Repetitions in beta-integers

L’. Balková, K. Klouda, E. Pelantová *

Doppler Institute and Department of Mathematics FNSPE,
Czech Technical University in Prague,
Trojanova 13, 120 00 Praha 2, Czech Republic

Abstract

Classical crystals are solid materials containing arbitrarily long periodic repetitions of a single motif. In this Letter, we study the maximal possible repetition of the same motif occurring in \( \beta \)-integers – one dimensional models of quasicrystals. We are interested in \( \beta \)-integers realizing only a finite number of distinct distances between neighboring elements. In such a case, the problem may be reformulated in terms of combinatorics on words as a study of the index of infinite words coding \( \beta \)-integers. We will solve a particular case for \( \beta \) being a quadratic non-simple Parry number.

1 Introduction

This Letter takes up the study of \( \beta \)-integers initiated by the investigation of their asymptotic properties in \[1\]. Similarly as in the previous Letter, we restrict our consideration to \( \beta \)-integers realizing only a finite number of distinct distances between neighbors; \( \beta \) is then called a Parry number. For Parry numbers, the set of \( \beta \)-integers forms a discrete aperiodic Delone set with a self-similarity factor \( \beta \) and of finite local complexity. It follows herefrom that \( \beta \)-integers are suitable for modeling materials with aperiodic long range order, the so-called quasicrystals \[2\]. Classical crystals are solid materials containing arbitrarily long periodic repetitions of a single motif. Quasicrystals do not share this property.

In this Letter, we are interested in the maximal possible repetition of one motif occurring in \( \beta \)-integers. It turns out to be suitable to reformulate and study this problem in terms of combinatorics on words.

For Parry numbers, coding distinct distances between neighboring nonnegative \( \beta \)-integers with distinct letters, one obtains a right-sided infinite word \( u_\beta \) over a finite alphabet. The reformulation of our task in the language of combinatorics on words has the following reading: For a given factor \( w \) of the infinite word \( u = u_\beta \), find the longest prefix \( v \) of the infinite periodic word \( w^\infty = \ldots vvvvvvv \) such that \( v \) occurs as a factor in \( u = u_\beta \). The ratio of the lengths of \( v \) and \( w \) is called the index of the factor \( w \) in \( u = u_\beta \) and is denoted by \( \text{ind}(w) \). Let us note that \( \text{ind}(w) \) is not necessarily an integer. Denote by \( k \) the lower integer part \( \lfloor \text{ind}(w) \rfloor \) of the index of \( w \), then the word \( w^k \), i.e., the concatenation of \( k \) words \( w \), is usually called the maximal integer power of \( w \).

The index of any infinite word \( u \) can be naturally defined as

\[
\text{ind}(u) = \sup \{ \text{ind}(w) \mid w \text{ factor of } u \}.
\]

Explicit values of the index are known only for few classes of infinite words. The index of Sturmian words has been studied in many papers \([3, 4, 5, 6]\), the complete solution to the problem was provided independently by Carpi and de Luca \[11\] and by Damanik and Lenz \[8\]. Recently, the index of infinite words has reinforced its importance: Damanik in \[9\] considers discrete one-dimensional Schrödinger operators with aperiodic potentials generated by primitive morphisms.

* e-mail: l.balkova@centrum.cz, karel@kloudak.eu, edita.pelantova@fjfi.cvut.cz
and he establishes purely singular continuous spectrum with probability one provided that the potentials (infinite words) have the index greater than three. Let us stress that infinite words $u_\beta$ associated with Parry numbers belong to the class of infinite words generated by primitive morphisms, too.

Here we study the index of infinite words $u_\beta$ for quadratic non-simple Parry numbers $\beta$. These words are determined by integer parameters $p, q$, where $p > q \geq 1$. We provide an explicit formula for $\text{ind}(u_\beta)$. In the particular case of $p = q + 1$, the infinite word $u_\beta$ is Sturmian and our result may be deduced also from the well-known formula for the index of Sturmian words. We have chosen the word $u_\beta$ associated with quadratic non-simple Parry numbers $\beta$ for our study of the index of non-Sturmian words since for such infinite words, we dispose of detailed knowledge on arithmetical properties of $\beta$-integers and combinatorial properties of the associated infinite words $u_\beta$ (\cite{10, 11}).

The Letter is organized in the following way. In Section 2, we introduce necessary notions from combinatorics on words and we cite a relevant result on the index of Sturmian words. In Section 3, we provide the background on infinite words from combinatorics on words and we cite a relevant result on the index of Sturmian words. In Section 4, we determine the maximal integer power occurring in $u_\beta$ (Theorem 4.6). Section 5 is devoted to the index of $u_\beta$ (Theorem 5.3) and to the comparison of our result with the formula for the index of Sturmian words.

2 Preliminaries

An alphabet $\mathcal{A}$ is a finite set of symbols, called letters. Throughout this paper, the binary alphabet $\mathcal{A} = \{0, 1\}$ is used. The string $w = w_1w_2 \ldots w_k$, where $w_i \in \mathcal{A}$ for each $i = 1, 2, \ldots, k$, is called a word of length $k$ on $\mathcal{A}$. The length of $w$ is then $k$ and it is denoted by $|w| = k$. The set of all finite words together with the operation of concatenation forms a monoid; its neutral element is the empty word $\varepsilon$. We denote this monoid $\mathcal{A}^*$. An infinite sequence $u = u_0u_1u_2 \ldots$ of symbols from the alphabet $\mathcal{A}$ is called an infinite word. A finite word $w$ is said to be a factor of the (finite or infinite) word $v$ if there exists a finite word $v'$ and a finite or infinite word $v''$ such that $v = v'vw''$. If $v'$ is the empty word, then $w$ is called a prefix of $v$, if $v''$ is the empty word, then $w$ is a suffix of $v$. If $w = v'w$, then $vw^{-1}$ is obtained from $v$ by erasing its suffix $w$. The set of all factors of an infinite word $u$ is said to be the language of $u$ and is denoted $\mathcal{L}(u)$. An infinite word $u$ is called recurrent if every of its factors occurs infinitely many times in $u$ and $u$ is called uniformly recurrent if for every of its factors $v$, the set of all factors in $\mathcal{L}(u)$ that do not contain $w$ as their factor is finite. In other words, every sufficiently long element of $\mathcal{L}(u)$ contains $w$ as its factor.

The number of factors of the infinite word $u$ gives us insight into its variability. The function $C_u : \mathbb{N} \rightarrow \mathbb{N}$ that to every $n$ associates the number of distinct factors of length $n$ occurring in $u$ is called the factor complexity of the infinite word $u$. An infinite periodic word $u = www \ldots$, where $w$ is a finite word, is usually denoted $w^\omega$. Its factor complexity $C_u$ is bounded; it is readily seen that $C_u(n) \leq |w|$. Similarly, the factor complexity of an eventually periodic word $u = w'w^\omega$, where $w', w$ are finite words, is bounded. The necessary and sufficient condition for an infinite word to be aperiodic is validity of the equation $C_u(n) \geq n + 1$ for all $n \in \mathbb{N}$. Infinite aperiodic words satisfying $C_u(n) = n + 1$ for all $n \in \mathbb{N}$ are called Sturmian words; these are thus infinite aperiodic words of the lowest possible factor complexity. Sturmian words are the best known aperiodic words; a survey on their properties may be consulted in \cite{12}. In particular, any Sturmian word is uniformly recurrent.

For determination of the factor complexity of an infinite word $u$, an essential role is played by special factors. We recall that a factor $w$ of an infinite word $u$ over a binary alphabet $\{0, 1\}$ is called left special if $0w$ and $1w$ are both factors of $u$, $w$ is called right special if $w0$ and $w1$ are both factors of $u$, and $w$ is said to be bispecial if $w$ is both left special and right special.

The crucial notion of our study is the index of a factor $w$ in a given infinite word $u$. Let us define first the integer power of $w$. For any $k \in \mathbb{N}$, the $k$-th power of $w$ is the concatenation of $k$ words $w$, usually denoted $w^k$. Analogously for $r \in \mathbb{Q}, r \geq 1$, we call a word $v$ the $r$-th power of the
word \( w \) if there exists a proper prefix \( w' \) of \( w \) such that

\[
v = \underbrace{w \ldots w}_r w' \quad \text{and} \quad r = \lfloor r \rfloor + \frac{|w'|}{|w|}.
\]

The \( r \)-th power of \( w \) is denoted \( w^r \).

Our aim is to find, for a given factor \( w \) of an infinite word \( u \), the highest power of \( w \) occurring in \( u \). We will be interested exclusively in aperiodic uniformly recurrent words. Such words contain for every factor \( w \) only \( r \)-th powers of \( w \) with bounded \( r \), see [13]. Therefore, for aperiodic uniformly recurrent words \( u \), it makes sense to define the index of \( w \) in \( u \) as

\[
\ind(w) = \max \{ r \in \mathbb{Q} \mid w^r \in \mathcal{L}(u) \}.
\]

The word \( w^{\ind(w)} \) is called the maximal power of \( w \) in \( u \), the word \( w^{\lfloor \ind(w) \rfloor} \) the maximal integer power of \( w \) in \( u \). The index of the infinite word \( u \) is defined as

\[
\ind(u) = \sup \{ \ind(w) \mid w \in \mathcal{L}(u) \}.
\]

Let us remark that an aperiodic uniformly recurrent word \( u \) can have an infinite index; even among Sturmian words, one can find words with an infinite index. The language of every Sturmian word is characterized by an irrational parameter \( \alpha \in (0, 1) \), called slope. If \( \alpha \) is the slope of a Sturmian word \( u \), then the word obtained by exchanging letters in \( u \) has the slope \( 1 - \alpha \) and has evidently the same index as \( u \). Consequently, we may assume without loss of generality that \( \alpha > \frac{1}{2} \). In order to determine the index of a Sturmian word, we need to express \( \alpha \) in the form of its continued fraction. Since \( \frac{1}{2} < \alpha < 1 \), the continued fraction of \( \alpha \) equals \([0; 1, a_2, a_3, \ldots]\). The results of [7] and [8] say that the index of a Sturmian word \( u \) with slope \( \alpha \) is equal to

\[
\ind(u) = 2 + \sup \left\{ a_n + \frac{q_n - 2}{q_{n+1}} \mid n \in \mathbb{N} \right\},
\]

where \( q_n \) is the denominator of the \( n \)-th convergent of \( \alpha \).

Now, let us describe a large class of uniformly recurrent words: fixed points of primitive morphisms. This class includes infinite words \( u_\beta \) associated with Parry numbers \( \beta \). The map \( \varphi : \mathcal{A}^* \mapsto \mathcal{A}^* \) is called a morphism if \( \varphi(wv) = \varphi(w)\varphi(v) \) for every \( w, v \in \mathcal{A}^* \). One may associate with \( \varphi \) the morphism matrix \( M_\varphi \) satisfying

\[(M_\varphi)_{ab} = \text{number of letters } b \text{ occurring in } \varphi(a),\]

for any pair of letters \( a, b \in \mathcal{A} \).

Knowing for any \( a \in \mathcal{A} \) the number of letters \( a \) occurring in a factor \( w \), we may obtain the same information for \( \varphi(w) \) by a simple formula. We mention the formula only for the binary alphabet \( \mathcal{A} = \{0, 1\} \) we are interested in. It follows straightforwardly from the definition of the morphism matrix that for every factor \( w \in \mathcal{A}^* \)

\[
(|\varphi(w)|_0, |\varphi(w)|_1) = (|w|_0, |w|_1)M_\varphi,
\]

where \( |v|_a \) denotes the number of letters \( a \) occurring in a word \( v \). Clearly, a similar formula holds for any finite alphabet.

A morphism is said to be primitive if a power of \( M_\varphi \) has all elements strictly positive. In other words, matrices of primitive morphisms fulfill the assumptions of the Perron-Frobenius theorem.

The action of the morphism \( \varphi \) may be naturally extended to an infinite word \( u = u_0u_1u_2 \ldots \) by the prescription

\[
\varphi(u_0u_1u_2 \ldots) = \varphi(u_0)\varphi(u_1)\varphi(u_2) \ldots
\]

An infinite word \( u \) is called a fixed point of \( \varphi \) if \( \varphi(u) = u \). It is known that any fixed point of a primitive morphism is uniformly recurrent and that any left eigenvector corresponding to the dominant eigenvalue of \( M_\varphi \) is proportional to the densities of letters in any fixed point of \( \varphi \).
3 Infinite words associated with $\beta$-integers

Here, we provide the description of infinite words $u_{\beta}$ associated with quadratic non-simple Parry numbers in terms of fixed points of primitive morphisms. We keep the notation from our precedent Letter [1], where the number theoretical background on $\beta$-integers and associated infinite words $u_{\beta}$ is at disposal. Nevertheless, to make this Letter self-contained, we will recall all notions needed for understanding of our results. In this section we also deduce some important properties of $u_{\beta}$, in particular, a transformation generating bispecial factors of $u_{\beta}$. Bispecial factors turn out to be essential for our main aim – determination of the maximal integer powers of factors and determination of the index of $u_{\beta}$.

A non-simple quadratic Parry number $\beta$ is the larger root of $x^2 - (p+1)x + p - q$, where $p > q \geq 1$. Its Rényi expansion of unity is $d_{\beta}(1) = pq^\omega$. The set of $\beta$ integers has two distances between neighbors: $\Delta_0 = 1$ and $\Delta_1 = \beta - p$. Consequently, the infinite word $u_{\beta}$ coding distances between neighboring $\beta$-integers is binary.

As we have already said, the reader may find the notions of the Rényi expansion of unity, $\beta$-integers, distances between neighbors in $\mathbb{Z}_\beta$ etc. in our precedent Letter [1]. However, in order to follow the ideas in the sequel, it is sufficient to know that $u_{\beta}$ is the unique fixed point of a morphism canonically associated with parameters $p, q$ characterizing non-simple quadratic Parry numbers $\beta$. Therefore we will use the result of [14] for an equivalent definition of $u_{\beta}$.

**Definition 3.1.** Let $\beta$ is the larger root of $x^2 - (p+1)x + p - q$, where $p > q \geq 1$. The unique fixed point of the morphism

$$\varphi(0) = 0^p1, \quad \varphi(1) = 0^q1$$

will be denoted $u_{\beta}$.

The infinite word $u_{\beta}$ starts as follows

$$u_{\beta} = \underbrace{0^p1 \ldots 0^p1}_{p \text{ times}} \underbrace{0^q1 \ldots 0^q1}_{q \text{ times}} \ldots$$

The morphism matrix $M_\varphi$ is $\left(\begin{array}{c} p \\ q \end{array}\right)$ and $\varphi$ is thus obviously primitive. Computing the left eigenvector of $M_\varphi$ corresponding to the dominant eigenvalue $\beta$, we get the densities $1 - \frac{1}{\beta}$ and $\frac{1}{\beta}$ of letters $0$ and $1$, respectively.

**Remark 3.2.** In the paper [10], it is shown that the factor complexity $C$ of $u_{\beta}$ satisfies

- if $p > q + 1$, then $\{C(n+1) - C(n) \mid n \in \mathbb{N}\} = \{1, 2\}$,
- if $p = q + 1$, then $\{C(n+1) - C(n) \mid n \in \mathbb{N}\} = \{1\}$.

Therefore, $u_{\beta}$ is Sturmian if and only if $p = q + 1$.

First of all, some simple, but very important properties of the morphism $\varphi$ are observed.

**Observation 3.3.** Let $10^k1$ be a factor of $u_{\beta}$, then $k = p$ or $k = q$.

**Observation 3.4.** Let $v$ be any factor of $u_{\beta}$ containing at least one 1. Then there exists $k_1, 0 \leq k_1 \leq p$, such that $0^{k_1}1$ is a prefix of $v$ and there exists $k_2, 0 \leq k_2 \leq p$, such that $10^{k_2}$ is a suffix of $v$. The fact that $\varphi(0)$ and $\varphi(1)$ end in 1 and contain only one letter 1 implies that there exists a unique word $w$ in $\{0, 1\}^*$ satisfying $v = 0^{k_1}1\varphi(w)0^{k_2}$. Clearly, $w$ is a factor of $u_{\beta}$.

Of significant importance is the map $T : \{0, 1\}^* \rightarrow \{0, 1\}^*$ defined by

$$T(w) = 0^q1\varphi(w)0^q.$$  

The map $T$ helps to generate bispecial factors that play a crucial role in the determination of the index of factors. Therefore, the rest of this section is devoted to the description of properties of $T$. 

4
Lemma 3.5. Let $T$ be the map defined in $\mathbb{L}$. 

1. For every $w \in \mathbb{L}(u_\beta)$, it holds that $T(w) \in \mathbb{L}(u_\beta)$.

2. Let $w$ be a factor of $u_\beta$ and let $a, b \in A$, then $awb \in \mathbb{L}(u_\beta)$ if and only if $aT(w)b \in \mathbb{L}(u_\beta)$.

3. Let $v$ be a bispecial factor of $u_\beta$ containing at least one letter 1, then there exists a unique factor $w$ such that $v = T(w)$.

4. Let $w, v$ be factors of $u_\beta$, then $w$ is a prefix of $v$ if and only if $T(w)$ is a prefix of $T(v)$.

5. Let $w, v$ be factors of $u_\beta$, then $w$ is a suffix of $v$ if and only if $T(w)$ is a suffix of $T(v)$.

Proof. 1. Take an arbitrary factor $w \in \mathbb{L}(u_\beta)$. Then $w$ is extendable to the right, and, since $u_\beta$ is recurrent, $w$ is also extendable to the left. In other words, there exists $a, b \in \{0, 1\}$ such that $awb$ is also a factor of $u_\beta$. As $u_\beta$ is a fixed point of $\varphi$, the image $\varphi(awb)$ belongs to $\mathbb{L}(u_\beta)$. Finally, $T(w)$ is a factor of $u_\beta$ because $T(w)$ is a subword of $\varphi(awb)$.

2. Let $1w1$ be a factor of $\mathbb{L}(u_\beta)$, then $01w1$ is as well a factor of $u_\beta$. Applying $\varphi$, we learn that $\varphi(01w1) = \varphi(0)T(w)1$ is a factor of $u_\beta$, which proves that $1T(w)1$ belongs to $\mathbb{L}(u_\beta)$. The other cases $0w0, 0w1, 1w0$ are analogous.

Let $0T(w)1 \in \mathbb{L}(u_\beta)$, using Observation 3.3 the word $v = 10^61\varphi(w)071$ is also a factor of $u_\beta$. Applying Observation 3.4, $v = 1\varphi(0w1)$ and $0w1$ is an element of $\mathbb{L}(u_\beta)$. All the other cases $0T(w)0, 1T(w)0, 1T(w)1$ are similar.

3. Observation 3.3 implies that each bispecial factor $v$ containing at least one letter 1 has the prefix $0^61$ and the suffix $10^4$. According to Observation 3.4 there exists a unique $w$ such that $v = T(w)$.

4. The implication $\Rightarrow$ is obvious noticing that $0^6$ is a prefix of $\varphi(a)$ for $a \in \{0, 1\}$. The opposite implication $\Leftarrow$ follows taking into account that $\varphi(1)$ is not a prefix of $\varphi(0)$ and $\varphi(0)$ is not a prefix of $\varphi(1a)$ for any $a \in \{0, 1\}$.

5. The implication $\Rightarrow$ is obvious noticing that $0^61$ is a suffix of $\varphi(a)$ for $a \in \{0, 1\}$. The opposite implication $\Leftarrow$ follows taking into account that $1\varphi(1)$ is not a suffix of $\varphi(0)$ and $\varphi(0)$ is not a suffix of $\varphi(x1)$ for any $x \in \{0, 1\}^*$.

\section{Integer powers in $u_\beta$}

Even if we want to describe the maximal integer powers of factors of $u_\beta$, it turns out to be useful to study first the relation between bispecial factors and the maximal rational powers of factors.

Lemma 4.1. Let $u$ be an infinite uniformly recurrent word over an alphabet $A$. Let $w^kw'$ be its factor for some proper prefix $w'$ of $w$ and some positive integer $k$. Let us denote by $P1, P2, P3$ the following statements:

$P1$ The factor $w$ has the maximal index in $u$ among all factors of $u$ with the same length $|w|$ and $w^kw'$ is the maximal power of $w$ in $u$.

$P2$ There exist $a, b \in \{0, 1\}$ such that

\[ aw^kw'b \in \mathbb{L}(u) \quad \text{and} \quad w'b \text{ is not a prefix of } w \quad \text{and} \quad a \text{ is not a suffix of } w. \]

$P3$ All the following factors are bispecial:

\[ w', ww', www', \ldots, w^{k-1}w'. \]

Then $P1$ implies $P2$ and $P2$ implies $P3$. 

Proof. P1 $\Rightarrow$ P2: As $u$ is recurrent, there exists $a$ such that $aw^k w'$ is a factor of $u$. Since $w^k w'$ is the maximal power of $w$, the letter $a$ is not a suffix of $w$, otherwise the factor $aw a^{-1}$ (usually called a conjugate of $w$) would have a larger index than $w$. On the other hand, if $w^k w' b$ is a factor of $w$, then $w' b$ is not a prefix of $w$, otherwise it contradicts the fact that $w^k w'$ is the maximal power of $w$ in $u$.

P2 $\Rightarrow$ P3: Since $w'$ is a proper prefix of $w$, there exists $x \in A$ such that $w' x$ is a prefix of $w$. Denote by $y$ the last letter of $w$. Obviously, $x \neq b$ and $y \neq a$. As $aw^j w' x$ is a prefix of $aw^k w' b$ and $yw^j w' b$ is a suffix of $aw^k w' b$, both $aw^j w' x$ and $yw^j w' b$ are in $L(u)$ for all $j$, $0 \leq j \leq k - 1$. It follows that all factors listed in P3 are bispecial.

In this section, our aim is to describe the maximal integer powers occurring in $u_\beta$. Since the letter 0 has the maximal index $p$, we may restrict our consideration to $k$-th powers of factors of $u_\beta$ with $k \geq p$. Crucial for the determination of the index of $u_\beta$ are Propositions 4.2 and 4.5.

Proposition 4.2. Let $p, q$ be integers, $p > q \geq 1$, and $u_\beta$ be the fixed point of the morphism $\beta$. Assume $p > 3$. Let $w$ be a factor of $u_\beta$ containing at least two 1s and $w'$ be a proper prefix of $w$. Denote $v = w^k w'$ for some $k \in \mathbb{N}$, $k \geq p$. If there exist $a, b \in \{0, 1\}$ so that

$$awb \in L(u_\beta) \text{ and } w' b \text{ is not a prefix of } w \text{ and } a \text{ is not a suffix of } w,$$

then there exist a unique $\tilde{w}$ of length $\geq 2$ and a proper prefix $\tilde{w}'$ of $\tilde{w}$ such that

$$w = 0^q 1 \varphi(\tilde{w})(0^q 1)^{-1} \text{ and } v = T(\tilde{v}) = T(\tilde{w}^k \tilde{w}').$$

moreover,

$$\tilde{a} \tilde{b} \in L(u_\beta) \text{ and } \tilde{w}' b \text{ is not a prefix of } \tilde{w} \text{ and } a \text{ is not a suffix of } \tilde{w}.$$

In order to prove Proposition 4.2, we will use the following lemma.

Lemma 4.3. Let $p, q$ be integers, $p > q \geq 1$, and $u_\beta$ be the fixed point of the morphism $\beta$. The following statements hold:

1. If $0(x1)^{\ell} x0 \in L(u_\beta)$ for some integer $\ell \geq 2$, then $\ell = 2$ and $x = 0^q$.
2. If $1(x0)^{\ell} x1 \in L(u_\beta)$ for some integer $\ell \geq p - 1$ and $p \leq 2q$, then $x$ is the empty word $\varepsilon$.

Proof. 1. At first, we exclude the case when $x$ contains a non-zero letter. Suppose that the letter 1 occurs in $x$. Since the factors $0x1$ and $1x0$ belong to $L(u_\beta)$, it follows that $x$ is bispecial. By Lemma 3.2, Item 6, $x$ starts in $0^q 1$ and ends in $0^q$. Therefore $10^q 10^q 1 \in L(u_\beta)$. As $10^q 10^q 1 = 1 \varphi(11)$, we have according to Observation 3.3 that $11 \in L(u_\beta)$ -- a contradiction.

Now consider $x = 0^q$ for some $s \in \mathbb{N}$. Then $0s1, 1x1 \in L(u_\beta)$, which implies by Observation 3.3 that $s = q$. If $\ell$ was at least 3, then $1x1x1 = 10^q 10^q 1 \in L(u_\beta)$, which leads to the same contradiction as before.

2. Again we start with the case of $x$ containing the letter 1. Since factors $1x0$ and $0x1$ belong to the language $L(u_\beta)$, it follows that $x$ is bispecial. Hence $x$ starts in $0^q 1$ and ends in $0^q$. Since $10^q 0^q 1 \in L(u_\beta)$, by Observation 3.3 we have $p = 2q + 1$. This contradicts the assumption $p \leq 2q$.

Suppose now that $x = 0^q$ for some $s \in \mathbb{N}$. Since $1(x0)^{\ell} x1 = 1(0^s 1)^{\ell} 0^q 1$ is a factor of $u_\beta$, Observation 3.3 gives that $(s + 1)\ell + s \leq p$, which is impossible if $s \geq 1$ and $\ell \geq p - 1$. Therefore $s = 0$ and $x$ is the empty word $\varepsilon$.

Proof of Proposition 4.2. The factor $w$ contains at least two 1s. Since $w$ and $v = w^k w'$ satisfy Item P2 of Lemma 3.1 both $ww'$ and $www'$ are bispecial, and therefore start in $0^q 1$ and end in $0^q$. Consequently, their form is $ww' = T(x)$ and $www' = T(y)$, where $x, y \neq \varepsilon$. According to Lemma 3.5, Item 5, $x$ is a suffix of $y$, i.e., $y = Sx$ for some $z \neq \varepsilon$. Observing $ww' = 0^q 1 \varphi(x)0^q$ and $www' = 0^q 1 \varphi(z) \varphi(x)0^q$, it follows directly that $w = 0^q 1 \varphi(z)(0^q 1)^{-1}$.

Let us, at first, show that
1. either $z$ is a prefix of $x$, 
2. or $z = x1$, 
3. or $z = x0$.

Assume $z = tdx'$ and $x = t(1-d)x'$ for a word $t \in \{0,1\}^*$ and for a letter $d$. Then $w = 0^q\varphi(t)\varphi(d)\varphi(z')(0^q)^{-1}$ is a prefix of $ww' = 0^q\varphi(t)\varphi(1-d)\varphi(x')0^q$. If $z' \neq z$, we have a contradiction immediately. If $z' = z$, then $t \neq z$ knowing that $z$ contains at least two letters ($w$ contains at least two $1$s).

- If $d = 1$, then $w = 0^q1\varphi(t)$ and $ww' = 0^q1\varphi(t)0^q1\varphi(x')0^q$, thus $w'$ starts in $0^q1$, which is not a prefix of $w$ – a contradiction.
- If $d = 0$, then $w = 0^q1\varphi(t)0^q-\varphi$ and $ww' = 0^q1\varphi(t)0^q-\varphi 0^q-1\varphi(x')0^q$, thus $w'$ starts in $0^q-1$, which is not a prefix of $w$ because $p \neq q$ – a contradiction again.

The situation $z = xz'$ for some $z'$ of length $\geq 2$ cannot occur because it implies $|w| > |ww'|$. Consequently, one of the situations 1., 2., 3. occurs.

1. If $z$ is a prefix of $x$, i.e., $x = zx''$, then $ww' = 0^q1\varphi(z)\varphi(x'')0^q$ and $w = 0^q1\varphi(z)(0^q)^{-1}$, thus $w' = 0^q1\varphi(z''')0^q$. Then $v = w^k w' = 0^q1\varphi(z^k x''')0^q$, therefore $\tilde{w} = z$, $\tilde{w}' = x'''$. As $w'$ is a proper prefix of $w$, it follows by Lemma 3.5 Item 4 that $\tilde{w}'$ is a proper prefix of $\tilde{w}$.

2. Assume $z = x1$, then $w = 0^q1\varphi(x)$ and $ww' = 0^q1\varphi(x)0^q$, thus $w' = 0^q$. Then $v = 0^q1\varphi((x1)^{k-1}x)0^q$. Since $w'^k$ is not a prefix of $w$ and $0$ is not a suffix of $w$, the assumptions imply that $0v0 \in L(u_\delta)$. By Observation 3.5, we have $1\varphi(0(x1)^{k-1}x0) \in L(u_\delta)$. Since $k-1 \geq 3$, we deduce by Lemma 4.3 that $0(x1)^{k-1}x0$ is not a factor of $u_\delta$. It contradicts Observation 3.4. Hence, the case $z = x1$ does not occur.

3. If $z = x0$, then $w = 0^q1\varphi(x)0^q-\varphi$ and $ww' = 0^q1\varphi(x)0^q$, thus $w' = 0^q-1$. Then $v = 0^q1\varphi((x0)^{k-1}x)0^q$. Since $w'$ is not a prefix of $w$ and $1$ is not a suffix of $w$, the assumptions imply that $1v1 \in L(u_\delta)$. Hence, $1\varphi(1(x0)^{k-1}x1) \in L(u_\delta)$. However, by Lemma 4.3, it follows that $x = \varepsilon$. Then $z = 0$, which contradicts the condition $|z| \geq 2$. Thus, the case $z = x0$ does not occur.

Let us finally note that the very last statement on the extensions of $\tilde{v}$ follows from Lemma 3.5 Items 3, 4, and 5.

**Remark 4.4.** Proposition 4.2 does not take into account $u_\delta$ given by parameters $p = 2, q = 1$, $p = 3, q = 2$, and $p = 3, q = 1$. In the two first cases, $u_\delta$ is a Sturmian word. Therefore, exclusion of the first two cases does not mean any loss. For the case of $p = 3, q = 1$, in the proof of Proposition 4.2, we cannot exclude the situation 2. In this case, Lemma 4.3 Item 1 implies either the validity of [8] or of $\text{w = }0^1\varphi(01)(0^1)^{-1}$ or $v = T(01010)$.

Proposition 4.2 thus claims that for every factor $w \in L(u_\delta)$ containing at least two 1s such that its $k$-th power $w^k$ is a factor of $u_\delta$ with $k \geq p$, there exists a shorter factor $\tilde{w}$ such that its $k$-th power $\tilde{w}^k$ is also in the language of $u_\delta$. As a consequence, in order to determine the maximal integer power present in $u_\delta$, it is sufficient to study the index of factors $w$ containing only one letter 1.

**Proposition 4.5.** Let $p, q$ be integers, $p > q \geq 1$, and $u_\delta$ be the fixed point of the morphism $\beta$. Let $w$ be a factor of $u_\delta$ containing one letter 1 and of the maximal index $\text{ind}(w)$ among all factors of length $|w|$ and such that $\text{ind}(w) \geq p \geq 3$. Denote $k := \lfloor \text{ind}(w) \rfloor$ and $v = w^k w'$ the maximal power of $w$. Then $w = 0^q1 \varphi(0)(0^q)^{-1}$ and $v = T(0^p)$ and $\text{ind}(w) = p + \frac{2q + 1}{p + 1}$.
Moreover, the maximal power of $u$ is $w^p$. Assume $w$ starts in $0^q1$. Therefore $w = 0^q10^s$ with $s \in \{0, p - q\}$ (Observation 3.3). The case $s = 0$ does not occur since $ww = 0^q10^s0^q10^s \in \mathcal{L}(u_\beta)$ and $0^q10^s0^q1 = 0^q1\varphi(11)$, but $11 \notin \mathcal{L}(u_\beta)$ — a contradiction to Observation 3.4. Hence $w = 0^q10^s = 0^q1\varphi(0)(0^q1)^{-1}$. It remains to determine the form of $v$. Again, since $ww$ is bispecial, $ww$ ends in $10^q$. As $w$ is a prefix of $w$, at most one 1 occurs in $w$.

- Suppose $w$ contains 1, then $w$ starts in $0^q1$ and ends in $10^q$, thus $w = 0^q10^q$. This is possible only in case when $p - q - 1 \geq q$, i.e., $p \geq 2q + 1$. Then, $v = (0^q10^{p-q})^k0^q10^q$. Consequently, $v = 0^q1\varphi(0^k)10^q = T(0^k)$. On one hand, Observations 3.3 and 3.4 imply that $k \leq p$. On the other hand, since $v$ is the maximal power of $w$, it follows that $k \geq p$.

- Assume $w$ does not contain 1. Since $ww$ ends in $10^q$ and $w = 0^q10^{p-q}$, the only possibility for $w$ is $w = 0^q10^q - p$. This comes in question only for $p \leq 2q$. Then $v = (0^q10^{p-q})^k0^q10^q = 0^q1\varphi(0^k-1)0^q = T(0^k-1)$. The same arguments as in the previous case imply that $k - 1 = p$.

Clearly, $\text{ind}(w) = \frac{|v|}{|w|} = p + \frac{2q+1}{p+1}$.

Let us state the main result of this section.

**Theorem 4.6.** Let $p, q$ be integers, $p > q \geq 1$, and $u_\beta$ be the fixed point of the morphism $\beta$. Assume $p \geq 3$.

- If $p \leq 2q$, then there exists a factor $w \neq \varepsilon$ satisfying $w^{p+1} \in \mathcal{L}(u_\beta)$ and no $(p + 2)$-nd power of any factor belongs to the language $\mathcal{L}(u_\beta)$.

- If $p > 2q$, then there exists a factor $w \neq \varepsilon$ satisfying $w^p \in \mathcal{L}(u_\beta)$ and no $(p + 1)$-st power of any factor belongs to the language $\mathcal{L}(u_\beta)$.

**Proof.** Proposition 4.2 implies that in order to determine the maximal integer power present in $u_\beta$, we can restrict our consideration to powers of factors containing only one letter 1. When we compute the integer part $\lfloor \text{ind}(w) \rfloor$ of such factors $w$ in Proposition 4.3, we find that the maximum is $p + 1$ if $p \leq 2q$ and $p$ otherwise.

---

5 **Index of $u_\beta$**

The task of this section is to compute the index of $u_\beta$, i.e.,

$$\text{ind}(u_\beta) = \sup \{ \text{ind}(w) \mid w \in \mathcal{L}(u_\beta) \}.$$ 

We already know that $\text{ind}(u_\beta) \geq p$. Using Lemma 4.1, it suffices to study rational powers $v = u^{(n)}w$ of factors $w$ with the property $\textbf{P2}$. As a direct consequence of Propositions 4.2 and 4.3, we have the following corollary.

**Corollary 5.1.** Let $p, q$ be integers, $p > q \geq 1$, and $u_\beta$ be the fixed point of the morphism $\beta$. Assume $p > 3$. The index of $u_\beta$ is given by the following formula

$$\text{ind}(u_\beta) = \sup \{ \text{ind}(w^{(n)}) \mid n \in \mathbb{N} \},$$

where

$$w^{(0)} = 0, \quad w^{(n+1)} = 0^q1\varphi(w^{(n)})(0^q1)^{-1}. \quad (7)$$

Moreover, the maximal power of $w^{(n)}$ is $v^{(n)}$, where

$$\begin{align*}
v^{(0)} &= 0^p, \quad v^{(n+1)} = T(v^{(n)}),
\end{align*} \quad (8)$$

In the sequel, let us determine the index of $w^{(n)}$ for every $n \in \mathbb{N}$. 


Lemma 5.2. The number of 0s and 1s in the words \( w(n) \) and \( v(n) \) satisfy

\[
(\left| w^{(n)} \right|_0, \left| w^{(n)} \right|_1) = (1,0)M^n_\varphi, \\
(\left| v^{(n)} \right|_0, \left| v^{(n)} \right|_1) = (p+1, 2q+1-p)M^n_\varphi - (1,2q+1-p),
\]

where \( M_\varphi = \left( \begin{array}{c} p+1 \\ p \end{array} \right) \) is the morphism matrix.

Proof. As \( w(n) \) is a conjugate of \( \varphi(w^{(n-1)}) \), the first formula holds by \([2]\). Let us show the second one by induction on \( n \).

For \( n = 0 \),

\[
(\left| v^{(0)} \right|_0, \left| v^{(0)} \right|_1) = (p,0) = (p+1, 2q+1-p) - (1,2q+1-p).
\]

For \( n > 0 \),

\[
(\left| v^{(n)} \right|_0, \left| v^{(n)} \right|_1) = (\left| v^{(n-1)} \right|_0, \left| v^{(n-1)} \right|_1)M_\varphi + (2q,1) = \\
\left| v^{(n-1)} \right|_0M_\varphi + (2q,1) = (p+1, 2q+1-p)M^{n-1}_\varphi - (1,2q+1-p).
\]

Since the eigenvalues \( \beta \) and \( \beta' \) of \( M_\varphi \) are roots of the Parry polynomial \( x^2 - (p+1)x + (p-q) \), it is straightforward to show that \( \vec{x}_1 = (\beta - 1,1) \) is a left eigenvector of \( M_\varphi \) corresponding to \( \beta \) and \( \vec{x}_2 = (\beta' - 1,1) \) is a left eigenvector of \( M_\varphi \) corresponding to \( \beta' \). The index of \( w(n) \) may be expressed as follows

\[
\text{ind}(w^{(n)}) = \frac{|v^{(n)}|}{|w^{(n)}|} = \frac{(p+1,0)M^n_\varphi \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right) + (0,2q+1-p)M^n_\varphi \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right) - (1,2q+1-p) \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right)}{(1,0)M^n_\varphi \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right)},
\]

where \( \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 = (0,1) \) and \( \gamma_1 \vec{x}_1 + \gamma_2 \vec{x}_2 = (1,0) \). Using the fact that \( \vec{x}_1 \) and \( \vec{x}_2 \) are eigenvectors of \( M_\varphi \), we have

\[
\text{ind}(w^{(n)}) = p + 1 + \frac{2q+1-p}{q} \frac{(\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2) - (\beta')}{(\gamma_1 \vec{x}_1 + \gamma_2 \vec{x}_2) - (\beta')},
\]

where \( \alpha_1 = \frac{1}{\beta' - \beta}, \ alpha_2 = \frac{\beta - 1}{\beta' - \beta}, \ gamma_1 = \frac{1}{\beta - \beta'}, \ gamma_2 = \frac{1}{\beta - \beta'}, \) and \( \vec{x}_1 \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right) = \beta, \vec{x}_2 \left( \begin{array}{c} \frac{1}{\beta} \\ \frac{1}{\beta'} \end{array} \right) = \beta' \). The final formula for the index of \( w(n) \) has the form

\[
\text{ind}(w^{(n)}) = p + 1 + \frac{2q+1-p}{q} \frac{(1 - \beta')(\beta^n - 1) - (1 - \beta)\beta^{n+1} - 2q+1-p(\beta - \beta')}{\beta^{n+1} - \beta^n} = \\
p + 1 + \frac{2q+1-p}{q} (1 - \beta') + \frac{\beta - \beta'}{q(\beta^{n+1} - \beta^n) (2q+1-p) \beta^{n+1} - (3q+1-p)}.
\]

Using the fact \( 0 < \beta' < 1 < \beta \), we determine the limit

\[
\lim_{n \to \infty} \text{ind}(w^{(n)}) = p + 1 + \frac{2q+1-p}{q} (1 - \beta') = p + 1 + \frac{2q+1-p}{\beta - 1}.
\]

This limit is the supreme of \( \{\text{ind}(w^{(n)}) \mid n \geq 0\} \) if and only if \( A(n) < 0 \) for all \( n \in \mathbb{N} \). It is an easy exercise to show that \( A(n) = (2q+1-p)(\beta^n - 1) - q < 0 \) for all \( n \in \mathbb{N} \) if and only if \( p \leq 3q + 1 \), otherwise \( A(n) > 0 \) for all sufficiently large \( n \).

Let us sum up the results in a theorem.
Theorem 5.3. Let \( p, q \) be integers, \( p > q \geq 1 \), and \( u_\beta \) be the fixed point of the morphism \( \beta \). Assume \( p > 3 \). Then the index of \( u_\beta \) satisfies

\[
\text{ind}(u_\beta) = p + 1 \frac{2q + 1 - p}{\beta - 1},
\]

otherwise there exists \( n_0 \in \mathbb{N} \) such that

\[
\text{ind}(u_\beta) = \text{ind}(w^{(n_0)}) > p + 1 \frac{2q + 1 - p}{\beta - 1}.
\]

Remark 5.4. Similarly as in the previous section, we have to treat the case of \( p = 3 \) and \( q = 1 \) separately. According to Remark 4.4, we have to determine the index of \((\hat{w}^{(n)})\) defined in (7), but moreover the index of \((\hat{w}^{(n)})\) defined recursively by

\[
\hat{w}^{(0)} = 01\varphi(01)(01)^{-1}, \quad \hat{w}^{(n+1)} = 0q\varphi(\hat{w}^{(n)})(0q1)^{-1}.
\]

Using the same technique as before, we obtain

\[
\sup\{|\text{ind}(\hat{w}^{(n)})| \mid n \in \mathbb{N}\} = \beta < 4 = \sup\{|\text{ind}(w^{(n)})| \mid n \in \mathbb{N}\}.
\]

Hence, Theorem 5.3 holds in fact also in this case.

At the conclusion, let us compare in case of Sturmian words \( u_\beta \) the formula for \( \text{ind}(w^{(n)}) \) with the formula (1) for the index of general Sturmian words. As we have already stated, \( u_\beta \) is Sturmian if and only if \( p = q + 1 \), i.e., \( \beta \) is the larger root of the polynomial \( x^2 - (p + 1)x + 1 \). For such \( \beta \), we have

\[
\text{ind}(u_\beta) = p + 1 \frac{2p - 1}{\beta - 1} = \beta + 1.
\]

In order to apply the formula from (1), we need to determine the slope \( \alpha \) of the Sturmian word \( u_\beta \). Since \( \alpha > \frac{1}{2} \) is the density of the more frequent letter, according to Section 3, \( \alpha = 1 - \frac{1}{\beta} \). Let us use some basic properties of continued fractions available at any book on Number Theory to determine the continued fraction of this value. Since

\[
1 - \frac{1}{\beta} = \frac{1}{1 + \frac{1}{\beta - 1}} = \frac{1}{1 + \frac{1}{p+1+\frac{1}{\beta}}},
\]

one obtains \( \beta = [0; 1, (p-1), 1, (p-1), \ldots] = [0; 1, (p-1)] \). Denominators \( q_n \) of the convergents of \( \beta \) fulfill therefore the following recurrent relations

\[
q_{2n+1} = (p-1)q_{2n} + q_{2n-1} \quad \text{and} \quad q_{2n} = q_{2n-1} + q_{2n-2}
\]

with initial values \( q_1 = 1, q_2 = p, q_3 = p + 1 \). By mathematical induction on \( n \), it may be shown easily that

\[
q_{2n-1} = \frac{1}{\beta - \beta'}(\beta^n - \beta'^n) \quad \text{and} \quad q_{2n} = \frac{1}{\beta - \beta'}((1 - \beta')\beta^{n+1} - (1 - \beta)\beta^{n+1}).
\]

As it holds for coefficients of the continued fraction of \( \beta \) that \( a_{2n-1} = 1 \) and \( a_{2n} = p - 1 \), it suffices to consider even \( n \) in (1). We obtain then finally

\[
a_{2n+2} = 2 + \frac{q_{2n} - 2}{q_{2n+1}} = p + 1 + \frac{(1 - \beta')\beta^{n+1} - (1 - \beta)\beta^{n+1} - 2(\beta - \beta')}{\beta^{n+1} - \beta^{n+1}},
\]

which is exactly \( \text{ind}(w^{(n)}) \). This result holds for all parameters \( p, q \) satisfying \( p = q + 1 \), even for \( p \leq 3 \). Consequently, Theorem 5.3 is in fact valid for all parameters \( p, q \) with \( p > q \geq 1 \).
Acknowledgements

The authors acknowledge financial support by Czech Science Foundation GA ČR 201/05/0169, by the grants MSM6840770039 and LC06002 of the Ministry of Education, Youth, and Sports of the Czech Republic.

References

[1] Balková, L’, Gazeau, J.-P., Pelantová, E.: Asymptotic behavior of beta-integers, Letters in Mathematical Physics 84 179-198 (2008)
[2] Frougny, Ch., Gazeau, J.-P., Krejcar, R.: Additive and multiplicative properties of point-sets based on beta-integers, Theor. Comp. Sci. 303 491-516 (2003)
[3] Berstel, J.: On the index of Sturmian words, in: Jewels are forever, Springer 287-294 (1999)
[4] Justin, J., Pirillo, G.: Fractional powers in Sturmian words, Theoret. Comput. Sci. 255 363-376 (2001)
[5] Mignosi, F., Pirillo, G.: Repetitions in the Fibonacci infinite word, RAIRO Inform. Theor. Appl. 26 199-204 (1992)
[6] Masáková, Z., Pelantová, E.: Relation between powers of factors and recurrence function characterizing Sturmian words, submitted, arXiv 0809.0603v2[math.CO]
[7] Carpi, A., de Luca, A.: Special factors, periodicity, and an application to Sturmian words, Acta Inform. 36 983-1006 (2000)
[8] Damanik, D., Lenz, D.: The index of Sturmian sequences, European J. Combin. 23 23-29 (2002)
[9] Damanik, D.: Singular continuous spectrum for a class of substitution Hamiltonians II, Letters in Mathematical Physics 54 25-31 (2000)
[10] Balková, L’, Masáková, Z.: Palindromic complexity of infinite words associated with non-simple Parry numbers, RAIRO - Theoret. Informatics Appl. 43 145-163 (2009)
[11] Balková, L’, Pelantová, E., Turek, O.: Combinatorial and arithmetical properties of infinite words associated with quadratic non-simple Parry numbers, RAIRO - Theoret. Informatics Appl. 41 307-328 (2007)
[12] Berstel, J.: Sturmian and episturmian words (a survey of some recent results), in: S. Bozapalidis, G. Rahonis (eds), Conference on Algebraic Informatics, Thessaloni, Lecture Notes Comput. Sci. 4728 23-47 (2007)
[13] Lothaire, M.: Algebraic combinatorics on words, volume 90 of Encyclopedia of Mathematics and its Applications, Cambridge University Press (2002)
[14] Fabre, S.: Substitutions et β-systèmes de numération, Theoret. Comput. Sci. 137 219-236 (1995)