BALANCES AND ABELIAN COMPLEXITY OF A CERTAIN CLASS OF INFINITE TERNARY WORDS

Ondřej Turek[1]

Abstract. A word u defined over an alphabet $\mathcal{A}$ is c-balanced ($c \in \mathbb{N}$) if for all pairs of factors v, w of u of the same length and for all letters $a \in \mathcal{A}$, the difference between the number of letters a in v and w is less or equal to c. In this paper we consider a ternary alphabet $\mathcal{A} = \{L, S, M\}$ and a class of substitutions φ_p defined by $\varphi_p(L) = L^p S$, $\varphi_p(S) = M$, $\varphi_p(M) = L^{p-1}S$ where $p > 1$. We prove that the fixed point of φ_p, formally written as $\varphi_p^\infty(L)$, is 3-balanced and that its Abelian complexity is bounded above by the value 7, regardless of the value of p. We also show that both these bounds are optimal, *i.e.* they cannot be improved.

Mathematics Subject Classification. 68R15.

1. Introduction

The balance property is a notion connected with Sturmian words from the very first beginning of their investigation. In [15], Sturmian words were defined as aperiodic words with the smallest possible factor complexity. Already in the same article, Hedlund and Morse observed that Sturmian words show also the smallest discrepancy in occurrences of letters. To quote precisely their result, let us denote by $|w|$ the length of the word w and by $|w|_a$ the number of occurrences of letter a in w. Hedlund and Morse proved that an infinite aperiodic word u over the alphabet $\{0, 1\}$ is Sturmian if and only if for all pairs w, v of factors of u with $|w| = |v|$ it holds $||w|_0 - |v|_0| \leq 1$. Let us note that the letter 0 is not preferred as in binary alphabet relations $|w| = |v|$ and $||w|_0 - |v|_0| \leq 1$ imply the inequality $||w|_1 - |v|_1| \leq 1$ as well.

During past 70 years many other characterizations of Sturmian words have appeared; for an overview see [13]. Each of these characterizations may serve and serves for generalization of Sturmian words to multiliteral alphabets, *cf.* [3]. Nevertheless, the balance property seems to be the most complicated to deal with. Only a few results are known about words satisfying the so-called c-balanced property.

Let us recall that an infinite word u over an alphabet $\mathcal{A}$ is c-balanced if for all letters $a \in \mathcal{A}$ and all pairs v, w of factors of u with $|v| = |w|$ it holds $||v|_a - |w|_a| \leq c$. Note that Sturmian words are 1-balanced in this terminology. The set of c-balanced words differs substantially from all other generalizations of Sturmian words. Neither generic Arnoux-Rauzy word nor generic word coding interval exchange transformation are c-balanced, see [7] and [1].

In [2], Adamczewski studies whether the fixed point of a primitive substitution is c-balanced for some constant c. He shows that the existence of such c depends only on the spectrum of the incidence matrix of the substitution. However, the minimal value of c cannot be deduced from the spectrum. In [19] and [4], the minimal value of c is determined for binary fixed points of canonical substitutions associated with quadratic Pisot numbers. The notion "canonical substitution associated with a number $\beta > 1$" comes from positional numeration systems with the base β, see [10]. Generally speaking, it is a very complicated problem to determine minimal value c for a ternary balanced word, let alone for words over alphabets of higher cardinalities. Despite a common belief that the Tribonacci word is 2-balanced, the first proof of this fact has appeared just one year ago in [16] (the Tribonacci word is the fixed point of the substitution $A \mapsto AB$, $B \mapsto AC$, $C \mapsto A$). In this article we provide minimal value of c for a certain class of ternary words, namely for fixed points of substitutions

$$L \mapsto L^p S, \quad S \mapsto M, \quad M \mapsto L^{p-1} S \qquad (1.1)$$

with the parameter $p > 1$. These substitutions are canonical substitutions associated with cubic Pisot numbers $\beta > 1$, roots of polynomials $x^3 - px^2 - x + 1$ (*cf.* [12]). Let us recall that the Tribonacci substitution is associated with a numeration system as well.

The definition of a 1-balanced word may be reformulated equivalently using Parikh vectors. Inspired by this fact, Richomme, Saari and Zamboni introduced the Abelian complexity $\mathrm{AC}(n)$ of infinite word. In their notation, Sturmian words are aperiodic words with $\mathrm{AC}(n) = 2$ for all $n \in \mathbb{N}$. The question on existence of words with constant Abelian complexity is natural. It was shown in [9] that for $k \geq 4$ no words with $\mathrm{AC}(n) = k$ exist. On the other hand, words with $\mathrm{AC}(n) = 3$ can be found in [17].

The relation between Abelian complexity and balance property is not straightforward. It is easy to see that an infinite word u is balanced if and only if its Abelian complexity is bounded. Moreover, if the Abelian complexity of u is bounded by k, then u is $k-1$ balanced. The Tribonacci case shows that the opposite implication is not valid: according to [16], the Abelian complexity of the Tribonacci word takes all values in the set $\{3, 4, 5, 6, 7\}$. The fixed point of any substitution (1.1) studied in this article has the same property.

2. Preliminaries

Let $\mathcal{A}$ be a finite alphabet. A concatenation of letters in $\mathcal{A}$ is called a *word*. The set $\mathcal{A}^*$ of all finite words over $\mathcal{A}$ equipped with the empty word ϵ and the operation of concatenation is a free monoid. The *length* of the word $w \in \mathcal{A}^*$, denoted by $|w|$, represents the number of its letters.

One may also consider infinite words $u = u_0 u_1 u_2 \cdots$ ($u_i \in \mathcal{A}$ for all $i \in \mathbb{N}_0$); the set of infinite words over the alphabet $\mathcal{A}$ is denoted by $\mathcal{A}^{\mathbb{N}}$.

A word w is called a *factor* of $v \in \mathcal{A}^*$ or $\mathcal{A}^{\mathbb{N}}$ if there exist words $w^{(1)} \in \mathcal{A}^*$ and $w^{(2)} \in \mathcal{A}^*$ or $w^{(2)} \in \mathcal{A}^{\mathbb{N}}$, respectively, such that $v = w^{(1)} w w^{(2)}$. The word w is called a *prefix* of v, if $w^{(1)} = \epsilon$. It is a *suffix* of v, if $w^{(2)} = \epsilon$.

Let $w \in \mathcal{A}^{\mathbb{N}}$. For $k \in \mathbb{N}$, the symbol w^k denotes the concatenation $\underbrace{ww \cdots w}_{k \text{ times}}$. We set $w^0 = \epsilon$. Let a word $v \in \mathcal{A}^*$ have the prefix w^k, $k \in \mathbb{N}$. Then the symbol $w^{-k}v$ denotes the word satisfying $w^k w^{-k} v = v$. Similarly, if a word $v \in \mathcal{A}^{\mathbb{N}}$ has the suffix w^k for a $k \in \mathbb{N}$, then vw^{-k} denotes the word with the property $vw^{-k}w^k = v$.

A *morphism* on the free monoid $\mathcal{A}^*$ is a map $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ satisfying $\varphi(vw) = \varphi(v)\varphi(w)$ for all $v, w \in \mathcal{A}^*$. Obviously, the morphism φ is determined if we define $\varphi(a)$ for all $a \in \mathcal{A}$.

A morphism φ is called a *substitution*, if $\varphi(a) \neq \epsilon$ for all $a \in \mathcal{A}$ and if there is an $a' \in \mathcal{A}$ such that $|\varphi(a')| > 1$. An infinite word u is said to be a *fixed point* of the substitution φ, or invariant under the substitution φ, if

$$\varphi(u_0)\varphi(u_1)\varphi(u_2)\cdots = u_0 u_1 u_2 \cdots . \tag{2.1}$$

If we naturally extend the action of φ to infinite words, we may rewrite (2.1) simply as $\varphi(u) = u$.

2.1. Balance properties

An infinite word u is c-balanced, if for every $a \in \mathcal{A}$ and for every pair of factors v, w of u such that $|v| = |w|$, it holds $||v|_a - |w|_a| \leq c$. This property determines the discrepancy of occurrences of letters in the word u. However, it turns out that if the cardinality of $\mathcal{A}$ is higher than two, it is useful to have more detailed information, namely what is the discrepancy of occurrences of each particular letter. For this purpose we introduce the following notion:

Definition 2.1. Let u be an infinite word over the alphabet $\mathcal{A}$ and let $a \in \mathcal{A}$. The word u is said to be *c-balanced with respect to the letter a*, if

$$|\, |v|_a - |w|_a \,| \leq c$$

for all pairs of factors v, w of u of the same length.

2.2. Abelian complexity

Let us consider an alphabet $\mathcal{A}$ with k elements, *i.e.* $\mathcal{A} = \{a_1, \ldots, a_k\}$, and an infinite word u over $\mathcal{A}$. For any factor w of u, its *Parikh vector* is the k-tuple $\Psi(w) = (|w|_{a_1}, \ldots, |w|_{a_k})$. Let the symbol $\mathcal{F}_u(n)$ denote the set of all factors of u of the length n. Then the *Abelian complexity* of the word u is a function $\mathrm{AC} : \mathbb{N} \to \mathbb{N}$ defined by

$$\mathrm{AC}(n) = \# \left\{ \Psi(w) \mid w \in \mathcal{F}_u(n) \right\}. \tag{2.2}$$

On the right hand side of (2.2) there is the cardinality of the set of Parikh vectors of all factors of u of the length n. In the sequel we will denote this set by $\mathcal{P}_u(n)$, *i.e.*

$$\mathcal{P}_u(n) = \left\{ \Psi(w) \mid w \in \mathcal{F}_u(n) \right\}.$$

2.3. On words studied in this paper

From now on, we will focus on a special class of substitutions on the ternary alphabet $\{L, S, M\}$. For any integer $p > 1$, we denote by φ_p the substitution given by

$$\begin{array}{rcl}
\varphi_p(L) & = & L^p S \\
\varphi_p(S) & = & M \\
\varphi_p(M) & = & L^{p-1} S.
\end{array} \tag{2.3}$$

The substitution φ_p has a unique fixed point, namely

$$u^{(p)} = \lim_{n \to \infty} \varphi_p^n(L).$$

If the results of [2] are applied on $u^{(p)}$, one finds out that there is a constant c such that the word $u^{(p)}$ is c-balanced, but it is not known what the value of c is and how it depends on p. This is the main aim of this paper – to determine c.

The fact that $u^{(p)}$ is balanced immediately implies that the Abelian complexity function of $u^{(p)}$ is bounded, see Introduction. The second aim of this paper is thus to find the optimal bound for $\mathrm{AC}(n)$.

Remark 2.2. The elements of $\mathcal{A}$ are usually denoted by numbers: $0, 1, 2$ etc. We have considered this notation, but we believe that the paper becomes more transparent if letters are used. The choice of L, S and M has its roots in the fact that the word $u^{(p)}$ is a fixed point of a substitution associated with a number $\beta > 1$, *cf.* Introduction. Let $\mathbb{Z}_\beta$ denote the set of numbers which can be written in the form $x = x_k \beta^k + \cdots + x_1 \beta + x_0$ for non-negative integers x_j. It can be shown (*cf.* [18]) that when the elements of $\mathbb{Z}_\beta$ are drawn on the real line, there are exactly three types of distances between neighbouring points. If we assign the letters L, M and S to the longest, the medium and the shortest distance, respectively, then the order of distances on the real line corresponds exactly to the order of the letters L, S, M in the infinite word $u^{(p)}$.

3. Main result and the proof outline

We begin by the formulation of the main result of the paper.

Theorem 3.1. *Let $u^{(p)}$ be the infinite word invariant under the morphism φ_p given by* (2.3). *Then $u^{(p)}$ is*

- *3-balanced with respect to the letter L,*
- *2-balanced with respect to the letter S,*
- *2-balanced with respect to the letter M,*

and none of these bounds can be improved.

This theorem has the following trivial consequence:

Corollary 3.2. *The infinite word $u^{(p)}$ is 3-balanced and this bound is optimal, i.e. it cannot be improved.*

Since the proof is long and slightly complicated, we will split it into four sections and proceed in the following way:

1. We prove that $u^{(p)}$ is 2-balanced with respect to the letter M.
2. We prove that $u^{(p)}$ is 2-balanced with respect to the letter S.
3. We prove that $u^{(p)}$ is 3-balanced with respect to the letter L.
4. We show that none of the bounds can be improved.

4. Properties of the word $u^{(p)}$

As we have explained in Preliminaries, the word $u^{(p)}$ is a fixed point of φ_p, *i.e.*

$$u^{(p)} = \varphi_p(u) = \varphi_p(u_0)\varphi_p(u_1)\varphi_p(u_2)\cdots$$

In this sense each letter of $u^{(p)}$ can be regarded as the image, or a factor of the image, of another letter of $u^{(p)}$. In view of the definition of φ_p, *cf.* (2.3), each segment $\varphi_p(u_j)$ has the structure $L^k Y$ for $k \in \{0, p-1, p\}$ and $Y \neq L$. The letters S and M are thus "terminating symbols" which cut $u^{(p)}$ to images of individual letters. This fact is particularly important when a factor v of $u^{(p)}$ is given and one needs to find a factor x of $u^{(p)}$ such that $\varphi_p(x) = v$. It holds:

Observation 4.1. *Let vY be a factor of $u^{(p)}$ such that $Y \in \{S, M\}$ and let one of the following conditions be satisfied:*

(i) *The first letter of v is M,*

(ii) *v has the prefix L^p,*

(iii) *SvY or MvY is a factor of $u^{(p)}$.*

Then there is a unique factor x of $u^{(p)}$ satisfying $\varphi_p(x) = vY$.

Proof. Any of the conditions (i), (ii) and (iii) together with $Y \in \{S, M\}$ ensures that vY is an image of certain factor x, and it is obvious from the definition of φ_p that $\varphi_p(x) = \varphi_p(y) \Rightarrow x = y$. $\qquad\square$

Observation 4.2. (i) Let X be a letter occurring in $\varphi_p(u_j)$ for a $j \in \mathbb{N}_0$. Then:

- If $X = M$, then $u_j = S$ and $\varphi_p(u_j) = X$.
- If $X = S$, then either $u_j = L$ and $\varphi_p(u_j) = L^p X$, or $u_j = M$ and $\varphi_p(u_j) = L^{p-1} X$.

(ii) If $X L^k Y$ is a factor of $u^{(p)}$ and $X \neq L$, $Y \neq L$, $k \neq 0$, then $Y = S$ and $k \in \{p-1, p\}$.

The following observation describes the possible neighbours of each of the letters L, S, M in the word $u^{(p)}$.

Observation 4.3. The sequence of letters in the word $u^{(p)}$ conform to these rules:

(i) Each letter S in $u^{(p)}$ is preceded by L and followed either by L or by M.

(ii) Each letter M in $u^{(p)}$ is preceded by S and followed by L.

Proof. (i) Each S is the last letter of $\varphi_p(u_j)$ for $u_j = L$ or $u_j = M$ according to Observation 4.2, *i.e.* it is the last letter of the block $L^p S$ or $L^{p-1} S$, thus is preceded by L. The letter S is followed by the first letter of $\varphi_p(u_{j+1})$, which can be either L or M (cf. the substitution rule φ_p).

(ii) Each M is equal to $\varphi_p(u_j)$ for $u_j = S$. We already know from (i) that $u_{j-1} = L$ and $u_{j+1} \in \{L, M\}$, therefore the M is preceded by the last letter of $\varphi_p(L)$ (which is S) and followed by the first letter of $\varphi_p(L)$ or $\varphi_p(M)$ (which is L). $\square$

In order to understand the structure of the word $u^{(p)}$, it is useful to describe possible segments z in factors of $u^{(p)}$ of the type SzS and MzM. This is done in the next two observations.

Observation 4.4. Let $Sz'S$ be a factor of $u^{(p)}$ such that $|z'|_S = 0$. Then one of the following equalities holds:

- $z' = L^p$,
- $z' = ML^p$,
- $z' = ML^{p-1}$.

Proof. Observations 4.1 and 4.2 imply that $z'S = \varphi_p(\check{z}')$, where $\check{z}'$ is a factor of $u^{(p)}$ the last letter of which is either L or M and which is preceded in $u^{(p)}$ by either L or M.

Since $\varphi_p(\check{z}')$ contains only one S, namely its last letter, all letters of $\check{z}'$ except the last one have to be different from L and M. Therefore $\check{z}' = S^k X$, where $X \in \{L, M\}$ and $k \geq 0$.

Taking into account Observation 4.3, we infer that $k = 0$ or $k = 1$. Therefore only four situations are possible: $\check{z}' = L$, $\check{z}' = M$, $\check{z}' = SL$, $\check{z}' = SM$. Moreover, since $\check{z}'$ is preceded by either L or M, it cannot hold $\check{z}' = M$, *cf.* Observation 4.3. Therefore $\check{z}'$ is equal to one of the factors L, SL, SM, which implies that z is equal to one of the factors L^p, ML^p, ML^{p-1}. $\square$

Observation 4.5. Let MzM be a factor of $u^{(p)}$ such that $|z|_M = 0$. Then one of the following equalities holds:

- $z = (L^p S)^p$,

- $z = L^{p-1}S(L^pS)^p$,
- $z = L^{p-1}S(L^pS)^{p-1}$.

Consequently, $p^2 + p - 1 \leq |z| \leq p^2 + 2p$.

Proof. It follows from Observations 4.1 and 4.2 that $MzM = \varphi_p(Sz'S)$, where $Sz'S$ is a factor of $u^{(p)}$ such that $|z'|_S = 0$. Therefore $z = \varphi_p(z')$ and, according to Observation 4.4, $z' \in \{\varphi_p(L^p), \varphi_p(ML^p), \varphi_p(ML^{p-1})\}$. $\square$

Many times we will need to compare the number of letters L, S, M in a factor of $u^{(p)}$ and in its image. The substitution rule (2.3) leads to the equalities

$$|\varphi_p(v)|_L = p|v|_L + (p-1)|v|_M, \quad |\varphi_p(v)|_S = |v|_L + |v|_M, \quad |\varphi_p(v)|_M = |v|_S,$$

which can be inverted subsequently:

Proposition 4.6. *For any factor v of $u^{(p)}$ it holds*

$$
\begin{aligned}
|v|_L &= |\varphi_p(v)|_L - (p-1)|\varphi_p(v)|_S, \\
|v|_S &= |\varphi_p(v)|_M, \\
|v|_M &= -|\varphi_p(v)|_L + p|\varphi_p(v)|_S, \\
|v| &= |\varphi_p(v)|_S + |\varphi_p(v)|_M.
\end{aligned}
$$

5. Balance bound with respect to the letter M

We begin the proof of Theorem 3.1 by its second statement, *i.e.* we show at first that $u^{(p)}$ is 2-balanced with respect to the letter M. As we will see, the determination of the balance bound with respect to the letter M is by far the most complicated part of the work.

Theorem 5.1. *Let v, w be factors of $u^{(p)}$ such that $|v| = |w|$. Then*

$$\bigl|\, |v|_M - |w|_M \,\bigr| \leq 2.$$

Proof of Theorem 5.1.

We will proceed by contradiction. Let us assume that there exist factors v, w of $u^{(p)}$ such that $|v| = |w| = n$ and

$$|v|_M - |w|_M > 2. \tag{5.1}$$

Let n be the *minimal* number with this property.

We denote $v = v_1 \cdots v_n$, $w = w_1 \cdots w_n$. The minimality of n implies

$$v_1 = M, \quad v_n = M, \tag{5.2}$$

$$w_1 \neq M, \quad w_n \neq M, \tag{5.3}$$

$$|v|_M - |w|_M = 3. \tag{5.4}$$

Stage 1: Introduction of f, g

If we apply Observations 4.1 and 4.2, equation (5.2) implies that the factor v is an image of certain factor of $u^{(p)}$ whose first and last letters equal S. This allows us to define a factor f of $u^{(p)}$ in this way:

$$\varphi_p(SfS) = v. \tag{5.5}$$

The factor w is not ready for a direct application of Observation 4.1 because of (5.3). For that reason we at first extend the factor w to both sides up to the closest letter M, *i.e.* we put

$$w' = Mw^{(1)}ww^{(2)}M, \tag{5.6}$$

where w' is a factor of $u^{(p)}$ and $|w^{(1)}|_M = |w^{(2)}|_M = 0$. Now we can define a factor g of $u^{(p)}$ by the relation

$$\varphi_p(SgS) = w'. \tag{5.7}$$

Let us show that the factor g is shorter than v and w:

Lemma 5.2. *It holds $|g| \leq n - 2(p^2 - 1)$.*

Proof. Since $|v|_M = |w|_M + 3$, the factor v contains at least 3 letters M, and thus $v = M \cdots M \cdots M$. Observation 4.5 then implies that $|v| \geq 1 + (p^2 + p - 1) + 1 + (p^2 + p - 1) + 1 = 2p^2 + 2p + 1$.

Since $|w| = |v|$, it holds $|w| \geq 2p^2 + 2p + 1$, hence necessarily $|w|_M \geq 1$ according to Observation 4.5. But then $|v|_M \geq 4$ and

$$v = Mz^{(1)}M \cdots Mz^{(2)}M, \tag{5.8}$$

where $|z^{(1)}|_M = |z^{(2)}|_M = 0$ and $|z^{(j)}| \geq p^2 + p - 1$ for $j = 1, 2$.

One more application of Observation 4.5 gives $|v| \geq 1 + (p^2 + p - 1) + 1 + (p^2 + p - 1) + 1 + (p^2 + p - 1) + 1 = 3p^2 + 3p + 1$, hence $|w| \geq 3p^2 + 3p + 1$ and necessarily $|w|_M \geq 2$, *i.e.*

$$w = \hat{w}^{(1)}M\hat{w}M\hat{w}^{(2)}, \tag{5.9}$$

where $|\hat{w}^{(1)}|_M = |\hat{w}^{(2)}|_M = 0$.

We deduce from equations (5.8), (5.9), from $|z^{(j)}| \geq p^2 + p - 1$ and from the minimality of n that $|\hat{w}^{(j)}| \geq p^2 + p$, $j = 1, 2$.

The factor w' (*cf.* (5.6) and (5.9)) is given by $w' = Mw^{(1)}\hat{w}^{(1)}M\hat{w}M\hat{w}^{(2)}w^{(2)}M$, hence Equality (5.7) together with Observation 4.2 imply

$$SgS = Sg^{(1)}S\hat{g}Sg^{(2)}S,$$

where $\varphi_p(g^{(1)}) = w^{(1)}\hat{w}^{(1)}$, $\varphi_p(g^{(2)}) = \hat{w}^{(2)}w^{(2)}$ and $\varphi_p(S\hat{g}S) = M\hat{w}M$. Hence $|g^{(1)}|_S = |g^{(2)}|_S = 0$, and consequently, using Observation 4.4, $|g^{(1)}| \leq p + 1$ and $|g^{(2)}| \leq p + 1$. The sought inequality $|g| \leq n - 2(p^2 - 1)$ follows from these relations:

$$n = |w| = |\hat{w}^{(1)}\varphi_p(S\hat{g}S)\hat{w}^{(2)}| = |\hat{w}^{(1)}| + |\varphi_p(S\hat{g}S)| + |\hat{w}^{(2)}| \geq 2(p^2 + p) + |S\hat{g}S|,$$

$$|g| = |g^{(1)}S\hat{g}Sg^{(2)}| = |g^{(1)}| + |S\hat{g}S| + |g^{(2)}| \leq 2(p + 1) + |S\hat{g}S|. \qquad \square$$

Lemma 5.3. *Factors f, g satisfy*

$$|f|_S - |g|_S = 1\,, \tag{5.10}$$
$$|f|_M - |g|_M \geq (p+1) \cdot (|f| - |g|) + 2 - p\,. \tag{5.11}$$

Proof. Both relations will be proved using the properties of v, w'. Let us begin with equation (5.10). Considering Proposition 4.6, we have

$$|v|_M = |\varphi_p(SfS)|_M = |SfS|_S = |f|_S + 2\,,$$
$$|w'|_M = |\varphi_p(SgS)|_M = |SgS|_S = |g|_S + 2\,.$$

Moreover, $|w'|_M = |Mw^{(1)}ww^{(2)}M|_M = |w|_M + 2$. Therefore, taking into account equation (5.4), we obtain

$$|f|_S - |g|_S = |v|_M - |w'|_M = |v|_M - |w|_M - 2 = 3 - 2 = 1\,.$$

Now we proceed to equation (5.11). It holds $|\varphi_p(z)| = (p+1)|z|_L + |z|_S + p|z|_M$ for any factor z (*cf.* the substitution rule (2.3)). The identity $|z| = |z|_L + |z|_S + |z|_M$ allows one to eliminate $|z|_L$, hence

$$|\varphi_p(z)| = (p+1)|z| - p|z|_S - |z|_M\,. \tag{5.12}$$

Since $|v| = |w|$ and $|w'| = |Mw^{(1)}ww^{(2)}M| \geq |w| + 2$, we have $|v| - |w'| \leq -2$, equivalently $|\varphi_p(SfS)| - |\varphi_p(SgS)| \leq -2$. If we apply (5.12) to the last inequality, we obtain

$$(p+1)(|f| - |g|) - p(|f|_S - |g|_S) - (|f|_M - |g|_M) \leq -2\,.$$

Now we substitute here from (5.10) which leads to (5.11). $\qquad\square$

There is another useful statement, namely Lemma 5.5, but to prove it we need one more observation which gives an estimate of the number of occurrences of the letter M in a factor of $u^{(p)}$ of a given length:

Observation 5.4. (i) For every factor $\hat{v}$ of $u^{(p)}$ it holds $|\hat{v}|_M \leq 1 + \frac{|\hat{v}|-1}{p^2+p}$.
(ii) If moreover $\hat{v}$ has a prefix or a suffix of the length greater or equal to Δ that does not contain M, it holds $|\hat{v}|_M \leq \left\lceil \frac{|\hat{v}|-\Delta}{p^2+p} \right\rceil$.

Proof. (i) Observation 4.5 says that if MzM is a factor of $u^{(p)}$ with $|z|_M = 0$, then $|zM| \geq p^2 + p$. Consequently $|\hat{v}|_M \leq \left\lceil \frac{|\hat{v}|}{p^2+p} \right\rceil$. The inequality $\left\lceil \frac{|\hat{v}|}{p^2+p} \right\rceil \leq 1 + \frac{|\hat{v}|-1}{p^2+p}$ holds trivially.
(ii) Let $\hat{v} = \hat{v}'\hat{v}''$ and

$$|\hat{v}'| = \Delta \quad \wedge \quad |\hat{v}'|_M = 0 \qquad \text{or} \qquad |\hat{v}''| = \Delta \quad \wedge \quad |\hat{v}''|_M = 0\,.$$

Then, employing the result of (i), one has $|\hat{v}|_M = |\hat{v}'|_M + |\hat{v}''|_M \leq 0 + \left\lceil \frac{|\hat{v}|-\Delta}{p^2+p} \right\rceil$. $\quad\square$

Lemma 5.5. *It holds* $|f| - |g| \leq 0$, *i.e. the factor f is not longer than g.*

Proof. Let us suppose that the contrary is true, *i.e.* $|f| - |g| = d > 0$. Let f' be the suffix of f of the length d and let $\hat{f} = f(f')^{-1}$. Then it holds $|\hat{f}| = |g| < n$ (the inequality "$< n$" is valid due to Lem. 5.2 and since $p \geq 2$) and $|\hat{f}|_M = |f|_M - |f'|_M$.

Our goal is to estimate $|\hat{f}|_M - |g|_M$ which is equal to $|f|_M - |g|_M - |f'|_M$. For this purpose an estimate for $|f|_M - |g|_M$ will be needed; we obtain it from (5.11): $|f|_M - |g|_M \geq (p+1) \cdot d + 2 - p$.

As for $|f'|_M$, since f is followed by S, the last letter of f' is L (*cf.* Observation 4.3). With regard to this fact, Observation 5.4 implies $|f'|_M \leq \left\lceil \frac{d-1}{p^2+p} \right\rceil$. Now we distinguish two cases:

- If $d = 1$, then $|f'|_M = 0$, hence

$$|\hat{f}|_M - |g|_M = |f|_M - |g|_M \geq (p+1) \cdot 1 + 2 - p \geq 3\,,$$

- if $d \geq 2$, then $|f'|_M \leq 1 + \frac{d-1}{p^2+p}$, hence

$$|\hat{f}|_M - |g|_M \geq |f|_M - |g|_M - \left(1 + \frac{d-1}{p^2+p}\right) \geq (p+1)d + 2 - p - \left(1 + \frac{d-1}{p^2+p}\right) =$$
$$\left(p+1 - \frac{1}{p^2+p}\right)(d-1) + 2 \geq p+1 - \frac{1}{p^2+p} + 2 > p+2 \geq 4\,.$$

We see that for any value of $d > 0$, the factors $\hat{f}$ and g of $u^{(p)}$ are of the same length less than n and satisfy (5.1). This is a contradiction with the minimality of n. □

Stage 2: Introduction of x, y

At this moment we define another pair of factors. Since SfS and SgS are factors of $u^{(p)}$, Observation 4.1 says that there exist factors x and y of $u^{(p)}$ such that

$$\varphi_p(x) = fS \tag{5.13}$$

and

$$\varphi_p(y) = gS\,. \tag{5.14}$$

Lemma 5.6. *It holds* $|x| \leq n - (2p^2 - 3)$ *and* $|y| \leq n - (2p^2 - 3)$, *i.e. both factors x, y are shorter than v, w.*

Proof. We use Lemmas 5.2 and 5.5:

$$|y| \leq |\varphi_p(y)| = |gS| = |g| + 1 \leq n - 2(p^2 - 1) + 1 = n - (2p^2 - 3)\,,$$
$$|x| \leq |\varphi_p(x)| = |fS| = |f| + 1 \leq |g| + 1 \leq n - (2p^2 - 3)\,.$$

We remark that one can achieve better estimates, but these will be sufficient. □

Lemma 5.7. *It holds*

$$|x|_M - |y|_M = -(|f| - |g|) + |f|_M - |g|_M + p + 1 \qquad (5.15)$$

$$|x| - |y| = 1 + |f|_M - |g|_M. \qquad (5.16)$$

Proof. The proof of both statements is straightforward using the definitions of x and y, Proposition 4.6, the identity $|v| = |v|_L + |v|_S + |v|_M$ holding for any factor v, and Equality (5.10). $\qquad \square$

Lemma 5.7 has an immediate corollary:

Lemma 5.8.

$$|x|_M - |y|_M = |x| - |y| - (|f| - |g|) + p \qquad (5.17)$$

$$|x|_M - |y|_M \geq |x| - |y| + p. \qquad (5.18)$$

Proof. If we substitute for $|f|_M - |g|_M$ from (5.16) to (5.15), we obtain (5.17). Equality (5.17) and Lemma 5.5 then give (5.18). $\qquad \square$

In what follows we split the proof according to the sign of $|x| - |y|$, and we show that whatever the value of $|x| - |y|$ is, it always contradicts the minimality of n.

Case $|x| - |y| > 0$

We set $|x| - |y| = d \geq 1$ and denote $\hat{y} = yy'$, where y' is a factor of $u^{(p)}$ of the length d such that yy' is a factor of $u^{(p)}$. Then $|x| = |\hat{y}|$. Moreover $|x| < n$ due to Lemma 5.6.

Our goal is to estimate $|x|_M - |\hat{y}|_M$ which is equal to $|x|_M - |y|_M - |y'|_M$. The difference $|x|_M - |y|_M$ can be estimated using Inequality (5.18): $|x|_M - |y|_M \geq d + p$.

Let us proceed to $|y'|_M$. Recall at first that $\varphi_p(y) = gS$ (*cf.* (5.14)). Thus the last letter of the factor y is either L or M, which implies, with respect to Observation 4.3, that the first letter of y' is either L or S, *i.e.* different from M. Observation 5.4 then gives $|y'|_M \leq \left\lceil \frac{d-1}{p^2+p} \right\rceil$. Now we distinguish two cases:

- If $d = 1$, then $|y'|_M = 0$, hence $|x|_M - |\hat{y}|_M = |x|_M - |y|_M \geq 1 + p \geq 3$;
- if $d \geq 2$, then $|y'|_M \leq 1 + \frac{d-1-1}{p^2+p}$, hence

$$|x|_M - |\hat{y}|_M \geq d + p - 1 - \frac{d-2}{p^2+p} \geq p + (d-1)\left(1 - \frac{1}{p^2+p}\right) + \frac{1}{p^2+p}$$

$$\geq p + 1 \cdot \left(1 - \frac{1}{p^2+p}\right) + \frac{1}{p^2+p} = p + 1 \geq 3.$$

The factors x and $\hat{y}$ are of the same length less than n and for any $d > 0$ satisfy $|x|_M - |\hat{y}|_M \geq 3$. In other words, they contradict the minimality of n.

Case $|x| - |y| = 0$

Equality (5.17) implies $|x|_M - |y|_M = -(|f| - |g|) + p$, and we know from Lemma 5.5 that $|f| - |g| \leq 0$. Therefore:

- If $|f| - |g| \leq -1$ or $p \geq 3$, we have $|x|_M - |y|_M \geq 3$. Since moreover $|x| = |y| < n$, we have reached a contradiction with the minimality of n.
- If $|f| - |g| = 0$ and $p = 2$, Equality (5.16) gives $|f|_M - |g|_M = -1$. This, however, contradicts Inequality (5.11).

The last case to deal with is $|x| - |y| < 0$. Since the situation $|x| - |y| = -1$ is very complicated to study, we will start with the case $|x| - |y| \leq -2$ and then deal with $|x| - |y| = -1$ separately.

Case $|x| - |y| \leq -2$

Let us set for simplicity $d = |x| - |y|$.

Equality (5.16) gives $|f|_M - |g|_M = d - 1 \leq -3$. We infer from here and from the minimality of n that $|f| < |g|$ (otherwise one could consider the prefix of f of the length $|g|$, denote it by $\check{f}$, then $|\check{f}| = |g| < n$, $|\check{f}|_M - |g|_M \leq |f|_M - |g|_M \leq -3$, which is a contradiction with the minimality of n).

Similarly, we infer from $|x| < |y| < n$ and from the minimality of n that $|x|_M - |y|_M \leq 2$ (otherwise we denote the prefix of y of the length $|x|$ by $\check{y}$, then $|x| = |\check{y}| < n$, $|x|_M - |\check{y}|_M \geq 3$). Equality (5.15) then gives an upper bound on $|g| - |f|$, namely $|g| - |f| \leq 2 - p - d$.

Let us denote $\hat{f} = Sff'$, where Sff' is a factor of $u^{(p)}$ and $|f'| = |g| - |f| - 1$. Then $|f'| \geq 0$ and $|\hat{f}| = |g| < n$, and it follows from (5.5) that if $f' \neq \epsilon$, the factor f' can be chosen such that its first letter is S, *i.e.* different from M.

Now we are going to express $|\hat{f}|_M - |g|_M$ which is equal to $|f|_M - |g|_M + |f'|_M$. Let us distinguish the cases $|f'| \in \{0, 1\}$ and $|f'| \geq 2$.

- If $|f'| \in \{0, 1\}$, we have $f' = \epsilon$ or $f' = S$, hence $|\hat{f}|_M - |g|_M = |f|_M - |g|_M = d - 1 \leq -3$.
- Let $|f'| \geq 2$. Note that at the same time it holds $|f'| = |g| - |f| - 1 \leq 2 - p - d - 1 \leq -d - 1$, thus necessarily $d \leq -3$. Observation 5.4 leads to

$$|f'|_M \leq \left\lceil \frac{|f'| - 1}{p^2 + p} \right\rceil \leq \left\lceil \frac{2 - p - d - 1 - 1}{p^2 + p} \right\rceil = \left\lceil \frac{-p - d}{p^2 + p} \right\rceil \leq 1 + \frac{-p - d - 1}{p^2 + p}.$$

Then

$$|\hat{f}|_M - |g|_M = |f|_M - |g|_M + |f'|_M = d - 1 + |f'|_M \leq d - 1 + 1 + \frac{-p - d - 1}{p^2 + p}$$

$$= d \cdot \frac{p^2 + p - 1}{p^2 + p} - \frac{p + 1}{p^2 + p} \leq -3 \cdot \frac{p^2 + p - 1}{p^2 + p} - \frac{p + 1}{p^2 + p} = -3 + \frac{2 - p}{p^2 + p} \leq -3.$$

We conclude that for any $d \leq -2$ it holds $|\hat{f}| = |g| < n$ and $|\hat{f}|_M - |g|_M \leq -3$. This is a contradiction with the minimality of n.

Case $|x| - |y| = -1$

In this case, Equality (5.16) implies $|f|_M - |g|_M = -2$ and Equality (5.17) implies $|x|_M - |y|_M = -1 - (|f| - |g|) + p$.

Since $|x| < |y| < n$, necessarily $|x|_M - |y|_M \leq 2$ (the contrary contradicts the minimality of n), hence (5.17) gives

$$|f| - |g| \geq p - 3. \tag{5.19}$$

We observe that this case, namely $|x| - |y| = -1$, can occur only for $p = 2$, because:

- If p was greater than 3, Inequality (5.19) would contradict the inequality $|f| - |g| \leq 0$, derived in Lemma 5.5.
- If p was equal to 3, Inequality (5.19) together with Lemma 5.5 would give $|f| - |g| = 0$, which would not conform to (5.11).

This allows us to restrict our considerations on the case $p = 2$. Inequality (5.19) implies $|f| - |g| \geq -1$, Inequality (5.11) implies $|f| - |g| \leq \frac{|f|_M - |g|_M}{3} = -\frac{2}{3}$, putting it together we infer $|f| - |g| = -1$. This allows us to use equation (5.15) to compute $|x|_M - |y|_M$: we obtain $|x|_M - |y|_M = 2$.

Let us sum up the relations between x and y and between f and g:

$$|x|_M - |y|_M = 2 \quad \wedge \quad |x| - |y| = -1, \tag{5.20}$$

$$|f|_M - |g|_M = -2 \quad \wedge \quad |f| - |g| = -1. \tag{5.21}$$

Our next goal is to prove this lemma:

Lemma 5.9. *It holds:*
- *The word Lx is a factor of $u^{(p)}$ and it holds $Lx = LLS \cdots ML$,*
- *the word MLy is a factor of $u^{(p)}$ and it holds $MLy = MLS \cdots SL$.*

Proof of Lemma 5.9.

The proof will be done in ten steps.

Step 1. (Possible prefixes and suffixes of v, w')
The following four statements hold:

$$
\begin{array}{llll}
\text{(i)} & v = MLSLLSM \cdots & \text{or} & v = MLLSLLSM \cdots, \\
\text{(ii)} & v = \cdots MLSLLSM & \text{or} & v = \cdots MLLSLLSM, \\
\text{(iii)} & w' = MLLSLLSM \cdots & \text{or} & w' = MLSLLSLLSM \cdots, \\
\text{(iv)} & w' = \cdots MLLSLLSM & \text{or} & w' = \cdots MLSLLSLLSM.
\end{array}
$$

Proof. We have found in the proof of Lemma 5.2 that

$$w' = Mw^{(1)}ww^{(2)}M = Mw^{(1)}\hat{w}^{(1)}M \cdots M\hat{w}^{(2)}w^{(2)}M,$$

where $|w^{(j)}\hat{w}^{(j)}|_M = 0$ and $|\hat{w}^{(j)}| \geq p^2 + p$ for $j = 1, 2$. Taking into account Observation 4.5, we deduce that for $p = 2$, the prefix of w' and its suffix can be equal to either $MLLSLLSM$ or $MLSLLSLLSM$, thus (iii) and (iv) are proved.

Statements (i) and (ii) say in fact, with regard to Observation 4.5, that v cannot have the segment $MLSLLSLLSM$ as its prefix and suffix, respectively. Let us suppose to the contrary that $e.g.$ $v = MLSLLSLLSM \cdots$. We define $\hat{v} = (MLSLLSLLS)^{-1}v$, then $|\hat{v}|_M = |v|_M - 1$. At the same time we denote $\hat{w} = \breve{w}^{-1}w$, where $\breve{w}$ is the prefix of w of the length 9 (thus $|\hat{w}| = |\hat{v}| < n$). Since $w = \hat{w}^{(1)}M \cdots$, it follows from Observation 4.5 that $\breve{w}$ contains exactly 1 letter M, hence $|\hat{w}|_M = |w|_M - 1$. Therefore $|\hat{v}|_M - |\hat{w}|_M = 3$, which is a contradiction with the minimality of n. $\qquad\square$

Step 2. (Possible prefixes and suffixes of f, g)
It holds

$$
\begin{array}{llll}
\text{(i)} & f = MLS \cdots & \text{or} & f = LLS \cdots, \\
\text{(ii)} & f = \cdots SML & \text{or} & f = \cdots SLL, \\
\text{(iii)} & g = LLS \cdots & \text{or} & g = MLLS \cdots, \\
\text{(iv)} & g = \cdots SLL & \text{or} & g = \cdots SMLL.
\end{array}
$$

Proof. It is a trivial consequence of Step 1 and of the definitions of f and g. $\quad\square$

Step 3. (The prefix and suffix of f)
It holds $f = LLS \cdots SLL$.

Proof. $\quad\bullet\quad$ We show at first that $f = \cdots SLL$.
Step 2 implies $f = \cdots SML$ or $f = \cdots SLL$. Let us suppose for a while that $f = \cdots SML$. We put $\hat{f} = Sf(ML)^{-1}$ (it is a factor of $u^{(p)}$, because Sf is a factor of $u^{(p)}$, $cf.$ Eq. (5.5)). Then $|\hat{f}| = |f| - 1$, $|\hat{f}|_M = |f|_M - 1$.

Since $g = \cdots LL$ according to Step 2, we are allowed to set $\hat{g} = g(LL)^{-1}$. Then it holds $|\hat{g}| = |g| - 2$, $|\hat{g}|_M = |g|_M$.

The factors $\hat{f}$ and $\hat{g}$ satisfy $|\hat{f}| = |\hat{g}| < n$ and $|\hat{f}|_M - |\hat{g}|_M = |f|_M - |g|_M - 1 = -3$, see (5.21). This is a contradiction with the minimality of n.

$\quad\bullet\quad$ The proof of $f = LLS \cdots$ is slightly simpler. We know from Step 2 that $f = MLS \cdots$ or $f = LLS \cdots$. However, if $f = MLS \cdots$, one can set $\hat{f} = M^{-1}f$, $\hat{g} = g(LL)^{-1}$, and end the proof by contradiction similarly as above.

$\quad\bullet\quad$ Finally, it follows from equation (5.10) and Step 2 that $f \neq LLSLL$. $\quad\square$

Step 4. (The prefixes and suffixes of v, w' and g)
It holds

$$
\begin{array}{ll}
\text{(i)} & v = MLLSLLSM \cdots MLLSLLSM, \\
\text{(ii)} & w' = MLSLLSLLSM \cdots MLSLLSLLSM, \\
\text{(iii)} & g = MLLS \cdots SMLL.
\end{array}
$$

Proof. Statement (i) is a straightforward consequence of Step 3. Statement (ii) follows from (i), which can be proven by contradiction using similar ideas as in the proof in Step 1. Statement (iii) is a consequence of (ii). □

Step 5. (The prefixes and suffixes of x and y)
It holds

$$
\begin{aligned}
\text{(i)} \quad & x = L \cdots LL \quad & \text{or} \quad & x = L \cdots ML, \\
\text{(ii)} \quad & y = SLLS \cdots \quad & \text{and} \quad & y = \cdots LSL.
\end{aligned}
$$

Proof. (i) Since $fS = \varphi_2(x)$ and $fS = LLS \cdots SLLS$ according to Step 3, obviously $x = L \cdots LL$ or $x = L \cdots ML$, *cf.* Observation 4.2 for $p = 2$.
(ii) It follows from Step 4 that $y = \cdots SL$, Observation 4.3 then gives $y = \cdots LSL$. Observation 4.4 and Step 4 imply $g = MLLSLLSM \cdots$, hence $y = SLLS \cdots$. □

Step 6. (Two letters after y)
The word yLS is a factor of $u^{(p)}$.

Proof. It is a consequence of Step 5 and Observation 4.4. □

Step 7. (The suffix of x)
It holds $x = \cdots ML$.

Proof. We consider the result of Step 5 and prove that $x = \cdots LL$ contradicts the minimality of n. Let $x = \cdots LL$. Observation 4.5 then implies that x is followed either by SM or by $SLLSM$. In the first case we define $\hat{x} = L^{-1}xSM$. Taking (5.20) into account, we see that the pair $\hat{x}$, y contradicts the minimality of n. In the case when x is followed by the group $SLLSM$, we define $\hat{x} = \breve{x}^{-1}xSLLSM$ where $\breve{x}$ is the prefix of x of the length 2 (*i.e.* $\breve{x} = LL$ or $\breve{x} = LS$). Also we define $\hat{y} = yLS$ (this is allowed due to Step 6). The pair $\hat{x}$, $\hat{y}$ now contradicts the minimality of n with regard to (5.20) (recall that $|y| + 2 < n$ by Lem. 5.6). □

Step 8. (The letter before x)
The word Lx is a factor of $u^{(p)}$.

Proof. Since SfS is a factor of $u^{(p)}$ and $fS = \varphi_p(x)$, it follows from Observation 4.2 that Lx or Mx is a factor of $u^{(p)}$. However, if Mx is a factor of $u^{(p)}$, then the factors Mx and y contradict the minimality of n with regard to (5.20). □

Step 9. (The prefix of x)
It holds $x = LS \cdots$.

Proof. Since Lx is a factor of $u^{(p)}$ according to Step 8 and $x = L \cdots$ according to Step 5, the word $Lx = LL \cdots$ is a factor of $u^{(p)}$. Observation 4.2 then gives $Lx = LLS \cdots$. □

Step 10. (Two letters before y)
The word MLy is a factor of $u^{(p)}$.

Proof. Step 5 and Observation 4.4 imply that LLy or MLy is a factor of $u^{(p)}$. Let us suppose for a while that $LLy = LLSLLS\cdots$ is a factor of $u^{(p)}$. Observation 4.5 implies immediately that $LLy = LLSLLSM\cdots$. We introduce the word $\hat{y} = (SLLSM)^{-1}yLS$ which is a factor of $u^{(p)}$ due to Step 6. Then we define $\hat{x} = (LS)^{-1}x$, this is a factor of $u^{(p)}$ due to Step 9. It follows from (5.20) that the factors $\hat{x}$, $\hat{y}$ satisfy $|\hat{x}| = |\hat{y}| < n$, $|\hat{x}|_M - |\hat{y}|_M = 3$, *i.e.* they contradict the minimality of n. $\square$

This finishes the proof of Lemma 5.9. The statement follows from the results of Steps 5, 7, 8, 9, 10.

Stage 3: Introduction of r, s

Lemma 5.9 together with Observation 4.1 allow one to define factors r and s of $u^{(p)}$ such that
$$\varphi_2(r)L = Lx, \qquad \varphi_2(s)L = Ly.$$
It is obvious that $|r| \le |x|$, $|s| \le |y|$, hence $|r| < n$, $|s| < n$, *cf.* Lemma 5.6.

Proposition 4.6 and relations (5.21), (5.20) enable us to compute $|r| - |s|$,

$$|r| - |s| = (|x|_S + |x|_M) - (|y|_S + |y|_M) = (|x|_S - |y|_S) + (|x|_M - |y|_M)$$
$$= (|fS|_M - |gS|_M) + (|x|_M - |y|_M) = (|f|_M - |g|_M) + (|x|_M - |y|_M) = -2+2 = 0,$$

and also $|r|_M - |s|_M$:

$$|r|_M - |s|_M = (-|x|_L + 2|x|_S) - (-|y|_L + 2|y|_S) = -(|x|_L - |y|_L) + 2(|x|_S - |y|_S)$$
$$= -[|x| - |y| - (|x|_S - |y|_S) - (|x|_M - |y|_M)] + 2(|x|_S - |y|_S)$$
$$= -(|x| - |y|) + 3(|x|_S - |y|_S) + (|x|_M - |y|_M)$$
$$= -(|x| - |y|) + 3(|fS|_M - |gS|_M) + (|x|_M - |y|_M) = -(-1) + 3 \cdot (-2) + 2 = -3.$$

The pair r, s thus satisfy $|r| = |s| < n$ and $|r|_M - |s|_M \le -3$. This is a contradiction with the minimality of n.

The proof of Theorem 5.1 is completed.

6. BALANCE BOUND WITH RESPECT TO THE LETTER S

Once we know that $u^{(p)}$ is 2-balanced with respect to the letter M, it is easy to prove that it is 2-balanced with respect to the letter S as well.

Theorem 6.1. *Let v, w be factors of $u^{(p)}$ such that $|v| = |w|$. Then*

$$\big| |v|_S - |w|_S \big| \le 2.$$

Proof. We will proceed by contradiction. Let us assume that there exist factors $|v|$, $|w|$ of $u^{(p)}$ such that $|v| = |w|$ and $|v|_S - |w|_S > 2$. Obviously one can suppose

that $|v|_S - |w|_S = 3$. Then, with regard to the substitution rule (2.3), one has

$$|\varphi_p(v)| = (p+1)|v|_L + |v|_S + p|v|_M$$
$$|\varphi_p(w)| = (p+1)|w|_L + |w|_S + p|w|_M \,,$$

hence, using the identities $|v|_L = |v| - |v|_S - |v|_M$, $|w|_L = |w| - |w|_S - |w|_M$,

$$\begin{aligned}
|\varphi_p(v)| - |\varphi_p(w)| &= (p+1)(|v|_L - |w|_L) + (|v|_S - |w|_S) + p(|v|_M - |w|_M) \\
&= (p+1)(|v| - |w|) - p(|v|_S - |w|_S) - (|v|_M - |w|_M) \,.
\end{aligned}$$

Since the factors v and w are of the same length, they satisfy $|v|_M - |w|_M \geq -2$ with regard to Theorem 5.1, hence

$$|\varphi_p(v)| - |\varphi_p(w)| \leq (p+1) \cdot 0 - p \cdot 3 - (-2) = 2 - 3p < 0 \,.$$

Furthermore, since $|\varphi_p(v)|_M = |v|_S$ and $|\varphi_p(w)|_M = |w|_S$, one has

$$|\varphi_p(v)|_M - |\varphi_p(w)|_M = |v|_S - |w|_S = 3 \,.$$

Let $\hat{w}'$ be a prefix of $\varphi_p(w)$ of the length $|\varphi_p(v)|$. Then $|\hat{w}'|_M \leq |\varphi_p(w)|_M$, hence

$$|\varphi_p(v)|_M - |\hat{w}'|_M \geq |\varphi_p(v)|_M - |\varphi_p(w)|_M \geq 3 \,,$$

and since it holds $|\hat{w}'| = |\varphi_p(v)|$, we have reached a contradiction with Theorem 5.1. $\qquad\square$

7. Balance bound with respect to the letter L

This section is devoted to the proof of the third part of Theorem 3.1. As we will see, we will use both facts proved in the previous two sections.

Theorem 7.1. *Let v, w be factors of $u^{(p)}$ such that $|v| = |w|$. Then*

$$\big|\, |v|_L - |w|_L \,\big| \leq 3 \,.$$

Proof. We will again proceed by contradiction. Let us assume that there exist factors $|v|$, $|w|$ of $u^{(p)}$ such that $|v| = |w| = n$ and

$$|v|_L - |w|_L > 3 \,. \tag{7.1}$$

Let n be the minimal number with this property.

We denote $v = v_1 \cdots v_n$, $w = w_1 \cdots w_n$. The minimality of n implies

$$v_1 = v_n = L, \tag{7.2}$$
$$w_1 \neq L, \quad w_n \neq L, \tag{7.3}$$
$$|v|_L - |w|_L = 4 \,. \tag{7.4}$$

The identity $|v| - |w| = |v|_L - |w|_L + |v|_S - |w|_S + |v|_M - |w|_M$ gives

$$|v|_S - |w|_S + |v|_M - |w|_M = -4 \,,$$

and since $|v|_S - |w|_S \geq -2$ and $|v|_M - |w|_M \geq -2$ by virtue of Theorems 5.1 and 6.1, respectively, we infer

$$|v|_S - |w|_S = -2 \qquad \text{and} \qquad |v|_M - |w|_M = -2 \,.$$

Note that since $v_1 = v_n = L$, Observation 4.2 implies that there are $k, \ell \in \{0, 1, \ldots, p - 1\}$ and $X \in \{S, M\}$ such that $XL^k v L^\ell S$ is a factor of $u^{(p)}$. We use these k and ℓ and define factors v' and w' of $u^{(p)}$ in the following way:

$$v' = L^k v L^\ell S \,,$$
$$w' = w_2 \cdots w_n \,.$$

Now we can apply Observation 4.1 which establishes the existence of factors x and y of $u^{(p)}$ satisfying

$$\varphi_p(x) = v' \,, \qquad \varphi_p(y) = w' \,;$$

obviously $|y| \leq |w'| = n - 1$.

We are going to compute $|x|_L - |y|_L$ and $|x| - |y|$. For that purpose the following relations will be useful:

$$
\begin{aligned}
&|v'|_L = |v|_L + k + \ell \,, &\qquad &|w'|_L = |w|_L \,, \\
&|v'|_S = |v|_S + 1 \,, &\qquad &|w'|_S - |w|_S = -|w_1|_S \in \{-1, 0\} \,, \\
&|v'|_M = |v|_M \,, &\qquad &|w'|_S + |w'|_M = |w|_S + |w|_M - 1 \,.
\end{aligned}
$$

Using these relations and Proposition 4.6, one can derive

$$
\begin{aligned}
|x|_L - |y|_L &= |v'|_L - (p - 1)|v'|_S - (|w'|_L - (p - 1)|w'|_S) \\
&= |v'|_L - |w'|_L - (p - 1)\left(|v'|_S - |w'|_S\right) \\
&= |v|_L + k + \ell - |w|_L - (p - 1)\left(|v|_S + 1 - |w|_S + |w_1|_S\right) \\
&= 4 + k + \ell - (p - 1)\left(-2 + 1 + |w_1|_S\right) \\
&\geq 4 + 0 + 0 - (p - 1)(-2 + 1 + 1) = 4
\end{aligned}
$$

and

$$
\begin{aligned}
|x| - |y| &= |v'|_S + |v'|_M - (|w'|_S + |w'|_M) \\
&= |v|_S + 1 + |v|_M - (|w|_S + |w|_M - 1) \\
&= |v|_S - |w|_S + |v|_M - |w|_M + 2 = -2 - 2 + 2 = -2 \,.
\end{aligned}
$$

Since x is shorter than y, we consider the prefix of length $|x|$ of y and denote it by $\hat{y}$; it obviously holds $|\hat{y}|_L \leq |y|_L$. Therefore

$$|x| = |\hat{y}| < n \quad \text{and} \quad |x|_L - |\hat{y}|_L \geq |x|_L - |y|_L = 4 \,;$$

in other words, factors x and $\hat{y}$ contradict the minimality of n. $\qquad\square$

8. Optimality of the balance bounds

To complete the proof of Theorem 3.1, we have to show that the bounds 3, 2 and 2 corresponding to L, S and M, respectively, are optimal. To demonstrate this fact it suffices to find three pairs of factors of $u^{(p)}$, let us denote them by $(v^{(M)}, w^{(M)})$, $(v^{(S)}, w^{(S)})$ and $(v^{(L)}, w^{(L)})$, such that

$$
\begin{array}{llll}
|v^{(M)}| = |w^{(M)}| & \text{and} & \left| |v^{(M)}|_M - |w^{(M)}|_M \right| = 2\,, \\
|v^{(S)}| = |w^{(S)}| & \text{and} & \left| |v^{(S)}|_S - |w^{(S)}|_S \right| = 2\,, \\
|v^{(L)}| = |w^{(L)}| & \text{and} & \left| |v^{(L)}|_L - |w^{(L)}|_L \right| = 3\,.
\end{array}
$$

There are many possibilities, one can take for example:

- $v^{(M)} = M\varphi_p^2(SM)$, $w^{(M)} = \varphi_p^2(SL)(L^{p-2}SM)^{-1}$
  Then $|v^{(M)}| = |w^{(M)}| = p^2 + p + 1$, $|v^{(M)}|_M - |w^{(M)}|_M = 2$.
- $v^{(S)} = L^{-p}\varphi_p^2(LSL)M^{-1}$, $w^{(S)} = \varphi_p^2(ML)L^p$
  Then $|v^{(S)}| = |w^{(S)}| = 2p^2 + 2p + 1$, $|v^{(S)}|_S - |w^{(S)}|_S = 2$.
- $v^{(L)} = \varphi_p^3(ML^{p-1})\varphi_p^2(LL)(L^{p-2}SM)^{-1}$, $w^{(L)} = SM\varphi_p^2(SM)\varphi_p^4(SM)$
  Then $|v^{(L)}| = |w^{(L)}| = p^4 + p^3 + 3p^2 + 1$, $|v^{(L)}|_L - |w^{(L)}|_L = 3$.

9. Abelian complexity

In this section we will determine the optimal bound for the Abelian complexity function $\mathrm{AC}(n)$ of the infinite word $u^{(p)}$, taking advantage of the results on optimal balance bounds that have been derived in the previous part of the paper.

From now on, we will drop the superscript (p) in symbols $\mathcal{F}_{u^{(p)}}(n)$, $\mathcal{P}_{u^{(p)}}(n)$.

Proposition 9.1. *For each $n \in \mathbb{N}$ there are numbers $s_n \in \mathbb{N}$ and $m_n \in \mathbb{N}$ such that $\mathcal{P}_u(n) \subset \{\Psi_1^{(n)}, \ldots, \Psi_9^{(n)}\}$, where*

$$
\begin{aligned}
\Psi_1^{(n)} &= (n - s_n - m_n + 2, s_n - 1, m_n - 1)\,, \\
\Psi_2^{(n)} &= (n - s_n - m_n + 1, s_n - 1, m_n)\,, \\
\Psi_3^{(n)} &= (n - s_n - m_n, s_n - 1, m_n + 1)\,, \\
\Psi_4^{(n)} &= (n - s_n - m_n + 1, s_n, m_n - 1)\,, \\
\Psi_5^{(n)} &= (n - s_n - m_n, s_n, m_n)\,, \\
\Psi_6^{(n)} &= (n - s_n - m_n - 1, s_n, m_n + 1)\,, \\
\Psi_7^{(n)} &= (n - s_n - m_n, s_n + 1, m_n - 1)\,, \\
\Psi_8^{(n)} &= (n - s_n - m_n - 1, s_n + 1, m_n)\,, \\
\Psi_9^{(n)} &= (n - s_n - m_n - 2, s_n + 1, m_n + 1)\,.
\end{aligned}
\tag{9.1}
$$

Consequently, $\mathrm{AC}(n) \leq 9$.

Proof. For each $n \in \mathbb{N}$, Theorem 6.1 implies that

$$\max\{\,|v|_S \mid v \in \mathcal{F}_u(n)\} - \min\{\,|v|_S \mid v \in \mathcal{F}_u(n)\} \leq 2\,,$$

thus there is an $s_n \in \mathbb{N}$ such that

$$v \in \mathcal{F}_u(n) \quad \Rightarrow \quad |v|_S \in \{s_n - 1, s_n, s_n + 1\}\,.$$

Similarly, using Theorem 5.1, one finds that there is an $m_n \in \mathbb{N}$ such that

$$v \in \mathcal{F}_u(n) \quad \Rightarrow \quad |v|_M \in \{m_n - 1, m_n, m_n + 1\}\,.$$

Since $|v|_L = |v| - |v|_S - |v|_M$, we deduce that the set $\mathcal{P}_u(n)$ is a subset of $\{\Psi_1^{(n)}, \dots, \Psi_9^{(n)}\}$, where $\Psi_j^{(n)}$, $j = 1, \dots, 9$, are given by (9.1). $\qquad\square$

Proposition 9.1 gives an upper bound of $\mathrm{AC}(n)$, namely $\mathrm{AC}(n) \leq 9$ for every $n \in \mathbb{N}$, but we can even say more. Indeed, Theorem 7.1 implies that $\mathcal{P}_u(n)$ cannot contain at the same time $\Psi_1^{(n)}$ and $\Psi_9^{(n)}$, hence $\mathrm{AC}(n) \leq 8$. In fact, the optimal bound is even lower, as we will show with the help of the following proposition.

Proposition 9.2. *There is no pair of factors v, w of $u^{(p)}$ such that their Parikh vectors satisfy $\Psi(v) - \Psi(w) = (3, -2, -1)$.*

Proof. We prove the statement by contradiction. Let us suppose that there are factors v and w of $u^{(p)}$ satisfying

$$|v|_L - |w|_L = 3\,, \quad |v|_S - |w|_S = -2\,, \quad |v|_M - |w|_M = -1\,. \tag{9.2}$$

The factors v and w are obviously of the same length which we denote by n. If $v = v_1 \cdots v_n$ and $w = w_1 \cdots w_n$, we may assume without loss of generality that $\{v_1, v_n\} \cap \{w_1, w_n\} = \emptyset$ (if *e.g.* $v_1 = w_n$, we can replace the pair (v, w) by $(v_2 \cdots v_n, w_1 \cdots w_{n-1})$). Having this assumption and equations (9.2), we deduce

$$w_1 \neq L \quad \text{and} \quad w_n \neq L\,, \tag{9.3}$$

because if *e.g.* $w_1 = L$, then the pair $v_2 \cdots v_n$ and $w_2 \cdots w_n$ contradicts Theorem 7.1. Similarly, Theorem 6.1 implies that

$$v_1 \neq S \quad \text{and} \quad v_n \neq S\,. \tag{9.4}$$

Let $k, \ell \geq 0$ and $X, Y \in \{S, M\}$ be such numbers and letters that the word $XL^k vL^\ell Y$ is a factor of $u^{(p)}$. It is a trivial fact that such k, ℓ, X, Y exist. Moreover, Observation 4.3 implies $Y = S$ and Observation 4.2 gives $k, \ell \leq p$. We take these k, ℓ and X and define a factor v' of $u^{(p)}$, and then define also a factor w' of $u^{(p)}$:

$$v' = L^k vL^\ell S,$$
$$w' = w_2 \cdots w_n.$$

Now we apply Observation 4.1 which says that there are factors x and y of $u^{(p)}$ such that

$$v' = \varphi_p(x), \qquad w' = \varphi_p(y).$$

Let us write down the following relations between v' and v and between w' and w,

$$
\begin{aligned}
&|v'|_L = |v|_L + k + \ell\,, && |w'|_L = |w|_L\,, \\
&|v'|_S = |v|_S + 1\,, && |w'|_S - |w|_S = -|w_1|_S = |w_1|_M - 1 \in \{-1, 0\}, \\
&|v'|_M = |v|_M\,, && |w'|_S + |w'|_M = |w|_S + |w|_M - 1,
\end{aligned}
$$

and use them to express $|x|_M - |y|_M$, applying Proposition 4.6:

$$
\begin{aligned}
|x|_M - |y|_M &= -|v'|_L + p|v'|_S - (-|w'|_L + p|w'|_S) \\
&= -\left(|v'|_L - |w'|_L\right) + p\left(|v'|_S - |w'|_S\right) \\
&= -\left(|v|_L + k + \ell - |w|_L\right) + p\left(|v|_S + 1 - |w|_S + 1 - |w_1|_M\right) \\
&= -(3 + k + \ell) + p\left(-2 + 2 - |w_1|_M\right) = -3 - k - \ell - p|w_1|_M, \quad (9.5)
\end{aligned}
$$

(note that $|x|_M - |y|_M \le -3$), and also to express $|x| - |y|$:

$$
\begin{aligned}
|x| - |y| &= |v'|_S + |v'|_M - \left(|w'|_S + |w'|_M\right) \\
&= |v|_S + 1 + |v|_M - \left(|w|_S + |w|_M - 1\right) \\
&= |v|_S - |w|_S + |v|_M - |w|_M + 2 = -2 - 1 + 2 = -1.
\end{aligned}
$$

Statement. It holds

(i) $y = M \cdots M$,

(ii) $w_1 = S$.

Both (i) and (ii) can be proved by contradiction:

(i) Suppose that *e.g.* the first letter of y is $X \ne M$. We define $\hat{y} = X^{-1}y$, then $|x| = |\hat{y}|$ and $|x|_M - |\hat{y}|_M = |x|_M - |y|_M \le -3$, which is a contradiction with Theorem 5.1. Similarly we prove that the last letter of y is M.

(ii) Suppose that $w_1 \ne S$. We know from (9.3) that also $w_1 \ne L$, hence $w_1 = M$. Then (9.5) gives $|x|_M - |y|_M \le -3 - p \le -5$. Let Y be the last letter of y and $\hat{y} = yY^{-1}$. Then it holds $|\hat{y}| = |x|$ and $|x|_M - |\hat{y}|_M = |x|_M - |y|_M + |Y|_M \le -5 + 1 = -4$, which is a contradiction with Theorem 5.1.

Statements (i) and (ii) can be used to determine the first letters of w:

$$w = w_1\varphi_p(y) = S\varphi_p(M \cdots M) = SL^{p-1}S \cdots L^{p-1}S\,.$$

However, Observation 4.4 implies that no factor of $u^{(p)}$ can contain the segment $SL^{p-1}S$. This is a contradiction, thus the proposition is proved. $\square$

Theorem 9.3. *(i) Let $u^{(p)}$ be the fixed point of the substitution φ_p defined in (2.3). Then its Abelian complexity satisfies $\mathrm{AC}(n) \le 7$ for all $n \in \mathbb{N}$ and for all $p > 1$.*
(ii) The bound 7 is optimal, i.e. it cannot be improved. Moreover, for any $p > 1$ there is infinitely many numbers $n \in \mathbb{N}$ such that $\mathrm{AC}(n) = 7$.

Proof. (i) Since $\Psi_1^{(n)} - \Psi_8^{(n)} = (3,-2,-1)$ and $\Psi_2^{(n)} - \Psi_9^{(n)} = (3,-2,-1)$, Proposition 9.2 implies that the set $\mathcal{P}_u(n)$ can contain at most one of the vectors $\Psi_1^{(n)}$, $\Psi_8^{(n)}$ and at most one of the vectors $\Psi_2^{(n)}$, $\Psi_9^{(n)}$. Therefore $\mathcal{P}_u(n)$ has at most 7 elements, hence $\mathrm{AC}(n) \le 7$.

(ii) We prove (ii) by showing that there are infinitely many values $n \in \mathbb{N}$ with this property: there are 7 factors of $u^{(p)}$ of the length n such that their Parikh vectors are mutually different.

For all $N \in \mathbb{N}$, we define two auxiliary words:

$$v^{(N)} = \varphi_p^{2N+2}(L)\varphi_p^{2N+1}(L),$$
$$w^{(N)} = \left(\varphi_p^{2N+1}(L)\right)^{-1} \varphi_p^{2N+2}(LS)\varphi_p^{2N+1}(L)\varphi_p^{2N}(L).$$

Statement. For each $N \in \mathbb{N}$ it holds:
  (i) $v^{(N)}$ and $L^p S M v^{(N)}$ are factors of $u^{(p)}$,
  (ii) $w^{(N)}$ and $SML^{p-1}Sw^{(N)}$ are factors of $u^{(p)}$,
  (iii) $v^{(N)}$ has the suffix $SML^{p-1}S$ and $w^{(N)}$ has the suffix $L^p SM$,
  (iv) $\Psi(v^{(N)}) = \Psi(w^{(N)})$.

To see (i), we observe that $v^{(N)}$ is a prefix of $\varphi_p^{2N+2}(LL)$, thus obviously a factor of $u^{(p)}$. Moreover, since MLL is a factor of $u^{(p)}$ (due to $\varphi_p^3(L) = ((L^p S)^p M)^p L^{p-1}S$ and $p \ge 2$), the word $\varphi_p^{2N+2}(MLL)$ is a factor of $u^{(p)}$ as well. Then it suffices to show that $\varphi_p^{2N+2}(M)$ has the suffix $L^p SM$ for each $N \in \mathbb{N}$, which can be done easily by induction.

The proof of (ii) is similar: since $\varphi_p^{2N+1}(L)w^{(N)}$ is a factor of $\varphi_p^{2N+2}(LSL)$, $w^{(N)}$ is a factor of $u^{(p)}$. To demonstrate that $SML^{p-1}Sw^{(N)}$ is a factor of $u^{(p)}$, it suffices to realize that $SML^{p-1}S$ is a suffix of $\varphi_p^{2N+1}(L)$ for all $N \in \mathbb{N}$, which can be shown by induction. (We use the symbol $\mathbb{N}$ for the set of positive integers.)

Let us proceed to (iii). By the definitions above, $v^{(N)}$ has the suffix $\varphi_p^{2N+1}(L)$, $w^{(N)}$ has the suffix $\varphi_p^{2N}(L)$. Now we can use the facts known from (i) and (ii), namely that $\varphi_p^{2N+1}(L) = \cdots SML^{p-1}S$ and $\varphi_p^{2N}(L) = \cdots L^p SM$ for all $N \in \mathbb{N}$.

The statement (iv) is a consequence of the following two equalities:

$$v^{(N)} = \varphi_p^{2N+2}(L)\varphi_p^{2N+1}(L) = \varphi_p^{2N+2}(L)\varphi_p^{2N}(L^p S) = \varphi_p^{2N+2}(L)\left(\varphi_p^{2N}(L)\right)^p \varphi_p^{2N}(S),$$
$$w^{(N)} = \left(\varphi_p^{2N+1}(L)\right)^{-1} \varphi_p^{2N+2}(L)\varphi_p^{2N}(L^{p-1}S)\varphi_p^{2N+1}(L)\varphi_p^{2N}(L)$$
$$= \left(\varphi_p^{2N+1}(L)\right)^{-1} \varphi_p^{2N+2}(L)\left(\varphi_p^{2N}(L)\right)^{p-1} \varphi_p^{2N}(S)\varphi_p^{2N+1}(L)\varphi_p^{2N}(L).$$

For each $N \in \mathbb{N}$, the auxiliary factors $v^{(N)}$ and $w^{(N)}$ are instrumental to define another set of words that we denote $f^{(1)}, \ldots, f^{(7)}$:

$$f^{(1)} = L^p SMv^{(N)}(SML^{p-1}S)^{-1}, \qquad f^{(2)} = Mv^{(N)}S^{-1},$$
$$f^{(3)} = LSw^{(N)}(SM)^{-1}, \qquad f^{(4)} = v^{(N)}, \qquad f^{(5)} = SMv^{(N)}(LS)^{-1},$$
$$f^{(6)} = Sw^{(N)}M^{-1}, \qquad f^{(7)} = SML^{p-1}Sw^{(N)}(L^p SM)^{-1}.$$

It follows from Statement above that all the words $f^{(1)}, \ldots, f^{(7)}$ are factors of $u^{(p)}$ of the same length $n_N = |v^{(N)}| = |w^{(N)}|$. If we compare the numbers of letters L, S and M in these factors, we find that their Parikh vectors are mutually different. This means that $\mathrm{AC}(n) = 7$ for every $n = |v^{(N)}|$, where $N \in \mathbb{N}$, *i.e.* the function AC attains the value 7 infinitely many times. $\qquad \square$

We finish this section by two statements describing the lower bound and the range of $\mathrm{AC}(n)$.

Proposition 9.4. *It holds* $\mathrm{AC}(n) \geq 3$ *for all* $n \in \mathbb{N}$ *and for all* $p > 1$.

Proof. At first we show that for any prefix v of $u^{(p)}$, both LSv and SMv are factors of $u^{(p)}$. This will be obvious from these three simple facts:
(i) Since LL is a factor of $u^{(p)}$, both $\varphi_p^{2N}(L)\varphi_p^{2N}(L)$ and $\varphi_p^{2N+1}(L)\varphi_p^{2N+1}(L)$ are factors of $u^{(p)}$ for any $N \in \mathbb{N}$.
(ii) It holds $\varphi_p^{2N}(L) = \cdots SM$ and $\varphi_p^{2N+1}(L) = \cdots LS$ for any $N \in \mathbb{N}$.
(iii) There is an $N \in \mathbb{N}$ such that v is a prefix of $\varphi_p^{2N}(L)$, and thus of $\varphi_p^{2N+1}(L)$.
Statements (i) and (ii) imply that both $SM\varphi_p^{2N}(L)v$ and $LS\varphi_p^{2N+1}(L)v$ are factors of $u^{(p)}$, and (iii) then implies that SMv and LSv are factors of $u^{(p)}$ as well.

Let now $v = u_0 u_1 \cdots u_{n-1}$ be the prefix of $u^{(p)}$ of the length n and let $\Psi(v)$ be its Parikh vector. We will show that then there are factors v' and v'' of $u^{(p)}$ such that $\Psi(v') \neq \Psi(v) \neq \Psi(v'')$ and $\Psi(v') \neq \Psi(v'')$. We distinguish three cases according to u_{n-1}:
• If $u_{n-1} = L$, then we set $v' = Su_0 u_1 \cdots u_{n-2}$, $v'' = Mu_0 u_1 \cdots u_{n-2}$.
• If $u_{n-1} = S$, it holds necessarily $u_{n-2} = L$ (*cf.* Observation 4.3). In this case we set $v' = Mu_0 u_1 \cdots u_{n-2}$, $v'' = SMu_0 u_1 \cdots u_{n-3}$.
• If $u_{n-1} = M$, it holds necessarily $u_{n-2} = S$ (*cf.* again Observation 4.3). We set $v' = Su_0 u_1 \cdots u_{n-2}$, $v'' = LSu_0 u_1 \cdots u_{n-3}$. $\qquad \square$

Remark 9.5. It can be demonstrated that for each value $k \in \{3, 4, 5, 6, 7\}$ there is an $n \in \mathbb{N}$ such that $\mathrm{AC}(n) = k$, but we omit the proof with regard to the length of the paper.

10. Conclusion

We have studied balance properties and the Abelian complexity of a certain class of infinite ternary words. We have found the optimal c such that these words are c-balanced, and also the optimal bound for their Abelian complexity functions. We have introduced a new notion, namely the property that a word is "c-balanced with respect to the letter a", which helped us to proceed more effectively from the knowledge of the balance properties to the estimate on $\mathrm{AC}(n)$.

The class of words studied in this paper has one parameter $p > 1$. However, it emerged from our calculations that all the three optimal bounds for balances with respect to particular letters, as well as the optimal bound for the Abelian complexity, are independent of the value of p.

The problem has one more aspect. So far the subword complexity and the Abelian complexity are considered as highly independent of each other (*cf. e.g.* the work of Richomme, Saari and Zamboni). However, our result can indicate that there are connections between them, for the present waiting for their discovery. It has been recently shown in [12] that a fixed point of a canonical substitution associated with a non-simple cubic Parry number has affine factor complexity if and only if it belongs just to the class with which we have dealt in this paper. Therefore, briefly speaking, "if the factor complexity is affine, then the Abelian complexity has the optimal bound 7" holds in the cubic non-simple Parry case. We remark that this sort of statement holds as well in the quadratic non-simple Parry case, although there it is a trivial fact.

References

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

[2] B. Adamczewski, Balances for fixed points of primitive substitutions. *Theoret. Comput. Sci.* **307** (2003) 47–75.

[3] Ľ. Balková, E. Pelantová and Š. Starosta, Sturmian Jungle (or Garden?) on Multiliteral Alphabets. *RAIRO: Theoret. Informatics Appl* (to appear).

[4] Ľ. Balková, E. Pelantová and O. Turek, Combinatorial and Arithmetical Properties of Infinite Words Associated with Quadratic Non-simple Parry Numbers. *RAIRO*: *Theoret. Informatics Appl.* **41 3** (2007) 307–328.

[5] V. Berthé and R. Tijdeman, Balance properties of multi-dimensional words. *Theoret. Comput. Sci.* **273** (2002) 197–224.

[6] J. Cassaigne, Recurrence in infinite words, in *Proc. STACS*, LNCS Dresden (Allemagne) **2010**, Springer (2001) 1–11.

[7] J. Cassaigne, S. Ferenczi and L.Q. Zamboni, Imbalances in Arnoux-Rauzy sequences. *Ann. Inst. Fourier* **50** (2000) 1265–1276.

[8] E.M. Coven and G.A. Hedlund, Sequences with minimal block growth. *Math. Syst. Th.* **7** (1973) 138–153.

[9] J. Currie and N. Rampersad, Recurrent words with constant Abelian complexity. *Adv. Appl. Math.* doi:10.1016/j.aam.2010.05.001 (2010).

[10] S. Fabre, Substitutions et β-systèmes de numération. *Theoret. Comput. Sci.* **137** (1995) 219–236.

[11] C. Frougny, J.P. Gazeau and J. Krejcar, Additive and multiplicative properties of point-sets based on beta-integers. *Theor. Comp. Sci.* **303** (2003) 491–516.

[12] K. Klouda and E. Pelantová, Factor complexity of infinite words associated with non-simple Parry numbers. *Integers – Electronic Journal of Combinatorial Number Theory* (2009) 281–310.

[13] M. Lothaire, *Algebraic combinatorics on words.* Cambridge University Press (2002).

[14] M. Morse and G.A. Hedlund, Symbolic dynamics. *Am. J. Math.* **60** (1938) 815–866.

[15] M. Morse and G.A. Hedlund, Symbolic dynamics II. Sturmian Trajectories. *Am. J. Math.* **62** (1940) 1–42.

[16] G. Richomme, K. Saari, L. Q. Zamboni, Balance and Abelian Complexity of the Tribonacci word. *Adv. Appl. Math.* **45** (2010) 212–231.

[17] G. Richomme, K. Saari, L. Q. Zamboni, Abelian Complexity in Minimal Subshifts. *J. London Math. Soc.* (to appear).

[18] W. Thurston, *Groups, tilings and finite state automata*. AMS Colloquium Lecture Notes (1989).

[19] O. Turek, Balance properties of the fixed point of the substitution associated to quadratic simple Pisot numbers. *RAIRO*: *Theoret. Informatics Appl.* **41** 2 (2007) 123–135.

[20] L. Vuillon, Balanced words. *Bull. Belg. Math. Soc. Simon Stevin* **10** (2003) 787–805.