily recurrent.

If $p = \Psi(p)$, the word p is a *Ψ-palindrome* or *pseudopalindrome*, if specification of the mapping Ψ is not needed. In the case $\Psi = R$, we say only palindrome instead of R-palindrome. The set of all Ψ-palindromes occurring as factors of a finite word w is denoted $\mathrm{Pal}^{\Psi}(w)$. The *Ψ-palindromic complexity* of an infinite word $\mathbf{u}$ is the mapping $P^{\Psi}_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$, defined by $\mathcal{P}^{\Psi}_{\mathbf{u}}(n) = \#\{p \in \mathcal{L}_n(\mathbf{u}) \colon p = \Psi(p)\}$.

A Ψ-palindrome w is *centered at* $x \in \mathcal{A} \cup \{\varepsilon\}$ if $w = vx\Psi(v)$ for some word v. If a Ψ-palindrome is centered at ε, then it is of even length.

3 G-defect and G-richness

First, we recall the definition of palindromic defect as it was introduced by Brlek, Hamel, Nivat and Reutenauer in [12]. This classical definition is based on the inequality

$$\#\mathrm{Pal}^{R}(w) \leq |w| + 1 \qquad \text{for all } w \in \mathcal{A}^*, \tag{4}$$

where $\mathrm{Pal}^{R}(w)$ is the set of all R-palindromic factors of w including the empty word.

The *R-defect* of a finite word w is

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

and R-defect of an infinite word $\mathbf{u}$ is

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

We prefer to use the name R-defect instead of the originally used "defect" because we will introduce an analogous notion for a general antimorphism Ψ as well. An infinite word $\mathbf{u}$ with $D^R(\mathbf{u}) = 0$ is called *R-full* or *R-rich*. If $D^R(\mathbf{u})$ is finite, we say that $\mathbf{u}$ is *almost R-rich*. In [13], Brlek and Reutenauer used the inequality (1) to introduce the value

$$T_{\mathbf{u}}(n) = \Delta\mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}^R_{\mathbf{u}}(n+1) - \mathcal{P}^R_{\mathbf{u}}(n) \qquad \text{for every } n \in \mathbb{N}$$

and they conjectured that if $\mathbf{u}$ is closed under reversal, then

$$2D^R(\mathbf{u}) = \sum_{n=1}^{\infty} T_{\mathbf{u}}(n)\,. \tag{5}$$

Their conjecture was proven in [8]. In particular, it means that $D^R(\mathbf{u})$ is finite if and only if there exists $N \in \mathbb{N}$ such that $T_{\mathbf{u}}(n) = 0$ for all $n \geq N$, or in other words in (1) the equality holds for all $n \geq N$.

To prove R-richness we will use the characterization of R-rich words given in [5]. It exploits the notion of the bilateral order $\mathrm{b}(w)$ of a factor w and the palindromic extension of a palindrome. The bilateral order was introduced in [16] as

$$\mathrm{b}(w) = \#\{awb \in \mathcal{L}(\mathbf{u})\colon a, b \in \mathcal{A}\} - \#\mathrm{Rext}(w) - \#\mathrm{Lext}(w) + 1. \tag{6}$$

The set of all palindromic extensions of a palindrome $w \in \mathcal{L}(\mathbf{u})$ is defined by

$$\mathrm{Pext}(w) = \{awa\colon awa \in \mathcal{L}(\mathbf{u}), a \in \mathcal{A}\}.$$

Theorem 1 ([6])**.** *Let $\mathbf{u}$ be an infinite word closed under reversal.*

1. *The word $\mathbf{u}$ is R-rich if and only if any bispecial factor w of $\mathbf{u}$ satisfies:*

$$\mathrm{b}(w) = \begin{cases} \#\mathrm{Pext}(w) - 1 & \textit{if } w \textit{ is a palindrome;} \\ 0 & \textit{otherwise.} \end{cases} \tag{7}$$

2. *If the word $\mathbf{u}$ is almost R-rich, then (7) is satisfied for all bispecial factors w up to finitely many exceptions.*

The first attempt to study the number of Ψ-palindromes for an involutory antimorphism Ψ was made in [9]. Blondin Massé, Brlek, Garon and Labbé considered the binary alphabet $\{0, 1\}$ and the antimorphism E. They showed that

$$\#\mathrm{Pal}^E(w) \leq |w| \quad \text{for all } w \in \mathcal{A}^* \setminus \{\varepsilon\}. \tag{8}$$

Starosta in [27] generalized this result for an arbitrary involutory antimorphism Ψ and arbitrary alphabet into the inequality

$$\#\mathrm{Pal}^\Psi(w) \leq |w| + 1 - \gamma_\Psi(w) \quad \text{for all } w \in \mathcal{A}^*\,, \tag{9}$$

where $\gamma_\Psi(w) = \#\big\{\{a, \Psi(a)\}\colon a \in \mathcal{A},\ a \text{ occurs in } w \text{ and } \Psi(a) \neq a\big\}$. Clearly, if $\Psi = E$ we have (8) as $\gamma_E(w) = 1$ for any $w \neq \varepsilon$, if $\Psi = R$ we have (4) as $\gamma_R(w) = 0$ for any w. Based on the inequality (9), the *Ψ-defect* of $w \in \mathcal{A}^*$ is defined by

$$D^\Psi(w) = |w| + 1 - \gamma_\Psi(w) - \#\mathrm{Pal}^\Psi(w). \tag{10}$$

The Ψ-defect of an infinite word $\mathbf{u}$ is defined analogously, i.e., $D^\Psi(\mathbf{u}) = \sup\{D^\Psi(w)\colon w \in \mathcal{L}(\mathbf{u})\}$.

Infinite words having finite Ψ-defect can be characterized by several properties, for more details about R-defect see [7] and about Ψ-defect see [27, 23]. In [23] we showed that there exists a very narrow connection between words with finite defect and words with zero defect. We proved that if $\mathbf{u}$ is closed under an involutory antimorphism Ψ and $D^\Psi(\mathbf{u})$ is finite, then $\mathbf{u}$ is a morphic image of a word $\mathbf{v}$ with

$D^\Phi(\mathbf{v}) = 0$ for some involutory antimorphism Φ. If moreover $\mathbf{u}$ is uniformly recurrent, then $\Phi = R$. In this sense, considering Ψ instead of R does not bring a broader variability into the concept of rich words.

The situation changes when we consider more antimorphisms. In [24] we defined a generalization of the notion of defect. In what follows, the symbol G stands for a finite group consisting of morphisms and antimorphisms over $\mathcal{A}^*$ and containing at least one antimorphism. The *orbit* of $w \in \mathcal{A}^*$ is the set

$$[w] = \{\mu(w) \colon \mu \in G\}. \tag{11}$$

We say that $\mathbf{u}$ is closed under G if $[w] \subset \mathcal{L}(\mathbf{u})$ for any $w \in \mathcal{L}(\mathbf{u})$. Word $p \in \mathcal{A}^*$ is a *G-palindrome* if $p = \Psi(p)$ for some antimorphism $\Psi \in G$. The generalization of the set of all palindromic factors of a word is a set consisting of palindromic orbits, namely the set

$$\mathrm{Pal}^G(w) = \left\{[p] \colon p \text{ occurs in } w \text{ and } p \text{ is a } G\text{-palindrome}\right\}.$$

Note that if $G = \{\mathrm{Id}, \Psi\}$ where Ψ is an involutory antimorphism, then $\mathrm{Pal}^\Psi(w)$ is in one-to-one correspondence with the set $\mathrm{Pal}^G(w)$ (the only difference is that the latter is a set of orbits instead of factors). Let us stress that in $\mathrm{Pal}^G(w)$ we count how many different orbits have a G-palindromic representative occurring in w.

Definition 2. Let w be a finite word. The *G-defect* of w is defined as

$$D^G(w) = |w| + 1 - \#\mathrm{Pal}^G(w) - \gamma_G(w),$$

where

$$\gamma_G(w) = \#\left\{[a] \colon a \in \mathcal{A}, a \text{ occurs in } w, \text{ and } a \neq \Psi(a) \text{ for every antimorphism } \Psi \in G\right\}.$$

A finite word is *G-rich* if its G-defect is 0. An infinite word is G-rich if all its factors are G-rich. In [24], a distinct and equivalent definition of G-richness is used: it is based on a specific structure of graphs representing the factors of same length of the word.

Example 3. We illustrate the previous notions on the Thue–Morse word $\mathbf{t}$, the fixed point of the morphism $0 \mapsto 01$ and $1 \mapsto 10$ starting with 0, i.e., $\mathbf{t} = 011010011001011010\cdots$. The word $\mathbf{t}$ is closed under R and E. Let $H = \{\mathrm{Id}, R, E, ER\}$. For the group H the value $\gamma_H(w) = 0$ for any $w \in \mathcal{A}^*$. Consider $w = 011010011001$, the prefix of $\mathbf{t}$ of length 12. We have

$$\begin{aligned}
\mathrm{Pal}^R(w) &= \{\varepsilon, 0, 1, 11, 00, 101, 010, 0110, 1001, 001100, 10011001\},\\
\mathrm{Pal}^E(w) &= \{\varepsilon, 01, 10, 0011, 1100, 1010, 110100, 001100, 01101001\},\\
\mathrm{Pal}^H(w) &= \{[\varepsilon], [0], [00], [01], [010], [0110], [0011], [1010], [110100],\\
&\qquad [100110], [001100], [10011001], [01101001]\}.
\end{aligned}$$

The corresponding defects of w are

$$\begin{aligned}
D^R(w) &= |w| + 1 - \#\mathrm{Pal}^R(w) = 2,\\
D^E(w) &= |w| - \#\mathrm{Pal}^E(w) = 3,\\
D^H(w) &= |w| + 1 - \#\mathrm{Pal}^H(w) = 0.
\end{aligned}$$

In fact, the Thue–Morse word is H-rich, whereas its R-defect and E-defect are both infinite, see Example 8 later.

For G-richness, theorems analogous to the theorems for the classical richness can be stated, c.f. [25]. The list of known G-rich words with G having at least two antimorphisms is modest. It contains the generalized Thue–Morse words $\mathbf{t}_{b,m}$. The word $\mathbf{t}_{b,m}$ is defined on the alphabet $\{0, \ldots, m-1\}$ for all $b \geq 2$ and $m \geq 2$ as

$$\mathbf{t}_{b,m} = (s_b(n) \mod m)_{n=0}^{+\infty},$$

where $s_b(n)$ denotes the sum of digits in the base-b representation of the integer n. See for instance [1] where this class of words is studied. The language of $\mathbf{t}_{b,m}$ is closed under a group isomorphic to the dihedral group of order $2m$, here denoted $I_2(m)$. In Section 5, we describe the group in details. In [28], the second author proved that $\mathbf{t}_{b,m}$ is $I_2(m)$-rich for any parameters $b \geq 2$ and $m \geq 2$.

In Corollary 14 we add to the list of H-rich words also complementary-symmetric Rote words. As already mentioned in Introduction, an infinite binary word $\mathbf{u}$ is a *complementary-symmetric Rote word* if its factor complexity satisfies $\mathcal{C}_{\mathbf{u}}(n) = 2n$ for all $n \geq 1$ and its language is closed under the exchange of the two letters E.

In this article, we focus on groups G acting on $\mathcal{A}^*$ for which the implication

$$\Psi_1(a) = \Psi_2(a) \quad \Longrightarrow \quad \Psi_1 = \Psi_2$$

is true for any letter $a \in \mathcal{A}$ and any pair of antimorphisms $\Psi_1, \Psi_2 \in G$. In [24], for such a group, the number 1 is called G-distinguishing, since the image of a single letter by an antimorphism from G allows to identify the antimorphism. For example, the number 1 is H-distinguishing for the group H used in Example 3. Also for the dihedral groups $I_2(m)$ studied in Section 5, the number 1 is $I_2(m)$-distinguishing.

If an infinite word $\mathbf{u}$ is closed under a group G and 1 is G-distinguishing, then

$$\Delta\mathcal{C}_{\mathbf{u}}(n) + \#G \;\geq \sum_{\Psi \in G^{(2)}} \Big(\mathcal{P}^{\Psi}_{\mathbf{u}}(n) + \mathcal{P}^{\Psi}_{\mathbf{u}}(n+1)\Big) \qquad \text{for all} \quad n \in \mathbb{N}, n \geq 1\,, \tag{12}$$

where $G^{(2)}$ denotes the set of all involutory antimorphisms from G, see [24]. Clearly, if G is generated by one antimorphism, say Ψ, then $\#G = 2$ and $G^{(2)} = \{\Psi\}$. The inequality (1) is the special case of (12). Similarly, the inequality (2) can be obtained from (12) if we put $G = H = \{\mathrm{Id}, R, E, ER\}$. The following G-analogue of the result obtained by Bucci, De Luca, Glen and Zamboni in [15] for the classical richness is proved in [25].

Theorem 4. *Let an infinite word $\mathbf{u}$ be closed under a group G such that the number 1 is G-distinguishing. The G-defect $D^G(\mathbf{u})$ is zero if and only if in (12) the equality holds for each $n \in \mathbb{N}, n \geq 1$.*

In [24] we have introduced also the notion almost G-rich word. A word $\mathbf{u}$ closed under a group G is *almost G-rich* if there exists $N \in \mathbb{N}$ such that the equality in (12) takes place for all integers $n \geq N$. An infinite word $\mathbf{u}$ is almost G-rich if and only if its G-defect

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

is finite.

Remark 5. In fact, in [25] the last statement is shown only for uniformly recurrent words. However, one can use the same argument we applied in proof of Theorem 2 in [8] and show that $D^G(\mathbf{u})$ is finite if and only if in (12) the equality takes place from some N on.

4 Binary words invariant under two involutory antimorphisms

4.1 G-richness in binary alphabet

In this section we suppose $\mathcal{A} = \{0, 1\}$. On binary alphabet we have only two antimorphisms R and E. Therefore, only the groups

$$\{\mathrm{Id}, R\}, \quad \{\mathrm{Id}, E\}, \quad \text{and} \quad H = \{\mathrm{Id}, R, E, ER\},$$

can be considered when inspecting the defect D^G. Let us start with examples of G-rich and almost G-rich words for these three groups.

Example 6. ($G = \{\mathrm{Id}, R\}$)
The classical richness has been studied very intensively and thus there are known many examples of binary R-rich words including Sturmian words [18], Rote Words [10], the period doubling word [3], etc. Plenty examples of binary almost R-rich words can be constructed by application of special standard P-morphisms to any rich word, see [20] for the definition of standard P-morphism and a proof.

Example 7. ($G = \{\mathrm{Id}, E\}$)
It can be easily seen, or shown using the results of [9], that there exist only two E-rich infinite words, namely the periodic word $\mathbf{u} = (01)^\omega$ and its shift $(10)^\omega$. The two mentioned words are also R-rich and H-rich as the equalities hold in (1) and (2) for all $n \in \mathbb{N}, n \geq 1$.

Examples of infinite words with finite E-defect are *E-standard words with seed* (see [14] for their definition and [23] for a proof). This class also includes very simple examples of words with finite E-defect: periodic words having the form w^ω with $w = E(w)$. One can easily show that in this case $D^E(w^\omega) = D^E(w^2)$ (see Corollary 8 in [12] for R-defect, a modification for E is straightforward).

Example 8. ($G = H = \{\mathrm{Id}, R, E, ER\}$)
The only so far known examples of H-rich words are given in [28]: they are the generalized Thue–Morse words $\mathbf{t}_{b,2}$.

If b is odd, then $\mathbf{t}_{b,2} = (01)^\omega$ and hence $\mathbf{t}_{b,2}$ is also R-rich and E-rich.

If b is even, the word is aperiodic and $D^R(\mathbf{t}_{b,2}) = D^E(\mathbf{t}_{b,2}) = +\infty$. To prove it for any even b we use the fact that $\mathbf{t}_{b,2}$ is a fixed point of the morphism φ determined by

$$\varphi: \quad 0 \mapsto (01)^{\frac{b}{2}} \quad \text{and} \quad 1 \mapsto (10)^{\frac{b}{2}}.$$

It is readily seen that the factor $w = (01)^{\frac{b}{2}}$ is *strong*, i.e., its bilateral order $b(w)$ is positive, specifically $b(w) = 1$, as all four words $0w1$, $0w0$, $1w1$, and $1w0$ belong to $\mathcal{L}(\mathbf{u})$. Moreover w is an E-palindrome. The form of the morphism ensures that

- $b(\varphi(v)) = 1$ for any strong factor $v \neq \varepsilon$;
- if v is an R-palindrome, then $\varphi(v)$ is an E-palindrome,
- if v is an E-palindrome, then $\varphi(v)$ is an R-palindrome,

These properties imply that for any $k \in \mathbb{N}$, the factor $\varphi^{2k}(w)$ is an E-palindrome and hence it is not an R-palindrome. Thus there exist infinitely many non-palindromic bispecial factors with non-zero bilateral order. Using Theorem 1 one may see that $D^R(\mathbf{t}_{b,2}) = +\infty$.

To prove that $D^E(\mathbf{t}_{b,2}) = +\infty$ we may proceed analogously. The factors $\varphi^{2k+1}(w)$ are R-palindromes but they are not E-palindromes for all $k > 0$. These factors are bispecial with the same bilateral order 1. A modification of Theorem 1 for the antimorphism E (which can be found in full generality in [25], Proposition 45) gives the result.

Now we look at the question whether a word can be simultaneously (almost) G-rich for two groups on the binary alphabet. We will discuss the connection between finiteness of defects D^R, D^E and D^H. In what follows we will consider words invariant under R and E simultaneously. First we study the relationship between R- and E-palindromes.

Lemma 9. *Let $p, q \in \mathcal{A}^*$ be R-palindromes such that the word pq is an E-palindrome, i.e.,*

$$pq = E(q)E(p). \tag{13}$$

There exist $c \in \mathcal{A}^*$ and $i, j \in \mathbb{N}$ such that $p = c\,(E(c)c)^i$ and $q = (E(c)c)^j\, E(c)$.

Proof. We will induce on the difference of $|p|$ and $|q|$. First, suppose that $|p| = |q|$, then (13) implies that $q = E(p)$ and it suffices to set $c = p$ and $i = j = 0$.

Suppose now that $|p| \neq |q|$. We can suppose without loss of generality that $|p| < |q|$. Set $q = q_1q_2$ with $|p| = |q_2|$. It follows from (13) that $pq_1q_2 = E(q_2)E(q_1)E(p)$, thus $p = E(q_2)$ and $q_1 = E(q_1)$. Therefore, q_2 is a palindrome. Since q is a palindrome, we have $R(q_1q_2) = R(q_2)R(q_1) = q_1q_2 = q_2R(q_1)$. We get

$$q_1q_2 = q_2R(q_1). \tag{14}$$

This equation on words, written in general as $xz = zy$, has a well-known solution: there exist words $u, v \in \mathcal{A}^*$ and $k \in \mathbb{N}$ such that $x = uv, y = vu$ and $z = (uv)^k u$. If the word z is palindrome, then the form of z implies that u and v are palindromes as well. To use the solution of $xz = zx$ to solve (14), we set $z = q_2$, $x = q_1$ and $y = R(q_1)$ and we get the solutions $q_1 = uv = E(uv)$ and $q_2 = (uv)^k u$. Since $|q_1| = |q| - |p| = |u| + |v|$, it follows that the difference of $|u|$ and $|v|$ is less than $|q_1| = |q| - |p|$. We apply the induction hypothesis on the palindromes u and v satisfying $E(uv) = uv$ and we get that $u = d(E(d)d)^m$ and $v = (E(d)d)^n E(d)$ for some $d \in \mathcal{A}^*$. Substituting for p and q one can find that it suffices to set $c = E(d)$ and the claim is proved. ☐

Corollary 10. *If p and q are palindromes such that $pq = E(pq)$, then there exists $c \in \mathcal{A}^*$ such that $pq = (cE(c))^j$ for some $j \in \mathbb{N}$.*

Proposition 11. *If an infinite recurrent word $\mathbf{u}$ has finite R-defect and finite E-defect, then $\mathbf{u}$ is periodic with a period conjugate to $rE(r)$, where r is an R-palindrome.*

Proof. Let $\mathbf{u}$ be an infinite recurrent word with finite R- and E-defects. Using Proposition 5 in [23], it follows that $\mathbf{u}$ is closed under R and E and there exists an integer h such that

$$\begin{aligned} \Delta\mathcal{C}_{\mathbf{u}}(n) + 2 &= \mathcal{P}^R(n+1) + \mathcal{P}^R(n) \qquad \text{and} \\ \Delta\mathcal{C}_{\mathbf{u}}(n) + 2 &= \mathcal{P}^E(n+1) + \mathcal{P}^E(n) \end{aligned}$$

for all $n \geq h$. Since $\mathbf{u}$ is also closed under all elements of the group H, combining the two previous equalities with (2) we get $0 \geq \Delta\mathcal{C}_{\mathbf{u}}(n)$ for all $n \geq h$, i.e., the word $\mathbf{u}$ is eventually periodic. Since $\mathbf{u}$ is recurrent and closed under R, the word $\mathbf{u}$ is purely periodic, i.e., $\mathbf{u} = w^\omega$. As $\mathbf{u}$ is closed under E, the word $E(w)$ is a factor of ww. It implies that $w = w_1w_2$ with $E(w_1) = w_1$ and $E(w_2) = w_2$. As the length of any E-palindrome is even, the concatenation of two E-palindromes is conjugate to an E-palindrome, in other words, the word w is conjugate to an E-palindrome, say v. Thus $\mathbf{u} = w^\omega = v'v^\omega$ for some v'. As v^ω has language closed under R as well, by the same reasoning we have $v = pq$, where $R(p) = p$ and $R(q) = q$. Applying Corollary 10 we get $v = pq = (cE(c))^j$ for some $j \in \mathbb{N}$. It is enough to set $r = c$. ☐

The following proposition treats another combination of two G-defects.

Proposition 12. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ be a word having its language closed under the group H and let $\Psi = R$ or $\Psi = E$. If $D^\Psi(\mathbf{u})$ is finite (resp. zero), then $D^H(\mathbf{u})$ is finite (resp. zero) as well.*

Before giving a proof of the last proposition, we recall Proposition 4.3 of [7] which will be needed.

Proposition 13. *Let $\mathbf{u}$ be an infinite word with language closed under reversal. Suppose that there exists an integer N such that for all $n \geq N$ the equality $\mathcal{P}^R_{\mathbf{u}}(n) + \mathcal{P}^R_{\mathbf{u}}(n+1) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2$ holds. The complete return words of any palindromic factor of length $n \geq N$ are palindromes.*

Proof of Proposition 12. Let us realize that closedness of $\mathbf{u}$ under R and E ensures that the numbers $\mathcal{P}^E_{\mathbf{u}}(n)$ and $\mathcal{P}^R_{\mathbf{u}}(n)$ are even. Indeed, if $w \in \mathcal{L}(\mathbf{u})$ is an E-palindrome of length n, then $R(w)$ is an E-palindrome as well, and analogously for R-palindromes.

First we consider $\Psi = R$. Let us suppose that there exists a positive integer N such that

$$\begin{aligned} \Delta\mathcal{C}_{\mathbf{u}}(n) + 2 &= \mathcal{P}^R_{\mathbf{u}}(n) + \mathcal{P}^R_{\mathbf{u}}(n+1) \qquad \text{for all } n \geq N \text{ and} \\ \Delta\mathcal{C}_{\mathbf{u}}(N) + 4 &> \mathcal{P}^R_{\mathbf{u}}(N) + \mathcal{P}^R_{\mathbf{u}}(N+1) + \mathcal{P}^E_{\mathbf{u}}(N) + \mathcal{P}^E_{\mathbf{u}}(N+1)\,. \end{aligned}$$

We will show that this assumption leads to a contradiction.

8

In particular the assumption yields the inequality $2 > \mathcal{P}^E_{\mathbf{u}}(N) + \mathcal{P}^E_{\mathbf{u}}(N+1)$, which implies that there is no E-palindrome of length at least N. Let $w \in \mathcal{L}(\mathbf{u})$ be an R-palindrome of length at least N. We say that a factor f has Property π if it satisfies all of the following:

1) w occurs in f exactly once,

2) $E(w)$ occurs in f exactly once,

3) w is a suffix or a prefix of f,

4) $E(w)$ is a suffix or a prefix of f.

Let u be a factor with Property π. Such factor must exist as $\mathcal{L}(\mathbf{u})$ is closed under E and thus $E(w) \in \mathcal{L}(\mathbf{u})$ as well. As w is an R-palindrome and $\mathbf{u}$ is closed under reversal, the factor $R(u)$ has Property π as well. Since $ER(w) = E(w)$, we can assume without loss of generality that u is the factor starting in w and ending in $E(w)$. Let us look at the complete return word of w, say p, with prefix u. The fact that the equality $\Delta\mathcal{C}_{\mathbf{u}}(n) + 2 = \mathcal{P}^R_{\mathbf{u}}(n) + \mathcal{P}^R_{\mathbf{u}}(n+1)$ is valid for all $n \geq N$ implies according to Proposition 13 that the complete return word p of w is an R-palindrome. Thus the factor $R(u)$ is a suffix of p. Moreover p contains only two factors (namely u and $R(u)$) with Property π.

We have shown for every factor u' with Property π that its closest right neighbor in $\mathbf{u}$ with Property π is its mirror image $R(u')$. Therefore, there exist only two factors with Property π, namely u and $R(u)$.

On the other hand, if u has Property π, then $E(u)$ has Property π as well and thus $E(u) \in \{u, R(u)\}$. As $E(w)$ is a suffix of u, the factor $E(u)$ has a prefix w. It implies that $E(u) = u$ which contradicts the fact that there is no E-palindrome longer than $|w|$.

We have shown that

$$\Delta\mathcal{C}_{\mathbf{u}}(n) + 2 = \mathcal{P}^R_{\mathbf{u}}(n) + \mathcal{P}^R_{\mathbf{u}}(n+1) \qquad \text{for all } n \geq N$$

implies

$$\Delta\mathcal{C}_{\mathbf{u}}(n) + 4 = \mathcal{P}^R_{\mathbf{u}}(n) + \mathcal{P}^R_{\mathbf{u}}(n+1) + \mathcal{P}^E_{\mathbf{u}}(n) + \mathcal{P}^E_{\mathbf{u}}(n+1) \qquad \text{for all } n \geq N.$$

If $\mathbf{u}$ is R-rich, then $N = 1$ and thus $\mathbf{u}$ is also H-rich. If its defect $D(\mathbf{u})$ is finite but nonzero, then $N > 1$ and $\mathbf{u}$ has finite H-defect.

In the case $\Psi = E$ the proof is analogous. □

Corollary 14. *Every complementary-symmetric Rote word is H-rich.*

Proof. In [10] Brlek, et al., proved that Rote words are R-rich. Since a complementary-symmetric Rote word is closed under H, the previous theorem proves the statement. □

Remark 15. Let us stress that the reverse implication in Proposition 12 does not hold. As shown in Example 8, the Thue–Morse word has $D^H(\mathbf{t}) = 0$, whereas $D^R(\mathbf{t}) = D^E(\mathbf{t}) = \infty$.

According to Proposition 11, the finiteness of both defects $D^E(\mathbf{u})$ and $D^R(\mathbf{u})$ forces the word $\mathbf{u}$ to be periodic. The Rote words illustrate that there exist aperiodic words with finite $D^H(\mathbf{u})$ and $D^R(\mathbf{u})$.

4.2 The mapping S on binary words

In this section we introduce and study the basic properties of the mapping $S : \mathcal{A}^* \setminus \{\varepsilon\} \to \mathcal{A}^*$ that is given by

$$S(u_0 \cdots u_n) = v_1 \cdots v_n, \quad \text{where} \quad v_i = (u_{i-1} + u_i) \bmod 2 \quad \text{for} \quad i = 1, \ldots, n.$$

In particular, $S(a) = \varepsilon$ for every $a \in \mathcal{A}$. The following list contains some elementary properties of S.

I. $SR = RS$, and $SE = SR$.

II. $S(w) = S(u)$ if and only if $w = u$ or $w = ER(u)$.

III. $S(w)$ is an R-palindrome if and only w is an R-palindrome or an E-palindrome.

Proof. Points I and II give

$$S(w) = R\big(S(w)\big) \iff S(w) = S\big(R(w)\big) \iff w = R(w) \text{ or } w = ER\big(R(w)\big) = E(w)\,.$$

□

The operation S is naturally extended to $\mathcal{A}^{\mathbb{N}}$ by setting

$$S(u_0u_1u_2\ldots) = v_1v_2\ldots, \quad \text{where} \quad v_i = (u_{i-1}+u_i) \bmod 2 \quad \text{for} \quad i \geq 1.$$

To describe the factor complexity of $S(\mathbf{u})$ we study special factors in

$$\mathcal{L}\big(S(\mathbf{u})\big) = \{S(v)\colon v \in \mathcal{L}(\mathbf{u})\}\,.$$

Lemma 16. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$. A factor $S(v)$ is right special in $\mathcal{L}(S(\mathbf{u}))$ if and only if one of the following occurs:*

a) v or $ER(v)$ is right special in $\mathcal{L}(\mathbf{u})$,

b) $\{v, ER(v)\} \subset \mathcal{L}(\mathbf{u})$, and $\{va, ER(va)\} \not\subset \mathcal{L}(\mathbf{u})$ for both $a \in \{0,1\}$.

Proof. Let $S(v)$ be right special in $\mathcal{L}(S(\mathbf{u}))$. Then $S(v0)$ and $S(v1)$ belong to $\mathcal{L}(S(\mathbf{u}))$. It may happen that either both $v0$ and $v1$ belong to $\mathcal{L}(\mathbf{u})$, which means that v is right special in $\mathcal{L}(\mathbf{u})$, or both $ER(v0)$ and $ER(v1)$ belong to $\mathcal{L}(\mathbf{u})$, which means that $ER(v)$ is right special in $\mathcal{L}(\mathbf{u})$.

Otherwise v and $ER(v)$ are not right special in $\mathcal{L}(\mathbf{u})$, but necessarily both belong to $\mathcal{L}(\mathbf{u})$. Let va and $ER(v)b$ be the unique right prolongations in $\mathcal{L}(\mathbf{u})$ of v and $ER(v)$ respectively. Since $S(va)$ and $S\big(ER(v)b\big)$ must be distinct right prolongations of $S(v) = S\big(ER(v)\big)$, we have $a \neq ER(b)$, i.e., $a = b$. Since $ER(v)$ has a unique extension to the right $ER(v)a$, we get $ER(v)(1-a) = ER(va) \notin \mathcal{L}(\mathbf{u})$. □

Lemma 17. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$. The word $\mathbf{u}$ is uniformly recurrent if and only if $S(\mathbf{u})$ is uniformly recurrent.*

Proof. ($\Rightarrow$): Let w be a factor of $S(\mathbf{u})$. Then $w = S(v)$ for some $v \in \mathcal{L}(\mathbf{u})$. The gaps between the neighboring occurrences of v in $\mathbf{u}$ are bounded by some constant. The gaps between the occurrences of w in $S(\mathbf{u})$ are bounded by the same constant.

($\Leftarrow$): Let v be a factor of $\mathbf{u}$. Then $w = S(v)$ is a factor of $S(\mathbf{u})$ and the gaps between the occurrences of w are bounded, say by K. If v is the only factor of $\mathbf{u}$ such that $w = S(v)$, i.e., v is the only preimage of w by S in $\mathbf{u}$, then the occurrences of v in $\mathbf{u}$ are bounded by K as well. Let us suppose that w has more preimages in $\mathbf{u}$. According to Property II, there are only two preimages of w, namely v and $ER(v)$. Let f be a factor of $\mathbf{u}$ such that v is a prefix of f and $ER(v)$ is a suffix of f and v and $ER(v)$ occur in f only once. Then $S(f)$ is a complete return word of $w = S(v) = S(ER(v))$. As $S(\mathbf{u})$ is uniformly recurrent, the gaps between the occurrences of the factor $S(f)$ are bounded, say by C. Both possible preimages of $S(f)$ in $\mathbf{u}$, namely f and $ER(f)$, contain v either as its prefix or its suffix. Thus the gaps between the occurrences of v in $\mathbf{u}$ are bounded by C as well. □

Lemma 18. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$. If $S(\mathbf{u})$ is closed under R, then $\mathbf{u}$ is closed under R or under E.*

Proof. Let v be a prefix of $\mathbf{u}$. The word $S(v)$ is a factor of $S(\mathbf{u})$. According to the assumption, $RS(v) = SR(v)$ belongs to $\mathcal{L}\big(S(\mathbf{u})\big)$ as well. Due to Property II, either $R(v)$ or $E\big(RR(v)\big) = E(v)$ belong to $\mathcal{L}(\mathbf{u})$. Thus

a) either there exist infinitely many prefixes $v \in \mathcal{L}(\mathbf{u})$ such that $R(v) \in \mathcal{L}(\mathbf{u})$;

b) or there exist infinitely many prefixes $v \in \mathcal{L}(\mathbf{u})$ such that $E(v) \in \mathcal{L}(\mathbf{u})$.

Let us suppose that a) happens. For any $w \in \mathcal{L}(\mathbf{u})$ we may find a prefix v such that $R(v) \in \mathcal{L}(\mathbf{u})$ and w is a factor of v. Thus, $R(w) \in \mathcal{L}(\mathbf{u})$ and we can conclude that $\mathbf{u}$ is closed under R.

The case b) is analogous. □

Example 19. The period doubling word is the fixed point of the primitive morphism

$$\varphi_{PD}: 0 \mapsto 11 \quad \text{and} \quad 1 \mapsto 10$$

Thus

$$\mathbf{u}_{PD} = 1011101010111011101110101 0\ldots$$

It is well-known that the period doubling word is the image of the Thue–Morse word $\mathbf{t}$ by S.

The word $\mathbf{u}_{PD} = S(\mathbf{t})$ is closed under R, the word $\mathbf{t}$ is closed under R and E. It illustrates that in the previous lemma the simultaneous closedness under R and E is not excluded.

The previous lemma guarantees that $\mathbf{u}$ is closed at least under one of the antimorphisms E and R. We now focus on a property of $S(\mathbf{u})$ that ensures that $\mathbf{u}$ is closed under both of them.

Lemma 20. *Let $\mathbf{v} = S(\mathbf{u}) \in \{0,1\}^{\mathbb{N}}$. The language $\mathcal{L}(\mathbf{u})$ contains infinitely many E-palindromes and R-palindromes if and only if $\mathcal{L}(\mathbf{v})$ contains infinitely many R-palindromes centered at the letter 1 and infinitely many R-palindromes not centered at the letter 1.*

Proof. Let u be a finite non-empty word and let $v = \mathcal{S}(u)$. It suffices to realize the following:

1. u is an E-palindrome if and only if v is an R-palindrome centered at the letter 1;
2. u is an R-palindrome of even length if and only if v is an R-palindrome centered at the letter 0;
3. u is an R-palindrome of odd length if and only if v is an R-palindrome of even length, i.e., centered at ε.

□

Corollary 21. *Let $\mathbf{v} = S(\mathbf{u}) \in \{0,1\}^{\mathbb{N}}$ be uniformly recurrent. If $\mathcal{L}(\mathbf{v})$ contains infinitely many R-palindromes centered at the letter 1 and infinitely many R-palindromes not centered at the letter 1, then $\mathbf{u}$ is closed under all elements of H.*

Proof. The previous lemma implies that $\mathcal{L}(\mathbf{u})$ contains infinitely many E-palindromes and R-palindromes. Let $w \in \mathcal{L}(\mathbf{u})$. Since $\mathcal{L}(\mathbf{u})$ contains R-palindromes of arbitrary length and $\mathbf{u}$ is uniformly recurrent by Lemma 17, the factor w is a factor of an R-palindromic factor of $\mathbf{u}$, thus $R(w)$ also occurs in $\mathbf{u}$. Analogously, $E(w)$ is factor of an E-palindromic factor and thus $E(w) \in \mathcal{L}(\mathbf{u})$. □

An example of application of the last corollary are Sturmian words. It is known that they contain infinitely many R-palindromes centered at 1 and 0, which implies that their preimages by S, namely the complementary-symmetric Rote words, have their language closed under H.

Another example is the period doubling word defined in Example 19. One can easily see can that given an R-palindrome w centered at $x \in \{0,1\}$, the word $\varphi_{PD}(w)1$ is also an R-palindrome centered at $1-x$. Therefore, the period doubling word satisfies the assumptions of the corollary and it follows that the language of one of its preimage by S, namely the Thue–Morse word, is closed under H.

The following definition is inspired by [4]. Given a finite word v and a letter a, the notation $|w|_a$ stands for the number of occurrences of the letter a in v.

Definition 22. We say that an infinite word $\mathbf{v} \in \{0,1\}^{\mathbb{N}}$ has *well distributed occurrences modulo* 2 (denoted WELLDOC(2)) if for every factor $w \in \mathcal{L}(\mathbf{v})$ we have

$$\Big\{\big(|v|_0, |v|_1\big) \bmod 2 \colon vw \text{ is a prefix of } \mathbf{v}\Big\} = \mathbb{Z}_2^2.$$

Proposition 23. *Let $\mathbf{v} \in \{0,1\}^{\mathbb{N}}$ have WELLDOC(2) and be closed under reversal. If $\mathbf{u}$ is a word such that $\mathbf{v} = S(\mathbf{u})$, then $\mathbf{u}$ is closed under all elements of H.*

Proof. Denote $\mathbf{v} = v_1v_2\dots$ and $\mathbf{u} = u_0u_1\dots$. Since $S(\mathbf{u}) = \mathbf{v}$, it follows that $u_0 + u_1 = v_1 \mod 2$, $u_1 + u_2 = v_2 \mod 2, \dots$ Summing first k equations we get

$$u_k = u_0 + \sum_{i=1}^{k} v_i.$$

It follows that

$$u_{k+j} = u_{k-1} + \sum_{i=k}^{k+j} v_i$$

for all $k > 0$ and $j \in \mathbb{N}$. Suppose that s and ℓ are two distinct occurrences of a factor $f \in \mathcal{L}(\mathbf{v})$ of length n. We have $\sum_{i=s}^{s+j} v_i = \sum_{i=\ell}^{\ell+j} v_i$ for all $j \in \{0,\dots,n-1\}$. If $u_{s-1} = u_{\ell-1}$, then we have $u_{s-1}\cdots u_{s+n} = u_{\ell-1}\cdots u_{\ell+n}$. On the other hand if $u_{s-1} \neq u_{\ell-1}$, then $u_{s-1}\cdots u_{s+n} = ER(u_{\ell-1}\cdots u_{\ell+n})$. Note that $u_{s-1} = u_0 + \sum_{i=1}^{s-1} v_i$ and $v_1\cdots v_{s-1}f$ is a prefix of $\mathbf{v}$, and analogously for the index ℓ. Since $\mathbf{v}$ has WELLDOC(2), we may choose the indices s and ℓ such that

$$\sum_{i=1}^{s-1} v_i = 0 \quad \text{and} \quad \sum_{i=1}^{\ell-1} v_i = 1.$$

It implies that with every factor $w \in \mathcal{L}(\mathbf{u})$, the factor $ER(w)$ also occurs in $\mathbf{u}$. As $\mathbf{v} = S(\mathbf{u})$ is closed under reversal, Lemma 18 implies that $\mathbf{u}$ is closed under R or E. This together with the closedness under ER already implies that $\mathbf{u}$ is closed under all elements of H. □

4.3 Richness of u versus richness of $S(\mathbf{u})$

This section is devoted to the study of images and preimages of almost rich words by the mapping S.

Theorem 24. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ be closed under all elements of $H = \{\mathrm{Id}, E, R, ER\}$. The word $\mathbf{u}$ is H-rich (resp. almost H-rich) if and only if $S(\mathbf{u})$ is R-rich (resp. almost R-rich).*

Proof. Let $\mathbf{v} = S(\mathbf{u})$. Since $\mathbf{u}$ is closed under ER, any factor of $\mathcal{L}(\mathbf{v})$ has two preimages in $\mathcal{L}(\mathbf{u})$. Moreover, v is right special in $\mathcal{L}(\mathbf{u})$ if and only if $ER(v)$ is right special in $\mathcal{L}(\mathbf{u})$ as well. Thus by Lemma 16, any right special factor in $\mathcal{L}(\mathbf{v})$ of length n is image of two right special factors of length $n+1$. According to (3) we get

$$2\Delta\mathcal{C}_{\mathbf{v}}(n) = \Delta\mathcal{C}_{\mathbf{u}}(n+1). \tag{15}$$

Analogously, v is an R- or E-palindrome in $\mathcal{L}(\mathbf{u})$ if and only if $ER(v)$ is an R- or E-palindrome in $\mathcal{L}(\mathbf{u})$. According to Property III we have

$$2\mathcal{P}^R_{\mathbf{v}}(n) = \mathcal{P}^R_{\mathbf{u}}(n+1) + \mathcal{P}^E_{\mathbf{u}}(n+1)\,.$$

Thus the equality

$$\Delta\mathcal{C}_{\mathbf{v}}(n) + 2 = \mathcal{P}^R_{\mathbf{v}}(n+1) + \mathcal{P}^R_{\mathbf{v}}(n),$$

testifying that $\mathbf{v}$ is R-rich, holds if and only if the equality

$$\Delta\mathcal{C}_{\mathbf{u}}(n+1) + 4 = \mathcal{P}^R_{\mathbf{u}}(n+1) + \mathcal{P}^E_{\mathbf{u}}(n+1) + \mathcal{P}^R_{\mathbf{u}}(n+2) + \mathcal{P}^E_{\mathbf{u}}(n+2),$$

testifying that $\mathbf{u}$ is H-rich, is satisfied. □

As already noted above, the word $\mathbf{t}_{b,2}$ is H-rich. Thus, using the last theorem with $\mathbf{u} = \mathbf{t}_{b,2}$ and (15) together with the equality

$$\mathcal{C}_{\mathbf{w}}(n) = 1 + \sum_{i=0}^{n-1} \Delta\mathcal{C}_{\mathbf{w}}(i)$$

valid for any infinite word $\mathbf{w}$, we obtain the following corollary:

Corollary 25. *For every integer b greater than 1 the word $\mathbf{v} = S(\mathbf{t}_{b,2})$ is R-rich. Its factor complexity satisfies*

$$\mathcal{C}_{\mathbf{v}}(n) = \frac{1}{2}(\mathcal{C}_{\mathbf{t}_{b,2}}(n) - 1).$$

Using the factor complexity of the word $\mathbf{t}_{b,2}$ described in [28], one can see that the binary R-rich word $S(\mathbf{t}_{b,2})$ is not Sturmian.

Remark 26. Since a complementary-symmetric Rote word $\mathbf{u}$ is closed under all elements of H and $S(\mathbf{u})$ is Sturmian, which is R-rich, Theorem 24 provides an alternative proof of Corollary 14 without exploiting the result that every Rote word $\mathbf{u}$ is R-rich.

Theorem 27. *Let $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ be a uniformly recurrent word. If $\mathbf{u}$ is almost R-rich, then the word $S(\mathbf{u})$ is almost R-rich.*

Proof. Since $\mathbf{u}$ is almost rich, its language contains infinitely many palindromes. This fact for uniformly recurrent words implies that $\mathbf{u}$ is closed under reversal. If $\mathbf{u}$ is closed under E as well, then according to Proposition 12 the word $\mathbf{u}$ is almost H-rich, and the claim follows from Theorem 24.

It is enough to consider $\mathbf{u}$ that is not closed under E. We will show that the set $\{w \in \mathcal{L}(\mathbf{u})\colon ER(w) \in \mathcal{L}(\mathbf{u})\}$ is finite. Assume the opposite. Let v be a factor of length n. As $\mathbf{u}$ is uniformly recurrent there exists a number $r(n)$ such that any factor of $\mathbf{u}$ longer than $r(n)$ contains all factors of length n. Since $\{w \in \mathcal{L}(\mathbf{u})\colon ER(w) \in \mathcal{L}(\mathbf{u})\}$ is not finite, there exists w belonging to this set and being longer than $r(n)$. And thus the factor v of length n occurs in w and $ER(v)$ occurs in $ER(w)$. Since both w and $ER(w)$ belong to $\mathcal{L}(\mathbf{u})$, the factor v and $ER(v)$ belongs to $\mathcal{L}(\mathbf{u})$ as well — a contradiction with assumption that $\mathbf{u}$ is not closed under E.

Let N be the maximal length of an element of the finite set $\{w \in \mathcal{L}(\mathbf{u})\colon ER(w) \in \mathcal{L}(\mathbf{u})\}$. Any factor of $S(\mathbf{u})$ longer than N has unique preimage in $\mathcal{L}(\mathbf{u})$. According to Lemma 16 for any $n > N$ there is one-to-one correspondence between $\mathcal{L}_n(S(\mathbf{u}))$ and $\mathcal{L}_{n+1}(\mathbf{u})$. Thus,

$$\Delta\mathcal{C}_{\mathbf{u}}(n+1) = \Delta\mathcal{C}_{S(\mathbf{u})}(n) \qquad \text{for all } n > N. \tag{16}$$

Moreover, there exists no E-palindrome of length $n > N$ and thus we have one-to-one correspondence between the set of all R-palindromes in $\mathcal{L}(S(\mathbf{u}))$ of length n and the set of all R-palindromes in $\mathcal{L}(\mathbf{u})$ of length $n+1$. It gives

$$\mathcal{P}^R_{\mathbf{u}}(n+1) = \mathcal{P}^R_{S(\mathbf{u})}(n) \qquad \text{for all } n > N. \tag{17}$$

Since $\mathbf{u}$ has language closed under reversal and is almost rich using (5) there exists a constant M such that

$$\Delta\mathcal{C}_{\mathbf{u}}(n) + 2 = \mathcal{P}_{\mathbf{u}}(n) + \mathcal{P}_{\mathbf{u}}(n+1) \quad \text{for all } n \geq M.$$

This equality and equalities (16) and (17) imply

$$\Delta\mathcal{C}_{S(\mathbf{u})}(n) + 2 = \mathcal{P}_{S(\mathbf{u})}(n) + \mathcal{P}_{S(\mathbf{u})}(n+1) \quad \text{for all } n > \max\{N, M\}.$$

It follows that the word $S(\mathbf{u})$ is almost R-rich. □

Corollary 28. *If $\mathbf{u}$ is a Sturmian word, then $S^k(\mathbf{u})$ is almost R-rich for all $k > 0$.*

We add two more examples related to images (Example 29) and preimages (Example 30) of words constructed by iterated operation S. However, we do not give any proofs of their properties and we just state them as hypotheses given by computer evidence.

Example 29. As stated in Corollary 25, the word $S(\mathbf{t}_{b,2})$ is R-rich. Theorem 27 then implies that $S^k(\mathbf{t}_{b,2})$ is almost R-rich for all $k > 0$. Our computer experiments suggest that in this case the word $S^k(\mathbf{t}_{b,2})$ is in fact R-rich.

As we have already mentioned, the list of known H-rich words is very modest: complementary-symmetric Rote words and generalized binary Thue–Morse words $\mathbf{t}_{b,2}$. Theorem 24, Proposition 23 and Corollary 21 give us a recipe for construction of an (almost) H-rich word: take a binary (almost) R-rich word with property WELLDOC(2) or a binary (almost) R-rich word with suitable structure of palindromes and find its preimage by the operation S. The complementary-symmetric Rote words were obtained by this procedure applied to the Sturmian words. The Thue–Morse word $\mathbf{t} = \mathbf{t}_{2,2}$ can be obtained by this procedure applied to the period doubling word.

Example 30. Let $\mathbf{u}$ be a Sturmian word. Let $\mathbf{u}^{(k)}$ be an infinite word such that $S^k(\mathbf{u}^{(k)}) = \mathbf{u}$ for all $k \in \mathbb{N}$. The word $\mathbf{u}^{(1)}$ is a complementary-symmetric Rote word which is, as already mentioned, H-rich and R-rich. According to our computer experiments, so is the word $\mathbf{u}^{(2)}$. The word $\mathbf{u}^{(3)}$ is not R-rich, but it is still H-rich. The word $\mathbf{u}^{(k)}$ for $k > 3$ is not H-rich nor R-rich. However, the symmetries of $\mathbf{u}^{(k)}$ are preserved: $\mathbf{u}^{(k)}$ is closed under all elements of H. This is witnessed by the following difference of its factor complexity

$$\Delta\mathcal{C}_{\mathbf{u}^{(k)}}(n) = 2^{n-1} \quad \text{for } 0 < n \le k \qquad \text{and} \qquad \Delta\mathcal{C}_{\mathbf{u}^{(k)}}(n) = 2^{k} \quad \text{for } n > k$$

which is suggested by our experiments.

5 The mapping S on multiliteral alphabets

In this section we study the mapping S acting on a larger alphabet $\mathbb{Z}_m = \{0, \ldots, m-1\}$. The mapping S is defined for every word $w = w_0 \cdots w_n$ with $w_i \in \mathbb{Z}_m$ by

$$S(w_0 w_1 \cdots w_n) = v_1 \cdots v_n, \tag{18}$$

where $v_i = (w_{i-1} + w_i) \bmod m$ for every $i \in \{1, \ldots, n\}$.

The alphabet $\mathbb{Z}_m$ allows many finite groups generated by involutory antimorphisms. We restrict our attention to groups isomorphic to groups of symmetries of a regular polyhedron. The reason is simple: we have examples of G-rich words only for such groups, namely the generalized Thue–Morse words. We demonstrate that at least for these words the mapping S transforms a G-rich word to an almost G'-rich word (cf. Theorems 24 and 27 for an analogue on the binary alphabet).

Let us describe the elements of the mentioned group explicitly. For all $x \in \mathbb{Z}_m$ denote by Ψ_x the antimorphism given by

$$\Psi_x(k) = x - k \quad \text{for all } k \in \mathbb{Z}_m$$

and by Π_x the morphism given by

$$\Pi_x(k) = x + k \quad \text{for all } k \in \mathbb{Z}_m.$$

The group $I_2(m)$ is the union of these antimorphisms and morphisms:

$$I_2(m) = \{\Psi_x : x \in \mathbb{Z}_m\} \cup \{\Pi_x : x \in \mathbb{Z}_m\}.$$

The definition of the generalized Thue–Morse words is recalled in Preliminaries. It is known that the word $\mathbf{t}_{b,m}$ is a fixed point of the morphism $\varphi_{b,m} : \mathbb{Z}_m^* \to \mathbb{Z}_m^*$ defined by

$$\varphi_{b,m} : \quad a \mapsto a(a+1)(a+2)\cdots(a+b-1) \quad \text{for all } a \in \mathbb{Z}_m\,.$$

Let us stress that all operations on letters in this section are taken modulo m. As shown in [28], the word $\mathbf{t}_{b,m}$ is closed under all elements of $I_2(m)$ and moreover $\mathbf{t}_{b,m}$ is $I_2(m)$-rich. We will focus on images of $\mathbf{t}_{b,m}$ by S with parameters $b \ge 3$ and $m \ge 3$. Let $I_2'(m)$ denote the group generated by antimorphisms $\{\Psi_{2y} : y \in \mathbb{Z}_m\}$; it can be easily seen that

$$I_2'(m) = \{\Psi_{2x} : x \in \mathbb{Z}_m\} \cup \{\Pi_{2x} : x \in \mathbb{Z}_m\}. \tag{19}$$

If m is odd, then $I_2'(m) = I_2(m)$, if m is even, then $I_2'(m)$ is isomorphic to $I_2(\frac{m}{2})$.

The aim of this section is to prove the following theorem.

Theorem 31. *Let $m, b \in \mathbb{Z}$ such that $m \geq 3$ and $b \geq 3$.*

1. *The word $S(\mathbf{t}_{b,m})$ is almost $I_2'(m)$-rich.*
2. *If m or b is odd, the word $S(\mathbf{t}_{b,m})$ is $I_2'(m)$-rich.*

The first part of Theorem 31 is a direct consequence of Proposition 34, the second part follows from Lemma 35 and the description of factors of $S(\mathbf{t}_{b,m})$ up to the length 3 presented at the end of this section.

Example 32. Let us consider the word $\mathbf{t}_{4,4}$. It starts with 0 and it is a fixed point of the morphism

$$\varphi_{4,4}: \quad 0 \mapsto 0123, \quad 1 \mapsto 1230, \quad 2 \mapsto 2301 \quad \text{and} \quad 3 \mapsto 3012\,.$$

$$\begin{aligned} \text{Thus} \quad \mathbf{t}_{4,4} &= 0123123023013012123023013012012323013012 0123\ldots \\ \text{and} \quad S(\mathbf{t}_{4,4}) &= 131031321310313331321310313213111310313213 1\ldots \end{aligned}$$

Now we consider the word $\mathbf{t}_{3,4}$. Its fixing morphism is

$$\varphi_{3,4}: \quad 0 \mapsto 012, \quad 1 \mapsto 123, \quad 2 \mapsto 230 \quad \text{and} \quad 3 \mapsto 301\,.$$

$$\begin{aligned} \text{Thus} \quad \mathbf{t}_{3,4} &= 012123230123230301230301012123230301230301 01\ldots \\ \text{and} \quad S(\mathbf{t}_{3,4}) &= 133311131311133313133311133311133313133311 13\ldots \end{aligned}$$

We start with a list of observations concerning properties of the mapping S. To deduce some of the observations we exploit a peculiar property of generalized Thue–Morse words. The form of morphism $\varphi_{b,m}$ forces the language of $\mathbf{u} = \mathbf{t}_{b,m}$ to have the following property

$$u_0u_1u_2u_3 \in \mathcal{L}(\mathbf{u}) \quad \Rightarrow \quad u_i - u_{i-1} = 1 \quad \text{for at least two indices } i \in \{1, 2, 3\}. \tag{20}$$

(A) $S\Psi_y = \Psi_{2y}S$ for any $y \in \mathbb{Z}_m$. If m is even, then $S\Psi_y = S\Psi_{y+\frac{m}{2}}$ for any $y \in \mathbb{Z}_m$.

(B) If $S(u_0 \cdots u_n) = S(v_0 \cdots v_n)$ with $u_i, v_j \in \mathbb{Z}_m$, then there exists $x \in \mathbb{Z}_m$ such that

$$v_0 \cdots v_n = (u_0 + x)(u_1 - x) \cdots (u_n + (-1)^n x).$$

(C) Let $\mathbf{u}$ be closed under all elements of $I_2(m)$ and satisfy (20). Consider $w = S(v)$ for some $v = v_0v_1 \cdots v_n \in \mathcal{L}(\mathbf{u})$ with $n \geq 3$.

If m is odd, then v is the only preimage of w by S in $\mathbf{u}$.

If m is even, then w has exactly two preimages by S in $\mathbf{u}$, namely $v_0v_1 \cdots v_n$ and $(v_0 + \frac{m}{2})(v_1 + \frac{m}{2}) \cdots (v_n + \frac{m}{2})$.

Proof. Let $S(u) = S(v)$ for a factor $u = u_0u_1 \cdots u_n \in \mathcal{L}(\mathbf{u})$. As $n \geq 3$, property (20) implies that there exists $j \in \{1, 2, 3\}$ such that $u_j - u_{j-1} = v_j - v_{j-1} = 1$. From Property (B), we obtain $v_j - v_{j-1} = u_j - u_{j-1} + (-1)^j 2x$, thus $2x = 0$. If m is odd, then necessarily $x = 0$. If m is even, then also $x = \frac{m}{2}$ satisfies $2x = 0$. As the morphism $\Psi_0\Psi_{\frac{m}{2}}$ maps a to $a + \frac{m}{2}$, the language of $\mathbf{u}$ is closed under addition of $\frac{m}{2}$ to all letters of any factor of $\mathbf{u}$, i.e., $(v_0 + \frac{m}{2})(v_1 + \frac{m}{2}) \cdots (v_n + \frac{m}{2}) \in \mathcal{L}(\mathbf{u})$. $\square$

(D) Let $\mathbf{u}$ be closed under all elements of $I_2(m)$ and satisfy (20). Consider $u \in \mathcal{L}(\mathbf{u})$.

(a) If u is a Ψ_y-palindrome, then $S(u)$ is a Ψ_{2y}-palindrome.

(b) If $S(u)$ is a Ψ_{2y}-palindrome with $|S(u)| \geq 3$ and m is odd, then u is a Ψ_y-palindrome.

(c) If $S(u)$ is a Ψ_{2y}-palindrome with $|S(u)| \geq 3$ and m is even, then u is a Ψ-palindrome for $\Psi = \Psi_y$ and $\Psi = \Psi_{y+\frac{m}{2}}$.

Proof. (Da): Applying S to $u = \Psi_y(u)$ and using (A), one has $S(u) = S(\Psi_y(u)) = \Psi_{2y}(S(u))$, i.e., $S(u)$ is a Ψ_{2y}-palindrome.

(Db) and (Dc): Using Property (A), we obtain $S(u) = \Psi_{2y}(S(u)) = S(\Psi_y(u))$. As $|S(u)| \geq 3$ implies $|u| \geq 4$ and (20) is satisfied, we may apply Property (C). For odd m, it implies $\Psi_y(u) = u$ as we want to show. For even m, we have also the second possibility $u = \Psi_y(\Psi_0\Psi_{\frac{m}{2}}(u))$. It is easy to check that $\Psi_y\Psi_0\Psi_{\frac{m}{2}} = \Psi_{y+\frac{m}{2}}$. □

To prove Theorem 31 we use the notion of complete G-return word of an orbit, as introduced in [25]. Let us recall that the orbit $[w]$ of a factor $w \in \mathcal{L}(\mathbf{u})$ is defined by (11).

A factor $v \in \mathcal{L}(\mathbf{u})$ is a *complete G-return word of* $[w]$ *in* $\mathbf{u}$ if

- $|v| > |w|$,
- a prefix and a suffix of v belong to $[w]$, and
- v contains no other elements of $[w]$.

Theorem 33 ([25])**.** *If* $\mathbf{u}$ *is an infinite word closed under all elements of* G*, then*

1. $\mathbf{u}$ *is G-rich if and only if for all* $w \in \mathcal{L}(\mathbf{u})$ *every complete G-return word of* $[w]$ *is a G-palindrome.*
2. $\mathbf{u}$ *is almost G-rich if and only if there exists and integer* N *such that for all* $w \in \mathcal{L}(\mathbf{u})$ *longer than* N *every complete G-return word of* $[w]$ *is a G-palindrome.*

Using the previous theorem, we can easily prove the proposition which directly implies the validity of the first part of Theorem 31 because the generalized Thue–Morse words satisfy its assumption.

Proposition 34. *Let* $\mathcal{A} = \mathbb{Z}_m$ *and let* $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ *be closed under all elements of* $I_2(m)$*. If* $\mathcal{L}(\mathbf{u})$ *satisfies* (20) *and* $\mathbf{u}$ *is* $I_2(m)$*-rich, then the word* $S(\mathbf{u})$ *is almost* $I_2'(m)$*-rich.*

Proof. To ease the notation put $G = I_2(m)$ and $G' = I_2'(m)$. If m is odd then $G = G'$. Otherwise, $\#G = 2\#G'$.

As $\mathbf{u}$ satisfies (20), we can apply Property (D) to each palindrome $S(u)$ in $S(\mathbf{u})$ with length $|S(u)| \geq 3$. Property (D) implies that $u \in \mathcal{L}(\mathbf{u})$ is a G-palindrome if and only if $S(u)$ is a G'-palindrome in $\mathcal{L}(S(\mathbf{u}))$. Let $S(u)$ be a G'-palindrome in $S(\mathbf{u})$ and $S(v)$ be a complete G'-return word of $[S(u)]$ in $S(\mathbf{u})$. Then v is a complete G-return word in $\mathbf{u}$ of $[u]$. The word $\mathbf{u}$ is G-rich and due to Theorem 33, the factor v is a G-palindrome. According to Property (D), the complete G'-return word $S(v)$ is a G'-palindrome. Thus $S(\mathbf{u})$ is almost G'-rich. □

In the remaining part of this section we focus on G'-richness of $S(\mathbf{t}_{b,m})$ in the case when m or b is odd.

Lemma 35. *Let* $G' = I_2'(m)$ *and* $\mathbf{v} = S(\mathbf{t}_{b,m})$*. If for* $n = 1$ *and* $n = 2$ *the equality*

$$\Delta\mathcal{C}_{\mathbf{v}}(n) + \#G' = \sum_{\substack{\Psi \in G' \\ \Psi \textit{ is an antimorphism}}} \left(\mathcal{P}_{\mathbf{v}}^{\Psi}(n) + \mathcal{P}_{\mathbf{v}}^{\Psi}(n+1)\right) \tag{21}$$

holds, then $\mathbf{v} = S(\mathbf{t}_{b,m})$ *is* G'*-rich.*

Proof. We combine the results of [25].

It is easy to see that $\Psi_{2x}(a) \neq \Psi_{2y}(a)$ for any $a \in \mathbb{Z}_m$ and any pair of distinct antimorphisms Ψ_{2x} and Ψ_{2y} from G'. This property guarantees that the number 1 is G'-distinguishing in the sense of Definition 7 of [25]. In the proof of Proposition 34 we verify that any complete G'-return word of $[w]$ in $\mathbf{v}$ is a G'-palindrome for each G'-palindrome $w \in \mathcal{L}(\mathbf{v})$ of length at least 3. As 1 is G'-distinguishing, Lemma 28 of [25] says that this fact implies equality (21) for all $n \geq 3$. According to Proposition 42 of [25] the word $\mathbf{v}$ is G'-rich if the equality in (21) holds for each $n \in \mathbb{N}, n \geq 1$. □

In the case m or b odd and $n = 1$ and $n = 2$, we will confirm the equality (21) in the next lemma. To ease the notation we put

$$F(n) = \sum_{\substack{\Psi \in G' \\ \Psi \text{ is an antimorphism}}} \mathcal{P}^{\Psi}_{\mathbf{v}}(n)$$

and $\rho : \mathbb{Z}_m \to \mathbb{Z}_m$ denotes the permutation

$$\rho(k) = k + b - 1 = \Pi_{b-1}(k).$$

We will use the following statements from [28]. Let q denote the order of ρ, i. e., the least positive integer q such that $q(b-1) \equiv 0 \pmod{m}$. The factors of length 2 of $\mathbf{t}_{b,m}$ are

$$\mathcal{L}_2(\mathbf{t}_{b,m}) = \{\rho^k(r-1)r \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\}$$

and of length 3

$$\mathcal{L}_3(\mathbf{t}_{b,m}) = \{\rho^k(r-1)r(r+1) \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\} \cup \{(r-1)r\rho^{-k}(r+1) \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\}.$$

It is easy to deduce the set of all factors of length 4:

$$\begin{aligned}\mathcal{L}_4(\mathbf{t}_{b,m}) = &\{\rho^k(r-1)r(r+1)(r+2) \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\} \\ &\cup \{(r-2)(r-1)r\rho^{-k}(r+1) \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\} \\ &\cup \{(r-1)r\rho^{-k}(r+1)(\rho^{-k}(r+1)+1) \colon r \in \mathbb{Z}_m, 0 \le k \le q-1\}.\end{aligned}$$

Lemma 36. *Let $\mathbf{v} = S(\mathbf{t}_{b,m})$. The numbers of factors of $\mathbf{v}$ of length n satisfy the following:*

n	m *odd*	m *even,* b *odd*	m *even,* b *even*
1	m	$\frac{m}{2}$	m
2	qm	$\frac{qm}{2}$	$\frac{3qm}{4}$
3	$3qm - 2m$	$\frac{3qm}{2} - m$	$\frac{3qm}{2} - m$

Proof. $n = 1$
It follows from the form of $\mathcal{L}_2(\mathbf{t}_{b,m})$ that $\mathcal{L}_1(\mathbf{v}) = \{\rho^k(r-1) + r \;:\; r \in \mathbb{Z}_m, 0 \le k \le q-1\}$. We have $\rho^k(r-1) + r = r - 1 + k(b-1) + r = 2r - 1 + k(b-1)$

If m is odd, then we have directly that $\mathcal{L}_1(\mathbf{v}) = \mathbb{Z}_m$ (for $k = 0$).

If m is even and b is odd, then $2r - 1 + k(b-1)$ is odd for every r and k, and thus $\mathcal{L}_1(\mathbf{v}) = \{2i + 1 \;:\; 0 \le i < \frac{m}{2}\}$.

If m is even and b is even, then for $k = 0$ the number $2r - 1$ is odd, and for $k = 1$ the number $2r - 1 + (b-1)$ is even, and we have $\mathcal{L}_1(\mathbf{v}) = \mathbb{Z}_m$.

$n = 2$
The structure of $\mathcal{L}_3(\mathbf{t}_{b,m})$ implies that the factors of $\mathbf{v}$ of length 2 are of the following forms:

1. $(\rho^k(r-1) + r)(2r+1) = (2r - 1 + k(b-1))(2r+1)$ for $r \in \mathbb{Z}_m$ and $0 \le k < q$, and
2. $(2r' - 1)(\rho^{k'}(r'+1) + r') = (2r' - 1)(2r' + 1 + k'(b-1))$ for $r' \in \mathbb{Z}_m$ and $0 \le k' < q$.

First, let us see the number of factors of type 1. Fix $r \in \mathbb{Z}_m$. Suppose $2r + 1 = 2\tilde{r} + 1$ for some $\tilde{r}$. This equation has 1 solution for m odd and 2 solutions for m even. It is easy to see that if m is odd, there are qm distinct factors of type 1 and if m is even their number is $\frac{qm}{2}$. The counts for the second type are exactly the same.

Let us now look how the two types of factors overlap. Fix r and k and suppose

$$\begin{aligned}2r - 1 + k(b-1) &= 2r' - 1, \\ 2r + 1 &= 2r' + 1 + k'(b-1).\end{aligned}$$

It follows that $2r' = 2r + k(b-1)$ which may not have a solution only if m is even and b is even, otherwise it has a solution and the two types overlap completely. If m is even and b is even, the two types overlap only if k is even. Thus, they overlap in $\frac{qm}{4}$ cases.

$n = 3$
It follows from $\mathcal{L}_4(\mathbf{t}_{b,m})$ that there are the following 3 types of factors of length 3:

1. $(\rho^k(r-1)+r)(2r+1)(2r+3)$ for $r \in \mathbb{Z}_m$, $0 \le k < q$;
2. $(2r'-3)(2r'-1)(r'+\rho^{-k'}(r'+1))$ for $r' \in \mathbb{Z}_m$, $0 \le k' < q$;
3. $(2r''-1)(r''+\rho^{-k''}(r''+1))(\rho^{-k''}(r''+1)+\rho^{-k''}(r''+1)+1))$ for $r'' \in \mathbb{Z}_m$, $0 \le k'' < q$.

Analogously to the previous case $n = 2$, it can be shown that there are qm distinct factors of each type if m is odd, and $\frac{qm}{2}$ distinct factors of each type if m is even.

When investigating the common factors of each type, one can show that each pair has m common factors if m is odd, and $\frac{m}{2}$ common factors otherwise. Overall, we find that $\#\mathcal{L}_3(\mathbf{v}) = 3qm - 2m$ if m is odd and $\#\mathcal{L}_3(\mathbf{v}) = \frac{3qm}{2} - m$ if m is even. □

Lemma 37. *Let $G' = I_2'(m)$ and $\mathbf{v} = S(\mathbf{t}_{b,m})$. The values of $F(n)$ for $n \in \{1,2,3\}$ are as follows:*

n	*m odd*	*m even, b odd*	*m even, b even*
1	m	$\frac{m}{2}$	m
2	qm	$\frac{qm}{2}$	$\frac{qm}{4}$
3	qm	$\frac{qm}{2}$	$\frac{qm}{2}$

Proof. $n = 1$:
It is not hard to show that every factor of length 1 is a Ψ-palindrome for a unique $\Psi \in G'$.

$n = 2$:
Suppose that the first type of factor of length 2 from the proof of Lemma 36 is a Ψ_ℓ-palindrome. We have

$$2r - 1 + k(b-1) = \Psi_\ell(2r+1) = \ell - 2r - 1,$$

which leads to $\ell = 4r - 2 + k(b-1)$, i.e., every factor of length 2 is a Ψ-palindrome for some $\Psi \in G'$ except for the case of m even and b even where one may find such ℓ only if k is even. As one can see in the proof of Lemma 36, the case of k even is when the two types of factors of length 2 overlap, thus the total number of G'-palindromes is $\frac{qm}{4}$ in this case.

$n = 3$:
We will refer to the 3 types of factors of length 3 as given in the proof of Lemma 36 above. The first two types can be a Ψ-palindrome for some $\Psi \in I_2(m)$ if and only if $k = 0$ or $k' = 0$. This case is also included in the third type for $k'' = 0$, so we need just to check this type.

Suppose that a factor of the third type is a Ψ_ℓ-palindrome:

$$2r'' - 1 = \Psi_\ell(\rho^{-k''}(r''+1) + \rho^{-k''}(r''+1)+1)) = \ell - (2r'' + 3 - 2k''(b-1)),$$
$$r'' + \rho^{-k''}(r''+1) = 2r'' + 1 - k''(b-1) = \Psi_\ell(r'' + \rho^{-k''}(r''+1)) = \ell - (2r'' + 1 - k''(b-1)).$$

Both equalities yield $\ell = 4r'' + 2 - 2k''(b-1)$. Thus, every factor of type 3 is a Ψ-palindrome for some $\Psi \in I_2(m)$. According to the proof of Lemma 36, there are qm such factors if m is odd, and $\frac{qm}{2}$ such factors otherwise. □

Proof of the second part of Theorem 31. Results of Lemmas 36 and 37 can be summarized into the following table:

	m odd	m even, b odd	m even, b even
$\Delta\mathcal{C}_{\mathbf{v}}(1)$	$(q-1)m$	$(q-1)\frac{m}{2}$	$\frac{3qm}{4}-m$
$F(1)+F(2)$	$(q+1)m$	$\frac{(q+1)m}{2}$	$\frac{qm}{4}+m$
$\Delta\mathcal{C}_{\mathbf{v}}(2)$	$2qm-2m$	$qm-m$	$\frac{3qm}{4}-m$
$F(2)+F(3)$	$2qm$	qm	$\frac{3qm}{4}$
$\#G'$	$2m$	m	m

Therefore the assumption of Lemma 35 is satisfied in the case when m or b is odd. Consequently, $S(\mathbf{t}_{b,m})$ is G'-rich. □

Corollary 38. *Let $b \in \mathbb{N}, b \geq 1$ and S be the operation defined by (18) for the alphabet $\mathbb{Z}_4$. We have*

- *$S(\mathbf{t}_{2b+1,4})$ is an infinite word over the binary alphabet $\{1,3\}$ and it is H-rich (here H stands for the group generated by the both involutory antimorphisms over the binary alphabet {1,3}).*
- *$S^2(\mathbf{t}_{2b+1,4})$ is an infinite word over the binary alphabet $\{0,2\}$ and it is R-rich (R stands for the reversal mapping over the binary alphabet $\{0,2\}$).*
- *$S^k(\mathbf{t}_{2b+1,4})$ is an infinite word over the binary alphabet $\{0,2\}$ and it is almost R-rich for any $k \in \mathbb{N}, k \geq 2$.*

Proof. The fact that the alphabet of $S(\mathbf{t}_{2b+1,4})$ is $\{1,3\}$ is shown in the proof of Lemma 36, where $\mathcal{L}_1(\mathbf{v})$ is described for any generalized Thue–Morse word $\mathbf{v}$. According to the second part of Theorem 31, the word $S(\mathbf{t}_{2b+1,4})$ is $I_2'(4)$-rich. The group $I_2'(4)$ defined by (19) is isomorphic to H.

It is easy to see that the operation S assigns to any word $\mathbf{u} \in \{1,3\}^{\mathbb{N}}$ the word over the alphabet $\{0,2\}$. The second part of the corollary follows from Theorem 24, the third one from Theorem 27. □

6 Comments and open questions

For infinite words over the binary alphabet $\{0,1\}$, we illustrated that the operation S puts into a broader context the classical richness (here usually referred to as R-richness) and H-richness. The main open question is which other operation acting on infinite words behaves analogously. Let us mention here some open questions connected with S.

- The operation S on $\{0,1\}$ applied to an almost R-rich word gives an almost R-rich word, see Theorem 27. In particular, any iteration of S applied to a Sturmian word $\mathbf{u}$ gives an almost rich word, cf. Corollary 28. Is the R-defect of $S^k(\mathbf{u})$ zero as suggested by our computer experiments?
- The operation S on $\{0,1\}$ applied to an H-rich word gives an R-rich word. In particular, $S(\mathbf{t}_{b,2})$ is rich for any generalized Thue–Morse word $\mathbf{t}_{b,2}$. Our computational experiments suggest that $S^k(\mathbf{t}_{b,2})$ is R-rich for any $k \in \mathbb{N}, k \geq 1$, see Example 29. Is it true?
- On the other hand, any preimage by S of each Sturmian word $\mathbf{u}$ is H-rich and R-rich simultaneously, in fact it is a complementary-symmetric Rote word. Our computer experiments suggest that even the second preimage $S^{-2}(\mathbf{u})$ is simultaneously H- and R-rich, whereas $S^{-3}(\mathbf{u})$ is only H-rich, but not R-rich, see Example 30. Is it true?

We have introduced the operation S over the alphabet $\mathbb{Z}_m$ with $m \geq 3$ as well. But our results on multiliteral alphabet are restricted to special groups and words.

- We have considered G-richness for $G = I_2(m)$ only. Proposition 34 connects $I_2(m)$-richness of $\mathbf{u}$ and $I_2'(m)$-richness of $S(\mathbf{u})$ for words $\mathbf{u}$ satisfying the assumption (20). Is the proposition valid without the assumption?

- It would be interesting to study behaviour of ternary episturmian words with respect to operation S on $\mathbb{Z}_3$. For example, which group of symmetries G has the preimage of the Tribonacci word by S? Is the preimage G-rich? Are images of the Tribonacci word by S still R-rich?

- Corollary 38 illustrates that the operation S over the alphabet $\mathbb{Z}_4$ can produce binary almost R-rich words as well. What is the R-defect of the words $S^k(\mathbf{t}_{2b+1,4})$?

The last comment we want to state here concerns the palindromic closure operator. It is used for construction of standard episturmian words. The construction is governed by a directive sequence of letters Δ. Any episturmian word $\mathbf{u}$ is closed under reversal and $\mathbf{u}$ is rich in the classical sense. In [17], de Luca and De Luca introduced the concept of generalized pseudopalindromic closure operator, where multiple involutory antimorphisms are used. It means that the construction is governed by two sequences: a directive sequence of letters Δ and a directive sequence of antimorphisms Θ. Let us denote the resulting infinite word by $\mathbf{u}(\Delta, \Theta)$.

In general, $\mathbf{u}(\Delta, \Theta)$ is closed under the group G generated by the involutory antimorphisms occurring infinitely many times in the directive sequence Θ, but the word $\mathbf{u}(\Delta, \Theta)$ need not to be G-rich. Nevertheless, several examples of G-rich words constructed by generalized pseudopalindromic closure operator are already known. De Luca and de Luca showed that the Thue–Morse word $\mathbf{t} = \mathbf{t}_{2,2}$ can be constructed in this way. In [22] the generalized Thue–Morse words $\mathbf{t}_{b,m}$ with the same property are characterized. The concept of generalized pseudopalindromic closure on binary alphabet is systematically studied by Blondin Massé, Paquin, Tremblay and Vuillon in [11]. In particular, they proved that any standard complementary-symmetric Rote word can be constructed by using generalized pseudopalindromic closure operator. Nevertheless, the question which pairs (Δ, Θ) produce H-rich words is open and requires a deeper study.

Acknowledgments

The first author acknowledges financial support from the Czech Science Foundation grant GAČR 13-03538S and the second author acknowledges financial support from the Czech Science Foundation grant GAČR 13-35273P. The computer experiments were performed using the open-source algebra system SAGE [30]. We would like to thank the anonymous referee who provided us with useful hints to improve the article.

References

[1] J.-P. Allouche and J. Shallit, *Sums of digits, overlaps, and palindromes*, Discrete Math. Theor. Comput. Sci., 4 (2000), pp. 1–10.

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

[3] L. Balková, *Beta-integers and Quasicrystals*, PhD thesis, Czech Technical University in Prague and Université Paris Diderot-Paris 7, 2008.

[4] L. Balková, M. Bucci, A. De Luca, J. Hladký, and S. Puzynina, *Aperiodic pseudorandom number generators based on infinite words*, preprint available at http://arxiv.org/abs/1311.6002, (2013).

[5] L. Balková, E. Pelantová, and Š. Starosta, *Palindromes in infinite ternary words*, RAIRO-Theor. Inf. Appl., 43 (2009), pp. 687–702.

[6] ——, *Sturmian jungle (or garden?) on multiliteral alphabets*, RAIRO-Theor. Inf. Appl., 44 (2010), pp. 443–470.

[7] ——, *Infinite words with finite defect*, Adv. in Appl. Math., 47 (2011), pp. 562–574.

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

[9] 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.

[10] A. Blondin Massé, S. Brlek, S. Labbé, and L. Vuillon, *Palindromic complexity of codings of rotations*, Theor. Comput. Sci., 412 (2011), pp. 6455–6463.

[11] A. Blondin Massé, G. Paquin, H. Tremblay, and L. Vuillon, *On generalized pseudostandard words over binary alphabets*, J. Integer Seq., 16 (2013).

[12] 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.

[13] S. Brlek and C. Reutenauer, *Complexity and palindromic defect of infinite words*, Theor. Comput. Sci., 412 (2011), pp. 493–497.

[14] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni, *On different generalizations of episturmian words*, Theor. Comput. Sci., 393 (2008), pp. 23–36.

[15] 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 (2009), pp. 60–74.

[16] J. Cassaigne, *Complexity and special factors*, Bull. Belg. Math. Soc. Simon Stevin 4, 1 (1997), pp. 67–88.

[17] A. de Luca and A. De Luca, *Pseudopalindrome closure operators in free monoids*, Theor. Comput. Sci., 362 (2006), pp. 282–300.

[18] X. Droubay, J. Justin, and G. Pirillo, *Episturmian words and some constructions of de Luca and Rauzy*, Theor. Comput. Sci., 255 (2001), pp. 539–553.

[19] X. Droubay and G. Pirillo, *Palindromes and Sturmian words*, Theor. Comput. Sci., 223 (1999), pp. 73–85.

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

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

[22] T. Jajcayová, E. Pelantová, and Š. Starosta, *Palindromic closures using multiple antimorphisms*, Theor. Comput. Sci., 533 (2014), pp. 37–45.

[23] E. Pelantová and Š. Starosta, *Almost rich words as morphic images of rich words*, Int. J. Found. Comput. Sci., 23 (2012), pp. 1067–1083.

[24] ——, *Languages invariant under more symmetries: overlapping factors versus palindromic richness*, Discrete Math., 313 (2013), pp. 2432–2445.

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

[26] G. Rote, *Sequences with subword complexity $2n$*, J. Number Th., 46 (1993), pp. 196–213.

[27] Š. Starosta, *On theta-palindromic richness*, Theor. Comput. Sci., 412 (2011), pp. 1111–1121.

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

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

[30] W. Stein et al., *Sage Mathematics Software (Version 6)*, The Sage Development Team, 2014. `http://www.sagemath.org`.

Chapter 7

Palindromic closures using multiple antimorphisms

Authors:	T. Jajcayová, E. Pelantová, and Š. Starosta
Publication reference:	Theor. Comput. Sci., 533 (2014), pp. 37–45
Cited as:	[VI]

Theoretical Computer Science 533 (2014) 37–45

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Palindromic closures using multiple antimorphisms

Tatiana Jajcayová [a], Edita Pelantová [b], Štěpán Starosta [c,*]

[a] Department of Applied Informatics, FMFI Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia
[b] Department of Mathematics, FNSPE, Czech Technical University in Prague, Trojanova 13, 120 00 Praha 2, Czech Republic
[c] Department of Applied Mathematics, FIT, Czech Technical University in Prague, Thákurova 9, 160 00 Praha 6, Czech Republic

ARTICLE INFO

Article history:
Received 28 June 2013
Received in revised form 12 February 2014
Accepted 16 March 2014
Communicated by D. Perrin

Keywords:
Palindromic closure
Generalized Thue–Morse word
Involutory antimorphism

ABSTRACT

A generalized pseudostandard word $\mathbf{u}$, as introduced in 2006 by de Luca and De Luca, is given by a directive sequence of letters from an alphabet $\mathcal{A}$ and by a directive sequence of involutory antimorphisms acting on $\mathcal{A}^*$. Prefixes of $\mathbf{u}$ with increasing length are constructed using a pseudopalindromic closure operator.
We show that generalized Thue–Morse words $\mathbf{t}_{b,m}$, with $b, m \in \mathbb{N}$ and $b,\ m \geqslant 2$, are generalized pseudostandard words if and only if $\mathbf{t}_{b,m}$ is a periodic word or $b \leqslant m$. This extends the result of de Luca and De Luca obtained for the classical Thue–Morse words.

1. Introduction

A *palindromic closure* of a finite word w is the shortest palindrome having w as a prefix. This concept was introduced in 1997 by Aldo de Luca for words over the binary alphabet. In [10], de Luca showed that any standard Sturmian word is a limit of a sequence of palindromes (w_n), where w_0 equals the empty word and w_{n+1} is the palindromic closure of $w_n\delta_{n+1}$ for some letter δ_{n+1} from the binary alphabet. And, vice versa, the limit of such a sequence is always a standard Sturmian word. This construction was extended to any finite alphabet $\mathcal{A}$ by Droubay, Justin and Pirillo in [13], and words arising by their construction are called *standard episturmian words*. The sequence $\delta_1\delta_2\delta_3\ldots$ of letters that are added successively at each step is referred to as the *directive sequence* of the standard episturmian word.

An important generalization of standard episturmian words appeared in [11], where the palindromic closure is replaced by ϑ*-palindromic closure* with ϑ an arbitrary involutory antimorphism of the free monoid $\mathcal{A}^*$. The corresponding words are called ϑ*-standard words*, or *pseudostandard words*.

A further generalization was provided in 2008 by Bucci, de Luca, De Luca and Zamboni in [8], where the sequence (w_n) is allowed to start with an arbitrary finite word w_0, and the limit word is a ϑ*-standard word with the seed* w_0. Words obtained by these generalizations are in some sense quite similar to the standard episturmian words: by results of Bucci and De Luca [9], any ϑ-standard word with a seed is a morphic image of a standard episturmian word.

The described constructions of (w_n) guarantee that the language of a standard episturmian word contains infinitely many palindromes.

Droubay, Justin and Pirillo in [13] deduced that any finite word w contains at most $|w|+1$ distinct palindromes ($|w|$ stands for the length of w). A word w with exactly $|w|+1$ palindromes is called *rich* (in [14]), or *full* (in [7]). An infinite word is *rich* if every finite factor of this word is rich. Examples of rich words include all episturmian words, see [13], two

* Corresponding author.

http://dx.doi.org/10.1016/j.tcs.2014.03.020

interval coding of rotations, see [3], words coding interval exchange transformation with symmetric permutation, see [2], etc.

The notion of rich word was generalized as well: the concept of ϑ*-rich* word was introduced in [18]. A ϑ-rich word is saturated by ϑ*-palindromes*, fixed points of the involutory antimorphism ϑ, up to the highest possible level. Another generalization of the concept of richness is a measure of how many palindromes are missing in a certain sense. This quantity is called the *palindromic defect* and it was first considered in [7].

Again, both generalizations of rich words are not too far from the original notion. In particular, any uniformly recurrent ϑ-rich word and any uniformly recurrent word with finite defect is just a morphic image of a rich word, see [15]. Nevertheless, the operation of palindromic closure enables us to construct a big class of rich words.

Up to this point, we were concerned with properties of words with respect to one fixed involutory antimorphism ϑ on $\mathcal{A}^*$.

In the last section of the paper [11], De Luca and de Luca introduced an even more general concept – generalized pseudostandard words. They considered a set $\mathcal{I}$ of involutory antimorphisms over $\mathcal{A}^*$ and beside a directive sequence $\Delta = \delta_1\delta_2\delta_3\ldots$ of letters from $\mathcal{A}$ also a directive sequence of antimorphisms $\Theta = \vartheta_1\vartheta_2\vartheta_3\ldots$ from $\mathcal{I}$. The construction of (w_n) starts with the empty word w_0 and recursively, w_n is the ϑ_n-palindromic closure of $w_{n-1}\delta_n$.

Then De Luca and de Luca focused on the prominent Thue–Morse word $\mathbf{u}_{TM}$. They showed that $\mathbf{u}_{TM}$ is a generalized pseudostandard word with the directive sequences $\Delta = 01111\ldots = 01^\omega$ and $\Theta = RERERE\ldots = (RE)^\omega$, where R denotes the mirror image operator, and E the antimorphism which exchanges the letters $0 \leftrightarrow 1$.

The example of the Thue–Morse word illustrates that generalized pseudostandard words substantially differ from the previous notions where only one antimorphism is used. This is due to the fact that the Thue–Morse word is not a morphic image of a standard episturmian word as all ϑ-standard words are. This can be seen when comparing the factor complexities of a standard episturmian word, which is of the form $an + b$ except for finitely many integers n (see [13]), and the factor complexity of the Thue–Morse word (see [6] or [12]).

The results of [5] confirm that the notion of the generalized pseudostandard word is very fruitful. In particular, Blondin-Massé, Paquin, Tremblay, and Vuillon showed that any standard Rote word is a generalized pseudostandard word. Since the Rote words are defined over the binary alphabet, their directive sequences Θ contain antimorphisms R and E only.

In this article we focus on the so-called generalized Thue–Morse words. Given two integers b and m such that $b > 1$ and $m > 1$, we denote the generalized Thue–Morse word by $\mathbf{t}_{b,m}$. The alphabet of $\mathbf{t}_{b,m}$ is $\mathcal{A} = \mathbb{Z}_m = \{0, \ldots, m-1\}$. For a given integer base b, the number $s_b(n)$ denotes the digit sum of the expansion of number n in the base b. The word $\mathbf{t}_{b,m}$ is defined

$$\mathbf{t}_{b,m} = \big(s_b(n) \bmod m\big)_{n=0}^{\infty}. \tag{1}$$

In this notation the classical Thue–Morse word equals $\mathbf{t}_{2,2}$. As shown in [19], the language of $\mathbf{t}_{b,m}$ is closed under a finite group containing m involutory antimorphisms. This group is isomorphic to the dihedral group $I_2(m)$. Our aim in this paper is to prove the following theorem:

Theorem 1.1. *The generalized Thue–Morse word $\mathbf{t}_{b,m}$ is a generalized pseudostandard word if and only if $b \leqslant m$ or $b - 1 = 0$ $(\bmod\ m)$.*

Unlike the case of the standard words with seed, very little is known about the properties of generalized pseudostandard words. In the last section, we propose several questions the answering of which would bring a better understanding of the structure of such words.

Our motivation for the study of generalized pseudostandard words stems from a desire to find G-rich words recently introduced in [16]. Words that are rich in the original sense are G-rich with respect to $G = \{R, Id\}$. In [19], the last author showed that the words $\mathbf{t}_{b,m}$ are $I_2(m)$-rich. In particular, the classical Thue–Morse word is H-rich with $H = \{E, R, ER, Id\}$. Using the result of [3], one can also show that the complementary symmetric Rote words are H-rich (for definition of the Rote words see [17]). These examples are almost all the examples of G-rich words we know for which the group G is not isomorphic to $\{R, Id\}$. We believe that generalized pseudostandard words can provide many other new examples.

2. Preliminaries

By $\mathcal{A}$ we denote a finite set of symbols usually called the *alphabet*. A *finite word* w over $\mathcal{A}$ is a string $w = w_0w_1\cdots w_{n-1}$ with $w_i \in \mathcal{A}$. Its *length* is denoted $|w|$ and equals n. The set of all finite words over $\mathcal{A}$, including the *empty word* ε, together with the operation of concatenation of words, form the free monoid $\mathcal{A}^*$. A *morphism* of $\mathcal{A}^*$ is a mapping $\varphi: \mathcal{A}^* \to \mathcal{A}^*$ satisfying $\varphi(wv) = \varphi(w)\varphi(v)$ for all finite words $w, v \in \mathcal{A}^*$. A morphism is uniquely given by the images $\varphi(a)$ of all letters $a \in \mathcal{A}$. If, moreover, there exists a letter $b \in \mathcal{A}$ and a non-empty word $w \in \mathcal{A}^*$ such that $\varphi(b) = bw$, then the morphism φ is called a *substitution*.

If a word $w \in \mathcal{A}^*$ can be written as a concatenation $w = uvz$, then v is called a *factor* of w. If u is the empty word, then v is called a *prefix* of w; if z is the empty word, then v is called a *suffix* of w. If φ is a substitution, then for every $n \in \mathbb{N}$ the word $\varphi^n(b)$ is a prefix of $\varphi^{n+1}(b)$.

An *infinite word* $\mathbf{u}$ over $\mathcal{A}$ is a sequence $u_0u_1u_2\ldots \in \mathcal{A}^{\mathbb{N}}$. The set of all factors of $\mathbf{u}$ is denoted by $\mathcal{L}(\mathbf{u})$ and referred to as the *language of* $\mathbf{u}$. The action of a morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ can be naturally extended to $\mathcal{A}^{\mathbb{N}}$ by $\varphi(\mathbf{u}) = \varphi(u_0u_1u_2\ldots) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots$. If $\varphi(\mathbf{u}) = \mathbf{u}$ for some infinite word $\mathbf{u}$, then the word $\mathbf{u}$ is a *fixed point* of the morphism φ. Every substitution φ has a fixed point, namely the infinite word which has prefix $\varphi^n(b)$ for every n; this infinite word is denoted $\varphi^\infty(b)$.

Example 2.1. The Thue–Morse word $\mathbf{u}_{TM}$ is a fixed point of the substitution φ_{TM} which maps $0 \mapsto \varphi_{TM}(0) = 01$ and $1 \mapsto \varphi_{TM}(1) = 10$. The substitution has two fixed points: the Thue–Morse word $\mathbf{u}_{TM} = \varphi^\infty(0)$ and the word $\varphi^\infty(1)$.

To manipulate a fixed point $\mathbf{u}$ of a substitution φ, we will need the notion of an ancestor: We say that a word $w = w_0w_1\cdots w_k$ is a *φ-ancestor* of a word $v \in \mathcal{L}(\mathbf{u})$ if the following three conditions are satisfied:

- v is a factor of $\varphi(w_0w_1\cdots w_k)$,
- v is not a factor of $\varphi(w_1\cdots w_k)$,
- v is not a factor of $\varphi(w_0w_1\cdots w_{k-1})$.

Example 2.2. Consider the Thue–Morse word

$$\mathbf{u}_{TM} = 011010011001011010010110 01\ldots = \varphi_{TM}(0)\varphi_{TM}(1)\varphi_{TM}(1)\varphi_{TM}(0)\varphi_{TM}(1)\varphi_{TM}(0)\ldots$$

The factor $v = 010011$ has an ancestor $w = 1101$, since v is a factor of $\varphi_{TM}(1101) = 10100110$, and w is neither a factor of $\varphi_{TM}(101) = 100110$ nor a factor of $\varphi_{TM}(110) = 101001$. In fact, w is the unique ancestor of 010011.

The factor $v = 010$ is a factor of $\varphi(11) = 1010$ and a factor of $\varphi(00) = 0101$. Thus $v = 010$ has two ancestors, namely 11 and 00.

Let us now define the key notion of this article. Any mapping $\Psi : \mathcal{A}^* \mapsto \mathcal{A}^*$ satisfying

$$\Psi(uv) = \Psi(v)\Psi(u) \quad \text{for any } u, v \in \mathcal{A}^* \quad \text{and} \quad \Psi^2 \text{ equal to the identity}$$

is an *involutory antimorphism*. Any antimorphism Ψ is determined by the images of the letters from $\mathcal{A}$. The restriction of an involutory antimorphism Ψ to the alphabet $\mathcal{A}$ is a permutation on $\mathcal{A}$ with cycles of length 1 or 2 only.

A word $w \in \mathcal{A}^*$ is a *Ψ-palindrome* if $\Psi(w) = w$. The notion *pseudopalindrome* is also used.

A *Ψ-palindromic closure* of a factor $w \in \mathcal{A}^*$ is the shortest Ψ-palindrome having w as a prefix. The Ψ-palindromic closure of w is denoted by w^Ψ. If $w = uq$ such that q is the longest Ψ-palindromic suffix of w, then

$$w^\Psi = uq\Psi(u).$$

Example 2.3. There are two distinct involutory antimorphisms on the alphabet $\mathcal{A} = \{0, 1\}$: R, the mirror image, and E, the antimorphism exchanging letters, i.e., $E(0) = 1$ and $E(1) = 0$. Put $u = 0110110$. Then u is an R-palindrome, and thus $u^R = u$. The word u is not an E-palindrome, since $E(u) = E(0110110) = 1001001 \neq u$. The E-palindromic closure of u is $u^E = 011011001001$ since the longest E-palindromic suffix of u is 10.

In [11], de Luca and De Luca generalized the notion of standard words considering the set $\mathcal{I}$ of all involutory antimorphisms on $\mathcal{A}^*$ instead of just one fixed antimorphism. We will denote by $\mathcal{I}^{\mathbb{N}}$ the set of all infinite sequences over $\mathcal{I}$.

Definition 2.4. Let $\Theta = \vartheta_1\vartheta_2\vartheta_3\ldots \in \mathcal{I}^{\mathbb{N}}$ and $\Delta = \delta_1\delta_2\delta_3\ldots \in \mathcal{A}^{\mathbb{N}}$. Denote

$$w_0 = \varepsilon \quad \text{and} \quad w_n = (w_{n-1}\delta_n)^{\vartheta_n} \quad \text{for any } n \in \mathbb{N}, n \geqslant 1.$$

The word

$$\mathbf{u}_\Theta(\Delta) = \lim_{n\to\infty} w_n$$

is called a *generalized pseudostandard word* with the directive sequence of letters Δ and the directive sequence of antimorphisms Θ.

Let us stress that the definition of $\mathbf{u}_\Theta(\Delta)$ is correct as w_n is a prefix of w_{n+1} for any n.

Example 2.5. Consider the directive sequence of letters $\Delta = 0(101)^\omega$ and the directive sequence of antimorphisms $\Theta = (RE)^\omega$. Then

$$\begin{aligned}
&w_0 = \varepsilon\\
&w_1 = 0^R = 0\\
&w_2 = (01)^E = 01\\
&w_3 = (010)^R = 010\\
&w_4 = (0101)^E = 0101\\
&w_5 = (01011)^R = 01011010\\
&w_6 = (010110100)^E = 010110100101\\
&w_7 = (0101101001011)^R = 0101101001011010\\
&w_8 = (01011010010110101)^E = 01011010010110101001011010010 1\\
&\vdots
\end{aligned}$$

The authors of [11] proved that the famous Thue–Morse word $\mathbf{u}_{TM}$ is a generalized pseudostandard word with directive sequences

$$\Delta = 01^\omega \quad \text{and} \quad \Theta = (ER)^\omega.$$

2.1. The generalized Thue–Morse words and their properties

In the Introduction, we defined a generalized Thue–Morse word by (1), i.e., its n-th letter is the digit sum of the expansion of n in base b taken modulo m. It can be shown that $\mathbf{t}_{b,m}$ is a fixed point of the substitution $\varphi_{b,m}$ over the alphabet $\mathbb{Z}_m$:

$$\varphi(k) = \varphi_{b,m}(k) = k(k+1)(k+2)\cdots(k+b-1) \quad \text{for every } k \in \mathbb{Z}_m \tag{2}$$

where letters are expressed modulo m. As already stated in [1], $\mathbf{t}_{b,m}$ is periodic if and only if $b = 1 \pmod m$.

(Note about our subsequent notation: When dealing with letters from $\mathbb{Z}_m$, we will consider all operations modulo m. We will denote the relation $x = y \pmod m$ by $x =_m y$ to ease the notation.)

The language of $\mathbf{t}_{b,m}$ has many symmetries: denote by $I_2(m)$ the group generated by antimorphisms Ψ_x defined for every $x \in \mathbb{Z}_m$ by

$$\Psi_x(k) = x - k \quad \text{for every } k \in \mathbb{Z}_m. \tag{3}$$

This group — usually called the *dihedral group of order* $2m$ — contains m morphisms and m antimorphisms. As shown in [19], if w is a factor of $\mathbf{t}_{b,m}$, then $\nu(w)$ is a factor of $\mathbf{t}_{b,m}$ for every element ν of the group $I_2(m)$.

Let us list some properties of the generalized Thue–Morse word we will use later. They are not hard to observe. (See also [19].)

Properties of $\mathbf{t}_{b,m}$

1. Let $b \neq_m 1$. If $v = v_0v_1\cdots v_{k-1}$ is a factor of $\mathbf{t}_{b,m}$ of length $k \geqslant 2b+1$, then there exists $j \in \{0, 1, \ldots, k-2\}$ such that $v_j + 1 \neq_m v_{j+1}$. Such index j will be called *jump* in v. It is important to note here that we always start indices from 0. Sometimes, when no confusion can occur, we will say that there is a jump between the letters v_j and v_{j+1}.
2. If $b \neq_m 1$, then a factor v of length at least $2b+1$ has uniquely determined φ-ancestors.
3. If $v = v_0v_1\cdots v_k$ is a Ψ-palindrome and an index j is a jump in v, then also the index $k-j$ is a jump in v.
4. $\Psi_x\varphi = \varphi\Psi_{x-b+1}$ for every $x \in \mathbb{Z}_m$.
5. For every $\Psi \in I_2(m)$ there exists a unique $\Psi' \in I_2(m)$ such that $\Psi\varphi = \varphi\Psi'$.
6. If $w \neq \varepsilon$ is a Ψ-palindrome for some $\Psi \in I_2(m)$, then for every antimorphism $\Psi' \in I_2(m)$ such that $\Psi' \neq \Psi$ we have $\Psi'(w) \neq w$.
7. If $\Psi \neq \Psi'$, then for every letter $a \in \mathcal{A}$, $\Psi(a) \neq \Psi'(a)$.

3. Proof of Theorem 1.1

Proof of Theorem 1.1 will be split into Propositions 3.1 and 3.9.

Proposition 3.1. *Let* $m, b \in \mathbb{Z}$. *Denote*

$$\Delta = 0\big(12\cdots(b-1)\big)^\omega \in \mathbb{Z}_m^{\mathbb{N}} \quad \textit{and} \quad \Theta = (\Psi_0\Psi_1\cdots\Psi_{m-1})^\omega \in I_2(m)^{\mathbb{N}}. \tag{4}$$

If $b \leqslant m$ *or* $b = 1 \pmod m$, *then the generalized pseudostandard word* $\mathbf{u}_\Theta(\Delta)$ *with directive sequences* Δ *and* Θ *equals* $\mathbf{t}_{b,m}$.

Axel Thue found the classical Thue–Morse word $\mathbf{t}_{2,2}$ when he searched for infinite words without overlapping factors, i.e., words without factors of the form $v = ws = pw$ such that $|w| > |s|$. The authors of [4] showed that the generalized Thue–Morse word $\mathbf{t}_{b,m}$ is overlap-free if and only if $b \leqslant m$. It is worth to mention that the same condition appears in our characterization of non-periodic words $\mathbf{t}_{b,m}$ which are the generalized pseudostandard words.

For parameters $b = m = 2$, Proposition 3.1 was shown in [11]. The following example illustrates that the assumption $b \leqslant m$ is crucial for validity of Proposition 3.1.

Example 3.2. Consider $b = 4$ and $m = 2$. On the alphabet $\mathcal{A} = \mathbb{Z}_2$, we have $\Psi_0(k) = 0 - k = k$ and $\Psi_1(k) = 1 - k$ for any letter k. In the notation of Example 2.3, it means $\Psi_0 = R$ and $\Psi_1 = E$. Therefore the sequences Δ and Θ from Proposition 3.1 coincide with sequences Δ and Θ from Example 2.5. The generalized Thue–Morse word $\mathbf{t}_{4,2}$ starts as

$$\mathbf{t}_{4,2} = 0101101001011010101001011010010101011010 0101\ldots$$

Note that, using the notation from Definition 2.4, w_8 is not a prefix of $\mathbf{t}_{4,2}$ and thus the generalized pseudostandard word $\mathbf{u}_\Theta(\Delta)$ from Proposition 3.1 does not correspond to $\mathbf{t}_{4,2}$.

Propositions 3.1 and 3.9 rely on several technical lemmas. The first one settles the case for periodic Thue–Morse words.

Lemma 3.3. *Let $b =_m 1$. The word $\mathbf{u}_\Theta(\Delta)$ with the directive sequences Δ and Θ given in (4) equals $\mathbf{t}_{b,m}$.*

Proof. Let $n = \sum_{i=0}^{k} a_i b^i$ be the expansion of the number n in the base b. The assumption $b =_m 1$ implies $b^i =_m 1$ for any $i \in \mathbb{N}$. With respect to (1), we can write

$$\mathbf{t}_{b,m}(n) =_m s_b(n) = \sum_{i=0}^{k} a_i =_m \sum_{i=0}^{k} a_i b^i = n.$$

Since $\Delta = 0(1 \cdots (b-1))^\omega$ equals $(01 \cdots (m-1))^\omega$, we have showed that $\mathbf{t}_{b,m} = \Delta$. Moreover, the sequence of antimorphisms Θ can be indexed by natural numbers as $\Theta = \Psi_0\Psi_1\Psi_2\Psi_3\ldots$ where $\Psi_n = \Psi_x$ for $n =_m x$. Clearly by (3)

$$\Psi_n(012\cdots n) = 012\cdots n = (012\cdots n)^{\Psi_n}. \tag{5}$$

Let the words w_n have the meaning as in Definition 2.4. We will show by induction that $w_{n+1} = 0123\cdots n$ for any $n \in \mathbb{N}$. We have $w_1 = (0)^{\Psi_0} = 0 = \Psi_0(0)$. Using definition of w_{n+1} and (5) we get

$$w_{n+1} = (w_n\delta_{n+1})^{\Psi_n} = \big((012\cdots(n-1))n\big)^{\Psi_n} = 012\cdots(n-1)n.$$

This means that $\mathbf{t}_{b,m} = \lim_{n\to\infty} w_n$, as desired. □

We can now concentrate on the non-periodic Thue–Morse words, i.e., on the case $b \neq_m 1$, which will be treated using several lemmas.

Lemma 3.4. *If $\Psi \in I_2(m)$ is an antimorphism and p is a Ψ-palindromic factor of $\mathbf{t}_{b,m}$ such that $\varphi(a_1a_2)a_3$ is a suffix of p for some letters $a_1, a_2, a_3 \in \mathcal{A}$, then there exists a word w of length at least 2 and antimorphism $\Psi' \in I_2(m)$ such that*

$$p = \Psi(a_3)\varphi(w)a_3, \quad \Psi'(w) = w \quad \textit{and} \quad \Psi\varphi = \varphi\Psi'.$$

Proof. Let $p = p_0p_1\cdots p_n$. Since p has a suffix $\varphi(a_1a_2)a_3$ of length $2b+1$, according to Property 1, p has a jump position. The jump position of p is either $n-1$ or $n-b-1$. As p is a Ψ-palindrome, the index 0 or b is a jump of p. Since every image of a letter by φ is of the same length b and contains no jump, a prefix of p is of the form $a_3'\varphi(a_2'a_1')$ for some letters a_1', a_2', a_3'. Thus $p = \Psi(a_3)\varphi(w)a_3$ for some word w with $|w| \geqslant 2$. As $p = \Psi(p)$, using Property 5, we get $\varphi(w) = \Psi(\varphi(w)) = \varphi\Psi'(w)$. Since φ is injective, we have $w = \Psi'(w)$. □

Lemma 3.5. *Fix $n \in \mathbb{N}$ and $k \in \{2, \ldots, b-1\}$. Put $\Psi = \Psi_{(b-1)n+k}$. The longest Ψ-palindromic suffix of the factor*

$$v = \varphi^n(0)\varphi^n(1)\varphi^n(2)\cdots\varphi^n(k-1)k$$

is $\Psi(k)\varphi^n(1)\varphi^n(2)\cdots\varphi^n(k-1)k$ and thus

$$v^\Psi = \varphi^n(0)\varphi^n(1)\varphi^n(2)\cdots\varphi^n(k-1)\varphi^n(k). \tag{6}$$

Proof. First we show that $u = \Psi(k)\varphi^n(1)\varphi^n(2)\cdots\varphi^n(k-1)k$ is a Ψ-palindromic suffix of the factor v. It is easy to check that the last letter of $\varphi^n(0)$ is the letter $(b-1)n$. In our notation $\Psi(k) = (b-1)n + k - k = (b-1)n$. Therefore $\Psi(k)$ is the last letter of $\varphi^n(0)$, and thus u is a suffix of v. To show that u is a Ψ-palindromic suffix, we need to show that

$$\Psi\big(\varphi^n(i)\big) = \varphi^n(k-i) \quad \text{for any } i = 1, 2, \ldots, k-1.$$

Using Property 4, we get

$$\Psi\varphi^n = \Psi_{(b-1)n+k}\varphi^n = \varphi^n\Psi_k,$$

and thus $\Psi(\varphi^n(i)) = \varphi^n(\Psi_k(i)) = \varphi^n(k-i)$ for all i, including $i = 0$ and $i = k$.

Now we show by contradiction that u is the longest Ψ-palindromic suffix of v. Consider the minimal n for which the statement is false, i.e., the longest Ψ-palindromic suffix of v — denote it by p — is longer than u. Since $01\cdots(k-1)k = (01\cdots(k-1)k)^{\Psi}$, the minimal n is greater than 0. As $|p| > |u|$, p has a suffix u and we can apply Lemma 3.4. Therefore p has the form $p = \Psi(k)\varphi(w)k$, where w is a Ψ'-palindrome and $\varphi\Psi' = \Psi\varphi$. According to Property 4, we have $\Psi' = \Psi_{(b-1)(n-1)+k}$. In particular, $\Psi'(k)wk$ is Ψ'-palindromic suffix of $\varphi^{n-1}(0)\varphi^{n-1}(1)\varphi^{n-1}(2)\cdots\varphi^{n-1}(k-1)k$.

As $|p| = 2 + |\varphi(w)| > |u|$, necessarily $|w| > |\varphi^{n-1}(1)\varphi^{n-1}(2)\cdots\varphi^{n-1}(k-1)|$. It means that $\varphi^{n-1}(0)\varphi^{n-1}(1)\varphi^{n-1}(2)\cdots\varphi^{n-1}(k-1)k$ has the longest Ψ'-palindromic suffix longer than $\Psi'(k)\varphi^{n-1}(1)\varphi^{n-1}(2)\cdots\varphi^{n-1}(k-1)k$—contradiction with the minimality of n. $\square$

Lemma 3.6. *Let v be a factor with the suffix $\varphi((a-1)a)1$ and let $b \neq_m 1$. Put $\Psi = \Psi_{a+b}$. Under these assumptions, the longest Ψ-palindromic suffix p of the factor v is of length at least 2. Moreover, for $|p|$ and the parameters a and b, the following holds:*

1. *if $|p| \geqslant b+1$, then $p = \Psi(1)\varphi(w)1$, where w is a Ψ_{a+1}-palindrome of length at least 2;*
2. *if $3 \leqslant |p| \leqslant b$, then $a + b =_m 1$ and $b > m$;*
3. *if $|p| = 2$, then either $a + b \neq_m 1$ or $a + b =_m 1$ with $b \leqslant m$.*

Proof. Since the last two letters of v are $(a+b-1)1$, and $\Psi(1) = a+b-1$, the word v has a palindromic suffix of length 2.

If p itself has the suffix $\varphi((a-1)a)1$, then the form of p is given by Lemma 3.4 as $p = \Psi(1)\varphi(w)1$. According to Properties 4 and 5 in Section 2.1 we have $\Psi_{a+b}\varphi = \varphi\Psi_{a+1}$ and thus w is a Ψ_{a+1}-palindrome.

Let p be shorter than the suffix $\varphi((a-1)a)1$. It means that p is a suffix of the factor $a(a+1)\cdots(a+b-2)a(a+1)\cdots(a+b-1)1$. Since $b-1 \neq_m 0$, we have a jump between letters $a+b-2$ and a. Let us discuss the following two cases separately:

(i) If $a + b \neq_m 1$, then the other jump is between the last two letters $a+b-1$ and 1. In the Ψ-palindrome, jump positions must be symmetric with respect to the center, and thus the only two candidates for the palindromic suffix are $(a+b-2)a(a+1)\cdots(a+b-1)1$ and $(a+b-1)1$. Since $\Psi(1) = \Psi_{a+b}(1) \neq_m a+b-2$, only the latter possibility $p = \Psi(1)1$ occurs.

(ii) If $a + b =_m 1$, then $a(a+1)\cdots(a+b-2)a(a+1)\cdots(a+b-1)1$ has only one jump, namely, as we mentioned above, between letters $a+b-2$ and a. Therefore the longest palindromic suffix p does not contain any jumps. It implies that p is a suffix of $a(a+1)\cdots(a+b-1)1$. Let $k \in \{0, 1, \ldots, m-1\}$ be a letter such that $p = (a+k)(a+k+1)\cdots(a+b-1)1$. Then $a + k = \Psi(1) = a+b-1 =_m 0$. Or equivalently, $k =_m b-1$.

If $b \leqslant m$, the equality $k =_m b-1$ has the only solution $k = b-1$, i.e., $p = (a+b-1)1 = \Psi(1)1$, as before.

If $b > m$, then the smallest $k \in \{0, 1, \ldots, b-1\}$ solving $k =_m b-1$, satisfies $k \leqslant b-1-m \leqslant b-3$, and as well $k > 0$ (since $k = b-1 \neq_m 0$). As $p = (a+k)(a+k+1)\cdots(a+b-1)1$ is of length $|p| = b-k+1$, we get $3 \leqslant |p| \leqslant b$. $\square$

The following claim addresses the question of the length of the longest Ψ-palindromic suffix of the factor $\varphi^n(0)1$.

Claim 3.7. *Let $b \neq_m 1$. Put $q = \min\{i \in \mathbb{N}: i > 0 \text{ and } i(b-1) =_m 0\}$ and for a fixed $n \in \mathbb{N}$ denote $\Psi = \Psi_{(b-1)n+1}$.*

1. *If $b \leqslant m$, then the longest Ψ-palindromic suffix of the factor $\varphi^n(0)1$ is of length 2.*
2. *If $b > m$ and $n < q$, then the longest Ψ-palindromic suffix of the factor $\varphi^n(0)1$ is of length 2.*
3. *If $b > m$ and $n = q$, then the longest Ψ-palindromic suffix of the factor $\varphi^n(0)1$ is of length greater than 2 and less than $b+1$.*

Proof. As $b-1 \neq_m 0$, we have $q \geqslant 2$. First we show that Claim holds for $n = 0$ and $n = 1$.

Consider $n = 0$. The factor 01 is a Ψ_1-palindrome as $\Psi_1(0) = 1$.

Consider $n = 1$. The factor $\varphi(0)1 = 01\cdots(b-1)1$ has only one jump, namely between two last letters $b-1$ and 1 because of $b-1 \neq_m 0$. Thus $\varphi(0)1$ cannot have Ψ-palindromic suffix longer than 2. Since $\Psi(b-1) = \Psi_b(b-1) = 1$, the factor $\varphi(0)1$ has the longest Ψ-palindromic suffix of length 2.

Now suppose that there exists an index n such that $\varphi^n(0)1$ has a Ψ-palindromic suffix of length at least 3. Consider the smallest such n. Obviously, $n \geqslant 2$. Denote by p the longest Ψ-palindromic suffix of $\varphi^n(0)1$. We will apply Lemma 3.6

with $v = \varphi^n(0)1 = \varphi(\varphi^{n-1}(0))1$. Since the last letter of $\varphi^{n-1}(0)$ equals $(n-1)(b-1)$, we denote $a = (n-1)(b-1)$. For this choice of a, the antimorphism $\Psi = \Psi_{(b-1)n+1} = \Psi_{a+b}$ is as required in Lemma 3.6. Since $|p| \geqslant 3$, only Cases 1 and 2 from Lemma 3.6 apply:

Case 1: $p = \Psi(1)\varphi(w)1 = (a+b-1)\varphi(w)1$ for some factor w of length $|w| \geqslant 2$ and w is a Ψ_{a+1}-palindrome.

Let us realize that $\varphi^n(0)1$ is a prefix of $\mathbf{t}_{b,m}$ for any n and thus $\varphi^n(0)\varphi(1)$ is its prefix as well. Since $(a+b-1)$ is the last letter of $\varphi(a)$, we can deduce that $\varphi(a)\varphi(w)\varphi(1)$ is a suffix of $\varphi^n(0)\varphi(1)$ and thus $aw1$ is a suffix of $\varphi^{n-1}(0)1$. Moreover, $aw1$ is a Ψ_{a+1}-palindrome of length $\geqslant 4$. This means that $\varphi^{n-1}(0)1$ has a Ψ'-palindromic suffix of length greater than 2, where $\Psi' = \Psi_{a+1} = \Psi_{(n-1)(b-1)+1}$. This is a contradiction with the minimality of n.

Case 2: $a + b =_m 1$, $b > m$.

Since we denoted $a = (n-1)(b-1)$, we have $n(b-1) =_m 0$. The smallest n satisfying this equality was denoted by q.

We can conclude: If $b \leqslant m$ then for all n, the longest Ψ-palindromic suffix of $\varphi^n(0)1$ is of length 2; if $b > m$ then for all $n < q$, the longest Ψ-palindromic suffix $\varphi^n(0)1$ is of length 2; if $b > m$ and $n = q$, then the longest Ψ-palindromic suffix $\varphi^n(0)1$ is of length greater than 2 and less than $b+1$. □

Lemma 3.8. *Let $b \neq_m 1$. Denote $q = \min\{i \in \mathbb{N}\colon\ i > 0$ and $i(b-1) =_m 0\}$. Fix $n \in \mathbb{N}$ and put $\Psi = \Psi_{(b-1)n+1}$.*

1. *If $b \leqslant m$, then $(\varphi^n(0)1)^{\Psi} = \varphi^n(0)\varphi^n(1)$.*
2. *If $b > m$ and $n < q$, then $(\varphi^n(0)1)^{\Psi} = \varphi^n(0)\varphi^n(1)$.*
3. *If $b > m$, then $(\varphi^q(0)1)^{\Psi}$ is not a prefix of $\mathbf{t}_{b,m}$.*

Proof. Let us denote by s the length of the longest Ψ-palindromic suffix of $\varphi^n(0)1$. We will apply Claim 3.7.

If $s = 2$, we clearly have $(\varphi^n(0)1)^{\Psi} = \varphi^n(0)\Psi(\varphi^n(0))$. According to Property 4, $\Psi\varphi^n = \varphi^n\Psi_1$ and thus $\Psi(\varphi^n(0)) = \varphi^n(\Psi_1(0)) = \varphi^n(1)$.

Consider now $s \in \{3, 4, \ldots, b\}$. Then $(\varphi^n(0)1)^{\Psi}$ is of length $2|\varphi^n(0)| + 2 - s$. From the form of the substitution φ and the fact that $\mathbf{t}_{b,m}$ is its fixed point, it follows that a jump in $\mathbf{t}_{b,m} = u_0u_1u_2\ldots$ can occur only on the indices $i - 1 =_b -1$. Since $\varphi^n(0)$ is a prefix of $\mathbf{t}_{b,m}$, the prefix of the palindrome $(\varphi^n(0)1)^{\Psi}$ of length $|\varphi^n(0)|$ has jumps on the positions $i - 1 =_b -1$. Jumps in any palindrome occur symmetrically with respect to the center of the palindrome. The length of the palindrome $(\varphi^n(0)1)^{\Psi}$ is $(2-s) \bmod b$. As $2 - s \neq_b 0$, jumps in the left part of $(\varphi^n(0)1)^{\Psi}$ are not compatible with the jump positions in $\mathbf{t}_{b,m}$ and thus $(\varphi^n(0)1)^{\Psi}$ cannot be a prefix of $\mathbf{t}_{b,m}$. □

Now we are ready to complete the proof of Proposition 3.1 for the non-periodic generalized Thue–Morse words.

Proof. From Lemma 3.8, Part 1, we get the first identity in the following list; the others follow from Lemma 3.5:

$$\begin{aligned}
&(1) \quad \left(\varphi^n(0)1\right)^{\Psi} = \varphi^n(0)\varphi^n(1) \quad \text{if } \Psi = \Psi_{(b-1)n+1} \text{ and } b \leqslant m;\\
&(2) \quad \left(\varphi^n(0)\varphi^n(1)2\right)^{\Psi} = \varphi^n(0)\varphi^n(1)\varphi^n(2) \quad \text{if } \Psi = \Psi_{(b-1)n+2};\\
&\quad\vdots\\
&(\text{b-1}) \quad \left(\varphi^n(0)\varphi^n(1)\cdots\varphi^n(b-2)(b-1)\right)^{\Psi} = \varphi^n(0)\varphi^n(1)\cdots\varphi^n(b-1) \quad \text{if } \Psi = \Psi_{(b-1)n+b-1}.
\end{aligned}$$

Since $\mathbf{t}_{b,m} = \lim_{n\to\infty}\varphi^n(0)$, this together with the simple fact

$$\varphi^n(0)\varphi^n(1)\cdots\varphi^n(b-1) = \varphi^n\big(\varphi(0)\big) = \varphi^{n+1}(0)$$

finishes the proof of Proposition 3.1. □

Proposition 3.9. *Let $m, b \in \mathbb{Z}$. If $b > m$ and $b \neq 1 \pmod m$, then $\mathbf{t}_{b,m}$ is not a generalized pseudostandard word.*

Proof. First, we show that a pseudopalindromic prefix of $\mathbf{t}_{b,m}$ whose length is greater than b is an image of a shorter pseudopalindromic prefix of $\mathbf{t}_{b,m}$.

Since the word $\mathbf{t}_{b,m} = 01\cdots(b-1)1\ldots$ has its first jump equal to $b-1$, every its pseudopalindromic prefix p longer than b has a jump $|p| - b$. This implies that $p = \varphi(p')$ for some prefix p'. Since $\Psi(p) = p$ for some antimorphism $\Psi \in I_2(m)$, according to Property 5, we have $\Psi(\varphi(p')) = \varphi(\Psi'(p')) = p = \varphi(p')$ for some antimorphism $\Psi' \in I_2(m)$. Since φ is injective, the last equality implies $\Psi'(p') = p'$, and thus p' is a Ψ'-palindromic prefix.

One can see that for all n, $\varphi^n(0)$ and $\varphi^n(01)$ are pseudopalindromic prefixes. Next, we show that for each $n \in \mathbb{N}$, the only pseudopalindromic prefix of $\mathbf{t}_{b,m}$ which is longer than $|\varphi^n(0)|$ and shorter than $2|\varphi^n(0)| + 2$ is the prefix $\varphi^n(0)\varphi^n(1)$.

This part of the proof will proceed by contradiction: Suppose that n is the minimal integer for which the claim does not hold. Clearly $n > 1$, since the claim can be easily verified for $n = 1$. Using the fact that every pseudopalindromic prefix of

$\mathbf{t}_{b,m}$ is a φ-image of a shorter one, we can immediately see that even for $n-1$ the statement does not hold, which is a contradiction with the minimality of n.

Since $b>2$, there is no pseudopalindromic prefix of length $|\varphi^n(0)|-1$. For the lengths of the words w_i from Definition 2.4, we have that $|w_{i+1}| \leqslant 2|w_i|+2$ for all i. Therefore, for each n, there exists an index i such that $w_i=\varphi^n(0)$ and $w_{i+1}=\varphi^n(0)\varphi^n(1)$. Let Ψ be the antimorphism which fixes w_{i+1}, i.e., $w_{i+1}=(w_i 1)^{\Psi}$. The lengths of w_i and w_{i+1} imply that the longest Ψ-palindromic suffix of $w_i 1$ is of length 2.

Since the last letter of $\varphi^n(0)$ is the letter $n(b-1)$, the antimorphism Ψ satisfies $\Psi(1)=n(b-1)$ and thus $\Psi=\Psi_{n(b-1)+1}$. Set $n=q$ where q is the order of $(b-1)$. It follows from Part 3 of Lemma 3.8 that the $\Psi_{q(b-1)+1}$-palindromic closure of $w_i 1$ is not a prefix of $\mathbf{t}_{b,m}$. $\square$

4. Comments and open questions

1. As shown in Proposition 3.1, the word $\mathbf{t}_{3,4}$ is a generalized pseudostandard word and its directive sequences are $\Delta=0(12)^{\omega}$ and $\Theta=(\Psi_0\Psi_1\Psi_2\Psi_3)^{\omega}$. One can easily check that the pairs
$$\Delta=0(21)^{\omega},\quad \Theta=(\Psi_1\Psi_2\Psi_3\Psi_0)^{\omega}\quad \text{and}\quad \Delta=01(12)^{\omega},\quad \Theta=\Psi_0\Psi_2\Psi_3(\Psi_0\Psi_1\Psi_2\Psi_3)^{\omega}$$
also correspond to the word $\mathbf{t}_{3,4}$.
The authors of [5] study this phenomenon for the generalized pseudostandard word on the binary alphabet, where $\Delta\in\{0,1\}^{\mathbb{N}}$ and $\Theta\in\{R,E\}^{\mathbb{N}}$. They defined the notion of a normalized bisequence and showed (Theorem 27 in [5]) that every pseudostandard word is generated by a unique normalized bisequence. Moreover, for any generalized pseudostandard word $\mathbf{u}_{\Theta}(\Delta)$, a simple algorithms which transforms the pair Δ, Θ into the normalized bisequence is given.
Question: Is it possible to generalize the notion of a normalized bisequence for the case of a multi-literal alphabet?
2. It is well known that the factor complexity of standard episturmian words is bounded by $(\#\mathcal{A}-1)n+1$. In particular, on a binary alphabet these words which are not periodic are precisely standard Sturmian words and their factor complexity is $\mathcal{C}(n)=n+1$.
In [5], the authors conjectured that generalized pseudostandard words on binary alphabet have their factor complexity bounded by $4n+const$.
The factor complexity of binary generalized Thue–Morse words can be found in [20]. The word $\mathbf{t}_{2k+1,2}$ is periodic, and thus its factor complexity is bounded by a constant. The word $\mathbf{t}_{2k,2}$ is aperiodic and its factor complexity is $\leqslant 4n$ for any parameter k. It means that even $\mathbf{t}_{4,2}$ and $\mathbf{t}_{6,2}$ (which are not generalized pseudostandard words) have a small complexity. Of course, it does not contradict the conjecture.
The factor complexity of generalized Thue–Morse words on any alphabet is deduced in [19]. If the word $\mathbf{t}_{b,m}$ is aperiodic, then
$$(qm-1)n\leqslant\mathcal{C}(n)\leqslant qmn,$$
where q is the order of $b-1$ in the additive group $\mathbb{Z}_m$, i.e. q is the minimal positive integer such that $q(b-1)=_m 0$.
The factor complexity of any infinite word can be derived from the knowledge of its bispecial factors. Each aperiodic standard episturmian word $\mathbf{u}$ has a nice structure of its bispecial factors. (A factor w is bispecial if and only if w is a palindromic prefix of $\mathbf{u}$.)
Question: Is it possible to describe the structure of bispecial factors for a generalized pseudostandard word?
3. It is known [13] that classical standard episturmian words with a periodic directive sequence $\Delta=(\delta_1\delta_2\cdots\delta_k)^{\omega}$ are invariant under a substitution. For example, the Tribonacci word has the directive sequence $\Delta=(012)^{\omega}$ and simultaneously, it is a fixed point of the substitution $\varphi: 0\mapsto 01, 1\mapsto 02, 2\mapsto 0$.
Let us denote $\mathbf{s}_{b,m}$ to be the generalized pseudostandard word with
$$\Delta=0\big(12\cdots(b-1)\big)^{\omega}\quad\text{and}\quad \Theta=(\Psi_0\Psi_1\cdots\Psi_{m-1})^{\omega}.$$
If $b\leqslant m$, then $\mathbf{s}_{b,m}=\mathbf{t}_{b,m}$ and obviously $\mathbf{s}_{b,m}$ is invariant under the substitution described in (2).
Question: Is the word $\mathbf{s}_{b,m}$ a fixed point of a substitution if $b>m$?

Acknowledgements

The first author acknowledges support from Project VEGA 1/0577/14, the second author acknowledges financial support from the Czech Science Foundation grant GAČR 13-03538S and the last author acknowledges financial support from the Czech Science Foundation grant GAČR 13-35273P.

References

[1] Jean-Paul Allouche, Jeffrey Shallit, Sums of digits, overlaps, and palindromes, Discrete Math. Theor. Comput. Sci. 4 (2000) 1–10.
[2] Peter Baláži, Zuzana Masáková, Edita Pelantová, Factor versus palindromic complexity of uniformly recurrent infinite words, Theoret. Comput. Sci. 380 (3) (2007) 266–275.

[3] A. Blondin Massé, S. Brlek, S. Labbé, L. Vuillon, Palindromic complexity of codings of rotations, Theoret. Comput. Sci. 412 (46) (2011) 6455–6463.
[4] Alexandre Blondin Massé, Srecko Brlek, Amy Glen, Sébastien Labbé, On the critical exponent of generalized Thue–Morse words, Discrete Math. Theor. Comput. Sci. 9 (1) (2007).
[5] Alexandre Blondin-Massé, Geneviève Paquin, Hugo Tremblay, Laurent Vuillon, On generalized pseudostandard words over binary alphabets, J. Integer Seq. 16 (2013), no. Article 13.2.11.
[6] Srecko Brlek, Enumeration of factors in the Thue–Morse word, Discrete Appl. Math. 24 (1–3) (1989) 83–96.
[7] Srecko Brlek, Sylvie Hamel, Maurice Nivat, Christophe Reutenauer, On the palindromic complexity of infinite words, Internat. J. Found. Comput. Sci. 15 (2) (2004) 293–306.
[8] Michelangelo Bucci, Aldo de Luca, Alessandro De Luca, On different generalizations of episturmian words, Theoret. Comput. Sci. 393 (1–3) (2008) 23–36.
[9] Michelangelo Bucci, Alessandro 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, Springer-Verlag, 2009, pp. 259–266.
[10] Aldo de Luca, Sturmian words: structure, combinatorics, and their arithmetics, Theoret. Comput. Sci. 183 (1) (1997) 45–82.
[11] Aldo de Luca, Alessandro De Luca, Pseudopalindrome closure operators in free monoids, Theoret. Comput. Sci. 362 (1–3) (2006) 282–300.
[12] Aldo de Luca, Stefano Varricchio, Some combinatorial properties of the Thue–Morse sequence and a problem in semigroups, Theoret. Comput. Sci. 63 (3) (1989) 333–348.
[13] Xavier Droubay, Jacques Justin, Giuseppe Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (1–2) (2001) 539–553.
[14] Amy Glen, Jacques Justin, Steve Widmer, Luca Q. Zamboni, Palindromic richness, European J. Combin. 30 (2) (2009) 510–531.
[15] Edita Pelantová, Štěpán Starosta, Almost rich words as morphic images of rich words, Internat. J. Found. Comput. Sci. 23 (05) (2012) 1067–1083.
[16] Edita Pelantová, Štěpán Starosta, Languages invariant under more symmetries: overlapping factors versus palindromic richness, Discrete Math. 313 (2013) 2432–2445.
[17] Günter Rote, Sequences with subword complexity $2n$, J. Number Theory 46 (1993) 196–213.
[18] Štěpán Starosta, On theta-palindromic richness, Theoret. Comput. Sci. 412 (12–14) (2011) 1111–1121.
[19] Štěpán Starosta, Generalized Thue–Morse words and palindromic richness, Kybernetika (Prague) 48 (3) (2012) 361–370.
[20] John Tromp, Jeffrey Shallit, Subword complexity of a generalized Thue–Morse word, Inform. Process. Lett. (1995) 313–316.

Chapter 8

An algorithm for enumerating all infinite repetitions in a D0L-system

Authors:	K. Klouda and Š. Starosta
Publication reference:	J. Discrete Algorithms, 33 (2015), pp. 130–138
Cited as:	[VII]

Journal of Discrete Algorithms 33 (2015) 130–138

Contents lists available at ScienceDirect

Journal of Discrete Algorithms

www.elsevier.com/locate/jda

An algorithm for enumerating all infinite repetitions in a D0L-system

Karel Klouda *, Štěpán Starosta

Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 13 December 2014
Received in revised form 17 February 2015
Accepted 20 March 2015
Available online 27 March 2015

Keywords:
D0L-system
Periodicity
Repetition

ABSTRACT

We describe a simple algorithm that finds all primitive words v such that v^k is a factor of the language of a given D0L-system for all k. It follows that the number of such words is finite. This polynomial-time algorithm can be also used to decide whether a D0L-system is repetitive.

1. Introduction

This work deals with repetitions occurring in words which are generated by a morphism. Repetitions in words were already studied by Thue [19] in 1906 and since then this topic has been addressed by many authors from many perspectives. To describe the words generated by a morphism, we use the concept of D0L-system (see [16]). A D0L-system is given by a morphism and a finite word; its language is the set of words which are generated by repeated application of this morphism to the given finite word. In particular, we focus on the case when a D0L-system is *repetitive*, meaning that for each positive k there is a word v such that its k-power v^k appears as a factor in the language. The other case, when the D0L-system is not repetitive, was profoundly studied by Krieger [10,11].

Ehrenfeucht and Rozenberg proved in [3] that a D0L-system is repetitive only if it is *strongly repetitive*, i.e., there is a word v such that v^k is a factor in the language for all $k \in \mathbb{N}$. Moreover, it follows from the work of Mignosi and Séébold in [14] that there is a constant M such that v^M being a factor in the language implies that v^k is a factor in the language for all $k \in \mathbb{N}$ (we prove a slightly more general version of this statement, see Lemma 18). Hence, to describe all k-powers occurring as factors in the language of a D0L-system for any k larger than M, we can limit ourselves to all finite words v having their ℓ-power v^ℓ in some element of the language for all positive ℓ. Our goal is to design an algorithm that returns all these words.

The first problem is to decide whether a given D0L-system is repetitive or not. Decidability of this problem was proved for the first time in [3] and then, using a different strategy, also in [14]. The algorithms are quite complicated and their complexity is unknown. Another algorithm working in polynomial time is given in [8] by Kobayashi and Otto. Their approach uses the notion of quasi-repetitive elements: a word v is a *quasi-repetitive element* for a morphism φ if there exist positive integers n and p with $p \neq 0$ such that $\varphi^n(v) = v_1^p$ where v_1 is a conjugate of v. They proved that repetitiveness is strongly related to the existence of quasi-repetitive factors (see Corollary 5.5 in [8]).

* Corresponding author.
E-mail address: karel.klouda@fit.cvut.cz (K. Klouda).

http://dx.doi.org/10.1016/j.jda.2015.03.006

All algorithms mentioned above treat separately pushy and non-pushy D0L-systems. A D0L-system is pushy if its language contains arbitrarily long factors over bounded letters (see the next section for definitions). Pushy D0L-systems are known to be repetitive due to [3], in fact it holds that long enough factor over bounded letters has a form of repetition except for some short prefix or suffix (see Theorem 12).

The case of words, whose k-power is a factor in the language for all k and that contain an unbounded letter, is more complicated. The algorithm describing all such words in the language of a D0L-system with a morphism φ is based on the fact that these words appear in the language if and only if there is a letter a (occurring in the language of the system) and a positive integer ℓ such that the fixed point of φ^ℓ starting in a is purely periodic (see Theorem 15). For given integer ℓ and letter a, this is easy to decide employing the algorithm introduced by Lando in [12]. The last ingredient for the algorithm is the notion of injective simplification used in [2] to deal with systems with non-injective morphisms.

As a consequence of this result, this non-surprising but so far not proven fact follows: the number of primitive factors v such that for all $k \in \mathbb{N}$ the word v^k is a factor in the language is finite for any D0L-system, see Corollary 20. This implies that the language of any D0L-system cannot be finitely factorized to palindromes, see [4] for a proof of the statement and other details.

As already mentioned, our algorithm needs to test whether φ^ℓ has a periodic fixed point for some ℓ. This question is completely solved in the binary case: in [17], a characterization of binary morphisms having purely periodic fixed point was given. The decidability of a more general problem whether φ^ℓ has an eventually periodic fixed point has been shown in [15] and [6]. This question is also solved for the binary case: complete characterization of morphisms over binary alphabet with eventually periodic fixed points was given in [9]. Recently, Honkala in [7] gave a simple algorithm to decide the problem over a general alphabet. However, the algorithm needs to compute a power of the morphism φ that is greater than $k!$ where k is the cardinality of the alphabet.

The paper is organized as follows: the next section contains needed notation and definitions. The third section contains the main results and their proofs. The last section contains the description of the algorithm.

2. Definitions and basic notions

An *alphabet* $\mathcal{A}$ is a finite set of *letters*. We denote by $\mathcal{A}^*$ the *free monoid* on $\mathcal{A}$. Its neutral element is the *empty word* and it is denoted ε. A subset of $\mathcal{A}^*$ is a *language* and its elements are *words*. We denote by $\mathcal{A}^+$ the set of all non-empty words. Let $v = v_0 \cdots v_{n-1} \in \mathcal{A}^*$ with $v_i \in \mathcal{A}$ for $0 \le i < n$. The *length* of v is n and is denoted by $|v|$. We denote by $\mathrm{first}(v)$ the first letter of the word $v \in \mathcal{A}^+$, i.e., here $\mathrm{first}(v) = v_0$. By repeating the word v k-times with $k \in \mathbb{N}$ we obtain the *k-power* of v denoted by $v^k = vv \cdots v$. Any infinite sequence of letters $\mathbf{u} = u_0u_1 \ldots$ is called an *infinite word* over $\mathcal{A}$. A word $v \in \mathcal{A}^*$ is a *factor* of a finite or infinite word u if there exist words x and y such that $u = xvy$; if x is empty (resp. y is empty), v is a *prefix* (resp. *suffix*) of u. A word w is a *conjugate* of $v \in \mathcal{A}^*$ if $w = yx$ and $v = xy$ for some $x, y \in \mathcal{A}^*$; the set of all conjugates of v is denoted by $[v]$. A word v is *primitive* if $v = z^k$ implies $k = 1$. The shortest word x such that $v = x^k$, $k \in \mathbb{N}^+$, is the *primitive root* of v. An infinite word $\mathbf{u}$ is *eventually periodic* if it is of the form $\mathbf{u} = xyyyy\ldots = xy^\omega$; it is *(purely) periodic* if x is empty and *aperiodic* if it is not eventually periodic.

Given two alphabets $\mathcal{A}$ and $\mathcal{B}$, any homomorphism φ from $\mathcal{A}^*$ to $\mathcal{B}^*$ is called a *morphism*. A non-empty word w is *mortal* with respect to a morphism φ if $\varphi^k(w) = \varepsilon$ for some k, otherwise it is *immortal*. A morphism φ over $\mathcal{A}$ is *non-erasing* if $\varphi(a)$ is non-empty for all $a \in \mathcal{A}$. An infinite word $\mathbf{u}$ is a *periodic point* of a morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ if $\varphi^\ell(\mathbf{u}) = \mathbf{u}$ for some $\ell \ge 1$. If $\mathbf{u}$ is a periodic point of φ starting with the letter a such that $\varphi^\ell(a) = av$, for some immortal word $v \in \mathcal{A}^+$ and $\ell \in \mathbb{N}$, then we have $\mathbf{u} = (\varphi^\ell)^\omega(a) = \lim_{k\to+\infty} \varphi^{k\ell}(a)$.

The triplet $G = (\mathcal{A}, \varphi, w)$, where $\mathcal{A}$ is an alphabet, φ a morphism on $\mathcal{A}^*$ and $w \in \mathcal{A}^+$, is a *D0L-system*. The *language* of G is the set $L(G) = \{\varphi^k(w) : k \in \mathbb{N}\}$. The system G is *reduced* if every letter of $\mathcal{A}$ occurs in some element of $L(G)$. In what follows, we implicitly suppose that we have a reduced system since if a system is not reduced, then we may consider a subset of the alphabet and a restriction of the morphism to get a reduced system with the same language. If the set $L(G)$ is finite, then G is *finite*. A letter $a \in \mathcal{A}$ is *bounded* if the D0L-system $(\mathcal{A}, \varphi, a)$ is finite, otherwise a is *unbounded*. The set of all bounded letters is denoted by $\mathcal{A}_0$. The D0L-system G is *non-erasing* (resp. *injective*) if the morphism φ is non-erasing (resp. injective). We denote by $S(L(G))$ the set of all factors of the elements of the set $L(G)$.

If for any $k \in \mathbb{N}^+$ there is a word v such that $v^k \in S(L(G))$, then G is *repetitive*; if there is a word v such that $v^k \in S(L(G))$ for all $k \in \mathbb{N}^+$, then G is *strongly repetitive*. By [3], all repetitive D0L-systems are strongly repetitive. An important class of repetitive D0L-systems are pushy D0L-systems: G is *pushy* if $S(L(G))$ contains infinitely many words over $\mathcal{A}_0$.

3. Infinite repetitions

As explained above, to describe all k-powers for large enough k that are factors in the language of a D0L-system G, it suffices to study finite words v such that $v^k \in S(L(G))$ for all $k \in \mathbb{N}$. Therefore we introduce the following notions.

Definition 1. Given a D0L-system G, we say that v^ω is an *infinite repetition* of G if v is a non-empty word and $v^k \in S(L(G))$ for all positive integers k.

We say that infinite repetitions v^ω and u^ω are *equivalent* if the primitive root of u is a conjugate of the primitive root of v. We denote the equivalence class containing v^ω by $[v]^\omega$.

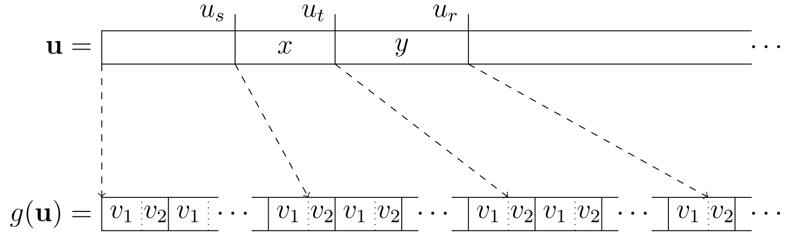

Fig. 1. The choice of prefixes u_0, u_1 and u_2.

3.1. Simplification

In what follows, some of the proofs are based on the assumption that the concerned D0L-system is injective. This does not mean any loss of generality of our results since any non-injective D0L-system is related to an injective one that has the same structure of infinite repetitions. This relation is given by simplification of morphisms used in [2].

Definition 2. Let $\mathcal{A}$ and $\mathcal{B}$ be two finite alphabets and let $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ and $\psi : \mathcal{B}^* \to \mathcal{B}^*$ be morphisms. We say that φ and ψ are *twined* if there exist morphisms $h : \mathcal{A}^* \to \mathcal{B}^*$ and $g : \mathcal{B}^* \to \mathcal{A}^*$ satisfying $g \circ h = \varphi$ and $h \circ g = \psi$. If $\#\mathcal{B} < \#\mathcal{A}$ and φ and ψ are twined, then ψ is a *simplification* (with respect to (g, h)) of φ.

If for D0L-system $G = (\mathcal{A}, \varphi, w)$ the morphism φ has no simplification, then G is called *elementary*. In [2], it is shown that elementary D0L-systems are injective and thus non-erasing.

Moreover, every non-injective morphism has a simplification that is injective. A simple algorithm to find an injective simplification of a morphism can be designed as follows. If the morphism is erasing, one can find a non-erasing simplification using Proposition 3.6 in [8]. Simplification of a non-erasing morphism is closely related to the defect theorem and one can adapt the algorithm described in [5] to find an injective simplification of a given non-erasing morphism.

Example 3. The morphism φ determined by $a \mapsto aca, b \mapsto badc, c \mapsto acab$ and $d \mapsto adc$ is not injective as $\varphi(ab) = \varphi(cd)$. Therefore, there must exist a simplification; let h be a morphism given by $a \mapsto x, b \mapsto yz, c \mapsto xy, d \mapsto z$ and g a morphism given by $x \mapsto aca$, $y \mapsto b, z \mapsto adc$. Since $\varphi = g \circ h$, φ is twined with the morphism $\psi = h \circ g$, which is determined by $x \mapsto xxyx$, $y \mapsto yz, z \mapsto xzxy$ and defined over the alphabet $\{x, y, z\}$. One can easily check that ψ has no simplification and hence it is elementary, and thus injective.

Let $G = (\mathcal{A}, \varphi, w)$ be a non-elementary D0L-system and let $\psi : \mathcal{B}^* \to \mathcal{B}^*$ be an injective simplification of φ with respect to (g, h). We say that the D0L-system $(\mathcal{B}, \psi, h(w))$ is an *injective simplification of G (with respect to* (g, h)*)*. The role of injective simplifications in the study of repetitions is given by the following lemma and its corollary.

Lemma 4. *Let $g : \mathcal{B}^* \to \mathcal{A}^*$ be an injective morphism. If $\mathbf{u}$ is an infinite word over $\mathcal{B}$ such that $g(\mathbf{u})$ is periodic, then $\mathbf{u}$ is eventually periodic.*

Proof. Let $g(\mathbf{u}) = v^\omega$ for some non-empty v. To obtain a contradiction suppose that $\mathbf{u}$ is not eventually periodic, i.e., is aperiodic. By the pigeonhole principle there exist words v_1 and v_2 such that $v = v_1 v_2$ and for all $i \in \mathbb{N}$ there exists a prefix u_i of $\mathbf{u}$ such that $|u_i| < |u_{i+1}|$ for all i and $g(u_i) = v^{k_i} v_1$ for some integers k_i.

Let $s, r, t \in \mathbb{N}$ be such that $s < t < r$. Denote $u_s x = u_t$ and $u_t y = u_r$. The situation is demonstrated in Fig. 1.

For any $w_1, w_2 \in \mathcal{A}^*$ the equality $w_1 w_2 = w_2 w_1$ implies that there exists a word z such that $w_1 = z^m$ and $w_2 = z^n$ for some integers m and n (see, e.g., Proposition 1.3.2 in [13]). Since $\mathbf{u}$ is aperiodic, this fact implies that we can choose the integers s, r and t such that $xy \neq yx$. We have $g(x) = v_2 v^k v_1$ and $g(y) = v_2 v^\ell v_1$ for some integers k and ℓ. This implies $g(xy) = g(yx)$ which is a contradiction. $\square$

The following corollary can be also deduced from Proposition 3.8 of [8]. We reformulate it here as needed later and give a short proof using our terminology.

Corollary 5. *Let $G = (\mathcal{A}, \varphi, w)$ be a non-elementary D0L-system and $G' = (\mathcal{B}, \psi, h(w))$ its injective simplification with respect to (g, h). Then for any infinite repetition v^ω of G there is an infinite repetition z^ω of G' such that $g(z)^\omega \in [v]^\omega$.*

Proof. Since $\psi = h \circ g$ is injective, the morphism g is injective as well. Using the definition of simplification, we obtain $\varphi^k(w) = (g \circ h)^k(w) = g(\psi^{k-1}(h(w)))$ for all $k > 0$, i.e., the language of G is the image by g of the language of G' (except for the word w). Therefore, there is an infinite word $\mathbf{u}$ such that all its prefixes belong to $S(L(G'))$ with $g(\mathbf{u}) = sv^\omega$ for some s suffix of v. By Lemma 4, the word $\mathbf{u}$ is eventually periodic. Let $\mathbf{u} = xz^\omega$. It follows that $g(z)^\omega$ is in $[v]^\omega$. $\square$

The last claim allows us to restrict our following considerations to injective D0L-systems.

3.2. Graph of infinite repetitions

If v^ω is an infinite repetition of $G = (\mathcal{A}, \varphi, w)$, then $(\varphi(v))^\omega$ is an infinite repetition, too. This gives us a structure that can be captured as a graph.

Definition 6. Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. The *graph of infinite repetitions* of G, denoted P_G, is a directed graph with loops allowed and defined as follows:

1. the set of vertices of P_G is the set
$$V(\mathrm{P}_G) = \left\{[v]^\omega : v^\omega \text{ is an infinite repetition of } G\right\};$$
2. there is a directed edge from $[v]^\omega$ to $[z]^\omega$ if and only if $\varphi(v)^\omega \in [z]^\omega$.

Obviously, the outdegree of any vertex of P_G is equal to one.

Lemma 7. *If $G = (\mathcal{A}, \varphi, w)$ is an injective D0L-system, then any vertex $[v]^\omega \in \mathrm{P}_G$ has indegree at least* 1.

Proof. Let $[v]^\omega \in V(\mathrm{P}_G)$. There exists an infinite word $\mathbf{u}$ such that all its prefixes belong to $S(L(G))$ and $\varphi(\mathbf{u}) = sv^\omega$ for some s suffix of v. By Lemma 4, the word $\mathbf{u}$ is eventually periodic, i.e., $\mathbf{u} = xz^\omega$ for some words x and z. It follows that $\varphi(z)^\omega$ is in $[v]^\omega$ and there is a directed edge in P_G from $[z]^\omega$ to $[v]^\omega$. $\square$

We have that the indegree of any vertex is at least one and the outdegree is equal to 1. It follows that every vertex is a vertex of a cycle or there is an infinite path in P_G that ends in this vertex.

In what follows, we distinguish two types of vertices based on whether the infinite repetition assigned to the vertex contains an unbounded letter or not. The next lemma states that this property is in fact a property of a component. In what follows, by a component of a directed graph we always mean a weakly connected component.

Lemma 8. *Let $G = (\mathcal{A}, \varphi, w)$ be an injective D0L-system. If $[v]^\omega$ and $[w]^\omega$ are two vertices of the same component in P_G, then v consists of bounded letters if and only if w consists of bounded letters.*

Proof. If $v \in \mathcal{A}_0^+$, then $\varphi^k(v) \in \mathcal{A}_0^+$ for all $k \in \mathbb{N}^+$. Similarly, if v contains an unbounded letter, then $\varphi^k(v)$ contains an unbounded letter as well for all $k \in \mathbb{N}^+$. $\square$

We first treat the case of components over bounded letters.

3.3. Components of P_G over bounded letters

If v^ω is an infinite repetition of a D0L-system G with v containing no unbounded letter, i.e. $v \in \mathcal{A}_0^*$, then G is pushy by definition. In [3], it is proved that it is decidable whether a D0L-system is pushy or not. In particular, the authors proved that a D0L-system is pushy if and only if it satisfies the *edge condition*: there exist $a \in \mathcal{A}$, $k \in \mathbb{N}^+$, $v \in \mathcal{A}^*$ and $u \in \mathcal{A}_0^+$ such that $\varphi^k(a) = vau$ or $\varphi^k(a) = uav$. We expand the idea and show how to describe all factors over $\mathcal{A}_0$.

Definition 9. Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. *The graph of unbounded letters of G to the right*, denoted UR_G, is the labeled directed graph defined as follows:

(i) the set of vertices is $V(\mathrm{UR}_G) = \mathcal{A} \setminus \mathcal{A}_0$,
(ii) there is a directed edge from a to b with the label $u \in \mathcal{A}_0^*$ if there exists $v \in \mathcal{A}^*$ such that $\varphi(a) = vbu$.

The graph of unbounded letters of G to the left UL_G is defined analogously: the only difference is that the roles of v and u in the definition of a directed edge are switched.

Clearly, the edge condition is satisfied if and only if one of the graphs UL_G and UR_G contains a cycle including an edge with an immortal label. Therefore, we have this:

Proposition 10. A D0L-system G is pushy if and only if one of the graphs UL_G and UR_G contains a cycle including an edge with an immortal label.

The following lemma will be used to detect infinite repetitions over bounded letters.

Lemma 11. *Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. If $u \in \mathcal{A}_0^+$ is immortal and $k \in \mathbb{N}^+$, then the infinite word*

$$\mathbf{u} = u\varphi^k(u)\varphi^{2k}(u)\varphi^{3k}(u)\dots$$

is eventually periodic.

Proof. Since u contains no unbounded letter, the sequence $(\varphi^j(u))_{j=0}^{+\infty}$ is eventually periodic. Let s and t be numbers such that $\varphi^s(u) = \varphi^{s+t}(u)$ with $t > 0$. Define ℓ_0 and ℓ_1 so that $\ell_0 k \geq s$ and $(\ell_1 - \ell_0)k$ is a multiple of t. We get that

$$\mathbf{u} = u\varphi^k(u)\cdots\varphi^{\ell_0 k}(u)\big(\varphi^{(\ell_0+1)k}(u)\varphi^{(\ell_0+2)k}(u)\cdots\varphi^{\ell_1 k}(u)\big)^\omega. \qquad \square$$

Since the outdegree of any vertex of the graphs UL_G and UR_G is 1 and the graphs are finite, they consist of components each containing exactly one cycle. Assume that a is a vertex of a cycle of UR_G that contains at least one edge with a non-empty immortal label, i.e., there exist $k \in \mathbb{N}^+$, $u_1, \dots, u_k \in \mathcal{A}_0^*$ and $a_1, \dots, a_{k-1} \in \mathcal{A} \setminus \mathcal{A}_0$ so that

$$a \xrightarrow{u_1} a_1 \xrightarrow{u_2} \cdots \xrightarrow{u_{k-1}} a_{k-1} \xrightarrow{u_k} a$$

is a cycle of UR_G with u_j being immortal for some j with $1 \leq j \leq k$. Let $u = u_k\varphi(u_{k-1})\cdots\varphi^{k-1}(u_1)$. For all $\ell \in \mathbb{N}^+$ the word $\varphi^{\ell k}(a)$ has a suffix

$$u\varphi^k(u)\varphi^{2k}(u)\cdots\varphi^{(\ell-1)k}(u) \in \mathcal{A}_0^+ .$$

Since $u \in \mathcal{A}_0^+$ is immortal, using Lemma 11, we obtain that the infinite word

$$u\varphi^k(u)\varphi^{2k}(u)\varphi^{3k}(u)\dots$$

is eventually periodic and thus its (infinite) suffix is an infinite repetition of G (over bounded letters).

Let FR_G be the set of all unbounded letters that are vertices in UR_G of components that contain a cycle with at least one edge having non-empty immortal label. Let Q_R be the mapping defined over FR_G such that for $b \in \mathrm{FR}_G$ we have $Q_R(b) = [v]^\omega$ where v^ω is the infinite repetition generated by the cycle in UR_G in the component of b.

Since the situation is analogous for the graph UL_G, the set FL_G and the mapping Q_L are defined analogously.

As the cycles of the graphs UR_G and UL_G are the only sources of factors over $\mathcal{A}_0$ of arbitrary length, we obtain the following theorem. This result has been already obtained in [1], Proposition 4.7.62, where one can find a distinct proof.

Theorem 12. *If $G = (\mathcal{A}, \varphi, w)$ is a pushy D0L-system, then there exist $L \in \mathbb{N}$ and a finite set $\mathcal{U}$ of non-empty words over $\mathcal{A}_0$ such that any factor from $\mathcal{S}(L(G)) \cap \mathcal{A}_0^+$ is of one of the following three forms:*

(i) w_1,
(ii) $w_1 u_1^{k_1} w_2$,
(iii) $w_1 u_1^{k_1} w_2 u_2^{k_2} w_3$,

where $u_1, u_2 \in \mathcal{U}$, $|w_j| < L$ for all $j \in \{1,2,3\}$, and $k_1, k_2 \in \mathbb{N}^+$.

Proof. We define several constants that we later use to specify L. To ease the notation, denote $\mathcal{A}_1 = \mathcal{A} \setminus \mathcal{A}_0$, i.e., the set of unbounded letters.

First, let S_I be the following subset of factors:

$$S_I = \big\{u : u \in \mathcal{A}_0^*, u \text{ is a factor of } w \text{ or } \varphi(c) \text{ for some } c \in \mathcal{A}_1\big\}.$$

The first constant is

$$C_I = \max\big\{|\varphi^j(u)| : j \in \mathbb{N}, u \in S_I\big\}.$$

Set

$$C_R = \max\big\{|x| : x \in \mathcal{A}_0^*, xey = \varphi^j(d) \text{ for some } j \in \mathbb{N}, e \in \mathcal{A}_1, d \in \mathcal{A}_1 \setminus \mathrm{FR}_G\big\}.$$

The constant C_L is defined analogously.

Let $d \in \mathrm{FR}_G$. Let z_d be a primitive word such that $z_d^\omega \in Q_R(d)$. By Lemma 11 we have that $\varphi^j(d) = p_j e_j x_j z_d^{i_j} y_j$ where $e_j \in \mathcal{A}_1$, $x_j, y_j \in \mathcal{A}_0^*$, $z_j \in [z_d]$, $i_j \in \mathbb{N}$, $p_j \in \mathcal{A}^*$. Moreover, suppose that i_j is maximal possible, i.e., x_j and y_j are taken the shortest possible. There exists a constant $K_R(d)$ such that $|x_j| < K_R(d)$ and $|y_j| < K_R(d)$ for all $j \in \mathbb{N}$. Set $K_R = \max\{K_R(d) : d \in \mathrm{FR}_G\}$. The constant K_L is defined analogously.

Set $L = C_R + K_R + C_I + C_L + K_L$.

Consider a factor $v = aub$ of $\varphi^n(w)$ such that $a, b \in \mathcal{A}_1$ and $u \in \mathcal{A}_0^+$. Let m be the maximal integer such that $m \le n$ and v is a factor of $\varphi^m(a_m u_m b_m)$ where $a_m, b_m \in \mathcal{A}_1$ and $u_m \in \mathcal{A}_0^*$ and v is not a factor of $\varphi^m(a_m u_m)$ nor of $\varphi^m(u_m b_m)$. Note that if u_m is not empty, then $a_m u_m b_m$ is a factor of the word w, i.e., $m = n$, or $a_m u_m b_m$ is factor of $\varphi(c)$ for some $c \in \mathcal{A}_1$. In both cases, $|\varphi^j(u_m)| \le C_I$ for all $j \in \mathbb{N}$.

If $a_m \notin \mathrm{FR_G}$ and $b_m \notin \mathrm{FL_G}$, then $|u| \le C_R + C_I + C_L$.

If $a_m \in \mathrm{FR_G}$ and $b_m \notin \mathrm{FL_G}$, then $v = axz^i yb$ with $|x| \le K_R$ and $|y| \le K_R + C_I + C_L$.

If $a_m \notin \mathrm{FR_G}$ and $b_m \in \mathrm{FL_G}$, then $v = axz^i yb$ with $|x| \le C_R + C_I + K_L$ and $|y| \le K_L$.

If $a_m \in \mathrm{FR_G}$ and $b_m \in \mathrm{FL_G}$, then $v = axz_1^i s z_2^j yb$ with $|x| \le K_R$, $|s| \le K_R + C_I + K_L$ and $|y| \le K_L$.

Since in all the cases, the words z, z_1 and z_2 can be chosen to be primitive and from $[w]$ such that $[w]^\omega \in Q_R(\mathrm{FR_G})$ or $[w]^\omega \in Q_L(\mathrm{FL_G})$, the finiteness of $\mathcal{U}$ follows. The special cases when $v = ub$ is a prefix of $\varphi^n(w)$ or $v = au$ is a suffix of $\varphi^n(w)$, where $a, b \in \mathcal{A}_1$ and $u \in \mathcal{A}_0^+$ can be treated in an analogous way. □

Example 13. Consider the D0L-system $G = (\mathcal{A}, \varphi, 0)$ with $\mathcal{A} = \{0, 1, 2\}$ and φ determined by $0 \mapsto 012$, $1 \mapsto 2$ and $2 \mapsto 1$. There are two bounded letters, $\mathcal{A}_0 = \{1, 2\}$, and one unbounded letter. The graphs UR_G and UL_G both contain one loop on the only vertex 0 labeled with 12 and the empty word, respectively. Therefore, we have $\mathrm{FR_G} = \{0\}$ and $\mathrm{FL_G} = \emptyset$. Thus, G is pushy and the class of infinite repetitions over bounded letters $Q_R(0)$ is generated by the label 12 as follows:

$$12\varphi(12)\varphi^2(12)\varphi^3(12)\ldots.$$

Namely, we obtain $Q_R(0) = [1221]^\omega$. In fact, the infinite repetition $(1221)^\omega$ is a suffix of the eventually periodic fixed point of φ starting in 0.

The D0L-system $H = (\mathcal{B}, \psi, 0)$ with $\mathcal{B} = \{0, 1, 2, 3\}$ and ψ determined by $0 \mapsto 0123$, $1 \mapsto 2$, $2 \mapsto 1$ and $3 \mapsto 123$ is also pushy. The graph UL_H contains two loops: one on the vertex 0 labeled with the empty word and one on the vertex 3 labeled with 12; thus we have $\mathrm{FL_H} = \{3\}$. The labels in UR_H are all equal to the empty word, thus $\mathrm{FR_H} = \emptyset$. Hence, there is a class of infinite repetitions over bounded letters $Q_R(3)$ and it is generated by the label 12 in the following way: the word

$$\varphi^n(12)\cdots\varphi^3(12)\varphi^2(12)\varphi(12)12$$

is a factor of all elements of $Q_R(3)$ for all n. We conclude that $Q_R(3) = [2112]^\omega$. In this case, the fixed point of ψ is not eventually periodic, but still all prefixes of $(2112)^\omega$ are its factors.

3.4. Components of P_G containing an unbounded letter

Now we address the other case: infinite repetitions containing an unbounded letter. Before proving the main result we need the following auxiliary lemma. Let us define the *graph of first letters* for a morphism φ over $\mathcal{A}$: the set of vertices is equal to $\mathcal{A}$ and there is a directed edge from a to b if $b = \mathrm{first}(\varphi(a))$.

Lemma 14. *Given an injective repetitive D0L-system $G = (\mathcal{A}, \varphi, w)$. If $[v]^\omega$ is a vertex of P_G such that v contains an unbounded letter, then v contains an unbounded letter $b \in \mathcal{A}$ such that b is a vertex of a cycle in the graph of first letters.*

Proof. Let v^ω be an infinite repetition of G such that v contains an unbounded letter. Set $L = \#\mathcal{A}(|v| + 1)$. According to Lemma 7, each vertex of P_G has its indegree at least one. Thus, there is an arbitrarily long walk ending in $[v]^\omega$. It follows that there is a vertex $[u]^\omega$ such that $(\varphi^L(u))^\omega \in [v]^\omega$.

Since v contains an unbounded letter, so does the word u. Let a_0 be the unbounded letter in u which occurs the first in u, i.e., $u = u_0 a_0 x_0$ for $u_0 \in \mathcal{A}_0^*$ and $x_0 \in \mathcal{A}^*$. Let $\varphi(a_i) = u_{i+1} a_{i+1} x_{i+1}$ for $i \ge 0$ where $u_{i+1} \in \mathcal{A}_0^*$, a_{i+1} is an unbounded letter and $x_{i+1} \in \mathcal{A}^*$. There exist positive integers s and t such that $a_s = a_{s+t}$ with $s < \#\mathcal{A}$ and $t \le \#\mathcal{A}$. Suppose that $u_j \ne \varepsilon$ for $s < j \le t$. Due to the choice of L, it follows that the word $\varphi^L(u)$ contains a prefix over $\mathcal{A}_0$ longer than $|v|$ which is a contradiction. Therefore, $u_{s+1} = \cdots = u_{s+t} = \varepsilon$. It follows that for all i greater than s the letter a_i is a vertex of the cycle in the graph of first letters. As v contains a_L, the claim is proven. □

Using the last lemma we can prove the main result of this paper saying that all infinite repetitions containing an unbounded letter correspond to periodic points of the morphism.

Theorem 15. *Let $G = (\mathcal{A}, \varphi, w)$ be an injective repetitive D0L-system. If v^ω is an infinite repetition of G such that v contains an unbounded letter, then there exist $b \in \mathcal{A}$ and $\ell \in \mathbb{N}$, $1 \le \ell \le \#\mathcal{A}$, such that $(\varphi^\ell)^\omega(b) = z^\omega$ for some $z \in [v]$.*

Proof. We can assume that v is primitive. Let ℓ be the least common multiple of all lengths of cycles in the graph of first letters. Let L be a multiple of ℓ such that $|\varphi(x)^L| > 2|v|$ for all unbounded letters x. There is a vertex $[u]^\omega$ such that $(\varphi^{L+\ell}(u))^\omega \in [v]^\omega$. Lemma 14 implies that there is an unbounded letter b in u that is a vertex of a cycle in the graph of first letters.

Due to the choice of ℓ and L, we obtain $\varphi^{L+\ell}(b) = z^k z'$ where $z \in [v]$, z starts with the letter b, $k \geq 2$ and z' is a prefix of z shorter than z.

Since $\varphi^L(b)$ is a prefix of $\varphi^{L+\ell}(b) = z^k z'$ and due to the choice of L, the word zz is a prefix of $\varphi^L(b)$. It follows that $\varphi^\ell(z)$ equals $z^j z''$ where $j \geq 1$ and z'' is a prefix of z shorter than z.

Suppose that z'' is non-empty. Let x be the suffix of z such that $z = z''x$. Since $\varphi^\ell(zz) = (z''x)^j z''(z''x)^j z''$ is also a prefix of $z^\omega = (z''x)^\omega$, we obtain $z = z''x = xz''$. It implies (using Proposition 1.3.2 in [13]) that z is not primitive which is a contradiction.

Hence, z'' is empty and we have $\varphi^\ell(z) = z^j$ and therefore $\varphi^{m\ell}(b)$ is a prefix of z^ω for all m. □

Example 16. Let $G = (\{0, 1, 2\}, \varphi, 1)$ with $\varphi : 0 \mapsto 12, 1 \mapsto 0$ and $2 \mapsto 120$. One readily verifies that φ has exactly two periodic points, one starting in 0 and one starting in 1. We obtain $\lim_{n\to+\infty} \varphi^{2n}(0) = (012)^\omega$ and $\lim_{n\to+\infty} \varphi^{2n}(1) = (120)^\omega$. We conclude that all infinite repetitions containing an unbounded letter are elements of $[012]^\omega$.

Let us also give a non-trivial example which covers both types of repetitions.

Example 17. Let $G = (\{0, 1, \ldots, 5\}, \varphi, 0)$ where φ is determined by

$$0 \mapsto 1230,$$
$$1 \mapsto 11,$$
$$2 \mapsto 1245,$$
$$3 \mapsto 4453,$$
$$4 \mapsto 5,$$
$$5 \mapsto 4.$$

We have $\mathcal{A}_0 = \{4, 5\}$. The morphism φ admits one periodic point containing an unbounded letter, namely the word 1^ω. By Theorem 15 it is the only infinite repetition containing an unbounded letter. We have $\mathrm{FR_G} = \{2\}$ and $\mathrm{FL_G} = \{3\}$. These two letters generate the infinite repetitions $(4554)^\omega$ and $(554445)^\omega$, respectively. We can also conclude that there are no other infinite repetitions in G (up to equivalency).

3.5. *Infinite repetitions*

As mentioned in the introduction, to be sure that infinite repetitions cover all k-powers for k arbitrarily large, we have the following lemma, which is a slight generalization of Proposition 1 of [14] where the authors consider only non-pushy systems.

Lemma 18. *Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. There is a constant M such that v^M being in $S(L(G))$ implies that v^k is a factor in the language for all $k \in \mathbb{N}$.*

Proof. If G is non-pushy, then existence of M follows from [14].

Suppose G is pushy. We can suppose that v is primitive.

Suppose $v \in \mathcal{A}_0^*$. Let $\mathcal{U}$ and L be the set and integer given by Theorem 12. Set $M_1 = 2L + 2\max\{|u| : u \in \mathcal{U}\}$. Theorem 12 implies that if $v^{M_1} \in S(L(G))$, then $v^k \in S(L(G))$ for all k.

Suppose that v contains an unbounded letter a. Since for all $\ell \in \mathbb{N}$ there is a letter $b \in \mathcal{A} \setminus \mathcal{A}_0$ such that $\varphi^\ell(v)$ contains b, we can repeat the proof of Proposition 1 of [14] with the constant $M_2 = 2|w|(2\|\varphi\|)^{\#\mathcal{A}}$ with $\|\varphi\| = \max\{|\varphi(x)| : x \in \mathcal{A}\}$. (The constant is denoted $\underline{n}$ in [14].)

Finally, $M = \max\{M_1, M_2\}$. □

Corollary 19. *Let $G = (\mathcal{A}, \varphi, w)$ be a D0L-system. There exists a constant M, a finite set of primitive words $\mathcal{F}$ such that $v^k \in S(L(G))$ for every $k \geq M$ and $v \in \mathcal{F}$.*

Proof. Let M be the constant given by Lemma 18. Let $\mathcal{F}$ be the set of primitive factors of $S(L(G))$ such that $v^M \in S(L(G))$. According to Lemma 18, it remains to show that $\mathcal{F}$ is finite.

Suppose that G is injective. Theorem 12 implies that $\mathcal{F} \cap \mathcal{A}_0^*$ is finite. The finiteness of $\mathcal{F} \setminus \mathcal{A}_0^*$ follows directly from Theorem 15 since the morphism φ has at most $\#\mathcal{A}$ periodic points.

Suppose that G is not injective. Let G' be its injective simplification. According to the claim for injective D0L-systems, there is a constant M' and a finite set $\mathcal{F}'$ of primitive words such that such that $v^k \in S(L(G'))$ for every $k \geq M'$ and $v \in \mathcal{F}'$. The finiteness of $\mathcal{F}$ follows from Corollary 5. □

The last corollary can be restated in the terms of infinite repetitions.

Input : Alphabet $\mathcal{A}$, morphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$.
Output: All primitive v such that v^k can be generated by φ for all k.

$\varphi \leftarrow$ injective simplification of φ with respect to (h, k);
get the list of bounded letters $\mathcal{A}_0$ and the graphs UL_G and UR_G;
OutputList $\leftarrow$ empty list;
foreach *cycle* c *in* UL_G **do**
  if c *contains an edge with non-empty immortal label* **then**
    u $\leftarrow u_k \varphi(u_{k-1}) \cdots \varphi^{k-1}(u_1)$ where $u_1, \ldots, u_k$ are the labels of edges in the cycle c;
    find the least s and t such that t > 0 and $\varphi^{\mathsf{s}}(\mathsf{u}) = \varphi^{\mathsf{s}+\mathsf{t}}(\mathsf{u})$;
    $\mathsf{l0} \leftarrow \lceil \mathsf{s}/k \rceil \cdot k$;
    $\mathsf{l1} \leftarrow \mathsf{l0} + \mathtt{LeastCommonMultiple}(\mathsf{t}, k)/k$;
    append `PrimitiveRoot` $(\varphi^{(\mathsf{l0}+1)k}(\mathsf{u})\varphi^{(\mathsf{l0}+2)k}(\mathsf{u}) \cdots \varphi^{\mathsf{l1}k}(\mathsf{u}))$ to OutputList;
  end
end
// do analogous procedure for each cycle in UR_G as for cycles in UL_G **foreach** *letter* a *in* $\mathcal{A} \setminus \mathcal{A}_0$ **do**
  find the least l such that $\mathtt{FirstLetter}(\varphi^{\mathsf{l}}(\mathsf{a})) = \mathsf{a}$ with $\mathsf{l} \le \#\mathcal{A}$;
  if l *exists* **then**
    find the least s such that $\varphi^{\mathsf{l}\cdot\mathsf{s}}(\mathsf{a})$ contains at least two occurrences of one unbounded letter;
    if $\varphi^{\mathsf{l}\cdot\mathsf{s}}(\mathsf{a})$ *contains at least two occurrences of* a **then**
      v $\leftarrow$ the longest prefix of $\varphi^{\mathsf{l}\cdot\mathsf{s}}(\mathsf{a})$ containing only one occurrence of a;
      if $\varphi^{\mathsf{l}}(\mathsf{v}) = \mathsf{v}^m$ *for some integer* $m \ge 2$ **then**
        append v to OutputList;
      end
    end
  end
end
OutputList $\leftarrow \{\mathtt{PrimitiveRoot}(\mathsf{k}(w)) : w \in \text{OutputList}\}$;
add conjugates to OutputList;
`return` OutputList;

Algorithm 1: Pseudocode for the main algorithm.

Corollary 20. *Any repetitive D0L-system G contains a finite number of primitive words v such that v^ω is an infinite repetition of G. Moreover, the graph of infinite repetitions consists only of cycles.*

4. Algorithm

Given an injective D0L-system $G = (\mathcal{A}, \varphi, w)$, Corollary 20 states that the number of infinite repetitions of G is finite. In this section, we give a simple algorithm that outputs all primitive words which generate all the infinite repetitions of G.

All infinite repetitions over bounded letters can be obtained from the cycles in graphs UL_G and UR_G. It remains unclear how to find infinite repetitions containing an unbounded letter. Theorem 15 says that equivalence classes $[v]^\omega$ that contain those infinite repetitions are in one-to-one correspondence with purely periodic[1] points of the morphism φ. It holds that $(\varphi^\ell)^\omega(a)$ is an infinite periodic point of φ if and only if $a \in \mathcal{A} \setminus \mathcal{A}_0$ is a vertex of a cycle of the graph of first letters (see above) of a length that divides ℓ. Therefore, we have only finitely many candidates $a \in \mathcal{A}$ and $\ell \in \mathbb{N}^+$ for which we need to verify whether $(\varphi^\ell)^\omega(a)$ is a periodic infinite word. In fact, it suffices to check only one letter from each cycle of the graph of first letters. Finally, to verify whether the word $(\varphi^\ell)^\omega(a)$ is periodic we can use the algorithm described in [12]. This algorithm is very effective, in short it works as follows:

1. Find the least $k \le \#\mathcal{A}$ such that $(\varphi^\ell)^k(a)$ contains at least two occurrences of one unbounded letter.
2. If $(\varphi^\ell)^k(a)$ does not contain two occurrences of a, then $(\varphi^\ell)^\omega(a)$ is not periodic. Otherwise denote by v the longest prefix of $(\varphi^\ell)^k(a)$ containing a only as the first letter.
3. Now, $(\varphi^\ell)^\omega(a)$ is periodic if and only if $\varphi^\ell(v) = v^m$ for some integer $m \ge 2$.

This, together with the algorithm to construct an injective simplification and Corollary 5, gives us an effective algorithm that decides whether a given D0L-system is repetitive and, moreover, returns the list of all infinite repetitions.

The described algorithm is given in pseudocode in Algorithm 1.

Acknowledgement

We acknowledge financial support by the Czech Science Foundation grant GAČR 13-35273P. The algorithm was implemented and tested in the open-source mathematical software `Sage` [18].

[1] Note that "periodic periodic" is not a typing error: it is a periodic point of a morphism, which happens to be purely periodic, i.e., of the form w^ω.

References

[1] J. Cassaigne, F. Nicolas, Factor complexity, in: V. Berthé, M. Rigo (Eds.), Combinatorics, Automata and Number Theory, in: Encycl. Math. Appl., vol. 135, Cambridge Univ. Press, 2010, pp. 163–239.
[2] A. Ehrenfeucht, G. Rozenberg, Elementary homomorphisms and a solution of the D0L sequence equivalence problem, Theor. Comput. Sci. 7 (1978) 169–183.
[3] A. Ehrenfeucht, G. Rozenberg, Repetition of subwords in D0L languages, Inf. Comput. 53 (1983) 13–35.
[4] A.E. Frid, S. Puzynina, L. Zamboni, On minimal factorizations of words as products of palindromes, http://arxiv.org/abs/1210.6179, 2012.
[5] T. Harju, J. Karhumäki, On the defect theorem and simplifiability, Semigroup Forum 33 (1986) 199–217.
[6] T. Harju, M. Linna, On the periodicity of morphisms on free monoids, RAIRO Theor. Inform. Appl. 20 (1986) 47–54.
[7] J. Honkala, Cancellation and periodicity properties of iterated morphisms, Theor. Comput. Sci. 391 (2008) 61–64.
[8] Y. Kobayashi, F. Otto, Repetitiveness of languages generated by morphisms, Theor. Comput. Sci. 240 (2000) 337–378.
[9] Y. Kobayashi, F. Otto, P. Séébold, A complete characterization of repetitive morphisms over the two-letter alphabet, in: T. Jiang, D. Lee (Eds.), Computing and Combinatorics, in: Lect. Notes Comput. Sci., vol. 1276, Springer, Berlin, Heidelberg, 1997, pp. 393–402.
[10] D. Krieger, On critical exponents in fixed points of non-erasing morphisms, Theor. Comput. Sci. 376 (2007) 70–88.
[11] D. Krieger, On critical exponents in fixed points of k-uniform binary morphisms, RAIRO Theor. Inform. Appl. 43 (2009) 41–68.
[12] B. Lando, Periodicity and ultimate periodicity of D0L systems, Theor. Comput. Sci. 82 (1991) 19–33.
[13] M. Lothaire, Combinatorics on Words, Encycl. Math. Appl., vol. 17, Addison-Wesley, Reading, Mass, 1983. Reprinted in the Cambridge Mathematical Library, Cambridge University Press, Cambridge UK, 1997.
[14] F. Mignosi, P. Séébold, If a D0L language is k-power free then it is circular, in: ICALP '93: Proceedings of the 20th International Colloquium on Automata, Languages and Programming, Springer-Verlag, London, UK, 1993, pp. 507–518.
[15] J.J. Pansiot, Decidability of periodicity for infinite words, RAIRO Theor. Inform. Appl. 20 (1986) 43–46.
[16] G. Rozenberg, A. Salomaa, The Mathematical Theory of L Systems, Academic Press, 1980.
[17] P. Séébold, An effective solution to the DOL periodicity problem in the binary case, Bull. Eur. Assoc. Theor. Comput. Sci. (1988) 137–151.
[18] W.A. Stein, et al., Sage Mathematics Software (Version 5), The Sage Development Team http://www.sagemath.org, 2013.
[19] A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr. I Mat. Nat. Kl. Chris. 7 (1906) 1–22.

Chapter 9

Characterization of circular D0L-systems

Authors:	K. Klouda and Š. Starosta
Publication reference:	submitted to Discrete Math., 2015
Cited as:	[VIII]

Characterization of circular D0L-systems

Karel Klouda[a], Štěpán Starosta[a]

[a]*Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00, Prague, Czech Republic*

Abstract

We give a characterization of circularity of a D0L-system. The characterizing condition is simple to verify and yields an efficient algorithm. To derive it, we prove that every non-circular D0L-system contains arbitrarily long repetitions. This result was already published in 1993 by Mignosi and Séébold, however their proof is only a sketch. We give a complete proof that, in addition, is valid for a slightly relaxed definition of circularity, called weak circularity.

Keywords: D0L-system, circular D0L-system, repetition
2000 MSC: 68R15

1. Introduction

An L-system consists of an alphabet, production rules and an initial word. The production rules are used to construct a language from the initial word. L-systems were primarily conceived as a model for plant growth, see [10] for an overview. For basic results on L-systems, namely the used mathematical tools, see [11].

In this paper we focus on a subclass of L-systems, so-called D0L-systems. In the perspective of general L-systems, D0L-systems (deterministic Lindenmayer system with context of length 0) can be seen as the simplest case of an L-system. We focus on repetitions in languages generated by such systems and their relation to circularity (see below for a definition). The study of repetitions dates back to the work of Axel Thue. In [3], the authors show that it is decidable whether a D0L language is k-power free, i.e., does not contain a k successive repetitions of the same word for some $k \in \mathbb{N}$. In [8], the authors show that if a PD0L language is k-power free for some integer k, then it is circular. Since this is a key result for our purposes and the authors give mostly only sketches of proofs, we give a sound proof of this claim. Moreover, we generalize the result as we prove it for non-injective PD0L-systems and slightly relaxed definition of circularity, called weak circularity. We use this result and results of [5, 4] to deduce a characterization of circular D0L-systems which is simple enough to verify so that an effective test can be designed.

The paper is organized as follows. The next section contains notation and basic definitions. Section 3 is dedicated to the definitions of circularity. Section 4 contains proofs and in the last section we deduce a characterization of circularity.

Preprint submitted to Elsevier *July 26, 2016*

2. Preliminaries

Let $\mathcal{A}$ be an *alphabet*: a finite set of *letters*. A finite sequence of letters from $\mathcal{A}$ is a *finite word*. The free monoid over $\mathcal{A}$, denoted $\mathcal{A}^*$, is the set of all finite words over $\mathcal{A}$ endowed with concatenation. The *empty word* is denoted ε. The set of all non-empty words over $\mathcal{A}$ is denoted $\mathcal{A}^+$. The *length* of $w \in \mathcal{A}^*$ is denoted $|w|$. Given a word $w \in \mathcal{A}^*$, we say that $u \in \mathcal{A}^*$ is a *factor* of w if there exist words p and s, possibly empty, such that $w = pus$. Such a word p is a *prefix* of w, and the word s is a *suffix* of w. If $|p| < |w|$, p is a *proper* prefix; if $|s| < |w|$, s is a *proper* suffix. Given a nonnegative integer k, we denote by w^k the word consisting of k successive repetitions of the finite word w. A word w is *primitive* if it is not a power of another word, i.e., $w = z^k$ implies $k = 1$.

The set $\mathcal{A}^{\mathbb{N}}$ is the set of all *infinite words* over $\mathcal{A}$, i.e., infinite sequences over $\mathcal{A}$. Given a word w, we denote by w^ω the infinite word $www\ldots$.

Let φ be an endomorphism of $\mathcal{A}^*$. We define

$$\|\varphi\| = \max\{|\varphi(a)|\colon a \in \mathcal{A}\}.$$

A triplet $G = (\mathcal{A}, \varphi, w)$ is a *D0L-system* if $\mathcal{A}$ is an alphabet, φ is an endomorphism of $\mathcal{A}^*$, and w is a non-empty word over $\mathcal{A}$. The word w is the *axiom* of G. The *sequence of* G is $E(G) = (w_i)_{i\geq 0}$ where $w_0 = w$ and $w_i = \varphi^i(w_0)$. The *language of* G is the set $L(G) = \{\varphi^n(w)\colon n \in \mathbb{N}\}$ and by $S(L(G))$ we denote the set of all factors appearing in the elements of $L(G)$. The alphabet is always considered to be the minimal alphabet necessary, i.e., $\mathcal{A} \cap S(L(G)) = \mathcal{A}$.

We say that a D0L-system $G = (\mathcal{A}, \varphi, w)$ is *injective on* $S(L(G))$ if for every $u, v \in S(L(G))$, the equality $\varphi(u) = \varphi(v)$ implies that $u = v$. It is clear that if φ is injective, then G is injective on $S(L(G))$. The converse is not true: consider $\varphi : a \mapsto abca, b \mapsto bca, c \mapsto a$, then φ is not injective as $\varphi(cb) = \varphi(a)$. On the other hand, we have $cb \notin S(L(G))$ and one can verify that $G = (\{a, b, c\}, \varphi, a)$ is indeed injective on $S(L(G))$. If φ is *non-erasing*, i.e., $\varphi(a) \neq \varepsilon$ for all $a \in \mathcal{A}$, then we speak about *propagating D0L-system*, shortly PD0L.

Given a D0L-system $G = (\mathcal{A}, \varphi, w)$ we say that the letter a is *bounded* (or also of *rank zero*) if the set $\{\varphi^n(a)\colon n \in \mathbb{N}\}$ is finite. If a letter is not bounded, it is *unbounded*. We denote the subset of all bounded letters by $\mathcal{A}_0$. A D0L-system G is *pushy* if $S(L(G))$ contains infinitely many factors over $\mathcal{A}_0$, otherwise, it is *non-pushy*.

A D0L-system is *repetitive* if for any $k \in \mathbb{N}$ there is a non-empty word w such that w^k is a factor. By [3], any repetitive D0L-system is *strongly repetitive*, i.e., there is a non-empty word w such that w^k is a factor for all $k \in \mathbb{N}$.

3. Definition of circularity

Two slightly different views of circularity can be found in the literature. Both these perspectives can be expressed in the terms of interpretations:

Definition 1. *Let $G = (\mathcal{A}, \varphi, w)$ be a PD0L-system and $u \in S(L(G))$. A triplet (p, v, s) where $p, s \in \mathcal{A}^*$ and $v = v_1 \cdots v_n \in S(L(G))$ with $n > 0$ is an interpretation of the word u if $\varphi(v) = pus$.*

2

The following definition of circularity is used in [8].

Definition 2. *Let $G = (\mathcal{A}, \varphi, w)$ be a PD0L-system and let (p, v, s) and (p', v', s') be two interpretations of a non-empty word $u \in S(L(G))$ with $v = v_1 \cdots v_n \in \mathcal{A}^n$, $v' = v'_1 \cdots v'_m \in \mathcal{A}^m$ and $u = u_1 \cdots u_\ell \in \mathcal{A}^\ell$.*

We say that G is circular *with* synchronization delay $D > 0$ *if whenever we have*

$$|\varphi(v_1 \cdots v_i)| - |p| > D \quad \text{and} \quad |\varphi(v_{i+1} \cdots v_n)| - |s| > D$$

for some i such that $1 \leq i \leq n$, then there exists j such that $1 \leq j \leq m$ and

$$|\varphi(v_1 \cdots v_{i-1})| - |p| = |\varphi(v'_1 \cdots v'_{j-1})| - |p'|$$

and $v_i = v'_j$ (see Figure 1).

This definition says that a long enough word from $S(L(G))$ has a unique φ-preimage in $S(L(G))$ except for some prefix and suffix shorter than a constant D. Note that if the set $S(L(G))$ contains arbitrarily long words with two different φ-preimages in $S(L(G))$ (i.e., for any $n > 0$ there are distinct words v and u in $S(L(G))$ longer than n with $\varphi(v) = \varphi(u)$), it is not circular.

In [1], circularity is defined for D0L-systems with injective endomorphism using the notion of synchronizing point (see Section 3.2 in [1] for details). We give here an equivalent definition of synchronizing point employing the notion of interpretation.

Definition 3. *Let $G = (\mathcal{A}, \varphi, w)$ be a PD0L-system. We say that two interpretations (p, v, s) and (p', v', s') of a word $u \in S(L(G))$ are synchronized at position k if there exist indices i and j such that*

$$\varphi(v_1 \cdots v_i) = pu_1 \cdots u_k \quad \text{and} \quad \varphi(v'_1 \cdots v'_j) = p'u_1 \cdots u_k$$

with $v = v_1 \cdots v_n \in \mathcal{A}^n$, $v' = v'_1 \cdots v'_m \in \mathcal{A}^m$ and $u = u_1 \cdots u_\ell \in \mathcal{A}^\ell$ (if $k = 0$, we put $u_1 \cdots u_k = \varepsilon$), see Figure 2. Two interpretations that are not synchronized at any position are called non-synchronized.

We say that a word $u \in S(L(G))$ has a synchronizing point at position k with $0 \leq k \leq |u|$ if all its interpretations are pairwise synchronized at position k.

By [1], a D0L-system G with injective endomorphism is circular if there is positive D such that any v from $S(L(G))$ longer than D has a synchronizing point. In this case of injective endomorphism, this definition is equivalent to Definition 2. However, the synchronizing point is defined for D0L-systems with just non-erasing (possibly non-injective) endomorphism. Thus, we can omit the assumption of injectiveness and obtain the following definition.

Definition 4. *A PD0L-system G is called* weakly circular *if there is a constant $D > 0$ such than any v from $S(L(G))$ longer that D has a synchronizing point.*

As indicated above, if G is injective on $S(L(G))$, weak circularity is equivalent to circularity. However, as the following example shows, this is not true for the non-injective case.

3

Example 5. Consider the D0L-system $G_1 = (\{a, b, c\}, \varphi_1, a)$ with the non-injective endomorphism $\varphi_1 : a \mapsto abca, b \mapsto bc, c \mapsto bc$. This system is not circular as for all $m \in \mathbb{N}$ the word $(bc)^{2m}$ has two distinct φ_1-preimages in $S(L(G_1))$, namely $(bc)^m$ and $(cb)^m$. The corresponding interpretations, however, have synchronizing points for $m > 1$ at positions $2k$ for all $0 \le k \le m$. Moreover, one can easily check that G_1 is weakly circular.

We conclude that circularity implies weak circularity but the converse is not true.

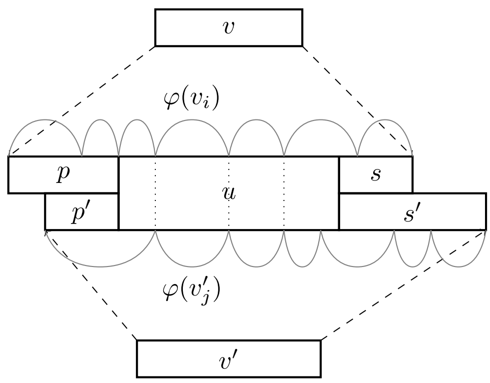

Figure 1: Two interpretations from Definition 2 with $v_i = v'_j$.

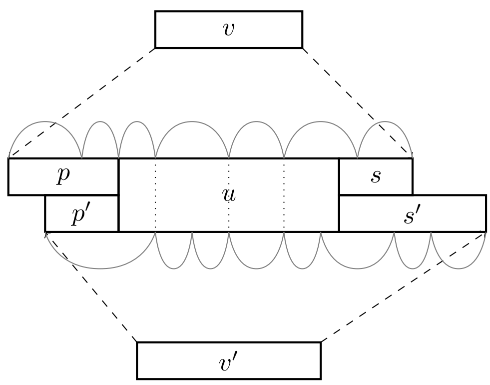

Figure 2: Two interpretations from Definition 3 synchronized at positions depicted by dotted lines.

4. Weak circularity is equivalent to non-repetitivenes

In this section we give a proof of the following theorem:

Theorem 6. *Any PD0L-system that is not weakly circular is repetitive.*

4

The two following lemmas will be used to prove this theorem. The first lemma and its proof are based on the ideas in the proof of Theorem 4.35 in [6].

Lemma 7. *Let $G = (\mathcal{A}, \varphi, w)$ be a PD0L-system. If there exists a sequence $(\epsilon(j))$ with $\lim\limits_{j \to +\infty} \epsilon(j) = +\infty$ such that for any $k \in \mathbb{N}$ there exist non-empty words u and v from $S(L(G))$, nonnegative integers m and n with $m > n$, and letters a and b such that the following conditions are satisfied*

(i) $|u| = k$;

(ii) the word $\varphi^i(u)$ is a factor of $\varphi^i(v)$ for each $i \in \{m, n\}$;

(iii) $aub \in S(L(G)$ and the word $\varphi^i(v)$ is a factor of $\varphi^i(aub)$ for each $i \in \{m, n\}$;

(iv) $\frac{|\varphi^m(u)|}{|\varphi^m(a)|} > \epsilon(k)$ *or* $\frac{|\varphi^m(u)|}{|\varphi^m(b)|} > \epsilon(k)$;

(v) for each $i \in \{m, n\}$ the factor $\varphi^i(u)$ has no synchronizing point: two non-synchronized interpretations are $(\varepsilon, \varphi^{i-1}(u), \varepsilon)$ and $(p_i, \varphi^{i-1}(v), s_i)$;

then G is repetitive.

Proof. Let S be the set of integers k such that $\frac{|\varphi^m(u)|}{|\varphi^m(a)|} > \epsilon(k)$ is true in requirement (iv). Suppose that S is infinite. If S is finite, then the other case $\frac{|\varphi^m(u)|}{|\varphi^m(b)|} > \epsilon(k)$ is true for infinitely many k and the proof is analogous.

Fix $k \in S$. It holds that

$$\varphi^m(v) = p_m \varphi^m(u) s_m = \varphi^{m-n}(\varphi^n(v)) = \varphi^{m-n}(p_n)\varphi^m(u)\varphi^{m-n}(s_n).$$

The fact that the interpretations $(\varepsilon, \varphi^{m-1}(u), \varepsilon)$ and $(p_m, \varphi^{m-1}(v), s_m)$ are not synchronized implies that $p_m \neq \varphi^{m-n}(p_n)$ (if $p_m = \varphi^{m-n}(p_n)$, the two interpretations of $\varphi^m(u)$ are synchronized at position 0, see Figure 3). Since $p_m \varphi^m(u) s_m = \varphi^{m-n}(p_n)\varphi^m(u)\varphi^{m-n}(s_n)$, the word p_m is a proper prefix of $\varphi^{m-n}(p_n)$ or vice versa. Moreover, p_m is not empty since it would again contradict requirement (v).

We denote by z the non-empty word such that $p_m z = \varphi^{m-n}(p_n)$ if p_m is a non-empty proper prefix of $\varphi^{m-n}(p_n)$; or such that $\varphi^{m-n}(p_n) z = p_m$ if $\varphi^{m-n}(p_n)$ is a non-empty proper prefix of p_m. In both cases, it implies that the word $\varphi^m(u)$ is a prefix of z^ω (see Figure 4 for the second case).

Since

$$\frac{|\varphi^m(u)|}{|z|} > \frac{|\varphi^m(u)|}{\max\{|p_m|, |\varphi^{m-n}(p_n)|\}} > \frac{|\varphi^m(u)|}{|\varphi^m(a)|} > \epsilon(k),$$

it follows that for all $k \in S$ there is a word z such that $z^{\lfloor \epsilon(k) \rfloor}$ is factor of $\varphi^m(u)$. As $\lim\limits_{k \to +\infty} \epsilon(k) = +\infty$, the D0L-system is repetitive. □

Lemma 8. *In any PD0L-system G there is a constant C such that all elements of $S(L(G))$ over bounded letters longer than C have a synchronizing point.*

5

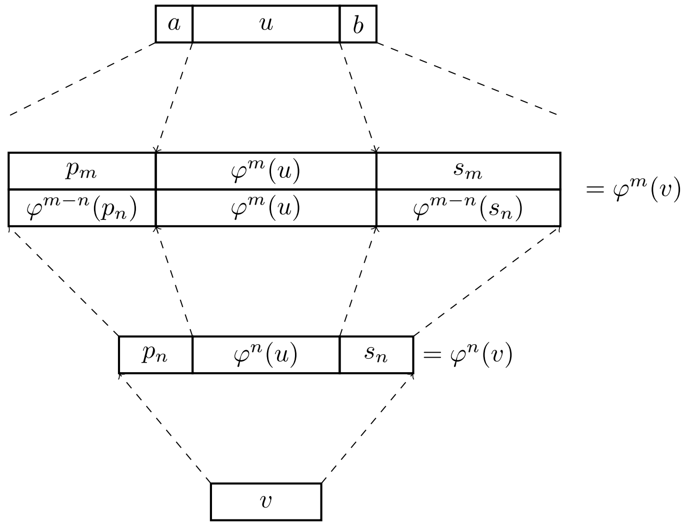

Figure 3: The first arrangement from the proof of Lemma 7.

Proof. Let $G = (\mathcal{A}, \varphi, w)$ be a PD0L-system. If G is non-pushy, then the set $S(L(G)) \cap \mathcal{A}_0$ is finite and the claim trivially follows.

Suppose G is pushy. There exist an integer n such that for all $c \in \mathcal{A}_0$ we have $|\varphi^m(c)| = |\varphi^{m+1}(c)|$ for every $m \geq n$. Let u be a factor over bounded letters only of length at least $C = 3\|\varphi^{n+1}\| \cdot |w|$. This implies that u appears as a factor in the sequence $E(G) = (w_i)_{i \geq 0}$ in w_k for some $k > n + 1$.

Let (p, z, s) be an interpretation of u. Since u is a factor of w_k such that $k > n + 1$ and $|w_k| > C$, there must be words $x, y \in \mathcal{A}^*$ and $v \in \mathcal{A}^+$ such that $z = x\varphi^n(v)y$ and

$$|\varphi(x)| - |p| < \|\varphi^{n+1}\| \quad \text{and} \quad |\varphi(y)| - |s| < \|\varphi^{n+1}\|.$$

As $\varphi^{n+1}(v)$ is a factor of u, it contains only bounded letters, and thus so does the word v. Moreover, by the definition of n, every letter c occurring in $\varphi^n(v)$ satisfies $|\varphi^n(c)| = |\varphi^{n+1}(c)|$.

It follows that any two interpretations (p, z, s) and (p', z', s') of the word u are synchronized at position $\|\varphi^{n+1}\|$. □

Proof of Theorem 6. Consider a PD0L-system $G = (\mathcal{A}, \varphi, w)$ with an infinite language. We define a partition of the alphabet $\mathcal{A} = \Sigma_0 \cup \Sigma_1 \cup \cdots \cup \Sigma_{m-1} \cup \Sigma_m$ as follows:

(i) $\Sigma_0 = \mathcal{A}_0$ is the set of bounded letters,

(ii) if x and y are from Σ_i, then the sequence $\left(\frac{|\varphi^n(x)|}{|\varphi^n(y)|}\right)_{n \geq 1}$ is $\Theta(1)$,

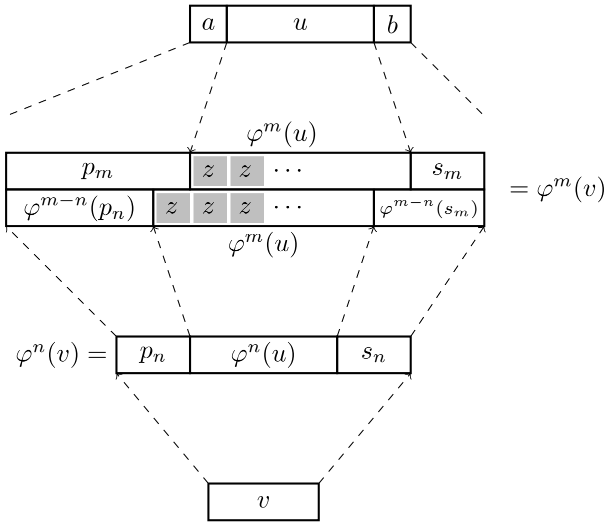

Figure 4: The second arrangement from the proof of Lemma 7.

(iii) for all $i = 1, \ldots, m$, if x is an element of Σ_i and y of Σ_{i-1}, then $\lim_{n\to+\infty} \frac{|\varphi^n(x)|}{|\varphi^n(y)|} = +\infty$.

In other words, we order the letters of $\mathcal{A}$ by their speed of growth (with respect to φ). This partition is well defined due to [12], where it is proved that for any $a \in \mathcal{A}$ there are numbers $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{R}_{\geq 1} \cup \{0\}$ such that $|\varphi^n(a)| = \Theta(n^\alpha \beta^n)$. Further we define for all $j = 0, 1, \ldots, m$ the sets

$$\mathcal{A}_j = \bigcup_{0\leq i\leq j} \Sigma_i.$$

Note that if we denote by $L_j = \{a \in \mathcal{A} \colon a \text{ is a factor of } \varphi(b) \text{ for } b \in \Sigma_j\}$, then $L_j \subset \mathcal{A}_j$ and $L_j \cap \Sigma_j \neq \emptyset$.

Since G is not weakly circular, there exists an element of $S(L(G))$ of arbitrary length that has no synchronizing point. Let j be an integer such that there is an element of $S(L(G))$ of arbitrary length over $\mathcal{A}_j$ that has no synchronizing point. As by Lemma 8 the factors without synchronizing point over $\mathcal{A}_0$ are bounded in length, we obtain $j > 0$.

In order to use Lemma 7, the existence of a factor of arbitrary length over $\mathcal{A}_j$ with no synchronizing point can be rephrased as follows. Let k be a positive integer. For any positive $\ell \in \mathbb{N}$ we can find words $u_\ell^{(k)} \in \mathcal{A}_j^*$ and $v_\ell^{(k)} \in \mathcal{A}_j^*$ and letters $a_\ell^{(k)} \in \mathcal{A}_j$ and $b_\ell^{(k)} \in \mathcal{A}_j$ such that

(a) $|u_\ell^{(k)}| = k$;

(b) $\varphi^\ell(u_\ell^{(k)})$ is a factor of $\varphi^\ell(v_\ell^{(k)})$;

7

(c) $a_\ell^{(k)} u_\ell^{(k)} b_\ell^{(k)} \in S(L(G))$ and $\varphi^\ell(v_\ell^{(k)})$ is a factor of $\varphi^\ell(a_\ell^{(k)} u_\ell^{(k)} b_\ell^{(k)})$;

(d) $\varphi^\ell(u_\ell^{(k)})$ has two non-synchronized interpretations

$$(\varepsilon, \varphi^{\ell-1}(u_\ell^{(k)}), \varepsilon) \quad \text{and} \quad (p_\ell^{(k)}, \varphi^{\ell-1}(v_\ell^{(k)}), s_\ell^{(k)})$$

where $p_\ell^{(k)} \varphi^\ell(u_\ell^{(k)}) s_\ell^{(k)} = \varphi^\ell(v_\ell^{(k)})$.

Since the length of $u_\ell^{(k)}$ is fixed, there exists an infinite set $E^{(k)} \subset \mathbb{N}$ such that $u_s^{(k)} = u_t^{(k)} = u^{(k)}$, $a_s^{(k)} = a_t^{(k)} = a^{(k)}$, and $b_s^{(k)} = b_t^{(k)} = b^{(k)}$ for all s, t from $E^{(k)}$.

If for each k there are indices m_k and n_k in $E^{(k)}$ such that $m_k > n_k$ and $v_{m_k}^{(k)} = v_{n_k}^{(k)} = v^{(k)}$, and if the number of letters from Σ_j in $u^{(k)}$ tends to infinity as $k \to +\infty$, then all the assumptions of Lemma 7 are satisfied and G is repetitive.

Assume that no such indices m_k and n_k exist for some k. It implies that $|v_\ell^{(k)}|$ tends to infinity as $\ell \to +\infty$. Since $u^{(k)} \in \mathcal{A}_j$, $|u^{(k)}| = k$ and $\varphi^\ell(u^{(k)})$ is a factor of $\varphi^\ell(v_\ell^{(k)})$, the number of letters from Σ_j in words $v_\ell^{(k)}$ is bounded (or even zero). Thus, there exists $j' \in \{0, 1, \ldots, j-1\}$ such that the number of letters from $\Sigma_{j'}$ in $v_\ell^{(k)}$ tends to infinity as $\ell \to +\infty$ and such that we may find a factor of arbitrary length without a synchronizing point over $\mathcal{A}_{j'}$. Moreover, $j' \neq 0$ as the factors of $v_\ell^{(k)}$ containing only letters from $\mathcal{A}_0$ are bounded in length (by Lemma 8), i.e., the case when no such indices m_k and n_k exist for some k is only possible if $j > 1$.

If such indices m_k and n_k exist for each k but the number of letters from Σ_j in $u^{(k)}$ is bounded as $k \to +\infty$, there is again some $j' \in \{0, \ldots, j-1\}$ such that the number of letters from $\Sigma_{j'}$ in $u^{(k)}$ tends to infinity as $k \to +\infty$ and there is again a factor of arbitrary length without a synchronizing point over $\mathcal{A}_{j'}$. As above, Lemma 8 implies $j' \neq 0$ and so this case is again possible only if $j > 1$.

Overall, if $j = 1$, then G is repetitive by Lemma 7. If $j > 1$, then either G is repetitive by Lemma 7 or we find a positive integer j' less than j such that there is a factor of arbitrary length without a synchronizing point over $\mathcal{A}_{j'}$ and we repeat the reasoning where we replace the value of j by j'. □

5. Simple criterion for circularity

In this section, we deduce a characterization of circularity which yields a criterion that can be effectively verified. To do that, we need some more notions.

Definition 9. *We say that a D0L-system G is unboundedly repetitive if there exists $w \in S(L(G))$ such that $w^k \in S(L(G))$ for all k and w contains at least one unbounded letter.*

In [2], the authors introduced the notion of simplification to study properties of a D0L-system. Given an endomorphism φ over $\mathcal{A}$, an endomorphism Ψ over $\mathcal{B}$ is its *simplification* if $\#\mathcal{B} < \#\mathcal{A}$ and there exist homomorphisms $h : \mathcal{A}^* \to \mathcal{B}^*$

and $k : \mathcal{B}^* \to \mathcal{A}^*$ such that $\varphi = kh$ and $\Psi = hk$. A corollary of the defect theorem (see [5]) is that every non-injective endomorphism has a simplification which is injective. A D0L-system $G' = (\mathcal{B}, \psi, w')$ is an *injective simplification* of $G = (\mathcal{A}, \varphi, w)$ if either $G = G'$ and φ is injective, or $\psi = hk : \mathcal{B}^* \to \mathcal{B}^*$ is an injective simplification of $\varphi = kh$ and $w' = h(w)$.

The following claim follows from [4].

Proposition 10. *A D0L-system G is unboundedly repetitive if and only if for some its injective simplification $G' = (\mathcal{B}, \psi, w'_0)$ there is a positive integer ℓ and $a \in \mathcal{B}$ such that*

$$\lim_{n \to +\infty} (\psi^\ell)^n(a) = w^\omega \qquad \textit{for some } w \in \mathcal{B}^+.$$

In fact, if the condition in the previous claim is satisfied for some injective simplification, then it is satisfied for all injective simplifications.

Using this proposition and the results of [4] we deduce the following theorem.

Theorem 11. *Let G be a repetitive D0L-system, then one of the following is true:*

(i) G is pushy,

(ii) G is unboundedly repetitive.

In the previous section we proved that any PD0L-system that is not weakly circular is repetitive. The next theorem gives a characterization of injective circular D0L-systems.

Theorem 12. *An D0L-system $G = (\mathcal{A}, \varphi, w)$ injective on $S(L(G))$ is not circular if and only if it is unboundedly repetitive.*

Proof. ($\Rightarrow$): Since injectiveness on $S(L(G))$ implies that φ is non-erasing, Theorem 6 implies that G is repetitive. Thus, by Theorem 11, G is pushy or unboundedly repetitive. Suppose it is pushy and not unboundedly repetitive. It implies that there exists an integer M (see Proposition 1 in [8] and Lemma 18 in [4] for the proof of existence of this constant) such that all ℓ-powers u^ℓ, where $\ell > M$ and $u \in S(L(G))$, are over bounded letters only, i.e., $u \in \mathcal{A}_0^+$. From the proof of Theorem 6 one can see that long enough non-synchronized factors contain longer and longer repetitions but these repetitions cannot be over bounded letters due to Lemma 8 — a contradiction.

($\Leftarrow$): Proposition 10 implies that there is a positive integer ℓ and a letter a such that $\lim_{n \to +\infty} (\varphi^\ell)^n(a) = w^\omega$ for some $w \in \mathcal{A}^+$. In [7] it is proved that the word w can be taken so that it contains the letter a only once at its beginning. It follows that $\varphi^\ell(w) = w^k$ for some $k > 1$.

Define $j_0 = \min\{j \mid \varphi^j(w) \text{ is not a primitive word}\}$, clearly j_0 exists and $1 \le j_0 \le \ell$. Denote $\varphi^{j_0 - 1}(w) = z$ and $\varphi(z) = x^m$ with $m > 1$. We must have $\varphi(p) \neq x^q$ for any proper prefix p of z and any integer q. Indeed, if $z = ps$ and $\varphi(p) = x^q$ for some integer q with $q < m$, then $\varphi(s) = x^{m-q}$ and $\varphi(ps) = \varphi(sp)$

9

and so $sp = ps$. It follows that there exist a word r and integers c and d such that $p = r^c$ and $s = r^d$. We obtain that the word $z = r^{c+d}$ is not primitive which is a contradiction.

It follows that the word x^{nm} has two non-synchronized interpretations $(\varepsilon, z^n, \varepsilon)$ and (x, z^{n+1}, x^{m-1}) for all $n \in \mathbb{N}$ and G is not circular. □

Remark 13. In the previous theorem, we cannot omit the assumption of injectiveness on $S(L(G))$ and replace circularity with weak circularity: consider again the D0L-system G_1 from Example 5. The conditions of Proposition 10 are satisfied for $\ell = 1$ and the letter b with $w = bc$ but still the corresponding D0L-system is weakly circular.

Since the existence of ℓ and a satisfying conditions of Proposition 10 can be tested by a simple and fast algorithm described in [7], we obtain a simple algorithm deciding circularity.

As a corollary of Theorem 12, we retrieve the following result of [9] for primitive endomorphisms. An endomorphism $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ is *primitive* if there exists an integer k such that for all letters $a, b \in \mathcal{A}$, the letter b appears in $\varphi^k(a)$. An infinite word $\mathbf{u}$ is a *periodic point* of an endomorphism φ if there exists an integer ℓ such that $\varphi^\ell(\mathbf{u}) = \mathbf{u}$. It is well known, that if φ is primitive, then there is always at least one periodic point and the sets of all factors of any two periodic points are equal. It follows that for any periodic point $\mathbf{u}$ of φ the set of all factors of $\mathbf{u}$ equals to $S(L(G))$ with $G = (\mathcal{A}, \varphi, a)$, where a is an arbitrary letter from $\mathcal{A}$.

Corollary 14 ([9])**.** *Let* $\mathbf{u}$ *is be an aperiodic periodic point of an endomorphism* φ *and* $G = (\mathcal{A}, \varphi, a)$ *a D0L-system. If* φ *is primitive and* G *injective on* $S(L(G))$*, then* G *is circular.*

Proof. Since any periodic point of a primitive endomorphism has the same set of factors as $\mathbf{u}$, every periodic point is aperiodic and by Proposition 10 the D0L-system G is not unboundedly repetitive. Theorem 12 yields the result. □

Acknowledgements

This work was supported by the Czech Science Foundation grant GAČR 13-35273P.

References

[1] J. Cassaigne, *An algorithm to test if a given circular HD0L-language avoids a pattern*, in IFIP World Computer Congress'94, Elsevier, 1994, pp. 459–464.

[2] A. Ehrenfeucht and G. Rozenberg, *Elementary homomorphisms and a solution of the D0L sequence equivalence problem*, Theoret. Comput. Sci., 7 (1978), pp. 169–183.

[3] ——, *Repetition of subwords in D0L languages*, Inform. Comput., 53 (1983), pp. 13–35.

[4] K. Klouda and Štěpán Starosta, *An algorithm enumerating all infinite repetitions in a D0L-system*, accepted in J. Discrete Algorithms, arXiv:1307.6408, (2015).

[5] Y. Kobayashi and F. Otto, *Repetitiveness of languages generated by morphisms*, Theoret. Comput. Sci., 240 (2000), pp. 337–378.

[6] P. Kůrka, *Topological and Symbolic Dynamics*, Societé Mathématique de France, 2003.

[7] B. Lando, *Periodicity and ultimate periodicity of D0L systems*, Theoret. Comput. Sci., 82 (1991), pp. 19–33.

[8] F. Mignosi and P. Séébold, *If a D0L language is k-power free then it is circular*, in ICALP '93: Proceedings of the 20th International Colloquim on Automata, Languages and Programming, London, UK, 1993, Spriger-Verlag, pp. 507–518.

[9] B. Mossé., *Reconnaissabilité des substitutions et complexité des suites automatiques*, Bull. Soc. Math. Fr., 124 (1996), pp. 329–346.

[10] P. Prusinkiewicz, *A look at the visual modelling of plants using L-systems*, Agronomie, 19 (1999), pp. 211–224.

[11] G. Rozenberg and A. Salomaa, *The mathematical theory of L systems*, Academic Press, 1980.

[12] A. Salomaa and M. Soittola, *Automata-Theoretic Aspects of Formal Power Series*, Monographs in Computer Science, Springer, 1st ed., 1978.

Chapter 10

The Rauzy Gasket

Authors:	P. Arnoux and Š. Starosta
Publication reference:	in Further Developments in Fractals and Related Fields, J. Barral and S. Seuret, eds., Trends in Mathematics, Springer Science+Business Media New York, 2013, pp. 1–23
Cited as:	[IX]

The Rauzy Gasket

Pierre Arnoux and Štěpán Starosta

Abstract We define the Rauzy gasket as a subset of the standard two-dimensional simplex associated with letter frequencies of ternary episturmian words. We prove that the Rauzy gasket is homeomorphic to the usual Sierpiński gasket (by a two-dimensional generalization of the Minkowski ? function) and to the Apollonian gasket (by a map which is smooth on the boundary of the simplex). We prove that it is also homothetic to the invariant set of the fully subtractive algorithm, hence of measure 0.

1 Introduction

Strict episturmian ternary words, also called Arnoux–Rauzy words, are a natural generalization of Sturmian words (see Sect. 2 for the definitions). Each such word is uniquely ergodic, and in particular, its letters have a well-defined frequency; one can prove that these frequencies completely define the minimal symbolic system associated with such a word.

These dynamical systems are associated with a particular family of interval exchange transformations (see [1]). It is known that some of these systems (in particular those defined by a substitution) can be represented by a toral rotation,

P. Arnoux (✉)
IML - Institut de Mathématiques de Luminy, UPR-Cnrs 9016, Campus de Luminy case 907, 13288 Marseille Cedex 9, France
e-mail: arnoux@iml.univ-mrs.fr

Š. Starosta
Department of Mathematics, FNSPE, Czech Technical University in Prague, Trojanova 13, 120 00 Prague 2, Czech Republic

Department of Theoretical Computer Science, FIT, Czech Technical University in Prague, Thákurova 9, 160 00 Prague 6, Czech Republic
e-mail: stepan.starosta@fit.cvut.cz

J. Barral and S. Seuret (eds.), *Further Developments in Fractals and Related Fields*, Trends in Mathematics, DOI 10.1007/978-0-8176-8400-6_1, 1

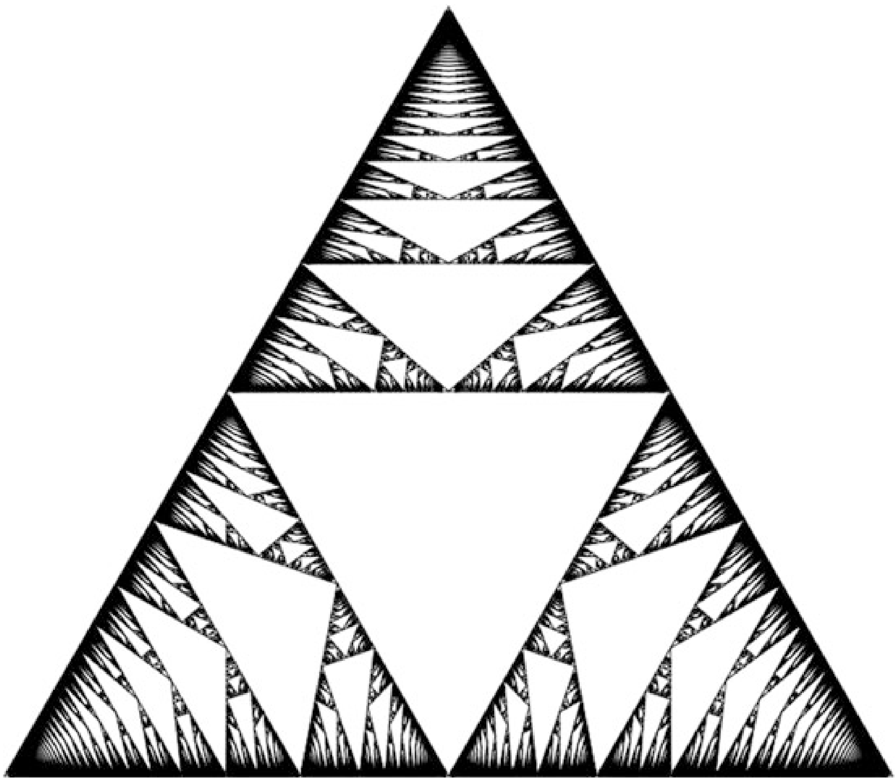

Fig. 1 The Rauzy gasket

and in particular, they have pure discrete spectrum (see [13]); on the other hand, it is known that some examples of Arnoux–Rauzy words are not balanced (see [3]); hence they cannot be represented by a toral rotation. It would be interesting to understand what is the general behavior; a preliminary step would be to find a "good" measure on the parameter set of the family of episturmian systems.

Arnoux–Rauzy words admit a natural renormalization process, which acts on the frequencies; this renormalization can be considered as a kind of generalized continued fraction; indeed, the equivalent renormalization on Sturmian words is a symbolic version of the classical additive continued fraction. It is then natural to look for an invariant measure (a Gauss measure) for the renormalization.

In this chapter, we define the Rauzy gasket as the set of admissible vectors of letter frequencies for episturmian ternary words; it is a compact subset of the two-dimensional simplex (see Fig. 1). The Rauzy gasket parametrizes the set of episturmian systems, and there is a generalized continued fraction algorithm acting on it. We would like to find a Gauss measure for this algorithm; however, since this set is fractal, one would first need to compute its Hausdorff dimension and Hausdorff measure.

We are far from reaching this goal, and some preliminary investigations are given in this chapter. We prove that the Rauzy gasket is homeomorphic to the usual Sierpiński gasket by a map which is a generalization of the Minkowski ? function; this map is not differentiable, and not absolutely continuous, so the Hausdorff dimensions of both sets have no reason to be the same. We also prove that the Rauzy gasket is homeomorphic to the Apollonian gasket by a quite regular map, since it is smooth on the boundary of the complementary triangles, although it is not a diffeomorphism on the rational points. We then show that the same set occurs in a classical two-dimensional continued fraction algorithm, the fully subtractive algorithm. We deduce from the proof of [11] that the Rauzy gasket has zero Lebesgue measure.

In Sect. 2, we give the necessary definitions for episturmian words and explain the origin of the problem. In Sect. 3, we define the Rauzy gasket and the related continued fraction. In Sect. 4, we prove that the Rauzy gasket is homeomorphic to the classical Sierpiński gasket by a generalization of the Minkowski ? function. In Sect. 5, we prove that it is homeomorphic to the Apollonian gasket. In Sect. 6, we show that the continued fraction associated with the Rauzy gasket is conjugate by a linear change of coordinates to the induction of the fully subtractive algorithm on the central part of the simplex, and we deduce that its Lebesgue measure is 0. In the last section, we give a few remarks and open questions.

2 Preliminaries

2.1 Background: Complexity and Sturmian Words

Let $\mathscr{A}$ denote an *alphabet*, a finite set of *letters*. A finite (infinite) sequence of letters is called a *finite (infinite) word*. We say that a finite word $\mathbf{w} = w_0w_1\dots w_n$, where $w_i \in \mathscr{A}$, is a *factor* of a word $\mathbf{v} = v_0v_1\dots$ (finite or infinite) if there exists an index k such that $w_0w_1\dots w_n = v_kv_{k+1}\dots v_{k+n}$. Furthermore, such an index k is called an *occurrence* of $\mathbf{w}$ in $\mathbf{v}$. By $|\mathbf{v}|_{\mathbf{w}}$ we denote the *number of occurrences* of $\mathbf{w}$ in $\mathbf{v}$.

The *language* of an infinite word $\mathbf{u}$ is the set of factors of $\mathbf{u}$. We say that this language is *closed under reversal* if, for any factor $\mathbf{w} = w_0w_1\dots w_n$ of $\mathbf{u}$, its *reverse* word $w_nw_{n-1}\dots w_0$ is also a factor of $\mathbf{u}$.

The *shift map* on $\mathscr{A}^\infty$ associates to any infinite word $\mathbf{u}$ the word $\mathbf{v}$ defined for all i by $v_i = u_{i+1}$. The dynamical system associated with a word $\mathbf{u}$ is the closure of its orbit by the shift; it is also the set of words whose language is contained in the language of $\mathbf{u}$. Hence it is completely determined by this language.

Let $\mathbf{u} = u_0u_1\dots$ be an infinite word over $\mathscr{A}$ and $\mathbf{w}$ be a factor of $\mathbf{u}$. Let $f_{\mathbf{u}}(\mathbf{w})$ denote the limit $\lim_{n\to+\infty}\frac{|u_0u_1\dots u_{n-1}|_{\mathbf{w}}}{n}$, if it exists. Such a number is then called the *frequency* of the factor $\mathbf{w}$ in $\mathbf{u}$.

The *factor complexity* of $\mathbf{u}$, denoted $\mathscr{C}(n)$, is the mapping which associates to an integer n the number of factors of $\mathbf{u}$ of length n. As usual, the *empty word* is counted as a factor and we have $\mathscr{C}(0) = 1$ and $\mathscr{C}(1) = \#\mathscr{A}$.

A factor $\mathbf{w}$ of $\mathbf{u}$ is said to be left (right) special, if there exist two distinct letters a and b such that $a\mathbf{w}$ and $b\mathbf{w}$ ($\mathbf{w}a$ and $\mathbf{w}b$) are factors of $\mathbf{u}$.

It is well-known that the complexity of a non eventually periodic word is strictly increasing; hence such a word has complexity $\mathscr{C}(n) \geq n+1$. Aperiodic words of minimal complexity are of particular interest.

Definition 1. An infinite word is called a *Sturmian word* if it is of minimal complexity $\mathscr{C}(n) = n + 1$.

Let $\lfloor x \rfloor$ (resp. $\lceil x \rceil$) denote the *floor* function, i.e., the largest integer $n \leq x$ (resp. the *ceiling* function, that is the smallest integer $n \geq x$).

Definition 2. An infinite word $\mathbf{u}$ is a *rotation word* if there exist $\alpha, \beta \in [0,1]$ such that for all n, $u_n = \lfloor (n+1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor$ or for all n, $u_n = \lceil (n+1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil$ (the difference between $\lfloor x \rfloor$ and $\lceil x \rceil$ is irrelevant unless $n\alpha + \beta$ takes integer values for some n).

One can prove that a rotation word is periodic if and only if α is rational, that the Sturmian words are the aperiodic rotation words, and that the closure in $\{0,1\}^{\mathbb{N}}$, for the product topology, of the set of Sturmian words is the set of rotation words.

Since, in a Sturmian word, there are three different factors of length 2, and words 10 and 01 must occur in a not eventually constant sequence, one of the words 00 or 11 does not occur. Hence, in a Sturmian word, one of the letters is *isolated*. Sturmian words admit a renormalization process by erasing the letter following the isolated letter, which gives a new Sturmian word; it is a symbolic counterpart of the classical continued fraction algorithm acting on the angle of the corresponding rotation. We will look more closely at this renormalization process in Sect. 4.1.

2.2 Arnoux–Rauzy Words and Episturmian Words: Definition

It is natural to look for good generalizations of Sturmian words, and one such family is the set of strict episturmian words:

Definition 3. Strict episturmian words on three letters, also called ternary Arnoux–Rauzy (AR for short) words, are infinite words of complexity $\mathscr{C}(n) = 2n + 1$ such that for each n there is only one left special factor and one right special factor of length n.

Since $\mathscr{C}(1) = 3$, they are words on three letters; they were first described in [14] and were further studied in [1], where a geometric representation was introduced. AR words code a specific family of six-interval exchange transformations.

One inconvenience of this definition is that the set of AR words is not closed in $\mathscr{A}^{\infty}$ for the natural topology; this is already the case for Sturmian words, whose closure contains periodic rotation words. It is then natural to consider the closure; the language of any AR word, like the language of a Sturmian word, is closed under reversal, and this leads to the definition of a wider family of infinite words known as *episturmian* words (see [4, 6, 7] for a survey) defined on an arbitrary alphabet.

Definition 4. Episturmian words are the infinite words having their language closed under reversal and at most one left special factor of each length.

It is an easy consequence of the definition that the set of episturmian words is closed in $\mathscr{A}^\infty$. Some episturmian words, those who can be extended on the left in several ways, are of particular interest.

Definition 5. An aperiodic episturmian word is *standard* (or *characteristic*) if each of its prefix is a left special factor.

Remark 1. Since every prefix of a left special factor is a left special factor, the definition implies that every aperiodic episturmian word has the same set of factors as a unique standard episturmian word. In other words, the subshift generated by an aperiodic episturmian word contains a unique standard word, which can be used as canonical representative of the system.

If an episturmian word is periodic, a factor longer than the period cannot be special, so the definition does not apply. In that case, we will say by abuse of language that a periodic episturmian word is standard if any prefix is a special left factor whenever there exists a special left factor of the same length. One can prove that the finite subshift generated by a periodic nonconstant episturmian word contains exactly two standard words with this definition.

By a result of Boshernitzan [2], episturmian words are uniquely ergodic; hence the frequencies of factors in episturmian words exist and are positive. On frequencies of AR words see more in [16].

The next section gives a classification of ternary episturmian words.

2.3 Ternary AR Words: Renormalization

In what follows we will consider the alphabet to be fixed $\mathscr{A} = \{1,2,3\}$. For all $i \in \mathscr{A}$, we define the morphism σ_i on the free monoid $\mathscr{A}^*$ by

$$\sigma_i(j) = ij \quad \text{if} \quad j \neq i \qquad \text{and} \qquad \sigma_i(i) = i.$$

We denote by $\mathscr{S}$ the set $\{\sigma_1, \sigma_2, \sigma_3\}$. The following is a restatement of claims in [14].

Proposition 1. *Let* $\mathbf{u}$ *be an AR word and* σ *a morphism from* $\mathscr{S}$. *Then* $\sigma(\mathbf{u})$ *is an AR word.*

This leads to the following claim using our notation.

Corollary 1. *Let* $\mathbf{u}$ *be an AR word and* $i \in \mathscr{A}$. *Then for all letter* $j \in \mathscr{A}$ *we have*

$$f_{\sigma_i(\mathbf{u})}(j) = \frac{f_{\mathbf{u}}(j)}{2 - f_{\mathbf{u}}(i)} \quad \textit{for} \quad i \neq j \qquad \textit{and} \qquad f_{\sigma_i(\mathbf{u})}(i) = \frac{1}{2 - f_{\mathbf{u}}(i)}.$$

What is less obvious is that, conversely, any standard AR word can be renormalized by using one of the morphisms in $\mathscr{S}$. This is the content of the next proposition, which is again a restatement of claims in [14] in our terms.

Proposition 2. *Let* $\mathbf{v}$ *be a standard AR word; then there exist an index* $i \in \mathscr{A}$ *and a standard AR word* $\mathbf{u}$ *such that* $\mathbf{v} = \sigma_i(\mathbf{u})$.

Furthermore, we have

$$f_{\mathbf{u}}(j) = \frac{f_{\mathbf{v}}(j)}{f_{\mathbf{v}}(i)} \quad \text{for} \quad i \neq j \quad \text{and} \quad f_{\mathbf{u}}(i) = \frac{2f_{\mathbf{v}}(i) - 1}{f_{\mathbf{v}}(i)}.$$

Let $(\zeta_1, \zeta_2, \zeta_3)$ denote the vector of frequencies of letters of the AR word $\mathbf{v}$. (Clearly, $\zeta_1 + \zeta_2 + \zeta_3 = 1$.) It follows from [1] that ζ_1 and ζ_2 are rationally independent irrational numbers. Furthermore, it follows from the previous proposition that one of the frequencies is always strictly greater than the sum of the two others, i.e., one letter is always *dominating* the word. Moreover, the dominating letter is *separating* (see [4], Lemma 4), i.e., all factors of $\mathbf{v}$ of length 2 contain at least one occurrence of the separating letter. The renormalization procedure is very simple: to obtain the renormalized word, it suffices to erase the letter following each non-separating letter. In other words, the index i from the last proposition is clearly the separating letter of $\mathbf{v}$ and $\mathbf{u}$ is the renormalized word. Moreover, the previous proposition ensures that, for a standard AR word, this procedure can be infinitely iterated.

In this way, one can associate to any AR subshift an infinite sequence of morphisms in $\mathscr{S}$; this sequence can be seen as a symbolic version of a generalized continued fraction expansion on the set of frequencies, as we show in the next section. This sequence is also the *$\mathscr{S}$-adic expansion* of the subshift (see [5]).

This construction has been extended to episturmian words; the following theorem summarizes results in [6, 7].

Theorem 1. *Let* $\mathbf{u}$ *be a standard episturmian word; then, one can find a standard episturmian word* $\mathbf{v}$ *and a morphism* $\sigma_i \in \mathscr{S}$ *such that* $u = \sigma_i(v)$. *Furthermore, the morphism* σ_i *is uniquely defined by the frequencies of the letters of* $\mathbf{u}$, *unless* $\mathbf{u}$ *is periodic of period* 2.

It follows immediately that, to any episturmian system, one can associate by iterating this construction a sequence σ_{i_n} *of morphisms (the word* $(i_n)_{n \in \mathbb{N}}$ *is called the directive word); three cases are possible:*

1. *Every letter in* $\mathscr{A}$ *occurs infinitely often in* (i_n)*; then the word* $\mathbf{u}$ *is an AR word and* (i_n) *is uniquely defined.*
2. *One letter occurs a finite number of times in* (i_n)*; then the word* $\mathbf{u}$ *is the image by a morphism (finite composition of elements of* $\mathscr{S}$*) of a Sturmian word and* (i_n) *is uniquely defined.*
3. *The word* (i_n) *is eventually constant; then the word* $\mathbf{u}$ *is periodic, and if the word* (i_n) *is not constant, there are two such possible words, one ending in* ij^∞ *and the other in* ji^∞.

The basic element of the proof is that the morphism σ_i is determined by the largest frequency ζ_i. Since this largest frequency satisfies $\zeta_i \geq \zeta_j + \zeta_k$, it is uniquely defined, unless we have $\zeta_k = 0$ and $\zeta_i = \zeta_j$, which corresponds, up to a permutation of indices, to the frequencies $(1/2, 1/2, 0)$ and a periodic word of period 2, and to the directive words 12^∞ and 21^∞, depending on whether we choose to erase letter 1 or 2. It is easy to check that this phenomenon occurs during the renormalization of a word if and only if the vector of frequencies is rational.

Remark 2. This phenomenon already occurs for rotation words; in that case, it is linked to the fact that an irrational number has only one continued fraction expansion, but a rational number has two finite continued fraction expansions.

3 The Rauzy Gasket

We are interested in the set of all possible vectors of letter frequencies $(\zeta_1, \zeta_2, \zeta_3)$ of episturmian words. For convenience, we will speak of the set of frequencies. We can now define the Rauzy gasket and show some of its properties.

Definition 6. The Rauzy gasket, denoted by $\mathbf{R}$, is the set of frequencies of episturmian words.

We will also be interested in the following subsets of the Rauzy gasket:

Definition 7. We denote by $\mathbf{R}_{\text{aper}}$ the set of frequencies of aperiodic episturmian words and by $\mathbf{R}_{\text{AR}}$ the set of frequencies of AR words.

Lemma 1. *We have* $\mathbf{R}_{\text{AR}} \subset \mathbf{R}_{\text{aper}} \subset \mathbf{R} = \overline{\mathbf{R}_{\text{AR}}}$.

Proof. The only nontrivial fact is that $\mathbf{R} \subset \overline{\mathbf{R}_{\text{AR}}}$. But every episturmian word can be approached arbitrarily close by an AR word: it suffices to take a long prefix of the directive word of the word and to compose it with the directive word $(123)^\infty$ to get an AR word. Hence the frequency vector of the given episturmian word can be approximated as closely as we want by the frequency vector of an AR word. □

The elements of $\mathbf{R}_{\text{aper}}$ are exactly the irrational elements of $\mathbf{R}$. Indeed, the frequencies completely characterize an episturmian system. If the word is periodic, its frequencies are rational. On the other hand, if the frequencies are rational, the height of the frequency vector (defined as the sum of coefficients of the smallest collinear integer vector) is strictly decreasing under renormalization, unless the two smallest coordinates are 0; this implies that, starting from any rational frequency, a finite number of renormalizations changes the word to a constant word, so an episturmian word with rational frequencies is periodic.

One can prove that the elements of $\mathbf{R}_{\text{aper}}$ which are not in $\mathbf{R}_{\text{AR}}$ are irrational vectors which satisfy one rational relation, since they are the images of a Sturmian word, with only two letters, by a morphism. One could conjecture that the elements of $\mathbf{R}_{\text{AR}}$ are completely irrational, but we do not know a proof of this.

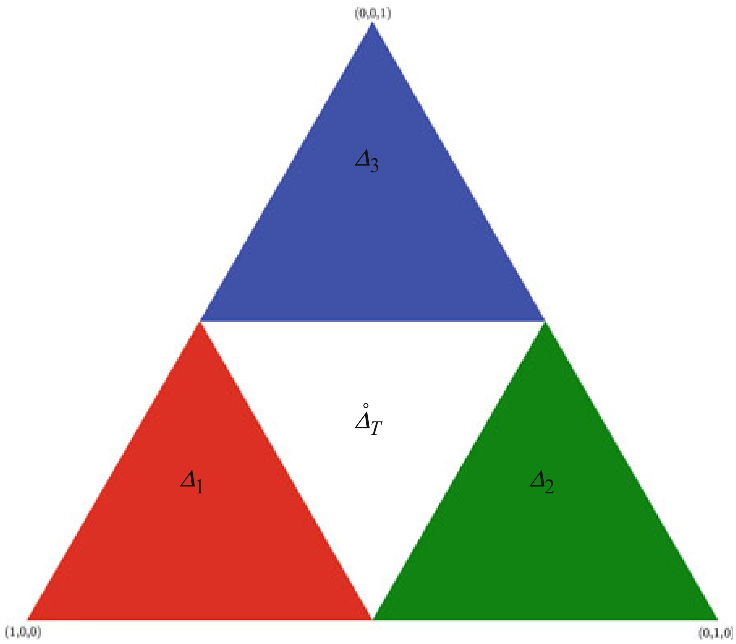

Fig. 2 The standard simplex Δ partitioned into four subsets $\Delta_1, \Delta_2, \Delta_3$, and $\mathring{\Delta}_T$

3.1 The Rauzy Gasket as an Iterated Function System

Let Δ denote the convex span of $\{e_1, e_2, e_3\}$, i.e.,

$$\Delta := \left\{(x_1,x_2,x_3) \in \mathbb{R}^3_{\geq 0} \mid \sum x_i = 1\right\}.$$

Let $\mathring{\Delta}_T$ denote the open set of triplets that satisfy the triangular inequalities $x_i < x_j + x_k$; it is the interior of the convex span of the centers of the sides of Δ:

$$\mathring{\Delta}_T := \left\{(x_1,x_2,x_3) \in \Delta \;\middle|\; \forall j, x_j < \sum_{i \neq j} x_i \right\}.$$

Furthermore, let us denote for all j

$$\Delta_j := \left\{(x_1,x_2,x_3) \in \Delta \;\middle|\; x_j \geq \sum_{i \neq j} x_i \right\}.$$

One has $\mathring{\Delta}_T = \Delta \setminus \bigcup \Delta_i$; see Fig. 2.

We consider the linear mapping $\tilde{F}$ defined on the set of strictly positive vectors which do not satisfy the triangle inequality by subtracting the two smaller coordinates form the larger one as follows:

$$\tilde{F}:(x_1,x_2,x_3)\mapsto\begin{cases}(x_1-x_2-x_3,x_2,x_3) & \text{if} \quad x_1\geq x_2+x_3,\\(x_1,x_2-x_1-x_3,x_3) & \text{if} \quad x_2\geq x_1+x_3,\\(x_1,x_2,x_3-x_1-x_2) & \text{if} \quad x_3\geq x_1+x_2,\end{cases}$$

Note that the definition is not consistent on the set of vectors of the form $(x,x,0)$, $(x,0,x)$, or $(0,x,x)$; we make an arbitrary choice in these three cases, which correspond to periodic words of period 2. As we have seen above, the renormalization operation is not well defined for these words.

This linear map, acting on the positive cone, gives rise to a projective map acting on $\Delta\setminus\mathring{\Delta}_T$; it will be denoted by $F:\Delta\setminus\mathring{\Delta}_T\mapsto\Delta$.

If $\mathbf{v}$ is any standard episturmian word and $\mathbf{u}$ is the corresponding renormalized word, one can now rewrite the formula from Proposition 2:

$$(f_{\mathbf{u}}(1),f_{\mathbf{u}}(2),f_{\mathbf{u}}(3))=F\left(f_{\mathbf{v}}(1),f_{\mathbf{v}}(2),f_{\mathbf{v}}(3)\right).$$

The map $\tilde{F}$ is 3-to-1, and the inversion $\tilde{F}^{-1}$ has three branches, denoted $\tilde{f}_i$, for $i\in\mathscr{A}$. These branches define projective maps $f_i:\Delta\to\Delta_i$. These maps correspond to linear maps given by matrices M_i, defined as:

$$M_1=\begin{pmatrix}1&1&1\\0&1&0\\0&0&1\end{pmatrix},\qquad M_2=\begin{pmatrix}1&0&0\\1&1&1\\0&0&1\end{pmatrix},\qquad M_3=\begin{pmatrix}1&0&0\\0&1&0\\1&1&1\end{pmatrix}.$$

We will use these inverse branches to find the set where F can be iterated infinitely many times; remark that $f_i(\Delta)=\Delta_i$ and $f_j(\Delta)=\Delta_j$, $i\neq j$ are disjoint except for one point.

It follows from Propositions 1, 2 that $\mathbf{R}$ satisfies

$$\mathbf{R}=\cup_{i\in\mathscr{A}}f_i(\mathbf{R})$$

so that $\mathbf{R}$ is the solution of an iterated function system. The function f_i being projective defined by positive matrix is contracting on the standard simplex; however, e_i, the ith vector of the standard basis of $\mathbb{R}^3$, is an indifferent fixed point for f_i, so this iterated function system is only weakly contracting, and a little care must be taken to apply the theorem of Hutchinson to prove unicity of the solution of this equation.

We define the family of maps $f_{i,j,n}$ by $f_{i,j,n}=f_i^n f_j$, for $i,j\in\mathscr{A}$, $i\neq j$, and $n\in\mathbb{N}^+$.

Lemma 2. *There exists a constant $c<1$ such that the maps $f_{i,j,n}$ are strict contractions with contraction ratio less than c.*

Proof. Since the f_i, defined by positive matrices on the positive simplex, are weak contractions, it is enough to prove it for $n=1$. Direct computation of the Jacobian matrix of f_1f_2 shows that, on the simplex, it is everywhere contracting by a contraction factor at least $\frac{2}{3}$. The result follows by symmetry for all f_if_j. □

Hence, the set of all the maps $f_{i,j,n}$ forms an infinite strictly contracting iterated function system; we can apply a modified version of Hutchinson's theorem. Let $H(\Delta)$ denote the set of all nonempty compact subsets of Δ, equipped with Hausdorff metric. Define $\Phi : H(\Delta) \mapsto H(\Delta)$ as

$$\Phi(X) = \overline{\bigcup_{i,j \in \mathscr{A}, i \neq j, n \in \mathbb{N}^+} f_{i,j,n}(X)}.$$

It is clear from the above lemma that Φ is a strict contraction on $H(\Delta)$; hence it has a unique fixed point, which is $\mathbf{R}$. Indeed, the analysis of the previous section showed that any nonconstant episturmian word $\mathbf{v}$, having a nonconstant directive word, can be renormalized as $\mathbf{v} = \sigma_i^n \sigma_j(\mathbf{u})$; restated in terms of frequency, this means that the Rauzy gasket satisfies:

$$\mathbf{R} = \{e_1, e_2, e_3\} \cup \bigcup_{i,j,n} f_{i,j,n}\mathbf{R}$$

from which it follows immediately that $\Phi(\mathbf{R}) = \mathbf{R}$.

We can now prove the main result of this section:

Theorem 2. *The Rauzy gasket* $\mathbf{R}$ *is the unique nonempty compact subset of the standard simplex which satisfies the equation:*

$$\mathbf{R} = \cup_{i \in \mathscr{A}} f_i(\mathbf{R}).$$

Proof. The only thing to prove is the uniqueness. Let X be another solution; from $X = \cup_{i \in \mathscr{A}} f_i(X)$, we obtain that $f_{i,j,n}(X) \subset X$; hence $\Phi(X) \subset X$. Let x be any element of X; by definition of X, we can find a sequence x_n in X, with $x_0 = x$, and an infinite word $(i_n)_{n \in \mathbb{N}}$, such that $x_n = f_{i_n}(x_{n+1})$. If the word is not constant, x is in some $f_{i,j,n}(X)$. If it is constant, x is in $\cap_{n \in \mathbb{N}} f_i^n(X)$; hence $x = e_i$. Hence $X = \{e_1, e_2, e_3\} \cup \left(\cup_{i,j \in \mathscr{A}, i \neq j, n \in \mathbb{N}^+} f_{i,j,n} X\right) \subset \Phi(X)$, so $\Phi(X) = X$ and $X = \mathbf{R}$. □

We note the following proposition, which might be useful to compute the Hausdorff dimension of $\mathbf{R}$:

Proposition 3. *The family* $\{f_{i,j,n}\}$ *satisfies the open set condition; that is, there is an open set* U *such that all the* $f_{i,j,n}(U)$ *are disjoint and contained in* U.

Proof. Let $\mathring{\Delta}$ be the interior of Δ. It is clear that $f_{i,j,n}(\mathring{\Delta}) \subset \mathring{\Delta}$; the explicit coordinates are easily computed, and the corresponding triangles are disjoint. Figure 3 shows the disposition of the triangles $f_{1,2,n}(\mathring{\Delta})$. □

3.2 Symbolic Dynamics for the Rauzy Gasket

The map F gives us a symbolic word associated with the elements of $\mathbf{R}$, unique except for the rational points. The easiest proof relies on the following lemma:

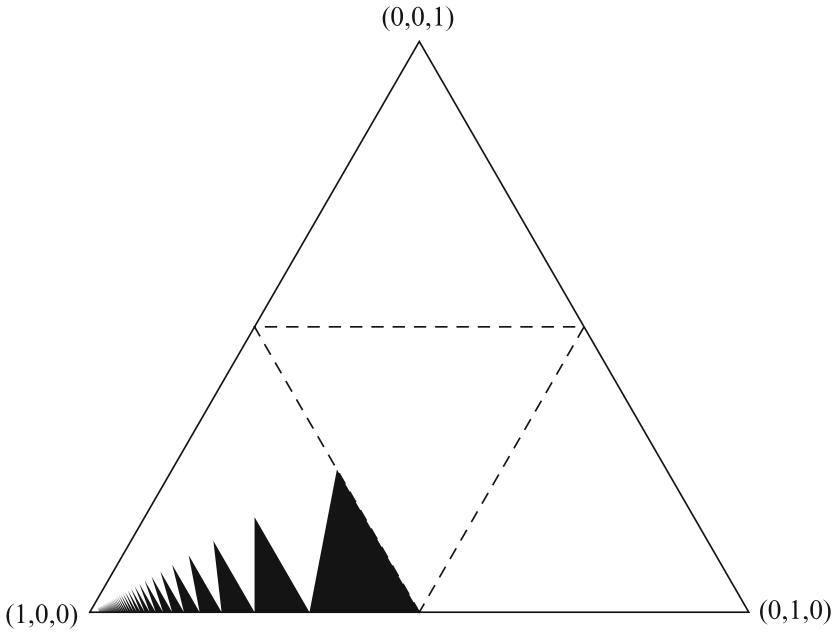

Fig. 3 The triangles $f_{1,2,n}(\mathring{\Delta})$

Lemma 3. *Let* (i_n) *be any infinite word in* $\mathscr{A}^\infty$*. The set* $\cap_{n\in\mathbb{N}} f_{i_0} f_{i_1} \dots f_{i_n}(\Delta)$ *contains exactly one point.*

Proof. Remark first that $f_1^n(\Delta)$ is the triangle with vertices $f_1^n(e_1) = (1,0,0)$, $f_1^n(e_2) = (\frac{n}{n+1}, \frac{1}{n+1}, 0)$, $f_1^n(e_3) = (\frac{n}{n+1}, 0, \frac{1}{n+1})$, whose diameter tends to 0; the intersection of these triangles is the point $(1,0,0)$. Similarly $f_j^n(\Delta)$ converges to $\{e_j\}$.

Hence, if the word (i_n) is eventually constant, i.e., for all $n > N$ we have $i_n = j$ for some $j \in \mathscr{A}$, the limit of the corresponding set is reduced to the point $f_{i_0} f_{i_1} \dots f_{i_N}(e_j)$.

If the word is not eventually constant, it can be decomposed in a unique way as a product of $f_{i,j,n}$; hence the diameter of the images goes to zero, so the intersection of this sequence of decreasing compact sets is reduced to a point. □

Definition 8. The symbolic coding for the Rauzy gasket is the map, $\pi_F : \mathscr{A}^\infty \to \mathbf{R}$, which associates to any infinite word (i_n) the unique point defined by the previous lemma.

This map is one-to-one, except on eventually constant words, corresponding to rational points, where it is 2-to-one. In the previous setting, the reciprocal map is easy to describe. Let ν be the coding map: $\cup_{i\in\mathscr{A}} \Delta_i \to \mathscr{A}$ defined by $\nu(x) = i$ if $x \in \Delta_i$ (this map is ill-defined on the three middle points, elements of $\Delta_i \cap \Delta_j$, $i \neq j$); the coding word of x is the word $(\nu(F^n(x)))_{n\in\mathbb{N}}$. It is well defined on the irrational points, and one could easily and tediously give the descriptions of the two coding words for the rational points.

Remark that π_F is obviously continuous; it is true, but less obvious, that the reciprocal map is continuous where it is well defined.

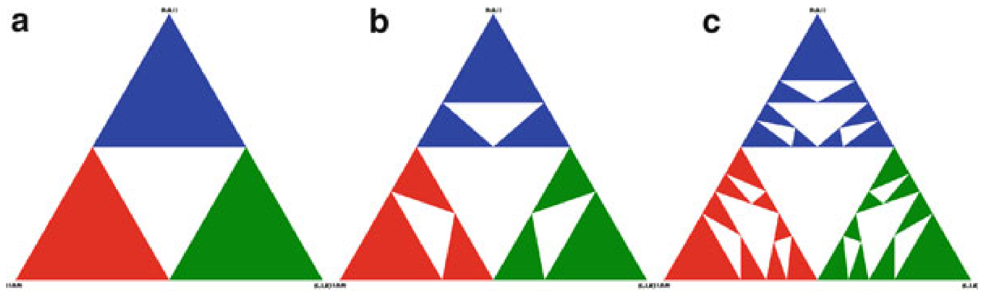

Fig. 4 (**a**) The set K_1. (**b**) The set K_2. (**c**) The set K_3

Remark 3. To get an explicit approximation of **R**, a first possibility is to start with $V_0 = \{e_1, e_2, e_3\}$ and to build an increasing sequence. Define $\Psi(X) = \cup_{i\in\mathscr{A}} f_i(X)$ and consider the recurrent sequence given by $V_{n+1} = \Psi(V_n)$; the union of V_n for all n is the set of rational points of **R**, and its closure is the Hausdorff limit **R** of the sequence V_n.

One can of course obtain a decreasing sequence by removing triangles. Denote $K_0 = \Delta$ and $K_n = \Psi^n(K_0)$. We get a decreasing sequence which converges to **R**. Figure 4 shows K_1, K_2, and K_3.

Each set K_n can bee seen as a union of 3^n triangles. The set V_n as defined above is the set of the vertices of those triangles.

One easily shows that the sets $\mathbf{R}_{\text{aper}}$ and $\mathbf{R}_{\text{AR}}$ are (noncompact) solutions to the equation $\Psi(X) = X$; they also satisfy the equation $X = \cup_{i,j,n} f_{i,j,n}(X)$ (without taking the closure here). One can also show that, if $\mathring{\Delta}$ is the interior of Δ, we have $\mathbf{R}_{\text{AR}} = \cap_n \Psi^n(\mathring{\Delta})$.

4 Relation with the Sierpiński Gasket and a Generalization of the Question Mark Function

The above properties, in particular the approximation by the sets K_n and the three types of points in **R** corresponding to periodic, non-strict episturmian, and AR words, are reminiscent of the topology of the Sierpiński gasket. We will show that this set is in fact homeomorphic to the Sierpiński gasket; we first recall basic facts about the Minkowski question mark function.

4.1 The Minkowski Question Mark Function

Dynamical systems generated by rotation words are completely determined by the frequency x of the letter 0. As we recalled in Sect. 2.1, such a system can be

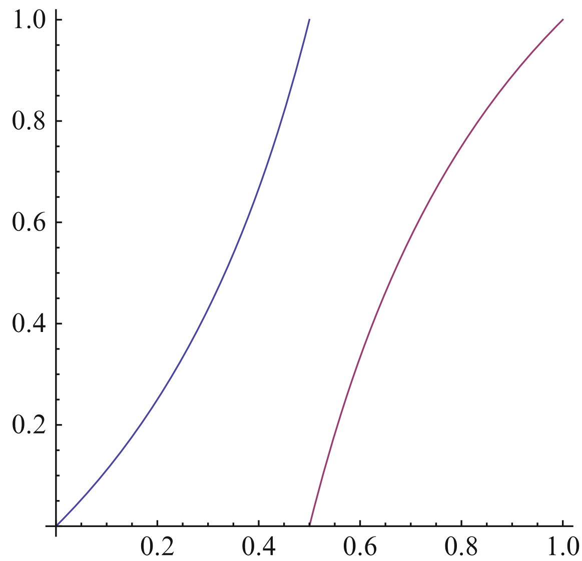

Fig. 5 The additive continued fraction map

renormalized by erasing any occurrence of the most frequent letter following the other letter; a simple computation shows that the frequency of 0 in the new system is $\phi(x)$, where ϕ is defined as

$$\phi : [0,1] \to [0,1] : \qquad x \mapsto \begin{cases} \frac{x}{1-x} & \text{if} \quad x < \frac{1}{2}, \\ 2 - \frac{1}{x} & \text{if} \quad x \geq \frac{1}{2}. \end{cases}$$

This map, represented in Fig. 5, is an exotic version of the usual additive continued fraction map; it has two indifferent fixed repelling points in 0 and 1, and it is ill-defined in $\frac{1}{2}$ (as we will see, the choice of the value 0 or 1 for $\phi(\frac{1}{2})$ is irrelevant; we have chosen here $\phi(\frac{1}{2}) = 0$).

By using the coding function ν defined by $\nu(x) = 0$ if $x < \frac{1}{2}$ and $\nu(x) = 1$ if $x \geq \frac{1}{2}$, we can associate to any $x \in [0,1]$ a coding word $\nu_\phi(x) = (\nu(\phi^n(x)))_{n\in\mathbb{N}}$. This defines a map $\nu_\phi : [0,1] \to \{0,1\}^{\mathbb{N}}$, which is one-to-one, and avoids only words which are eventually constant of value 1.

It is easy to prove that this map is increasing for the usual lexicographic order on $\{0,1\}^{\mathbb{N}}$, and that there is a reciprocal function which is increasing and one-to-one except on the set of eventually constant words.

One can do exactly the same thing with the function $\gamma : [0,1] \to [0,1] : \; x \mapsto 2x \mod 1$ and define a coding map $\nu_\gamma : [0,1] \to \{0,1\}^{\mathbb{N}} : \; x \mapsto (\nu(\psi^n(x)))_{n\in\mathbb{N}}$. This is the usual binary expansion of real numbers, which is also increasing and whose reverse map is defined, for any binary sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, by $\nu_\gamma^{-1}(\varepsilon) = \sum_{n=0}^{\infty} \frac{\varepsilon_n}{2^{n+1}}$.

Definition 9. The question mark function ? of Minkowski is defined by ?: $[0,1] \to [0,1] \quad x \mapsto ?(x) = \nu_\gamma^{-1}\left(\nu_\phi(x)\right)$.

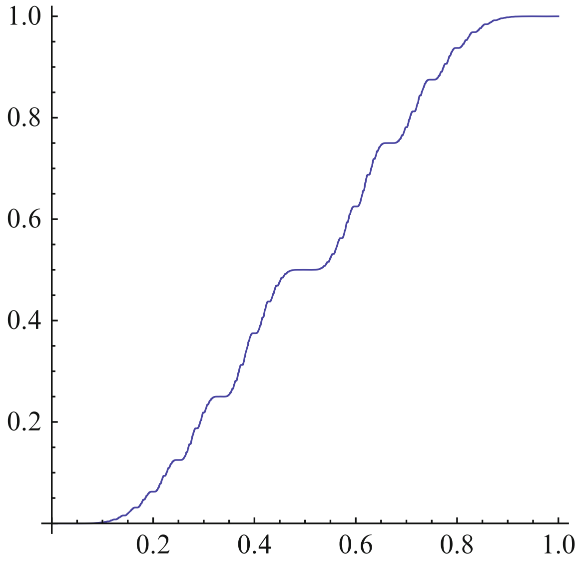

Fig. 6 The Minkowski ? function, also known as the slippery devil's staircase

The graph of this function is given in Fig. 6. The following properties of the function ? are easy to prove:

- The function ? is an increasing homeomorphism from $[0,1]$ to itself.
- It takes all rational numbers to dyadic numbers.
- It takes all quadratic numbers to rational numbers.
- It conjugates ϕ and γ: $\phi = ? \circ \gamma \circ ?^{-1}$.
- It has derivative 0 in 0 and in all rational numbers.

One can also prove that it is not absolutely continuous. Another way to define ? is to send the Farey set of order n to the set of all dyadic rationals between 0 and 1 such that their denominator in completely reduced form equals 2^k for some $0 \le k \le n$, preserving the order, and to show that this extends by continuity to an homeomorphism of $[0,1]$ to itself.

4.2 The Sierpiński Gasket

We consider the iterated function system $\{g_1, g_2, g_3\}$, where g_i is defined on Δ by $g_i(x_1,x_2,x_3) = \frac{(x_1,x_2,x_3)+e_i}{2}$.

The g_i are strict contractions with factor $\frac{1}{2}$, so the operator $H(\Delta) \to H(\Delta)$ given by $X \mapsto \bigcup_{i\in\mathscr{A}} g_i(X)$ has a unique fixed point, which is called the Sierpiński gasket; see Fig. 7. It will be denoted by $\mathbf{S}$.

Let G be the map defined on $\cup_{i\in\mathscr{A}}\Delta_i$ by $G(x_1,x_2,x_3) = 2(x_1,x_2,x_3) - e_i$ if $(x_1,x_2,x_3) \in \Delta_i$; it is 3-to-1, and its reciprocal branches are the contracting maps g_i. The Sierpiński gasket is the set on which G can be iterated infinitely, and by the dynamical system of the Sierpiński gasket we understand the dynamical system $(\mathbf{S}, G)$.

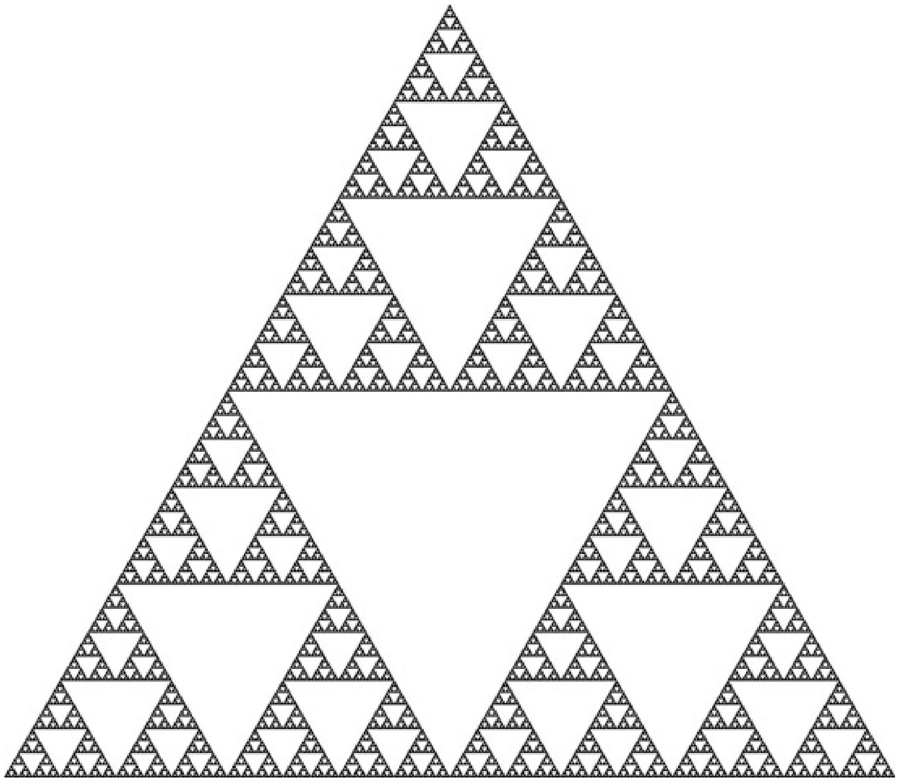

Fig. 7 The Sierpiński gasket

4.3 A Generalization of the Minkowski Question Mark Function

We can define a coding for the Sierpiński gasket, as for the Rauzy gasket, by $\nu_G(x) = (\nu(G^n(x))_{n\in\mathbb{N}}$. It is easy to prove that this coding is well defined, except for points with dyadic coordinates, where there are two possible codings, and that the map is continuous except for these points with dyadic coordinates; the reverse map π_G associates to any symbolic sequence (i_n) of elements of $\mathscr{A}$ the unique point in $\cap_{n\in\mathbb{N}} g_{i_0} g_{i_1} \dots g_{i_n}(\Delta)$. It is continuous.

We can now define a generalization of the Minkowski question mark function.

Proposition 4. *The map $\Theta = \pi_G \circ \nu_F : \mathbf{R} \to \mathbf{S}$ is well defined and continuous.*

Proof. The map is clearly well defined, except for rational points, which may have two codings. A direct study shows that the point $(\frac{1}{2}, \frac{1}{2}, 0)$ admits the codings 21^∞ and 12^∞ and that these two codings have the same image (again $(\frac{1}{2}, \frac{1}{2}, 0)$) by π_G, so the image does not depend on the choice of the coding and is well defined. This property easily extends to all rational points.

Continuity is clear for the irrational points, since ν_F is continuous in these points. A local study shows that symbolic coding of points close to $(\frac{1}{2}, \frac{1}{2}, 0)$ must have a long prefix common with one of the two possible codings for this point; hence their images by Θ are close, which proves the continuity at this point. A similar proof works for any rational point. □

Proposition 5. *The dynamical systems* $(\mathbf{S}, G)$ *and* $(\mathbf{R}, F)$ *are conjugate by* Θ.

Proof. This is an immediate consequence of the fact that π_F and π_G conjugate, respectively, except for a countable set, $(\mathbf{S}, G)$ and $(\mathbf{R}, F)$ to the shift on $\mathscr{A}^\infty$. □

Proposition 6. *The restriction of* Θ *to the segment of the boundary of* Δ *joining* e_1 *and* e_2 *is the Minkowski ? function.*

Proof. It suffices to remark that the restriction of F to this segment is exactly the function ϕ and the restriction of G is the function γ, so the conjugacy must be the question mark function. □

Remark 4. Another higher dimensional generalization of the question mark function has been described in [12].

5 The Apollonian Gasket

The Apollonian gasket $\mathbf{A}$ can be described as follows: consider three pairwise tangent circles, which define a curvilinear triangle in the complex plane. Remove from this triangle the unique disk which is tangent to the three circles; we obtain three smaller triangles, each delimited by three pairwise tangent circles, and we can iterate the procedure. The limit set is the Apollonian gasket.

Although it might seem to depend on the initial configuration of circles, there is only one Apollonian gasket up to conjugacy by a Möbius transformation. Indeed, the triangle of tangency points completely determines the centers of the three circles, which are on the tangents to the circumscribed circle at the tangency points; but the group of Möbius transformations acts transitively on the set of triangles; since Möbius transformations preserve circles, this action extends to the family of Apollonian gaskets.

It will be convenient to take as tangency points $0, 1$, and $\frac{1+i}{2}$, so the circles are the circles C_1, C_2 with radius $\frac{1}{2}$ and respective centers $1+\frac{i}{2}$ and $\frac{i}{2}$ and the horizontal axis, which is the generalized circle C_3 with infinite radius (see Fig. 8). We will call $\mathbf{A}$ the subset of $\mathscr{C}$ defined by this gasket.

The tangent circle to these three circles in the bounded region is the circle C of center $\frac{1}{2}+\frac{i}{8}$ and radius $\frac{1}{8}$. Note that C, C_j, C_k also define a new version of the Apollonian gasket, which we will denote by $\mathbf{A}_i$. We can find a Möbius transformation h_i which preserves C_j and C_k and sends C_i to C.

We denote by h_i the corresponding matrix in $SL(2,\mathscr{C})$. Computation shows that

$$h_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad h_2 = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}, \qquad h_3 = \begin{pmatrix} i & 1 \\ 2i & 2-i \end{pmatrix}.$$

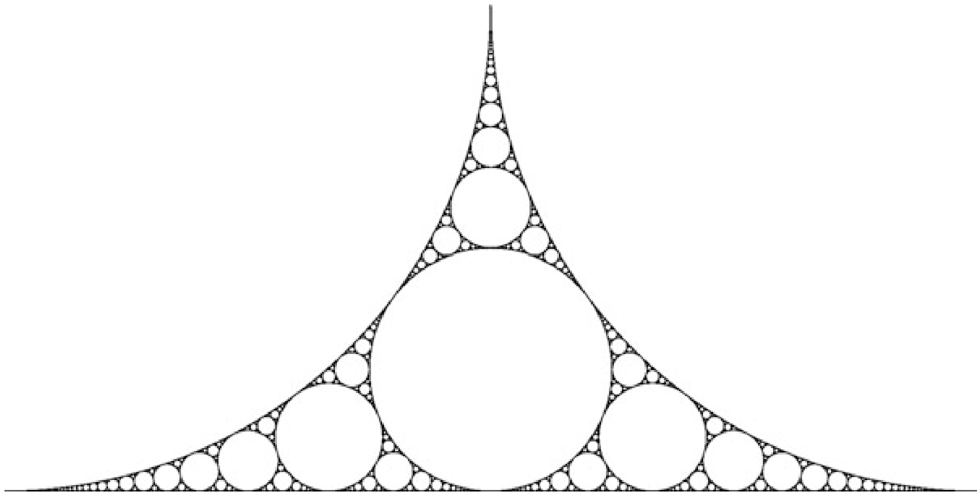

Fig. 8 The Apollonian gasket

Since the map h_i sends $\mathbf{A}$ to $\mathbf{A}_i$, the Apollonian gasket satisfies $\mathbf{A} = \cup_{i\in\mathscr{A}} h_i\mathbf{A}$; it is the solution of a conformal parabolic IFS. It has been thoroughly investigated (see for example [9, 10, 15]). Its exact Hausdorff dimension is not known, but it has been proved that its Hausdorff measure is finite.

By taking the inverse of the h_i, one can define a map $H : \mathbf{A} \to \mathbf{A}$ which is 3-to-1 and a coding map ν_H. Exactly as in the previous section, one can prove the following proposition:

Proposition 7. *There exists a homeomorphism* $\mathbf{R} \to \mathbf{A}$ *which conjugates the dynamical system* $(\mathbf{R}, F)$ *to* $(\mathbf{A}, H)$.

This map is certainly not a diffeomorphism, since it takes an equilateral triangle to a curvilinear triangle with angles 0; for the same reason, it cannot be a diffeomorphism in any rational point of the Rauzy gasket. It is however more regular than the conjugacy defined in the previous section.

Proposition 8. *The restriction of the conjugacy to the lower boundary of the Rauzy gasket in the identity.*

Proof. We already remarked in the proof of Proposition 6 that the restriction of F to this lower boundary was the function ϕ defined by $\phi(x) = \frac{x}{1-x}$ if $x < \frac{1}{2}$ and $\phi(x) = 2 - \frac{1}{x}$ if $x \geq \frac{1}{2}$; but the formulas above show that the restriction of H to the segment $[0,1]$ is given by the same formula, hence the result. $\square$

It follows immediately that the restriction of the map to the boundary of any triangle in the complement of the Rauzy gasket is smooth; that is, the restriction to the irrational points of the complement of $\mathbf{R}_{\mathrm{AR}}$ is smooth.

6 Relation with the Fully Subtractive Algorithm

6.1 The Fully Subtractive Algorithm

The fully subtractive algorithm has been treated for instance in [8,11]. We first recall its definition and some results.

The fully subtractive algorithm is defined on the positive cone $\mathbb{R}^3_{\geq 0}$; it subtracts the smallest number from the two others, i.e., it is given by the map $\tilde{S} : \mathbb{R}^3_{\geq 0} \mapsto \mathbb{R}^3_{\geq 0}$ defined by

$$\tilde{S} : (x_1, x_2, x_3) \mapsto \begin{cases} (x_1, x_2 - x_1, x_3 - x_1) & \text{if} \quad x_1 \leq x_2, x_1 \leq x_3, \\ (x_1 - x_2, x_2, x_3 - x_2) & \text{if} \quad x_2 \leq x_1, x_2 \leq x_3, \\ (x_1 - x_3, x_2 - x_3, x_3) & \text{if} \quad x_3 \leq x_1, x_3 \leq x_2, \end{cases}$$

Note that the definition is again not completely consistent; for the set of vectors having two coordinates equal to each other we make an arbitrary choice. Since the algorithm is clearly equivariant under permutations of coordinates, the algorithm is often defined on the quotient space given by $x_1 \leq x_2 \leq x_3$, with a reordering of the image vector; this removes the problem, but makes the geometry less clear.

By considering the action of $\tilde{S}$ on projective space, we can define a map $S : \Delta \to \Delta$, with barycentric coordinates. If one coordinate is 0, the point is fixed. Thus, the set of fixed points of S is the boundary of Δ. The map S is 3-to-1; its restriction to the set Γ_i defined by $x_i \leq \inf(x_j, x_k)$ is a homeomorphism from Γ_i to Δ.

Computation shows that the segment $x_i = \frac{1}{2}$ is invariant by S. Indeed, if $z < y < \frac{1}{2}$ and $y + z = \frac{1}{2}$, we have $\tilde{S}(\frac{1}{2}, y, z) = (\frac{1}{2} - z, y - z, z)$, so that after renormalization $S(\frac{1}{2}, y, z) = (\frac{1}{2}, \frac{y-z}{1-2z}, \frac{z}{1-2z})$. It follows that the restriction of S to Δ_i preserves this triangle.

The restriction of S to the segment $(\frac{1}{2}, \frac{1}{2} - z, z)$ is conjugate by $(\frac{1}{2}, \frac{1}{2} - z, z) \mapsto 2z$ to the map ϕ of Sect. 4.1. The points $(\frac{1}{2}, \frac{1}{2} - z, z)$, with $z \in \mathbb{Q}$ of this segment, are sent by a power of S first to $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ and then, after an arbitrary choice (since S is ill-defined in this point) to one of the endpoints $(\frac{1}{2}, \frac{1}{2}, 0)$ or $(\frac{1}{2}, 0, \frac{1}{2})$ which are fixed by S. The other points have their orbit contained in the interior of the segment. By linearity, the map S sends the segment joining the fixed point e_i to a point P on $x_i = \frac{1}{2}$ to the segment joining e_i to $S(P)$. Hence the orbit of any point in the interior of Δ_i either ends in a finite number of steps on the boundary (if it is on the segment joining e_i to a rational point) or tends to the vertex e_i, since computation shows that in that case the coordinate x_i tends to a limit which can only be 1; this last set has obviously full measure in Δ_i.

Hence there are three attractors of the system (Δ, S)—the vertices of the standard simplex. Figure 9 shows the action of S on Γ_3, with the preimages by S of the four triangles Δ_T and Δ_i, for $i \in \mathscr{A}$. Figure 10 shows the three basins of attraction, distinguished by different colors.

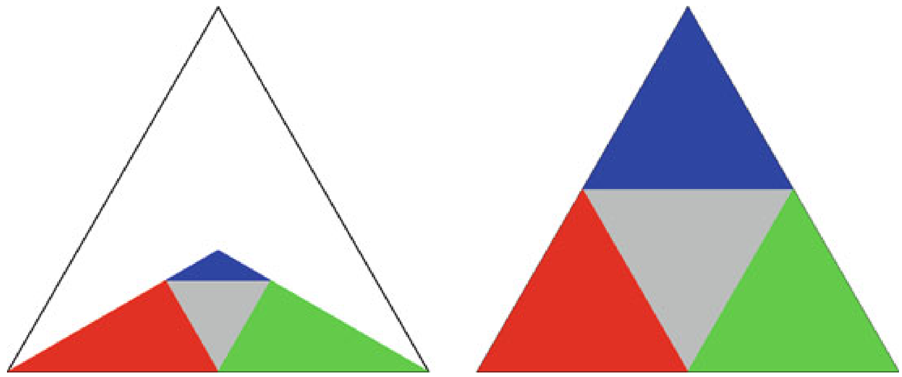

Fig. 9 The action of S on Γ_3

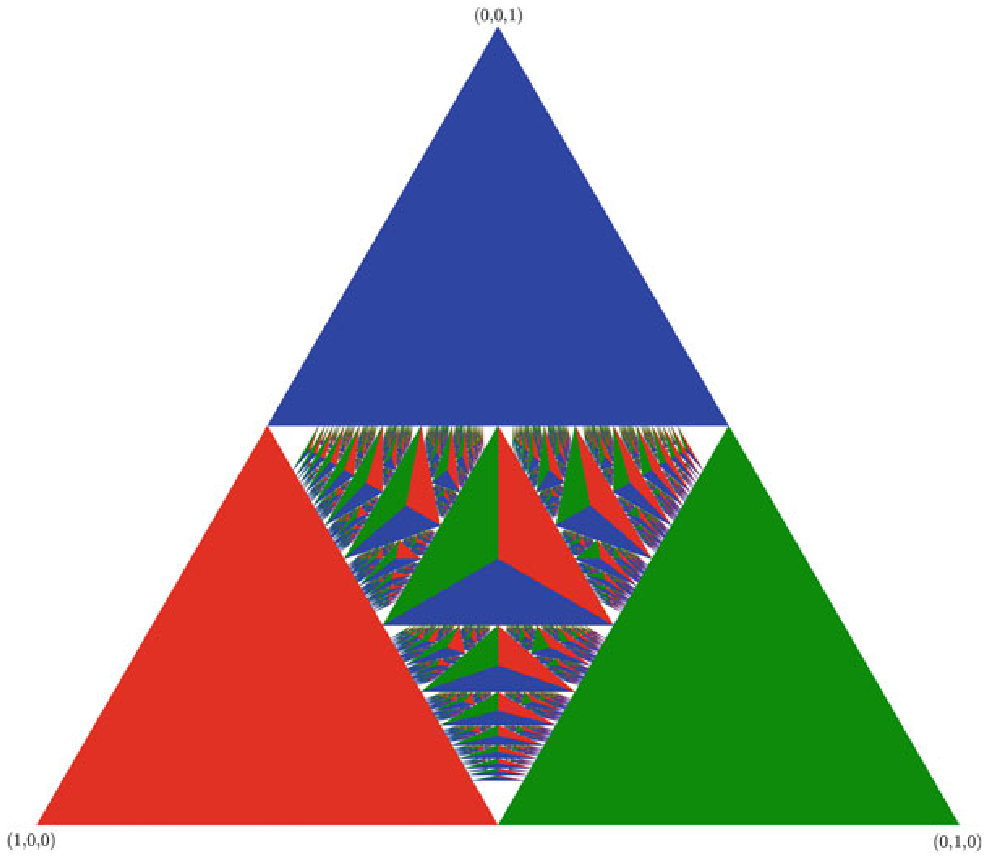

Fig. 10 Basins of attraction of the dynamical system (Δ, S). (More precisely, the set $S^{-8}(\Delta_i)$ is depicted for all i and colored by a different color)

6.2 The Fully Subtractive Algorithm as an Extension of the Rauzy Gasket

The map F associated with the Rauzy gasket was not defined on the central triangle; we will now extend it on all of Δ, by enlarging the set of definition to include points with negative coordinates. Let Δ' denote the convex span of $\{(-1,1,1),(1,-1,1),(1,1,-1)\}$, and let $\Delta'_i = \{(x_1,x_2,x_3) \in \Delta' \mid x_i \geq x_j, x_k\}$. Note that, with the notations of Sect. 3, Δ'_i contains Δ_i; we extend F to a map F' on Δ' by extending to Δ'_i the formula on Δ_i. It is the projective map associated with the piecewise linear map $\tilde{F}'$:

$$\tilde{F}' : (x_1,x_2,x_3) \mapsto \begin{cases} (x_1-x_2-x_3,x_2,x_3) & \text{if} \quad x_1 \geq x_2, x_1 \geq x_3, \\ (x_1,x_2-x_1-x_3,x_3) & \text{if} \quad x_2 \geq x_1, x_2 \geq x_3, \\ (x_1,x_2,x_3-x_1-x_2) & \text{if} \quad x_3 \geq x_1, x_3 \geq x_2. \end{cases}$$

Proposition 9. *(Δ',F') is conjugate to (Δ,S).*

Proof. It is enough to prove it for the piecewise linear maps $\tilde{F}'$ and $\tilde{S}$. Let P be the matrix that sends the canonical basis to the vertices of Δ'; let A (resp. B) be the matrix of $\tilde{F}'$ on the cone on Δ'_1 (resp. the matrix of $\tilde{S}$ on the cone on Γ_1). We have

$$\tilde{P} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Computation shows that $P(\Gamma_i) = \Delta'_i$ and that $B = P^{-1}AP$. □

In Fig. 10, the Rauzy gasket appears as the complement of the three basins of attraction; the dynamical system of the Rauzy gasket is the chaotic part of the fully subtractive algorithm.

6.3 Two Properties of the Rauzy Gasket

It is known that the fully subtractive algorithm is, in continued fraction terms, not convergent: for almost every point, the symbolic dynamics does not define the point. More precisely, we have the following:

Theorem 3 ([8,11], Theorem 1). *For almost all $x \in \mathbb{R}^3_{\geq 0}$ we have*

$$\lim_{j\to+\infty} \tilde{S}^j(x) \neq (0,0,0).$$

We can restate the theorem in terms of the projective map on the simplex:

Corollary 2. *Almost any point of Δ tends to one of the three vertices under the action of S.*

Since the map P sends the Rauzy gasket to the complement of the attraction basins, we obtain the following:

Corollary 3. *The set* **R** *has zero Lebesgue measure.*

Remark 5. We could give a direct proof of this corollary along the line of the proof of [11]. We consider the restriction of F' on Δ, and we want to prove that, almost surely, it cannot be iterated infinitely inside Δ. Since $F'(\mathring{\Delta}_T)$ is disjoint from Δ, it suffices to consider the restriction to one of the Δ_i, say Δ_1. The main problem is the indifferent fixed point in e_1; to avoid this, we define, for any $x \in \Delta_1$, the integer n_x which is the smallest integer such that $F'^{n_x}(x)$ is not in Δ_1, and we consider the map $x \to F'^{n_x}(x)$. One computes explicitly the continuity domain of this map and its reciprocal branches; one then shows that the branches are uniformly contracting and that the Jacobian has a bounded distortion property. We can then consider the cylinders defined by the symbolic dynamics associated with this map and prove that the proportion of any cylinder which goes to Δ_T and leaves the simplex under the next iteration is bounded from below, which implies the corollary.

Proposition 10. *For all $y \in P^{-1}\mathbf{R}$, $\delta > 0$ and $i \in \mathscr{A}$, there exists $y_i \in \mathring{\Delta}$, $|y - y_i| < \delta$ such that*

$$\lim_{j \to +\infty} S^j(y_i) = e_i.$$

In other words, any uncolored point in $\mathring{\Delta}$ in Fig. 10 has all three colors in any of its neighborhood.

Proof. Let us first denote A_i the basin of attraction to the attractor e_i:

$$A_i = \left\{ x \in \mathring{\Delta} \;\middle|\; \lim_{j \to +\infty} S^j(x) = e_i \right\}.$$

Using Remark 3 one can see that $\mathbf{R} = \lim_{j \to +\infty}(V_j)$ where V_j is the set of antecedents of order j by F of the three vertices. Thus it suffices to show that for all j that every point $x \in P^{-1}(V_j)$ lies on all three boundaries of the sets A_i for all i. The proof is by induction on j.

The point $P^{-1}e_1 = (0, \frac{1}{2}, \frac{1}{2})$ is on the boundary of Δ_2 and Δ_3; it is also the limit of the points $\left(\frac{1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1}\right)$ which are preimages of e_1 under S; hence $P^{-1}V_0$ lies on the boundary of the sets A_i.

Suppose it is true for $j < N$. We apply S^{-1} to both sides of the equation $P^{-1}(V_{N-1}) \subset \bigcap_{i \in \mathscr{A}} \overline{A_i}$. On the left-hand side we have $S^{-1}P^{-1}(V_{N-1}) = P^{-1}F'^{-1}(V_{N-1}) = P^{-1}(V_N)$. And on the right-hand side we have $S^{-1}\bigcap_{i \in \mathscr{A}} \overline{A_i} = \bigcap_{i \in \mathscr{A}} \overline{A_i}$.

□

7 Final Remarks

We have restricted ourselves to ternary Arnoux–Rauzy words. However, the definition of episturmian words immediately extends to any finite alphabet, with the same renormalization procedure related to the fully subtractive algorithm, and we can define a Rauzy gasket in dimension d. Since the result for the fully subtractive algorithm in [8] is valid for any dimension, we can use it to prove that the Rauzy gasket in dimension d has Lebesgue measure 0.

The Rauzy gasket can be seen as a generalized Julia set for the dynamical system associated to the subtractive algorithm, and it shares some properties of a Julia set. One would like to know more about the Hausdorff dimension of **R** and the invariant measure of the underlying dynamical system; a first step should be to understand better the conjugacy with the Apollonian gasket: can we extend the regularity found on the boundary? Does it preserve Hausdorff dimension and measure? This would not completely solve the problem, since the Hausdorff dimension of the Apollonian gasket is not exactly known, but it is known (see [9, 15]) that its Hausdorff measure is finite.

It is a curious fact that the map S is dual (in the linear algebra sense) of F; this can be used to give a natural extension of the dynamical system of the Rauzy gasket as the skew product:

$$\overline{F} : \mathbf{R} \times P^{-1}\mathbf{R} \to \mathbf{R} \times P^{-1}\mathbf{R} \qquad (x,y) \mapsto (f_i^{-1}(x),\ {}^t f_i(y)) \text{ if } x \in \Delta_i,$$

where f_i is branch of F^{-1} such that $f_i(\Delta) = \Delta_i$. This remark might be useful to study the invariant measures for F.

Acknowledgments The second author acknowledges 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. Arnoux, P., Rauzy, G.: Représentation géométrique de suites de complexité $2n+1$. Bull. Soc. Math. France **119**, 199–215 (1991)
2. Boshernitzan, M.: A unique ergodicity of minimal symbolic flows with linear block growth. J. Anal. Math. **46**, 196–213 (1993)
3. Cassaigne, J., Ferenczi, S., Zamboni, L.Q.: Imbalances in Arnoux-Rauzy sequences. Ann. Inst. Fourier **50**, 1265–1276 (2000)
4. Droubay, X., Justin, J., Pirillo, G.: Episturmian words and some constructions of de Luca and Rauzy. Theoret. Comput. Sci. **255**(1-2), 539–553 (2001)
5. Fogg, N.P.: Substitutions in Arithmetics, Dynamics and Combinatorics, 1st edn. Springer, New York (2002)
6. Glen, A., Justin, J.: Episturmian words: a survey. Theoret. Inf. Appl. **49**(3), 403–442 (2009)
7. Justin, J., Pirillo, G.: Episturmian words and episturmian morphisms. Theoret. Comput. Sci. **276**(1-2), 281–313 (2002). doi:http://dx.doi.org/10.1016/S0304-3975(01)00207-9
8. Kraaikamp, C., Meester, R.: Ergodic properties of a dynamical system arising from percolation theory. Ergod. Theor. Dyn. Syst. **15**, 653–661 (1995)
9. Mauldin, R.D., Urbański, M.: Dimension and measures for a curvilinear Sierpiński gasket or Apollonian packing. Adv. Math. **136**(1), 26–38 (1998)

10. McMullen, C.T.: Hausdorff dimension and conformal dynamics III: computation of dimension. Am. J. Math. **120**(4), 691–72 (1998)
11. Meester, R.W.J., Nowicki, T.: Infinite clusters and critical values in two dimensional circle percolation. Isr. J. Math. **68**(1), 63–81 (1989)
12. Panti, G.: Multidimensional continued fractions and a Minkowski function. Monatsh. Math. **154**, 247–264 (2008)
13. Rauzy, G.: Nombres algébriques et substitutions. Bull. Soc. Math. France **110**, 147–178 (1982)
14. Rauzy, G.: Suites à termes dans un alphabet fini. Séminaire de Théorie des Nombres de Bordeaux **Année 1982–1983**(exposé 25) (1983)
15. Sullivan, D.: Entropy, Hausdorff measure old and new, and limit sets of geometrically finite Kleinian groups. Acta Math. **153**, 259–277 (1984)
16. Wozny, N., Zamboni, L.Q.: Frequencies of factors in Arnoux-Rauzy sequences. Acta Arith. **96**, 261–278 (2001)

Chapter 11

Markov constant and quantum instabilities

Authors:	E. Pelantová, Š. Starosta, and M. Znojil
Publication reference:	J. Phys. A: Math. Theor., 49 (2016), p. 155201
Cited as:	[X]

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **49** (2016) 155201 (19pp) doi:10.1088/1751-8113/49/15/155201

Markov constant and quantum instabilities

Edita Pelantová[1], Štěpán Starosta[2] and Miloslav Znojil[3]

[1] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic
[2] Faculty of Information Technology, Czech Technical University in Prague, Prague, Czech Republic
[3] Nuclear Physics Institute ASCR, Řež, Czech Republic

E-mail: stepan.starosta@fit.cvut.cz

Received 8 October 2015, revised 26 January 2016
Accepted for publication 4 February 2016
Published 1 March 2016

Abstract

For a qualitative analysis of spectra of certain two-dimensional rectangular-well quantum systems several rigorous methods of number theory are shown productive and useful. These methods (and, in particular, a generalization of the concept of Markov constant known in Diophantine approximation theory) are shown to provide a new mathematical insight in the phenomenologically relevant occurrence of anomalies in the spectra. Our results may inspire methodical innovations ranging from the description of the stability properties of metamaterials and of certain hiddenly unitary quantum evolution models up to the clarification of the mechanisms of occurrence of ghosts in quantum cosmology.

Keywords: renormalizable quantum theories with ghosts, Pais–Uhlenbeck model, singular spectra, square-well model, number theory analysis, physical applications, metamaterials

1. Introduction

The main *mathematical* inspiration of our present physics-oriented paper may be traced back to the theory of Diophantine approximations in which an important role is played by certain sets of real numbers possessing an accumulation point called Markov constant [1]. The related ideas and techniques (to be shortly outlined below) are transferred to an entirely different context. Briefly, we show that and how some of the results of number theory may appear applicable in an analysis of realistic quantum dynamics.

The sources of our *phenomenological* motivation are more diverse. Among them, a distinct place is taken by the problems of quantum stability which are older than the quantum theory itself. Their profound importance already became clear in the context of the Niels Bohr's model of atom [2]. In this light one of the main achievements of the early quantum theory may be seen precisely in the explanation of the well verified experimental observation that many quantum systems (like hydrogen atom, etc) are safely stable.

During the subsequent developments of the quantum theory, the rigorous mathematical foundation of the concept of quantum stability found its safe ground in the spectral theory of self-adjoint operators in Hilbert space [3]. Although it may sound like a paradox, a similar interpretation of the *loss* of quantum stability is much less developed at present. This does not imply that the systematic study of instabilities would be less important. The opposite is true because the majority of existing quantum systems ranging from elementary particles to atomic nuclei and molecules are unstable.

In this direction of study one could only feel discouraged by the fact that the existing theoretical descriptions of quantum instabilities require complicated mathematics, be it in quantum field theory, in statistical quantum physics or, last but not least, in the representations of quantum models using non-selfadjoint operators [4]. For this reason we believe that our present approach combining a sufficiently rigorous level of mathematics with a not too complicated exemplification of quantum systems might offer a fresh and innovative perspective to quantum physics and, in particular, to some of its stability and instability aspects.

It is certainly encouraging for us to notice that a combination of Diophantine analysis with phenomenological physics already appeared relevant in the context of study of certain stable quantum systems controlled by point interactions and living on rectangular lattices [5] or on hexagonal lattices [6], where typically, the band spectra may depend on certain number-theoretical characteristics of the system. In what follows, we intend to turn our attention from complicated quantum graphs to a maximally elementary and exactly solvable model in which the hyperbolic partial differential equation

$$\Box f(x, y) = \lambda f(x, y), \quad \Box = \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}, \quad f|_{\partial R} = 0 \tag{1}$$

is studied and in which the instability is immanently present, in a way to be discussed below, via the unboundedness of the spectrum from below.

In our model the eigenfunctions are required to satisfy the most common Dirichlet boundary conditions, i.e., they are expected to vanish along the boundary of the two-dimensional rectangle

$$R = \{(x, y) : 0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b\}. \tag{2}$$

In sections 2–5 we describe and prove rigorous results of analysis of such a model. After a systematic presentation of these mathematical observations we return, in sections 6 and 7, to the problem of their various potential connections with physics. We also list there a few not entirely artificial samples of placing the Klein–Gordon-resembling equation (1) into a broader phenomenological context.

2. Spectral problem

2.1. Separation of variables

Our present analysis is fully concentrated upon the properties of spectra of hyperbolic partial differential operators $\Box$ of equations (1) + (2) which act upon twice differentiable functions $f(x, y)$ of two real variables. Setting $f(x, y) = g(x)h(y)$ we find that the eigenvalue problem is easily solvable by separation of variables, i.e., that there exist constants C and D such that

$$\frac{\Box f(x, y)}{f(x, y)} = \frac{\frac{\partial^2 g(x)}{\partial x^2}}{g(x)} - \frac{\frac{\partial^2 h(y)}{\partial y^2}}{h(y)} = C - D = \lambda.$$

The solution of the corresponding ordinary differential equation for unknown $g(x)$ (and, *mutatis mutandis*, for $h(y)$) yields

$$g(x) = \alpha \sin(\sqrt{-C}x) + \beta \cos(\sqrt{-C}a)$$

for $C < 0$

$$g(x) = \alpha x + \beta$$

for $C = 0$ and

$$g(x) = \alpha e^{-\sqrt{C}x} + \beta e^{\sqrt{C}x}$$

for $C > 0$. Under our Dirichlet boundary conditions, a non-zero solution is obtained only for $C < 0$. We obtain

$$a\sqrt{-C} = m\pi$$

for $m \in \mathbb{Z}$. Analogously, we obtain

$$b\sqrt{-D} = k\pi$$

for $k \in \mathbb{Z}$. Since $\lambda = C - D$, we have, finally

$$\lambda_{k,m} = \frac{k^2\pi^2}{a^2} - \frac{m^2\pi^2}{b^2} = \frac{\pi^2 m^2}{a^2}\left(\frac{k^2}{m^2} - \frac{a^2}{b^2}\right) = \frac{\pi^2 m^2}{a^2}\left(\frac{k}{m} - \frac{a}{b}\right)\left(\frac{k}{m} + \frac{a}{b}\right)$$

for all $k, m \in \mathbb{Z}$. Thus, the spectrum equals the closure of the set of all $\lambda_{k,m}$:

$$\sigma(\Box) = \overline{\{\lambda_{k,m} : k, m \in \mathbb{Z}\}}.$$

2.2. The number theory approach

Up to a multiplicative factor, the singular part of the spectrum $\sigma(\Box)$ coincides with the set

$$\mathcal{S}(\alpha) = \text{set of all accumulation points of } \left\{m^2\left(\frac{k}{m} - \alpha\right) : k, m \in \mathbb{Z}\right\},$$

where the ratio $\alpha = a/b$ is a dynamical parameter of the model. The structure of such sets is well understood in the theory of Diophantine approximations. In particular, the smallest accumulation point of the displayed set—the so-called Markov constant of α—is in the center of interest of many mathematicians.

This observation is in fact a methodical starting point of our present paper. In essence, our analysis of stability/instability issues are mainly inspired by the results of the existing number-theory literature on Markov constant.

3. Simple properties of $\mathcal{S}(\alpha)$

Assume $\alpha \in \mathbb{R}$. As the set $\mathbb{Z}^2$ is countable, the set $\left\{m^2\left(\frac{k}{m} - \alpha\right) : k, m \in \mathbb{Z}\right\}$ can be viewed as the range of a real sequence. Let us rephrase the definition of $\mathcal{S}(\alpha)$: a number x belongs to $\mathcal{S}(\alpha)$ if there exist strictly monotone sequences of integers (k_n) and (m_n) such that $x = \lim_{n\to\infty} m_n^2\left(\frac{k_n}{m_n} - \alpha\right)$.

We list several simple properties of $\mathcal{S}(\alpha)$.

(1) Since the set of accumulation points of any real sequence is closed, the set $\mathcal{S}(\alpha)$ is a topologically closed subset of $\mathbb{R}$.
(2) $\mathcal{S}(\alpha)$ is closed under multiplication by z^2 for each $z \in \mathbb{Z}$.

Proof. If $x \in \mathcal{S}(\alpha)$, i.e., $m_n^2\left(\frac{k_n}{m_n} - \alpha\right) \to x$, then $(m_n z)^2\left(\frac{k_n z}{m_n z} - \alpha\right) \to xz^2$, thus $xz^2 \in \mathcal{S}(\alpha)$. □

(3) If $\alpha \in \mathbb{Q}$, then $\mathcal{S}(\alpha)$ is empty.

Proof. If $\alpha = \frac{r}{s}$ with $r, s \in \mathbb{Z}$, then $m^2\left(\frac{k}{m} - \frac{r}{s}\right) = \frac{t}{s}$ for some $t \in \mathbb{Z}$. It means that $\left\{m^2\left(\frac{k}{m} - \alpha\right) : k, m \in \mathbb{Z}\right\}$ is a subset of the discrete set $\frac{1}{s}\mathbb{Z}$. □

(4) If $\alpha \notin \mathbb{Q}$, then $\mathcal{S}(\alpha)$ has at least one element in the interval $[-1, 1]$.

Proof. According to Dirichlet's theorem, there exist infinitely many rational numbers $\frac{k}{m}$ such that $\left|\frac{k}{m} - \alpha\right| < \frac{1}{m^2}$. □

In order to present another remarkable property of $\mathcal{S}(\alpha)$ we exploit simple rational transformations connected with

$$G = \{g \in \mathbb{Z}^{2\times 2} : \det(g) \neq 0\} \quad \text{and} \quad \mathrm{SL}_2(\mathbb{Z}) = \{g \in G : \det(g) = 1\}.$$

Note that G is a monoid, whereas $\mathrm{SL}_2(\mathbb{Z})$ is a group. We define the action of $g = \begin{pmatrix} c & d \\ e & f \end{pmatrix} \in G$ on the set $\mathbb{R}$ by $\alpha \mapsto g\alpha = \frac{c\alpha + d}{e\alpha + f}$.

Proposition 1. *Let $\alpha \in \mathbb{R}$ and $g \in G$. We have*

$$\det(g)\mathcal{S}(\alpha) \subset \mathcal{S}(g\alpha).$$

In particular, $\mathcal{S}(g\alpha) = \mathcal{S}(\alpha)$ if $g \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. Let $g = \begin{pmatrix} c & d \\ e & f \end{pmatrix} \in G$. Let $x \in \mathcal{S}(\alpha)$ and let (k_n) and (m_n) be sequences such that $m_n^2\left(\frac{k_n}{m_n} - \alpha\right) \to x$. We set

$$k_n' = ck_n + dm_n \quad \text{and} \quad m_n' = ek_n + fm_n.$$

We obtain

$$\begin{aligned}\left(\frac{k_n'}{m_n'} - g\alpha\right)m_n'^2 &= \left(\frac{ck_n + dm_n}{ek_n + fm_n} - \frac{c\alpha + d}{e\alpha + f}\right)(ek_n + fm_n)^2 \\ &= \frac{k_n(cf - de) - \alpha m_n(cf - de)}{(e\alpha + f)(ek_n + fm_n)}(ek_n + fm_n)^2\end{aligned}$$

$$= \frac{(k_n - \alpha m_n)(cf - de)}{e\alpha + f}(ek_n + fm_n)$$
$$= \det(g) m_n^2 \left(\frac{k_n}{m_n} - \alpha\right) \frac{e\frac{k_n}{m_n} + f}{e\alpha + f} \underset{n\to+\infty}{\longrightarrow} \det(g)\, x$$

as $\frac{k_n}{m_n} \to \alpha$. It means that $\det(g)x$ belongs to $\mathcal{S}(g\alpha)$.

If $\det g = 1$, then $\mathcal{S}(\alpha) \subset \mathcal{S}(g\alpha)$ and $g^{-1} \in \mathrm{SL}_2(\mathbb{Z})$ as well. Therefore, $\mathcal{S}(g\alpha) \subset \mathcal{S}(g^{-1}g\alpha) = \mathcal{S}(\alpha)$, too. □

In the sequel, $\lfloor x \rfloor$ stands for the integer part of x, i.e., the largest integer n such that $n \leqslant x$. Since $\alpha - \lfloor \alpha \rfloor = g\alpha$, where $g = \begin{pmatrix} 1 & \lfloor \alpha \rfloor \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, the previous proposition immediately implies the following corollary.

Corollary 2. *For any $\alpha \in \mathbb{R}$ we have $\mathcal{S}(\alpha) = \mathcal{S}(\alpha - \lfloor \alpha \rfloor)$.*

Note that $g\alpha = \alpha$ for any $g = \begin{pmatrix} z & 0 \\ 0 & z \end{pmatrix}$ with $z \in \mathbb{Z}$. Proposition 1 implies $z^2\mathcal{S}(\alpha) \subset \mathcal{S}(\alpha)$, as already observed.

4. Continued fractions and convergents

The theory of continued fractions plays a crucial role in Diophantine approximation, i.e., in approximation of an irrational number by a rational number. The definition of $\mathcal{S}(\alpha)$ indicates that the quality of an approximation of α by fractions $\frac{k}{m}$ governs the behavior of $\mathcal{S}(\alpha)$. The continued fraction of an irrational number x is a coding of the orbit of x under a transformation T defined by

$$T : \mathbb{R}\backslash\mathbb{Q} \to (1, +\infty)\backslash\mathbb{Q} \qquad \text{and} \qquad T(x) = \frac{1}{x - \lfloor x \rfloor}.$$

Definition 3. Let $x \in \mathbb{R}\backslash\mathbb{Q}$. The *continued fraction* of x is the infinite sequence of integers $[a_0, a_1, a_2, a_3, \ldots]$, where

$$a_i = \lfloor T^i(x) \rfloor \quad \text{for all} \quad i = 0, 1, 2, 3, \ldots.$$

Clearly, for all $i \geqslant 1$ the coefficient a_i is a positive integer. Only the coefficient a_0 takes values in the whole range of integers.

If α is an irrational number, then $T(\alpha) = g\alpha$ with $g = \begin{pmatrix} 0 & 1 \\ 1 & -\lfloor \alpha \rfloor \end{pmatrix} \in G$. As $\det(g) = -1$, proposition 1 implies $\mathcal{S}(T(\alpha)) = -\mathcal{S}(\alpha)$. A number x is usually identified with its continued fraction and we also write $x = [a_0, a_1, a_2, a_3, \ldots]$. Using this convention, the previous fact can be generalized as follows.

Corollary 4. *Let* $[a_0, a_1, a_2, a_3, \ldots]$ *be a continued fraction. We have*

$\mathcal{S}([a_{n+k}, a_{n+1+k}, a_{n+2+k}, \ldots]) = (-1)^k \mathcal{S}([a_n, a_{n+1}, a_{n+2}, \ldots])$ *for any* $k, n \in \mathbb{N}$.

The knowledge of the continued fraction of x allows us to find the best rational approximations, in a certain sense, of the number x. To describe these approximations, we use the following notation: $[a_0, a_1, a_2, a_3, \ldots, a_n]$, where $a_0 \in \mathbb{Z}$ and $a_1, \ldots, a_n \in \mathbb{N}\{0\}$, denotes the fraction

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}.$$

The number a_i is said to be the ith *partial quotient* of x.

Definition 5. Let x be an irrational number, $[a_0, a_1, a_2, a_3, \ldots]$ its continued fraction and let $N \in \mathbb{N}$. Let $p_N \in \mathbb{Z}$ and $q_N \in \mathbb{N}\setminus\{0\}$ denote coprime numbers such that $\frac{p_N}{q_N} = [a_0, a_1, a_2, a_3, \ldots, a_N]$. The fraction $\frac{p_N}{q_N}$ is called the Nth-convergent of x.

We list the relevant properties of convergents of an irrational number α. They can be found in any textbook of number theory, for example [1].

(1) We have $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$ and $q_1 = a_1$. For any $N \in \mathbb{N}$, we have

$$p_{N+1} = a_{N+1} p_N + p_{N-1} \quad \text{and} \quad q_{N+1} = a_{N+1} q_N + q_{N-1}. \tag{3}$$

(2) For $N \in \mathbb{N}$, set $\alpha_{N+1} = [a_{N+1}, a_{N+2}, a_{N+3}, \ldots]$. We have

$$\frac{p_N}{q_N} - \alpha = \frac{(-1)^{N+1}}{q_N(\alpha_{N+1} q_N + q_{N-1})} \quad \text{and in particular} \quad \left| \frac{p_N}{q_N} - \alpha \right| < \frac{1}{q_N^2}. \tag{4}$$

(3) For $N \in \mathbb{N}$ and $a \in \mathbb{Z}$ satisfying $1 \leqslant a \leqslant a_{N+1} - 1$ we have

$$\frac{a p_N + p_{N-1}}{a q_N + q_{N-1}} - \alpha = \frac{(-1)^{N+1}(\alpha_{N+1} - a)}{(a q_N + q_{N-1})(\alpha_{N+1} q_N + q_{N-1})}. \tag{5}$$

These rational approximations are known as *secondary convergents* of α.

Corollary 6. *Let* α *be an irrational number and* I *be an interval. There exists* $\beta \in I$ *such that* $\mathcal{S}(\alpha) = \mathcal{S}(\beta)$.

Proof. Without loss of generality, let I be an open interval and $\gamma \in I$ be an irrational number. Let $[a_0, a_1, a_2, \ldots]$ and $[c_0, c_1, c_2, \ldots]$ be continued fractions of α and γ, respectively. Find $\varepsilon > 0$ such that $(\gamma - 2\varepsilon, \gamma + 2\varepsilon) \subset I$. In virtue of (4) one can find an integer N such that the Nth-convergent $\frac{p'_N}{q'_N}$ of γ satisfies $\left| \gamma - \frac{p'_N}{q'_N} \right| < \frac{1}{(q'_N)^2} < \varepsilon$. Define

$$\beta = [c_0, c_1, \ldots, c_N, a_{N+1}, a_{N+2}, \ldots].$$

As the Nth-convergents of β and γ coincide and due to (4), we have $|\gamma - \beta| < \frac{2}{(q_N')^2} < 2\varepsilon$ and thus $\beta \in I$. Corollary 4 implies that the sets $\mathcal{S}(\alpha)$ and $\mathcal{S}(\beta)$ coincide as well. □

Theorem 7. *Let α be an irrational number and $\left(\frac{p_N}{q_N}\right)_{N\in\mathbb{N}}$ be the sequence of its convergents. If x belongs to $\mathcal{S}(\alpha) \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$, then x is an accumulation point of the sequence*

$$\left(q_N^2\left(\frac{p_N}{q_N} - \alpha\right)\right)_{N\in\mathbb{N}}. \tag{6}$$

Proof. The theorem is a direct consequence of Legendre's theorem (see for instance [1], theorem 5.12): Let α be an irrational number and $\frac{p}{q} \in \mathbb{Q}$. If $\left|\frac{p}{q} - \alpha\right| < \frac{1}{2q^2}$, then $\frac{p}{q}$ is a convergent of α. □

Therefore, we start to investigate the accumulation points of the sequence (6).

Lemma 8. *Let α be an irrational number and $\left(\frac{p_N}{q_N}\right)_{N\in\mathbb{N}}$ be the sequence of its convergents. For any $N \in \mathbb{N}$ we have*

$$q_N^2\left(\frac{p_N}{q_N} - \alpha\right) = (-1)^{N+1}([a_{N+1}, a_{N+2}, \ldots] + [0, a_N, a_{N-1}, \ldots, a_1])^{-1}.$$

In particular, for any $N \in \mathbb{N}$

$$\frac{1}{2 + a_{N+1}} < \left| q_N^2\left(\frac{p_N}{q_N} - \alpha\right) \right| < \frac{1}{a_{N+1}}.$$

Proof. Using (4) we obtain

$$q_N^2\left(\frac{p_N}{q_N} - \alpha\right) = \frac{(-1)^{N+1}}{\alpha_{N+1} + \dfrac{q_{N-1}}{q_N}}.$$

By definition $\alpha_{N+1} = [a_{N+1}, a_{N+2}, \ldots]$. It remains to show that $\frac{q_{N-1}}{q_N} = [0, a_N, a_{N-1} \ldots, a_1]$. We exploit the recurrent relation (3) for (q_N). We proceed by induction:

If $N = 1$, then $q_0 = 1$ and $q_1 = a_1$. Clearly $\frac{q_0}{q_1} = \frac{1}{a_1} = [0, a_1]$.

If $N > 1$, then

$$\frac{q_{N-1}}{q_N} = \frac{q_{N-1}}{a_N q_{N-1} + q_{N-2}} = \frac{1}{a_N + \frac{q_{N-2}}{q_{N-1}}}. \tag{7}$$

The number $\beta \in (0, 1)$ has its continued fraction in the form $[0, b_1, b_2, \ldots]$. If $1 \leqslant B \in \mathbb{Z}$, then the algorithm for construction of continued fraction assigns to the number $\frac{1}{B+\beta}$ the continued fraction $[0, B, b_1, b_2, \ldots]$. We apply this rule and the induction assumption to (7) with $B = a_N$ and $\beta = \frac{q_{N-2}}{q_{N-1}} = [0, a_{N-1}, a_{N-2}, \ldots, a_1]$. □

4.1. Spectra of quadratic numbers

A famous theorem of Lagrange says that an irrational number α is a root of the quadratic polynomial $Ax^2 + Bx + C$ with integer coefficients A, B, C if and only if the continued fraction of α is eventually periodic, i.e., $\alpha = [a_0, a_1, \dots, a_s, (a_{s+1}, \dots, a_{s+\ell})^\omega]$, where v^ω denotes the infinite string formed by the repetition of the finite string v.

Theorem 9. *Let α be a quadratic number and $\left(\frac{p_N}{q_N}\right)$ be the sequence if its convergents. Let ℓ be the smallest period of the repeating part of the continued fraction of α. The sequence $\left(q_N^2\left(\frac{p_N}{q_N} - \alpha\right)\right)_{N\in\mathbb{N}}$ has at most*

- *ℓ accumulation points if ℓ is even;*
- *2ℓ accumulation points if ℓ is odd.*

Moreover, at least one of the accumulation points belongs to the interval $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Proof. According to corollary 4 we can assume that the continued fraction of α is purely periodic, i.e., $\alpha = [(a_0, a_1, \dots, a_{\ell-1})^\omega]$ for some $\ell > 0$, and that the first digit satisfies $a_0 = \max\{a_0, a_1, \dots, a_{\ell-1}\}$. Let D denote the set of the accumulation points of $\left(q_N^2\left(\frac{p_N}{q_N} - \alpha\right)\right)_{N\in\mathbb{N}}$.

Suppose ℓ is even. By lemma 8 and since $\alpha = [(a_0, a_1, \dots, a_{\ell-1})^\omega]$, it follows that all the elements of D are the limit-points of the sequences $(c_k^{(j)})_{k\in\mathbb{N}}$, where

$$c_k^{(j)} = (-1)^{j+k\ell-1}([a_{j+k\ell}, a_{j+k\ell+1}, \dots] + [0, a_{j+k\ell-1}, \dots, a_1])^{-1}$$

for each j with $0 \leqslant j < \ell$. As ℓ is even, the term $(-1)^{j+k\ell-1}$ equals $(-1)^{j-1}$ and a limit exists. Thus, $\#D \leqslant \ell$.

If ℓ is odd, we define the number $c_k^{(j)}$ for $0 \leqslant j < 2\ell$ in the same way and the elements of D are exactly the limit-points of the sequences $(c_{2k}^{(j)})_{k\in\mathbb{N}}$. The term $(-1)^{j+2k\ell-1}$ in the expression of $c_{2k}^{(j)}$ equals again $(-1)^{j-1}$ and a limit exists for all j. Thus, $\#D \leqslant 2\ell$.

If $a_0 = 1$, then $\alpha = [1^\omega]$, i.e., it is the golden ratio. We have

$$\lim_{k\to+\infty} c_{2k}^{(0)} = -([1^\omega] + [0, 1^\omega])^{-1} = -\left(\frac{1+\sqrt{5}}{2} + \frac{2}{1+\sqrt{5}}\right)^{-1} = -\frac{1}{\sqrt{5}} \geqslant -\frac{1}{2}.$$

Thus, in this case, $D \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$ is not empty.

If $a_0 = 2$, then

$$\begin{aligned} |c_k^{(0)}| &= |[2, a_{k\ell+2}, \dots] + [0, a_{k\ell-1}, \dots, a_1]|^{-1} \\ &= |2 + [0, a_{k\ell+2}, \dots] + [0, a_{k\ell-1}, \dots, a_1]|^{-1} \leqslant \frac{1}{2}. \end{aligned}$$

It implies that $D \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$ is not empty. □

We add some remarks on the last theorem. The following observation follows from the last proof: if ℓ is odd, then D is symmetric around 0.

Let $\eta^{(j)} = [(a_j, \dots, a_{j+\ell-1})^\omega]$. The number $\eta^{(j)}$ is a reduced quadratic surd and its conjugate $\tilde{\eta}^{(j)}$ satisfies

$$-\frac{1}{\tilde{\eta}^{(j)}} = [(a_{j+\ell-1}, \dots, a_j)^\omega].$$

Therefore,

$$\lim_{k\to+\infty} c_{2k}^{(j)} = \frac{(-1)^{j-1}}{\eta^{(j)} - \tilde{\eta}^{(j)}}. \tag{8}$$

As follows from the last proof, the bound of theorem 9 is tight. On the other hand, there exist quadratic numbers such that the bound is not attained. It suffices to set $\alpha = [(1,\ 2,\ 1,\ 1)^\omega]$. We have

$$[(1,\ 2,\ 1,\ 1)^\omega] = \frac{2}{5}\sqrt{6} + \frac{2}{5} \quad \text{and}$$
$$[(1,\ 1,\ 1,\ 2)^\omega] = \frac{2}{5}\sqrt{6} + \frac{3}{5}.$$

Using (8) we obtain $\#D < \ell = 4$. In fact, $D = \left\{-\frac{5}{4\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{3}{4\sqrt{6}}\right\}$.

5. Well and badly approximable numbers

The search for the best rational approximation of irrational numbers motivates the notion of Markov constant.

Definition 10. Let α be an irrational number. The number

$$\mu(\alpha) = \inf\left\{c > 0 : \left|\alpha - \frac{k}{m}\right| < \frac{c}{m^2} \text{ has infinitely many solutions } k, m \in \mathbb{Z}\right\}$$

is the *Markov constant of* α.

The number α is said to be *well approximable* if $\mu(\alpha) = 0$ and *badly approximable* otherwise.

We give several comments on the value $\mu(\alpha)$:

(1) Theorem of Hurwitz implies $\mu(\alpha) \leqslant \frac{1}{\sqrt{5}}$ for any irrational real number α.

(2) A pair (k, m) which is a solution of $\left|\alpha - \frac{k}{m}\right| < \frac{c}{m^2}$ with $c \leqslant \frac{1}{\sqrt{5}}$ satisfies $k = \|m\alpha\|$, where we use the notation $\|x\| = \min\{|x - n| : n \in \mathbb{Z}\}$. Therefore

$$\mu(\alpha) = \liminf_{m\to+\infty} m\|m\alpha\| \qquad \text{and} \qquad \mu(\alpha) = \min|\mathcal{S}(\alpha)|$$

as the set $\mathcal{S}(\alpha)$ is topologically closed.

(3) Due to the inclusion $\det(g)\mathcal{S}(\alpha) \subset \mathcal{S}(g\alpha)$ for $g \in G$ we can write

$$|\det(g)|\ \mu(\alpha) \geqslant \mu(g\alpha).$$

(4) The inequality in lemma 8 implies

$$\mu(\alpha) = 0 \quad \Longleftrightarrow \quad (a_N) \quad \text{is not bounded} \quad \Longleftrightarrow \quad 0 \in \mathcal{S}(\alpha).$$

In other words, an irrational number α is well approximable if and only if the sequence (a_N) of its partial quotients is bounded.

5.1. Badly approximable numbers

As noted above, quadratic irrational numbers serve as an example of badly approximable numbers. The spectrum $\mathcal{S}(\alpha)$ of such a number has only finite number of elements in the interval $\left(-\frac{1}{2}, \frac{1}{2}\right)$. Theorems 11 and 13 give two examples of spectra of badly approximable numbers of different kinds.

Theorem 11. *There exists an irrational number* α *such that* $\mathcal{S}(\alpha) = (-\infty, -\varepsilon] \cup [\varepsilon, +\infty)$, *where* $\varepsilon = \frac{\sqrt{2}}{8} \sim 0.18$.

We first recall that the natural order on $\mathbb{R}$ is represented by an alternate order in continued fractions. More precisely, let x and y be two irrational numbers with the continued fractions $[x_0, x_1, \ldots]$ and $[y_0, y_1, \ldots]$, respectively. Set $k = \min\{i \in \mathbb{N} : x_i \neq y_i\}$. We have $x < y$ if and only if

$$(k \ \text{ is even and } \ x_k < y_k) \quad \text{or} \quad (k \ \text{ is odd and } \ x_k > y_k).$$

To study the numbers with bounded partial quotients we define the following sets:

$$F(r) = \{[t, a_1, a_2, \ldots] : t \in \mathbb{Z}, 1 \leqslant a_i \leqslant r\}$$

and

$$F_0(r) = \{[0, a_1, a_2, \ldots] : 1 \leqslant a_i \leqslant r\}.$$

These sets are 'sparse' and they are Cantor sets: perfect sets that are nowhere dense (see for instance [7]). For example, the Hausdorff dimension of $F(2)$ satisfies $0.44 < \dim_{\mathrm{H}}(F(2)) < 0.66$ (see example 10.2 in [8]). Taking into account the alternate order, the maximum and minimum elements of $F_0(r)$ can be simply determined. Thus, $\max F_0(r) = [0, 1, r, 1, r, 1, r, \ldots]$ and $\min F_0(r) = [0, r, 1, r, 1, r, 1, \ldots]$. A crucial result which enables us to prove theorem 11 is due to [9] (see also [7]):

$$F(4) + F(4) = \mathbb{R}. \tag{9}$$

It is worth mentioning that $r = 4$ is the least integer for which $F(r) + F(r) = \mathbb{R}$, i.e., in particular, $F(3) + F(3) \neq \mathbb{R}$ (see [10]). Applying Theorem 2.2 and Lemma 4.2 of [7] we obtain the following modification of (9):

$$F_0(4) + F_0(4) = [2 \min F_0(4), 2 \max F_0(4)] = [\sqrt{2} - 1, 4(\sqrt{2} - 1)]. \tag{10}$$

We use the last equality to construct the number α for the proof of theorem 11. The construction is based on the following observation.

Lemma 12. *Let* $\mathbf{a} = a_0 a_1 a_2 \ldots$ *be an infinite word over the alphabet* $\mathcal{A} = \{1, 2, \ldots, r\}$ *such that any finite string* $w_1 w_2 \ldots w_k$ *over the alphabet* $\mathcal{A}$ *occurs in* $\mathbf{a}$, *i.e., there exists index* $n \in \mathbb{N}$ *such that* $a_n a_{n+1} \ldots a_{n+k-1} = w_1 w_2 \ldots w_k$. *Any number* $z \in \mathcal{A} + F_0(r) + F_0(r)$ *is an accumulation point of the sequence* (S_{2N}) *and the sequence* (S_{2N+1}) *with*

$$S_N = [a_{N+1}, a_{N+2}, \ldots] + [0, a_N, a_{N-1}, \ldots, a_1]. \tag{11}$$

Proof. Let $x = [0, x_1, x_2, x_3, \ldots]$, $y = [0, y_1, y_2, y_3, \ldots] \in F_0(r)$ and $b \in \mathcal{A}$. For any string $w_1 w_2 \ldots w_k$ there exist infinitely many finite strings $u_1 u_2 \ldots u_{h-1} u_h$ such that $w_1 w_2 \ldots w_k$ is a prefix and a suffix of $u_1 u_2 \ldots u_{h-1} u_h$. According to our assumptions each of them occurs at least once in $\mathbf{a}$. It means that any string $w_1 w_2 \ldots w_k$ occurs in $\mathbf{a}$ infinitely many times on both odd and even positions. In particular, for any n there exists infinitely many odd and infinitely many even indices N such that

$$a_{N-n+1}\dots a_N a_{N+1}\dots a_{N+n} = x_n x_{n-1}\dots x_1 b y_1 y_2\dots y_{n-2} y_{n-1}.$$

Obviously, the number S_N given by (11) equals

$$b + [0, y_1, y_2, \dots, y_{n-1}, a_{N+n}, a_{N+n+1}, \dots] + [0, x_1, x_2, \dots, x_n].$$

As $b + y + x$ is the limit of the previous sequence, it is an accumulation point of the sequence (11). □

We can complete the proof of theorem 11.

Proof of theorem 11. We construct an infinite word **a** with letters in $\{1, 2, 3, 4\}$ satisfying the assumptions of lemma 12. We define a sequence $(u_n)_{n=0}^{+\infty}$ recursively as follows: u_0 is the empty word and $u_n = u_{n-1}v_n$, where v_n is the word which the concatenation of all words over $\{1, 2, 3, 4\}$ of length n ordered lexicographically. We have

$$\begin{aligned} u_1 &= 1\ 2\ 3\ 4 \qquad \text{and} \\ u_2 &= 1\ 2\ 3\ 4\ 11\ 12\ 13\ 14\ 21\ 22\ 23\ 24\ 31\ 32\ 33\ 34\ 41\ 42\ 43\ 44. \end{aligned}$$

As u_{n-1} is a prefix of u_n, we can set **a** to be the unique infinite word which has a prefix u_n for any $n \in \mathbb{N}$. One can easily see that **a** satisfies the assumptions of lemma 12.

Let α be the number with the continued fraction $[0, a_1, a_2, a_3, \dots]$, where $a_1a_2a_3\dots = \mathbf{a}$. Combining lemmas 8 and 12 and the equality (10) we obtain that $\frac{1}{z}$ and $-\frac{1}{z}$ belong to $\mathcal{S}(\alpha)$ for any $z \in [b + \sqrt{2} - 1, b + 4(\sqrt{2} - 1)]$ with $b \in \{1, 2, 3, 4\}$. Overall, we obtain

$$\left[-\tfrac{1}{\sqrt{2}}, -\tfrac{1}{4\sqrt{2}},\right] \bigcup \left[\tfrac{1}{4\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right] \subset \mathcal{S}(\alpha).$$

The property that $\mathcal{S}(\alpha)$ is closed under multiplication by z^2 for each positive integer z, in particular under multiplication by 4, already proves theorem 11. □

Theorem 13. *There exists an irrational number α such that the Hausdorff dimension of $\mathcal{S}(\alpha) \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$ is positive but less than 1. In particular, $\mathcal{S}(\alpha) \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$ is an uncountable set and its Lebesgue measure is* 0.

Proof. Let **a** be an infinite word with letters in $\{4, 5\}$ such that it contains any finite string over $\{4, 5\}$ infinitely many times. A word with such properties can be constructed in the same way as in the proof of theorem 11.

In accordance with the previous notation we set

$$F_0(\{4, 5\}) = \{[0, a_1, a_2, \dots] : a_i \in \{4, 5\}\}.$$

To simplify, we write $F = F_0(\{4, 5\})$. Theorem 1.2 in [7] implies that

$$\dim_H (F + F) \geqslant 0.263\dots.$$

To obtain an upper bound on the Hausdorff dimension of $F + F$, we first give a construction of F. Let I denote the interval $I = [\min F, \max F]$. Clearly, $F \subset I$.

For both letters $z = 4$ and $z = 5$ we define $f_z : I \to I$ as follows:

$$f_z(x) = \frac{1}{z + x} \text{ for all } x \in I.$$

Using the mean value theorem, one can easily derive that

$$\frac{|f_z(x) - f_z(y)|}{|x - y|} \leqslant \max_{\xi \in I} |f'(\xi)| \leqslant L := \frac{1}{(\min F + 4)^2}$$

for all $x, y \in I, x \neq y$. Thus, the mappings f_4 and f_5 are contractive and one can see that F is the fixed point of the iterated function system generated by these mappings. In other words, we have

$$F = \lim_{n \to +\infty} Z_n \quad \text{with} \quad Z_n = \bigcup_{a_1 a_2 \ldots a_n \in \{4,5\}^n} f_{a_1} f_{a_2} \cdots f_{a_n}(I).$$

Let us stress that $\lim_{n \to +\infty}$ on the previous row is defined via the Hausdorff metric on the space of compact subsets of $\mathbb{R}$.

Let $n \in \mathbb{N}$. It follows that there exists a covering of the set Z_n consisting of 2^n intervals of length at most $|I| \cdot L^n$. Similarly, the set $Z_n + Z_n$ can be covered by 4^n intervals of length at most $|I| \cdot L^n$. Since $F + F = \lim_{n \to +\infty} Z_n + \lim_{n \to +\infty} Z_n = \lim_{n \to +\infty}(Z_n + Z_n)$ and $Z_{n+1} \subset Z_n$, we can use this covering to estimate the Hausdorff dimension of $F + F$ (see [8], proposition 4.1) as follows:

$$\dim_H(F + F) \leqslant \lim_{n \to +\infty} \frac{\log 4^n}{-\log(|I| \cdot L^n)} = -\frac{\log 4}{\log L} = \frac{\log 2}{\log(4 + \min F)}.$$

As $\min F = [0, (5, 4)^\omega]$, we obtain $\min F = \frac{1}{5 + \frac{1}{4 + \min F}}$. Thus $\min F = 2\left(\sqrt{\frac{6}{5}} - 1\right)$ and we deduced the upper bound

$$\dim_H(F + F) \leqslant \frac{\log 2}{\log 2 + \log\left(\sqrt{\frac{6}{5}} + 1\right)} < \frac{1}{2}.$$

The rest of the proof is analogous to the end of the proof of theorem 11. We use lemma 8 and an analogous modification of lemma 12 for the alphabet $\mathcal{A} = \{4, 5\}$ to obtain that

$$\pm\frac{1}{x} \in \mathcal{S}(\alpha) \quad \text{for each} \quad x \in \{4, 5\} + F + F.$$

By theorem 7 we have

$$\left\{\frac{1}{x} : |x| \in \{4, 5\} + F + F\right\} \bigcap \left(-\frac{1}{2}, \frac{1}{2}\right) = \mathcal{S}(\alpha) \bigcap \left(-\frac{1}{2}, \frac{1}{2}\right).$$

Clearly, the union of the four sets $4 + F + F, 5 + F + F, -4 - F - F$ and $-5 - F - F$ with the same Hausdorff dimension is a set of the same dimension. Moreover, the Hausdorff dimensions of $f(M)$ and M coincide for any continuous mapping f, in particular for $f(x) = \frac{1}{x}$. It implies that the estimates on the Hausdorff dimension of $F + F$ are valid also for $\mathcal{S}(\alpha) \cap \left(-\frac{1}{2}, \frac{1}{2}\right)$. □

5.2. Well approximable numbers

We consider $\alpha = [a_0, a_1, a_2, \ldots]$ with unbounded partial quotients. Using second convergents defined in (5) we can write for any $N \in \mathbb{N}$ and $a \in \mathbb{N}$ with $1 \leqslant a < a_{N+1}$

$$(aq_N + q_{N-1})^2\left(\frac{ap_N + p_{N-1}}{aq_N + q_{N-1}} - \alpha\right) = (-1)^{N+1}\left(a + \frac{q_{N-1}}{q_N}\right)\frac{\alpha_{N+1} - a}{\alpha_{N+1} + \frac{q_{N-1}}{q_N}}. \tag{12}$$

Recall that $\alpha_{N+1} = [a_{N+1}, a_{N+2}, a_{N+3}, \ldots]$. Let (i_N) be a strictly increasing sequence of integers such that $\lim_{N\to+\infty} a_{1+i_N} = +\infty$. Clearly, $\lim_{N\to+\infty} \alpha_{1+i_N} = +\infty$. Let us fix $a \in \mathbb{N}$ and put $k_N = ap_{i_N} + p_{i_N-1}$ and $m_N = aq_{i_N} + q_{i_N-1}$, we have

$$m_N^2 \left| \frac{k_N}{m_N} - \alpha \right| = \left(a + \frac{q_{i_N-1}}{q_{i_N}}\right)E_N,$$

where we set $E_N = \frac{\alpha_{i_N+1} - a}{\alpha_{i_N+1} + \frac{q_{i_N-1}}{q_{i_N}}}$. Obviously, $\lim\limits_{N\to+\infty} E_N = 1$. Since the sequence (q_N) is a strictly increasing sequence of integers, the ratio $\frac{q_{i_N-1}}{q_{i_N}}$ belongs to $(0, 1)$. This implies that the sequence $m_N^2 \left| \frac{k_N}{m_N} - \alpha \right|$ has at least one accumulation point in the interval $[a, a + 1]$. Therefore we can conclude the next lemma.

Lemma 14. *Let α be an irrational well approximable number. For any $n \in \mathbb{N}$ the interval $[n, n + 1]$ or the interval $[-n - 1, -n]$ has a non-empty intersection with $\mathcal{S}(\alpha)$.*

Example 15. Unlike the number π, the continued fraction of the Euler constant has a regular structure:

$$\mathrm{e} = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, \ldots].$$

In general, for $\mathrm{e} = [2, a_1, a_2, \ldots]$ we have

$$a_{3n+1} = 1, \quad a_{3n+2} = 2(n + 1) \quad \text{and} \quad a_{3n+3} = 1 \quad \text{for any } n \in \mathbb{N}.$$

We demonstrate that

$$\left| q_{3N}^2\left(\frac{p_{3N}}{q_{3N}} - e\right) \right| \to \tfrac{1}{2}.$$

By lemma 8 we need to show

$$A_{3N} := [a_{3N+1}, a_{3N+2}, a_{3N+3}, \ldots] + [0, a_{3N}, a_{3N-1}, \ldots, a_1] \to 2.$$

Using the simple estimate valid for any continued fraction

$$b_0 + \frac{1}{b_1 + \frac{1}{b_2}} \ < \ [b_0, b_1, b_2, b_3, \ldots] < \ b_0 + \frac{1}{b_1 + \frac{1}{1 + b_2}}$$

we obtain the following bounds:

$$1 + \frac{1}{2(N + 1) + 1} + \frac{1}{1 + \frac{1}{2N}} < A_{3N} < 1 + \frac{1}{2(N + 1) + \frac{1}{2}} + \frac{1}{1 + \frac{1}{2N + 1}}.$$

Both bounds have the same limit, namely 2, as we wanted to show. Analogously one can deduce that

$$\left| q_{3N-1}^2\left(\frac{p_{3N-1}}{q_{3N-1}} - \mathrm{e}\right) \right| \to \tfrac{1}{2} \quad \text{and} \quad \left| q_{3N+1}^2\left(\frac{p_{3N+1}}{q_{3N+1}} - \mathrm{e}\right) \right| \to 0.$$

Since $(-1)^{3N}$ takes positive and negative signs, the values $0, \pm\frac{1}{2}$ belong to the spectrum of e and moreover

$$\left(-\tfrac{1}{2}, \tfrac{1}{2}\right) \bigcap \mathcal{S}(\mathrm{e}) = \{0\}.$$

As $a_{3N+2} = 2N > 1$, we can use the second convergents as well and for any fixed $a \in \mathbb{N}$ and any N such that $a < a_{3N+2}$ we obtain

$$(aq_{3N+1} + q_{3N})^2 \left| \mathrm{e} - \frac{ap_{3N+1} + p_{3N}}{aq_{3N+1} + q_{3N}} \right| = \left(a + \frac{q_{3N}}{q_{3N+1}}\right) E_{3N},$$

where $\lim_{N\to\infty} E_{3N} = 1$, see (12). By the proof of lemma 8, we have

$$\frac{q_{3N}}{q_{3N+1}} = [0, a_{3N+1}, a_{3N}, a_{3N-1}, \ldots, a_1] \to \tfrac{1}{2}.$$

We conclude for the spectrum of the Euler number satisfies

$$\{0\} \bigcup \{a + \tfrac{1}{2} : a \in \mathbb{Z}\} \subset \mathcal{S}(\mathrm{e}).$$

Of course, the inclusion cannot be replaced by an equality. The reason is simple; the spectrum is closed under multiplication by the factor 4, and thus

$$\{4a + 2 : a \in \mathbb{Z}\} \subset \mathcal{S}(\mathrm{e})$$

as well.

Theorem 16. *There exists an irrational number* α *such that* $\mathcal{S}(\alpha) = \mathbb{R}$.

Proof. Suppose that $\mathbf{a} = a_1 a_2 \ldots$ is an infinite word such that any sequence of the form $w_1 w_2 \ldots w_k N w_{k+1} w_{k+2} \ldots w_{2k}$, where the symbols w_i are from the alphabet $\{1, 2, 3, 4\}$ and $N > 1$, $N \in \mathbb{Z}$ occurs in $\mathbf{a}$ infinitely many times. The same reasoning as in the proofs of lemma 12 and theorem 11 together with the equality (10) imply the statement of the theorem. Therefore, it is enough to describe $\mathbf{a}$.

Fix $n \in \mathbb{N}$ and consider a word $w = w_1 w_2 \ldots w_n$ of length n over the alphabet $\{1, 2, 3, 4\}$. Copy(w) denotes the concatenation of n words of length $(n+1)$, each in the form $wh = w_1 w_2 \ldots w_n h$ with $h = 1, 2, \ldots, n$. Thus Copy(w) is a word of length $n(n+1)$. The word v_n is created by concatenation of Copy(w) for all words w of length n over the alphabet $\{1, 2, 3, 4\}$. In particular, the length of v_n is $4^n n(n+1)$.

The infinite word $\mathbf{a}$ is given by its prefixes (u_n) which are constructed recursively: u_0 is the empty word and $u_n = u_{n-1} v_n$. □

Remark 17. We note that the behavior of $\alpha = [a_1, a_2, \ldots]$ defined in the proof of the previous theorem is typical. In [11], Bosma, Jager and Wiedijk described the distribution of the sequence $q_n|p_n - \alpha q_n|$. A direct consequence of their result is that $\mathcal{S}(\alpha) = \mathbb{R}$ for almost all $\alpha \in [0, 1]$. Thus, it is also true for almost all $\alpha \in \mathbb{R}$.

6. Discussion and remarks

In the context of physics it must be emphasized that our choice of the elementary model (1) + (2) is motivated not only by its appealing number-theoretical properties but also by its possible straightforward phenomenological applicability. We feel motivated by the persuasion that the related *constructive* exemplification of certain spectral anomalies might prove attractive even from the point of view of a physicist who need not necessarily care about the deeper mathematical subtleties.

Using our purely mathematical tools we are able to arrive at a better understanding of certain purely formal connections between various structural aspects of the spectra, with the main emphasis put on its unboundedness from below (which could result into instabilities under small perturbations) in an interplay with the emergence of accumulation points in the point spectrum (in the latter case it makes sense to keep in mind the existing terminological ambiguities [12]).

Needless to add that the phenomenological role of the spectral accumulation points remains strongly model-dependent (see the rest of this section for a few samples). In the most elementary quantized hydrogen atom, for example, such a point represents just an entirely innocent lower bound of the continuous spectrum. A more interesting interpretation of these points is obtained in the case of the so called Efimov three-body bound states [13, 14], etc.

6.1. The context of systems with position-dependent mass

Irrespectively of the concrete physical background of quantum stability [15], its study encounters several subtle mathematical challenges [16–19]. In our present hyperbolic-operator square-well model living on a compact domain R, a number of interesting spectral properties is deduced and proved by the means and techniques of mathematical number theory, without any recourse to the abstract spectral theory. Still, the standard spectral theory is to be recalled. For example, once we return to the explicit units we may reinterpret our present *hyperbolic* partial differential operator $\Box$ in equation (1) as a result of a drastic deformation of an *elliptic* non-equal-mass Laplacean

$$\triangle = \frac{1}{2m_x}\frac{\partial^2}{\partial x^2} + \frac{1}{2m_y}\frac{\partial^2}{\partial y^2} \tag{13}$$

or rather of an even more general kinetic-energy operator

$$T(x, y) = \frac{1}{2m_x(x, y)}\frac{\partial^2}{\partial x^2} + \frac{1}{2m_y(x, y)}\frac{\partial^2}{\partial y^2} \tag{14}$$

containing the position-dependent positive masses. In the ultimate and decisive step one simplifies the coordinate dependence in the masses $m_{x,y}(x, y)$ (say, to piecewise constant functions) and, purely formally, allows one of them to become negative.

In such a setting our present mathematical project is also guided by the specific position-dependent mass physical projects of [20, 21] inspired, in their turn, by the non-Hermitian (a.k.a. $\mathcal{PT}-$ symmetric [22]) version of quantum Kepler problem. In these papers the mass $m(x)$ is allowed to be complex and, in particular, negative. In [20] the onset of the spectral instability is analyzed as an onset of an undesirable unboundedness of the discrete spectrum from below. A return to a stable system with vacuum is then shown controllable only via an energy-dependent mass $m(x, E)$, i.e., via an ad hoc spectral cut-off (see also [23]).

6.2. The context of generalized quantum waveguides

Before one recalls the boundary conditions (2), the majority of physicists would perceive our hyperbolic partial differential equation (1) as the Klein–Gordon equation describing the free relativistic one-dimensional motion of a massive and spinless point particle. Whenever one adds an external (say, attractive Coulomb) field, the model becomes realistic (describing, say, a pionic atom). Now, even if we add the above-mentioned Dirichlet boundary conditions

$f|_{\partial R} = 0$, a certain physical interpretation of the spectrum survives the characterization of, say, the bound states in a 'relativistic quantum waveguide'.

One of the most interesting consequences of the latter approach may be seen in the possibility of a collapse of the system in a strong field. The most elementary illustrations of such a type of instability may even remain non-relativistic: Landau and Lifshitz [24] described the phenomenon in detail. Another, alternative, type of quantum instabilities connected with the emergence of spectral accumulation points occur also in Horava–Lifshitz gravity with ghosts [25, 26] or in the conformal theories of gravity [27–29] etc.

Our present choice of the elementary illustrative example with compact and rectangular R changes the physics and becomes more intimately related to the problems of the so called quantum waveguides with impenetrable walls [30]. Most of the mathematical problems solved in the latter context are very close to the present ones. Typically, they concern the possible relationship between the spectra and geometry of the spatial boundaries. In this setting, various transitions to the infinitely thin and/or topologically non-trivial domains R (one may then speak about quantum graphs) and, possibly, also to the various anomalous point-interaction forms of the interactions are being also studied [6].

Up to now, people only very rarely considered a replacement of the positive-definite kinetic-energy operator (i.e., Laplacean) by its hyperbolic alternative. Thus, in spite of some progress [31], such a 'relativistic' generalization of the concept of quantum waveguide and/or of quantum graph still remains to be developed.

6.3. The context of classical optical systems with gain and loss

One of the most characteristic features of modern physics may be seen in the multiplicity of overlaps between its apparently remote areas. *Pars pro toto* let us mention here the unexpected productivity of the transfer of several quantumtheoretical concepts beyond the domain of quantum theory itself [32]. One of the best known recent samples of such a transfer starts in quantum field theory [33] and ends up in classical electrodynamics [34]. A common mathematical background consists in the requirements of the Krein-space self-adjointness [35] *alias* parity-times-time-reversal symmetry ($\mathcal{PT}$-symmetry).

It is worth adding that the latter form of a transfer of ideas already proceeded in both directions. The textbook formalism of classical electrodynamics based on Maxwell equations was enriched by the mathematical techniques originating in spectral theory of quantum operators in Hilbert space (see e.g., section 9.3 of the review paper [36] for more details). In parallel, the $\mathcal{PT}-$ symmetry-related version of quantum theory (see also its older review [22]) took an enormous profit from the emergence and success of its experimental tests using optical metamaterials [37]. People discovered that the time is ripe to think about non-elliptic versions of Maxwell equations reflecting the quick progress in the manufacture of various sophisticated metamaterials which possess non-real elements of the permittivity and/or permeability tensors [38–41].

Naturally, the mutual enrichments of the respective theories would not have been so successful without the progress in experimental techniques, and vice versa. In fact, the availability of the necessary optical metamaterials (which can simulate the $\mathcal{PT}$-symmetry of quantum interactions via classical gain-loss symmetry of prefabricated complex refraction indices) was a highly non-trivial consequence of the quick growth of the know-how in nanotechnologies [42, 43]. In opposite direction, the experimental simulations of various quantum loss-of-stability phenomena in optical metamaterials encouraged an intensification of the related growth of interest in the questions of stability of quantum systems with respect to perturbations [44–46].

6.4. The context of unbounded spectra

Our last comment on the possible phenomenological fructification of our present study of the toy model (1) + (2) concerns its possible, albeit purely formal, connection to the traditional Pais–Uhlenbeck (PU) oscillator [47]. The idea itself is inspired by the Smilga's paper [48] which provides us with a compact review of the appeal of the next-to-elementary PU model in physics.

We imagine, first of all, that the unboundedness of the spectrum of the PU oscillators parallels the same 'threat of instability' feature of our rectangular model. At the same time, in the broad physics community, the PU oscillator is much more widely accepted as a standard model throwing a new light on several methodical aspects of the loss of stability, especially in the context of quantum cosmology and quantization of gravity (see also [49–51]). In particular, the PU model contributes to the understanding of the role of renormalizability in higher-order field theories [25, 52], etc.

For these reasons we skip the problems connected with the ambiguity of transition from Lagrangians to Hamiltonians [53] and we restrict our attention just to one of the specific, PU-related quantum Hamiltonian(s), viz., to the operator picked up for analysis, e.g., in [48],

$$H = (-\partial_x^2 + \Omega_x^2 x^2) - (-\partial_y^2 + \Omega_y^2 y^2). \tag{15}$$

In a way resembling our present results, the related quantum dynamics looks pathological because even the choice of the incommensurable oscillator frequencies Ω_x and Ω_y leads to a quantum system in which the bound-state energy spectrum (i.e., in the language of mathematics, point spectrum—see a comment Nr. 2 in [48]) is real but dense and unbounded

$$E_{nm} = \left(n + \frac{1}{2}\right)\Omega_x - \left(m + \frac{1}{2}\right)\Omega_y, \qquad n, m = 0, 1, 2, \ldots. \tag{16}$$

In the related literature (see e.g., [54–56]) several remedies of the pathologies are proposed ranging from the use of the Wick rotation of $y \to \mathrm{i}y$ [57] up to a suitable modification of the Hamiltonian as performed already before quantization, on classical level [58–60].

This being said, an independent disturbing feature of the PU toy model (15) may be seen in an abrupt occurrence of a set of spectral accumulation points in the equal-frequency limit $\Omega_x - \Omega_y \to 0$ [61, 62]. The emergent new technical difficulty originates from the fact that the resulting Hamiltonian becomes non-diagonalizable, acquiring a rather peculiar canonical matrix structure of an infinite-dimensional Jordan block.

This is one the most dangerous loss-of-quantum-meaning aspects of the model. Its serious phenomenological consequences are discussed, e.g., in the scalar field cosmology (see the freshmost papers [63, 64] with further references). In a narrower context of specific pure fourth-order conformal gravity, such a spectral discontinuity cannot be circumvented at all [65].

7. Summary

The aim of this paper is to demonstrate the variability of spectra in dependence on the number-theoretical properties of the ratio $\alpha = a/b$ of the sides of the rectangular R. In particular, we show that in an arbitrarily short interval $I \subset \mathbb{R}$ one can find numbers $\alpha, \beta, \gamma, \delta, \varepsilon$ such that the spectrum of $\mathcal{S}(\alpha)$ is empty, the spectrum of $\mathcal{S}(\beta)$ forms an infinite discrete set, the spectrum $\mathcal{S}(\gamma) = \mathbb{R}$ covers the whole real line, the spectrum $\mathcal{S}(\delta) = \mathbb{R}\backslash(-a, a)$ has a 'hole' with some positive real $a = a(\delta)$. Finally the spectrum $\mathcal{S}(\varepsilon)$ has zero Lebesgue measure, it is uncountable, but it has a positive Hausdorff dimension which is less than 1. It means that a small change of the dynamical parameter $\alpha = a/b$ dramatically influences the spectrum.

Although we give just an extremely elementary example for the detailed and rigorous analysis, we would like to emphasize that our present approach proves productive in spite of lying far beyond the standard scope and methods of spectral analysis. A non-trivial insight in the underlying physics is provided purely by the means of number theory.

From the point of view of number theory, various results on the Markov constant, i.e., $\min\left\{m^2\left|\frac{k}{m}-\alpha\right| : k, m \in \mathbb{Z}\right\}$, may be found. In the present article, we provide some insight into the behavior of all the accumulation points of the concerned set. Since we restrict ourselves to some special cases, naturally, the next step would be to fully investigate the properties of $\mathcal{S}(\alpha)$.

Acknowledgments

EP acknowledges financial support from the Czech Science Foundation grant 13-03538S and ŠS acknowledges financial support from the Czech Science Foundation grant 13-35273P. MZ was supported by RVO61389005 and acknowledges financial support from the Czech Science Foundation grant 16-22945S.

References

[1] Burger E B 2000 *Exploring the Number Jungle: A Journey Into Diophantine Analysis* (Providence, RI: American Mathematical Society)
[2] Bohr N 1913 *Phil. Mag.* **26** 476
[3] Reed M and Simon B 1980 *Methods of Modern Mathematical Physics* vol 1–4 (New York: Academic)
[4] Bagarello F, Gazeau J P, Szafraniec F H and Znojil M (ed) 2015 *Non-Selfadjoint Operators in Quantum Physics: Mathematical Aspects* (New York: Wiley)
[5] Exner P and Gawlista R 1996 *Phys. Rev.* B **53** 7275–86
[6] Exner P and Turek O 2015 *Integr. Equat. Oper. Th.* **81** 535–57
[7] Astels S 1999 *Trans. Am. Math. Soc.* **352** 133–70
[8] Falconer K 2003 *Fractal Geometry: Mathematical Foundations and Applications* 2nd edn (New York: Wiley)
[9] Hall M Jr 1947 *Ann. of Math.* **48** 966–93
[10] Diviš B 1973 *Acta Arith.* **22** 157–73
[11] Bosma W, Jager H and Wiedijk F 1983 *Indagat. Math.* **86** 281–99
[12] Krejčiřík D and Siegl P 2015 Elements of spectral theory without the spectral theorem *Non-Selfadjoint Operators in Quantum Physics: Mathematical Aspects* 1st edn ed F Bagarello *et al* (New York: Wiley) pp 241–92
[13] Efimov V 1970 *Phys. Lett.* B **33** 563–4
[14] Tung S K, Jiménez-García K, Johansen J, Parker C V and Chin C 2014 *Phys. Rev. Lett.* **113** 240402
[15] Li J Q, Li Q and Miao Y G 2012 *Commun. Theor. Phys.* **58** 497–503
[16] Behrndt J and Krejčiřík D 2014 *J. Anal. Math.* arXiv:1407.7802 to appear
[17] Kostrykin V 2012 The div(a grad) operator without ellipticity self-adjointness and spectrum *Geometric Aspects of Spectral Theory,* 33/2012 Mathematisches Forschungsinstitut Oberwolfach p 2061 joint work with Amru Hussein, David Krejcirik, Stepan Schmitz
[18] Hussein A 2013 *PhD Thesis* Johannes Gutenberg-Universität Mainz
[19] Schmitz S 2014 *PhD Thesis* Johannes Gutenberg-Universität Mainz
[20] Znojil M and Lévai G 2012 *Phys. Lett.* A **376** 3000–5
[21] Lévai G, Siegl P and Znojil M 2009 *J. Phys. A: Math. Theor.* **42** 295201
[22] Bender C M 2007 *Rep. Prog. Phys.* **70** 947
[23] Chen T J, Fasiello M, Lim E A and Tolley A J 2013 *J. Cosmol. Astropart. Phys.* JCAP(2013)042

[24] Landau L D and Lifshitz L M 1977 *Quantum Mechanics (Non-Relativistic Theory)* (Oxford: Elsevier)
[25] Hawking S W and Hertog T 2002 *Phys. Rev.* D **65** 103515
[26] Wang A and Maartens R 2010 *Phys. Rev.* D **81** 024009
[27] Ilhan I B and Kovner A 2013 *Phys. Rev.* D **88** 044045
[28] Mannheim P D 2012 *Found. Phys.* **42** 388–420
[29] Klusoň J, Oksanen M and Tureanu A 2014 *Phys. Rev.* D **89** 064043
[30] Exner P and Kovarik H 2015 *Quantum Waveguides* (Berlin: Springer)
[31] Krejčiřík D private communication; see also http://www.math.unipa.it/~phhqp15/pdf/krejcirik.pdf
[32] Znojil M 2015 Non-selfadjoint operators in quantum physics: ideas people and trends *Non-Selfadjoint Operators in Quantum Physics: Mathematical Aspects* 1st edn ed F Bagarello *et al* (New York: Wiley) pp 7–58
[33] Bender C M and Milton K A 1997 *Phys. Rev.* D **55** R3255–9
[34] Mostafazadeh A and Loran F 2008 *Europhys. Lett.* **81** 10007
[35] Albeverio S and Kuzhel S 2015 PT-symmetric operators in quantum mechanics: Krein spaces methods *Non-Selfadjoint Operators in Quantum Physics: Mathematical Aspects* 1st edn ed F Bagarello *et al* (New York: Wiley) pp 293–344
[36] Mostafazadeh A 2010 *Int. J. Geom. Methods. M* **07** 1191–306
[37] Smith D R, Pendry J B and Wiltshire M C K 2004 *Science* **305** 788–92
[38] Depine R and Lakhtakia A 2004 *Microw. Opt. Technol. Lett.* **41** 315–6
[39] Kriegler C E, Rill M S, Linden S and Wegener M 2010 *IEEE J. Sel. Top. Quantum* **16** 367–75
[40] Tanaka T and Ishikawa A 2014 Fabrication of isotropic infrared metamaterials *Advanced Electromagnetic Materials in Microwaves and Optics (METAMATERIALS)* pp 418–20
[41] Soukoulis C M and Wegener M 2011 *Nature Photon.* **5** 523–30
[42] Bastard G 1988 *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Ulis: Les Editions de Physique)
[43] Shelby R A, Smith D R and Schultz S 2001 *Science* **292** 77–9
[44] Davies E B 2007 *Linear Operators and Their Spectra* (Cambridge: Cambridge University Press)
[45] Chen J H, Pelantová E and Znojil M 2008 *Phys. Lett.* A **372** 1986–9
[46] Znojil M 2012 *J. Phys. A: Math. Theor.* **45** 444036
[47] Pais A and Uhlenbeck G 1950 *Phys. Rev.* **79** 145–65
[48] Smilga A V 2009 *SIGMA* **5** 017
[49] Lopez-Sarrion J and Reyes C M 2013 *Eur. Phys. J.* C **73** 2391
[50] Jimenez J B, Di Dio E and Durrer R 2013 *J. High Energy Phys.* JHEP04(2013)030
[51] Kaparulin D S, Lyakhovich S L and Sharapov A A 2014 *Eur. Phys. J.* C **74** 3072
[52] Kaparulin D S and Lyakhovich S L 2015 *Russ. Phys.* J **57** 1561–5
[53] Bagchi B, Modak S, Panigrahi P K, Ruzicka F and Znojil M 2015 *Mod. Phys. Lett.* A **30** 1550213
[54] Bagarello F 2011 *Int. J. Theor. Phys.* **50** 3241–50
[55] Bolonek K and Kosinski P 2005 *Acta Phys. Pol.* B **36** 2115–31
[56] Masterov I 2015 *Mod. Phys. Lett.* A **30** 1550107
[57] Bender C M and Mannheim P D 2008 *Phys. Rev. Lett.* **100** 110402
[58] Nucci M C and Leach P G L 2010 *Phys. Scripta* **81** 055003
[59] Nucci M C and Leach P G L 2009 *J. Math. Phys.* **50** 113508
[60] Bagchi B, Choudhury A G and Guha P 2013 *Mod. Phys. Lett.* A **28** 1375001
[61] Mostafazadeh A 2010 *Phys. Lett.* A **375** 93–8
[62] Mostafazadeh A 2011 *Phys. Rev.* D **84** 105018
[63] Pulgar G, Saavedra J, Leon G and Leyva Y 2015 *J. Cosmol. Astropart. Phys.* JCAP15(2015)046
[64] Jizba P, Kleinert H and Scardigli F 2015 *Eur. Phys. J.* C **75** 245
[65] Mannheim P D 2006 *Prog. Part. Nucl. Phys.* **56** 340–445