FIXED POINTS OF MORPHISMS AMONG BINARY GENERALIZED
PSEUDOSTANDARD WORDS

LUBOMÍRA DVOŘÁKOVÁ AND TEREZA VELKÁ

Abstract. We introduce a class of fixed points of primitive morphisms among aperiodic binary
generalized pseudostandard words. We conjecture that this class contains all fixed points of
primitive morphisms among aperiodic binary generalized pseudostandard words that are not
standard Sturmian words.

1. Introduction

This note is devoted to binary infinite words generated by a construction called generalized
pseudopalindromic closure, known as generalized pseudostandard words. Concerning this topic,
the following facts are known so far: In [3] the generalized pseudostandard words were defined
and it was proved there that the famous Thue–Morse word is an example of such words. The
authors of [5] characterized generalized pseudostandard words in the class of generalized Thue–
Morse words. A necessary and sufficient condition on periodicity of binary and ternary generalized
pseudostandard words was provided in [1]. The authors of [2] focused on binary generalized pseu-
dostandard words and obtained several interesting results: an algorithm for a so-called normaliza-
tion, an effective algorithm for generation of such words, description of generalized pseudostandard
words among Rote words.

In this paper, we introduce a new class of aperiodic binary generalized pseudostandard words
being fixed points of primitive morphisms. We moreover conjecture that this is the only class
among aperiodic binary generalized pseudostandard words except standard Sturmian words con-
sisting of fixed points of primitive morphisms.

2. Preliminaries

Throughout the paper, $\mathbb{N}$ denotes natural numbers, i.e., $\mathbb{N} = \{1, 2, 3, \ldots\}$, while $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. We restrict ourselves to the binary alphabet $\{0, 1\}$, we call 0 and 1 letters. A (finite) word $w$ over $\{0, 1\}$ is any finite binary sequence. Its length $|w|$ is the number of letters $w$ contains. The empty word – the neutral element for concatenation of words – is denoted by $\varepsilon$ and its length is set $|\varepsilon| = 0$. The set of all finite binary words is denoted by $\{0, 1\}^*$. If a finite word $w = pvs$, where $p, v, s \in \{0, 1\}^*$, then $p$ is called a prefix of $w$ and $s$ is called a suffix of $w$. An infinite word $u$ over $\{0, 1\}$ is any binary infinite sequence. The set of all infinite words is denoted by $\{0, 1\}^\mathbb{N}$. A finite word $w$ is a factor of the infinite word $u = u_0u_1u_2 \ldots$ with $u_i \in \{0, 1\}$ if there exists an index $i \geq 0$ such that $w = u_iu_{i+1} \ldots u_{i+|w|-1}$. Such an index is called an occurrence of $w$ in $u$.

An infinite word $u$ is called recurrent if each of its factors occurs infinitely many times in $u$. It is said to be uniformly recurrent if for every $n \in \mathbb{N}$ there exists a length $r(n)$ such that every factor of length $r(n)$ of $u$ contains all factors of length $n$ of $u$. We say that an infinite word $u$ is eventually periodic if there exists $v, w \in \{0, 1\}^*$ such that $u = uv^\omega$, where $\omega$ denotes an infinite repetition. If $w = \varepsilon$, we call $u$ (purely) periodic. If $u$ is not eventually periodic, $u$ is said to be aperiodic. It is not difficult to see that if an infinite word is recurrent and eventually periodic, then it is necessarily purely periodic.

1
A morphism is a map \( \varphi : \{0,1\}^* \to \{0,1\}^* \) such that for every \( v, w \in \{0,1\}^* \) we have \( \varphi(vw) = \varphi(v)\varphi(w) \). It is clear that in order to define a morphism, it suffices to provide letter images. Application of a morphism \( \varphi \) may be naturally extended to an infinite word \( u = u_0u_1u_2\ldots \in \{0,1\}^\mathbb{N} \) as \( \varphi(u) = \varphi(u_0)\varphi(u_1)\varphi(u_2)\ldots \) If \( u = \varphi(u) \) for some infinite or finite word \( u \), we call \( u \) a fixed point of \( \varphi \). Let us make a trivial observation: any periodic infinite word \( u = \omega^\infty \) is a fixed point of a morphism (for instance, it suffices to set every letter image equal to \( \epsilon \)).

A morphism \( \varphi \) is said to be primitive if there exists \( n \in \mathbb{N} \) such that both \( \varphi^n(0) \) and \( \varphi^n(1) \) contain both letters 0 and 1. It is obvious that if an infinite word \( u \in \{0,1\}^\mathbb{N} \) satisfies \( \varphi(u) = u \) for some primitive morphism \( \varphi \), then there exists a letter \( a \in \{0,1\} \) such that \( \varphi(a) \) starts in \( a \) and \( |\varphi(a)| \geq 2 \). We say that \( \varphi \) is prolongable on \( a \). The fixed point \( u \) starting in \( a \) has evidently \( \varphi^n(a) \) as its prefix for all \( n \in \mathbb{N} \). We sometimes write \( u = \lim_{n \to \infty} \varphi^n(a) \). It is known that fixed points of primitive morphisms are uniformly recurrent.

An involutory antimorphism is a map \( \vartheta : \{0,1\}^* \to \{0,1\}^* \) such that for every \( v, w \in \{0,1\}^* \) we have \( \vartheta(vw) = \vartheta(w)v\vartheta(v) \) and moreover \( \vartheta^2 \) equals identity. There are only two involutory antimorphisms over the alphabet \( \{0,1\} \): the reversal (mirror) map \( R \) satisfying \( R(0) = 0, R(1) = 1 \), and the exchange antimorphism \( E \) given by \( E(0) = 1, E(1) = 0 \). We use the notation \( 0 = 1 \) and \( 1 = 0 \). \( E = R \) and \( R = E \). A finite word \( w \) is a palindrome (an \( R \)-palindrome) if \( w = R(w) \), and \( w \) is an \( E \)-palindromic (pseudopalindromic) if \( w = E(w) \). The palindromic closure \( w^R \) of a word \( w \) is the shortest palindrome having \( w \) as prefix. Similarly, the pseudopalindromic closure \( w^E \) of \( w \) is the shortest \( E \)-palindrome having \( w \) as prefix.

2.1. Generalized pseudostandard words. Let us underline that we again restrict ourselves only to the binary alphabet \( \{0,1\} \).

**Definition 1.** Let \( \Delta = \delta_1\delta_2\ldots \in \{0,1\}^\mathbb{N} \) and \( \Theta = \vartheta_1\vartheta_2\ldots \in \{E,R\}^\mathbb{N} \). The infinite word \( u(\Delta, \Theta) \) generated by the generalized pseudopalindromic closure (or generalized pseudostandard word) is the word whose prefixes \( w_n \) are obtained from the recurrence relation

\[
w_{n+1} = (w_n\delta_{n+1})^{\vartheta_{n+1}},
\]

\[
w_0 = \epsilon.
\]

The sequence \( \Lambda = (\Delta, \Theta) \) is called the directive bi-sequence of the word \( u(\Delta, \Theta) \).

If \( \Theta = R^\infty \), the word \( u(\Delta, \Theta) \) is called \( R \)-standard. If it is moreover aperiodic, it is called standard Sturmian. If \( \Theta = E^\infty \), the word \( u(\Delta, \Theta) \) is called \( E \)-standard or pseudostandard.

It is readily seen that generalized pseudostandard words are uniformly recurrent.

The sequence \( (w_k)_{k \geq 0} \) of prefixes of a generalized pseudostandard word \( u(\Delta, \Theta) \) does not have to contain all \( E \)-palindromic and \( R \)-palindromic prefixes of \( u(\Delta, \Theta) \). Blondin Massé et al. [2] introduced the notion of normalization of the directive bi-sequence.

**Definition 2.** A directive bi-sequence \( \Lambda = (\Delta, \Theta) \) of a generalized pseudostandard word \( u(\Delta, \Theta) \) is called normalized if the sequence of prefixes \( (w_k)_{k \geq 0} \) of \( u(\Delta, \Theta) \) contains all \( E \)-palindromic and \( R \)-palindromic prefixes of \( u(\Delta, \Theta) \).

The authors of [2] proved that every directive bi-sequence \( \Lambda \) can be normalized, i.e., transformed to such a form \( \Lambda \) that the new sequence \( (\tilde{w}_k)_{k \geq 0} \) contains already every \( E \)-palindromic and \( R \)-palindromic prefix and \( \Lambda \) generates the same generalized pseudostandard word as \( \Lambda \).

**Theorem 3.** Let \( \Lambda = (\Delta, \Theta) \) be a directive bi-sequence. Then there exists a normalized directive bi-sequence \( \tilde{\Lambda} = (\tilde{\Delta}, \tilde{\Theta}) \) such that \( u(\Delta, \Theta) = u(\tilde{\Delta}, \tilde{\Theta}) \).

Moreover, in order to normalize the sequence \( \Lambda \), it suffices firstly to execute the following changes of its prefix (if it is of the corresponding form):
\[ (\alpha \alpha, \theta \theta) \rightarrow (\alpha \alpha \alpha, \theta \theta \theta), \]

\[ (a^i, R^{i-1}E) \rightarrow (a^i \alpha, R^i E) \text{ for } i \geq 1, \]

\[ (a^i \alpha \alpha, R^i E E) \rightarrow (a^i \alpha \alpha \alpha, R^i E E) \text{ for } i \geq 1, \]

and secondly to replace step by step from left to right every factor of the form:

\[ (a^i, \vartheta \vartheta) \rightarrow (a \alpha, \vartheta \vartheta \vartheta), \]

where \( a, b \in \{0, 1\} \) and \( \vartheta \in \{E, R\} \).

A necessary and sufficient condition for periodicity of binary generalized pseudostandard words was found in [1].

**Theorem 4.** A binary generalized pseudostandard word \( u(\Delta, \Theta) \), where \( \Delta = \delta_1 \delta_2 \ldots \in \{0, 1\}^\mathbb{N} \) and \( \Theta = \theta_1 \theta_2 \ldots \in \{E, R\}^\mathbb{N} \), is periodic if and only if the directive bi-sequence \((\Delta, \Theta)\) satisfies the following condition:

\[ (\exists a \in \{0, 1\})(\exists \theta \in \{E, R\})(\exists n_0 \in \mathbb{N})(\forall n > n_0, n \in \mathbb{N})(\delta_{n+1} = a \Leftrightarrow \vartheta_n = \theta). \]

3. **Fixed points of morphisms among generalized pseudostandard words**

The aim of this section is to introduce a new class of aperiodic binary generalized pseudostandard words being fixed points of morphisms. The only known aperiodic binary examples are so far:

1. The Thue–Morse word which is the fixed point of the morphism \( 0 \mapsto 01, 1 \mapsto 10 \) (or \( 0 \mapsto 0110, 1 \mapsto 1001 \)) and also the pseudostandard word with the directive bi-sequence \((\Delta, \Theta) = (01^\omega, R(ER)^\omega)\), as described in [3].
2. R-standard words were studied in [4]. An R-standard word \( u(\Delta, R^\omega) \) is periodic if and only if the sequence \( \Delta \) is eventually constant, i.e., \( \Delta = v a^\omega \), where \( v \in \{0, 1\}^* \) and \( a \in \{0, 1\} \). An aperiodic R-standard (standard Sturmian) word \( u(\Delta, R^\omega) \) is a fixed point of a morphism if and only if the sequence \( \Delta \) is periodic.

3.1. **A new class of fixed points of morphisms among aperiodic binary generalized pseudostandard words.** We will now study morphisms \( \varphi_k \), for \( k \in \mathbb{N} \), of the following form:

\[ \varphi_k : \begin{array}{ll}
0 & \mapsto 0(110)^k, \\
1 & \mapsto 1(001)^k.
\end{array} \]

Such a morphism \( \varphi_k \) has evidently two fixed points, \( \lim_{n \to \infty} \varphi_k^n(0) \) and \( \lim_{n \to \infty} \varphi_k^n(1) \). We will prove that the first fixed point is a generalized pseudostandard word whose directive bi-sequence equals \((\Delta, \Theta) = (01^\omega, R(ER^k)^\omega)\) and the second fixed point has the directive bi-sequence \((\Delta, \Theta) = (10^\omega, R(ER^k)^\omega)\).

First of all, given the fact that \( \varphi_k(0) = \overline{\varphi_k(1)} \), the fixed point \( \lim_{n \to \infty} \varphi_k^n(0) \) equals \( \lim_{n \to \infty} \varphi_k^n(1) \) and therefore, without loss of generality, we will study the fixed point \( u = \lim_{n \to \infty} \varphi_k^n(0) \) and then easily generalize the results for both cases.

Let us consider the directive bi-sequence \((\Delta, \Theta) = (01^\omega, R(ER^k)^\omega)\). Note that this directive bi-sequence is normalized by Theorem 3. Remark also that for every \( k \in \mathbb{N} \) the corresponding word \( u = u(\Delta, \Theta) \) is aperiodic by Theorem 4. We will show in two steps a relation between the prefixes \( w_k \) obtained by the generalized pseudopalindromic closure and the morphism \( \varphi_k \).

Firstly, we will show a recursive relation concerning the prefixes \( w_k \) of the word \( u \). Let \( K \) denote the length of the shortest period of \( \Theta \), i.e., \( K = |RER^{k-1}| = k + 1 \). We will show that the construction of these prefixes can be expressed depending on \( K \).
Lemma 5. Let \((\Delta, \Theta) = (01^\omega, R(ER^k)^\omega)\) be the directive bi-sequence of the word \(u\). Let \(w_i\), \(i \in \mathbb{N}\), be its prefixes obtained by the generalized pseudopalindromic closure and \(K = k + 1\) the length of the shortest period of \(\Theta\). Then for all \(l \in \mathbb{N}_0\) the following holds:

\[
\begin{align*}
w_{l \cdot K + 2} &= w_{l \cdot K + 1} \cdot E(w_{l \cdot K + 1}) = w_{l \cdot K + 1} w_{l \cdot K + 1}^{R}\quad (E\text{-palindrome}) \\
w_{l \cdot K + 3} &= w_{l \cdot K + 2} \cdot R(w_{l \cdot K + 2}) = w_{l \cdot K + 2} w_{l \cdot K + 2}^{R}\quad (R\text{-palindrome}) \\
w_{l \cdot K + 4} &= w_{l \cdot K + 3} \cdot w_{l \cdot K + 1}^{R} w_{l \cdot K + 3}\quad (R\text{-palindrome}) \\
w_{l \cdot K + 5} &= w_{l \cdot K + 4} \cdot w_{l \cdot K + 3}^{R} w_{l \cdot K + 4}\quad (R\text{-palindrome}) \\
&\vdots \quad \vdots \quad \vdots \\
w_{l \cdot K + K} &= w_{l \cdot K + (K - 1)} \cdot w_{l \cdot K + (K - 2)}^{R} w_{l \cdot K + (K - 1)}\quad (R\text{-palindrome}) \\
w_{(l + 1) \cdot K + 1} &= w_{(l + 1) \cdot K} \cdot w_{l \cdot K + (K - 1)}^{R} w_{(l + 1) \cdot K}\quad (R\text{-palindrome})
\end{align*}
\]

Remark 6. The reason why we start from \(l \cdot K + 2\) is that \(w_1\) is the only prefix for which the lemma does not hold. Naturally, if \(K = k + 1 = 2, 3\) or \(4\), only the first \(2, 3\) or \(4\) lines hold and then another cycle is started with \(l = l + 1\).

Proof. Let \(\Delta = \delta_1 \delta_2 \ldots\) and \(\Theta = \vartheta_1 \vartheta_2 \ldots\). We will prove the lemma directly using the construction of pseudostandard words. In each case, we will construct \(w_{i+1} = (w_i \delta_{i+1})^{\vartheta_{i+1}}\) after finding the longest \(\vartheta_{i+1}\)-palindromic suffix of \(w_i \delta_{i+1}\). We will split the proof into several cases:

1. Case of \(w_{l \cdot K + 2}\)

If we show that the longest \(E\)-palindromic suffix of \(w_{l \cdot K + 1}\) is \(01\), then it is clear that \(w_{l \cdot K + 2} = w_{l \cdot K + 1} \cdot E(w_{l \cdot K + 1}) = w_{l \cdot K + 1} w_{l \cdot K + 1}^{R}\). The last equality holds because using the form of \(\Theta = R(ER^k)^\omega\) one can see that \(w_{l \cdot K + 1}\) is an \(R\)-palindrome, and applying \(E\) on an \(R\)-palindrome is equivalent to the exchange of zeros and ones.

The longest \(E\)-palindromic suffix is at least \(01\) because \(w_{l \cdot K + 1}\) is an \(R\)-palindromic and thus ends with a 0. It cannot be longer than \(01\) because that would mean there exists an \(E\)-palindromic prefix of \(w_{l \cdot K + 1}\) followed by a 0 (the prefix \(w_{l \cdot K + 1}\) is an \(R\)-palindromic and the reverse image of its \(E\)-palindromic suffix preceded by a 0 is an \(E\)-palindromic prefix followed by a 0) and this is not possible since every \(E\)-palindromic prefix is followed by 1 due to the fact that \(\Delta\) is normalized and \(\Delta = 01^\omega\).

2. Case of \(w_{l \cdot K + 3}\)

Similarly to the case above, if we show that the longest \(R\)-palindromic suffix of \(w_{l \cdot K + 2}\) is \(11\), it directly follows that \(w_{l \cdot K + 3} = w_{l \cdot K + 2} \cdot R(w_{l \cdot K + 2}) = w_{l \cdot K + 2} w_{l \cdot K + 2}^{R}\). The last equality holds because again, applying the antimorphism \(R\) on an \(E\)-palindrome is equivalent to the exchange of zeros and ones.

The factor \(11\) is an \(R\)-palindromic suffix of \(w_{l \cdot K + 2}\) because \(w_{l \cdot K + 2}\) ends with a 1 since it is an \(E\)-palindromic that begins with 0. Moreover, every \(R\)-palindromic suffix of \(w_{l \cdot K + 2}\) is preceded by a 0 because they are \(E\)-images of \(R\)-palindromic prefixes of \(w_{l \cdot K + 2}\) and all non-empty \(R\)-palindromic prefixes are preceded by 0 because of the form of the directive bi-sequence. Hence, we cannot find a longer \(R\)-palindromic suffix of \(w_{l \cdot K + 2}\) preceded by 1 and therefore \(11\) is the longest palindromic suffix of \(w_{l \cdot K + 2}\).

3. Case of \(w_{l \cdot K + 4}\)

In this case, we want to prove that \(1w_{l \cdot K + 4}\) is the longest \(R\)-palindromic suffix of \(w_{l \cdot K + 3}\). Then, we have \(w_{l \cdot K + 4} = w_{l \cdot K + 3} \cdot w_{l \cdot K + 3}^{R}\). Now, let us find the longest \(R\)-palindromic suffix of \(w_{l \cdot K + 3}\) preceded by a 1. Because of the normalization of \(\Delta\), we know the only pseudopalindromic suffixes are images of the
prefixes \( w_i \). It cannot be \( w_{l \cdot K + 3} \) since it is too long. It cannot be also \( R(w_{l \cdot K + 2}) \) since it is an \( E \)-palindrome. The longest possibility is now \( R(w_{l \cdot K + 1}) = w_{l \cdot K + 1} \) and it is the correct one because it is preceded by 1 since \( w_{l \cdot K + 3} \) is an \( R \)-palindrome and the prefix \( w_{l \cdot K + 1} \) is followed by 1 according to the form of the directive bi-sequence.

(4) Cases of \( w_{l \cdot K + 5} \) to \( w_{(l+1) \cdot K + 1} \)

In these last cases, we will proceed analogously to the previous case. We want to find the longest \( R \)-palindromic suffix of \( w_{l \cdot K + (i-1) \cdot 1} \), \( i \in \{5, \ldots, K + 1 \} \). It can be easily seen that it is \( 1 w_{l \cdot K + (i-2) \cdot 1} \) because of the normalization of \( \Delta \). It cannot be \( 1 w_{l \cdot K + (i-1) \cdot 1} \) (it is too long). The next longest \( R \)-palindromic suffix of \( w_{l \cdot K + (i-1)} \) preceded by a 1 is \( w_{l \cdot K + (i-2)} \) and therefore \( 1 w_{l \cdot K + (i-2) \cdot 1} \) is the longest \( R \)-palindromic suffix of \( w_{l \cdot K + (i-1) \cdot 1} \). As in the previous case, it follows that \( w_{l \cdot K + i} = w_{l \cdot K + (i-1)} w_{l \cdot K + (i-2)}^{-1} w_{l \cdot K + (i-1)} \).

\[ k = 1: \quad \Theta = R(ER^1)\omega \]
\[ k = 2: \quad \Theta = R(ER^2)\omega \]
\[ k = 3: \quad \Theta = R(ER^3)\omega \]

Table 1. The first seven prefixes \( w_i \) of the word \( u = u(01^\omega, R(ER^k)\omega) \) for \( k = 1 \) (the Thue–Morse word), 2 and 3.

Secondly, we will prove a proposition showing a relation between the prefixes of \( u = u(\Delta, \Theta) \) and the prefixes of the fixed point of \( \varphi_k \). The fact that \( u \) is a fixed point of \( \varphi_k \) follows then immediately from this proposition.

**Proposition 7.** Let \( (\Delta, \Theta) = (01^\omega, R(ER^k)\omega) \) be the directive bi-sequence of the word \( u \). Let \( w_i \), \( i \in \mathbb{N} \), be its prefixes obtained by the generalized pseudopalindromic closure and \( K = k + 1 \) the length of the shortest period of \( \Theta \). Let \( \varphi_k \) be the morphism defined in (1). Then for every \( l \in \mathbb{N} \):

\[ w_{l \cdot K + r} = \varphi_k(w_{(l-1) \cdot K + r}) \]

for all \( r \in \{1, 2, \ldots, K \} \).

**Proof.** We proceed by induction on \( l \).

First, we will prove that if the equality (2) holds for some \( l - 1 \), it also holds for \( l \). The proof for \( l = 1 \) will be then easier because during the proof of the induction step, we will show that for a fixed \( l \), if the equality holds for \( r = 1 \), then it easily follows that (2) holds for all \( r \in \{2, \ldots, K \} \).

We rewrite the prefixes \( w_i \) using Lemma 5 and we apply the fact that \( \varphi_k(w) = \overline{\varphi_k(w)} \).
Conjecture 9. Let \( \Phi_{k} \) be an aperiodic binary generalized pseudostandard word, not standard Sturmian, a fixed point of a primitive morphism. Then \( \Phi_{k} \equiv \Phi_{k}(u) \) for some \( k \in \mathbb{N} \), where \( \Phi_{k} \) is the morphism defined in (1).

4. Open problems

According to our computer experiments, it seems that the morphisms defined in (1) are the only primitive morphisms whose fixed points are aperiodic binary generalized pseudostandard words that are not standard Sturmian words. Let us state thus the following conjecture.

Conjecture 9. Let \( \Phi_{k} \) be aperiodic binary generalized pseudostandard word, not standard Sturmian, being a fixed point of a primitive morphism. Then \( \Phi_{k} \equiv \Phi_{k}(u) \) for some \( k \in \mathbb{N} \), where \( \Phi_{k} \) is the morphism defined in (1).
In this note, we focused only on binary words. Generalized pseudostandard words are defined over any alphabet \( \{0, 1, \ldots, m - 1\} \) for \( m > 1, m \in \mathbb{N} \) [3]. To our knowledge, the only fact known about fixed points of primitive morphisms in a multiliteral case is a result from [5]: The generalized Thue–Morse words are defined for \( m, b \in \mathbb{N}, m > 1, b > 1 \) as follows:

\[
t_{b,m} = \left( s_b(n) \mod m \right)_{n=0}^\infty
\]

where \( s_b(n) \) denotes the digit sum of the expansion of number \( n \) in the base \( b \). Such words are fixed points of morphisms. A generalized Thue–Morse word \( t_{b,m} \) is a generalized pseudostandard word if and only if \( b \leq m \) or \( b - 1 = 0 \pmod{m} \). Note that the Thue–Morse word is a special case of \( t_{b,m} \) for \( b = m = 2 \). For the form of morphisms whose fixed points are the generalized Thue–Morse words and for the form of their directive bi-sequence \( (\Delta, \Theta) \) and other properties, see [5]. Hence, it is an open problem to detect fixed points of primitive morphisms over larger alphabets in the class of aperiodic generalized pseudostandard words.

ACKNOWLEDGEMENTS

We acknowledge the financial support of Czech Science Foundation GAČR 13-03538S.

REFERENCES


