

Avoiding Letter Patterns in Ternary Square-Free Words

Elena A. Petrova

Department of Mathematics and Computer Science
Ural Federal University
Ekaterinburg, Russia

`Elena.Petrova@urfu.ru`

Submitted: Jul 20, 2015; Accepted: Jan 17, 2016; Published: Feb 5, 2016
Mathematics Subject Classification: 68R15

Abstract

We consider special patterns of lengths 5 and 6 in a ternary alphabet. We show that some of them are unavoidable in square-free words and prove avoidability of the other ones. Proving the main results, we use Fibonacci words as codes of ternary words in some natural coding system and show that they can be decoded to square-free words avoiding the required patterns. Furthermore, we estimate the minimal local (critical) exponents of square-free words with such avoidance properties.

Keywords: square-free ternary words, pattern avoidability, Fibonacci words.

1 Introduction

Repetition-free words and morphisms are among the most important objects of study in combinatorics on words and formal language theory. At the beginning of the 20th century Axel Thue constructed an infinite square-free word over ternary alphabet [18] and an infinite binary cube-free (and, moreover, overlap-free) word [19]. Since Thue, the most popular constructions for infinite repetition-free words were based on repetition-free morphisms, intensively studied in many works; see, e.g., the book [8], and also papers [4, 7, 16]. Sometimes, more general substitutions were used instead of morphisms, for example in the Arshon words [1].

Constructing repetition-free words with additional restrictions forms a significant share of important and interesting tasks in the study of repetition-free words. In [19] Thue posed a question: which of the three-letter words over ternary alphabet are avoidable by square-free words? He showed that the word abc and thus any word obtained from it by permuting the alphabet is unavoidable. Also he considered the pairs of words, where the

first one is from the set $\{aba, bcb, cac\}$ and the other is from $\{bab, cbc, aca\}$, and proved that all these pairs are avoidable. Related research was also done for other repetition-free words, like the binary cube-free words. A full description of binary patterns avoidable by these words was obtained in [10]. Another interesting restriction was offered in [14] where the authors constructed an infinite binary cube-free word with squares of length at most 4. In [2, 3], it was shown that there exist infinite words over a k -letter alphabet, where $k \geq 3$, containing only a finite number of distinct factors of exponent $RT(k)$, which is the repetition threshold from Dejean's conjecture [6] equal to the infimum of avoidable powers over the k -letter alphabet.

In this paper, we continue the investigations of Thue and consider the avoidability of more general patterns in ternary square-free words. These patterns are words over an alphabet of variables $\{x, y, z\}$, where each variable stands for one letter from $\{a, b, c\}$ and different variables denote different letters. For example, the pattern $xyxzx$ represents the set of words $\{abaca, acaba, babcb, bcbab, cacbc, cbcac\}$; to prove that this pattern is avoidable, we need to build an infinite square-free word over $\{a, b, c\}$ containing no factors from this set. We call such patterns "letter patterns". It follows immediately from the results of Thue that square-free letter patterns of length 3 and 4 are unavoidable. We consider all square-free letter patterns of lengths 5 and 6 and clarify their avoidability status, proving the following

Theorem 1. *The following ternary square-free letter patterns are avoidable by ternary square-free words: (a) $xyxzx, xyzxy$; (b) $xyxzyz$ and all patterns of length 6 containing a pattern from (a). All other such patterns of length ≤ 6 are unavoidable by ternary square-free words.*

To construct square-free words for avoidable letter patterns and to prove that the other ones are unavoidable we use an idea by Pansiot [12] who proposed, in relation with Dejean's conjecture, a binary encoding for k -ary words avoiding "local" repetitions. Pansiot used a morphism to generate an infinite binary word which "decodes" into a quaternary word avoiding all powers greater than $RT(4) = 7/5$. The approach based on Pansiot's encoding was used in all later papers devoted to the proof of different cases of Dejean's conjecture. For example, Rao [15] built the appropriate binary codewords as morphic images of the Thue-Morse word (which is itself generated by a morphism).

Shur developed the idea of Pansiot's encoding for the case of ternary square-free words [17]. Namely, in this case the Pansiot codeword can be represented by a walk in a weighted $K_{3,3}$ graph, where each vertex has edges of weights 1, 2, and 3. Due to symmetry, such a walk is just a ternary sequence of weights, called a *codewalk*. In the cited paper, codewalks generated by means of morphisms were used to generate circular square-free words. It appears that letter patterns of lengths 5 and 6 have clear representations in terms of codewalks. The only three avoidable letter patterns correspond to codewalks containing just two weights of the three available. Also by means of codewalks it is easy to prove that the remaining letter patterns are unavoidable by square-free words. In each case of an avoidable pattern the famous Fibonacci word (also generated by a morphism) is used as the codewalk. After proving that the Fibonacci word decodes to a square-free

word in all cases, we describe more precisely the fractional powers avoided by each of obtained square-free words.

2 Preliminaries

2.1 Notation and definitions

An *alphabet* Σ is a nonempty finite set, the elements of which are called *letters*. We consider finite and infinite sequences of letters both called *words* over the binary alphabet $\{a, b\}$, the ternary alphabet $\{a, b, c\}$, and some auxiliary alphabets.

The *empty word* is denoted by λ . We write $|W|$ for the length of a word W . The letters of nonempty finite and infinite words are numbered from 1; thus, $W = W[1..|W|]$ for a finite word.

We use standard definitions of factors, prefixes, and suffixes of a word. Words U and V are called *conjugates* if there exist two words X and Y such that $U = XY$ and $V = YX$. We also call V a *cyclic shift* of U . A positive integer $p \leq |W|$ is a *period* of a word W if $W[1..|W|-p] = W[1+p..|W|]$. If p is the minimal period of W , we use a standard notation $W = U^k$, where $U = W[1..p]$ and $k = |W|/p$. In this case, we call U the *root* of W and k the *exponent* of W (denoted by $\text{exp}(W)$). Words of exponent 2 and 3 are called *squares* and *cubes*, respectively. A square is *minimal* if it does not contain shorter squares as factors. The *local exponent* of a word is the number $\text{lexp}(W) = \sup\{\text{exp}(V) \mid V \text{ is a factor of } W\}$. Local exponents of infinite words are also called *critical exponents*.

A word W is β -free [β^+ -free] if $\text{lexp}(W) < \beta$ [respectively, $\text{lexp}(W) \leq \beta$]. The 2-free words are called *square-free*. It is obvious that a word is square-free if and only if it contains no minimal squares as factors.

2.2 Fibonacci words

Consider the *Fibonacci morphism* ϕ , defined over the binary alphabet by the equalities $\phi(a) = ab, \phi(b) = a$. The iteration of this morphism on the letter b gives the *Fibonacci words*: $f_{-1} = b, f_0 = a, f_1 = ab, f_2 = aba, f_3 = abaab, f_4 = abaababa$, and so on. Since f_n is a prefix of f_{n+1} for any $n \in \mathbb{N}$, one can consider the infinite word $f = \lim_{n \rightarrow \infty} \{f_n\}$ which is the fixed point of ϕ . We call f the *Fibonacci ω -word* to distinguish it from finite Fibonacci words (notice that in [9] the term ‘‘Fibonacci word’’ is used only for this infinite word). We will use the following properties of the Fibonacci words:

1. f is a *Sturmian* word, i.e. it has exactly $n + 1$ different factors of length n [9].
2. $f_n = f_{n-1}f_{n-2}$ for all $n \in \mathbb{N}$.
3. The length of the n th Fibonacci word is the n th Fibonacci number Φ_n for all $n \in \mathbb{N}$ (assuming $\Phi_0 = 1, \Phi_1 = 2$; follows from property 2).
4. The Fibonacci ω -word does not contain factors aaa and bb (words with such factors have no preimages under ϕ).

5. If a factor of the Fibonacci ω -word has nontrivial periods, then its minimal period is a Fibonacci number [5].
6. $\text{lexp}(f) = 2 + \rho$, where ρ is the golden ratio; this supremum is not reached, so the local exponent of any Fibonacci word is smaller than $2 + \rho$ [11].
7. If u^k is a factor of f , where $u \neq \lambda, k > (2 + \rho)/2$, then there exists $n \geq 1$ such that u is a conjugate of f_n and, moreover, each occurrence of u^k is contained in a maximal one of f_n^s for some $s \in [2, 2 + \rho]$ [13].
8. The length of a factor in f whose period is Φ_n is at most $\Phi_{n+1} + 2\Phi_n - 2$ [11].

2.3 Codewords and codewalks

Any ternary word U of length ≥ 3 containing no squares of letters (in particular, any ternary square-free word) can be encoded by a binary *Pansiot codeword* $\text{cwd}(U)$ of length $|U| - 2$.

$$\text{cwd}(U)[i] = \begin{cases} 0 & \text{if } U[i] = U[i+2], \\ 1 & \text{otherwise;} \end{cases}$$

for example,

$$\begin{array}{rcccccccc} U & = & a & b & c & b & a & c & b & c & \dots, \\ \text{cwd}(U) & = & 1 & 0 & 1 & 1 & 1 & 0 & & & \dots. \end{array}$$

This type of encoding was proposed in [12] for bigger alphabets and studied in [17] for the ternary alphabet. We recall some facts from [17].

The codewords of square-free words are also called square-free. Let us consider them. They do not contain the factors 00 and 1111 encoding the squares of period 2 and 3, respectively. Zeroes in a codeword correspond to the “jumps” of one letter over another letter in the encoded word. There are six such jumps, represented by the factors aba , $bc b$, cac , aca , bab , and cbc . We call the first three jumps *right* and the remaining jumps *left*. A right jump in a square-free word is always followed by a left jump and vice versa. The next jump is obtained from the previous one by

- changing the central letter (e. g., $aba \leftrightarrow aca$) if the 0's are separated by 1;
- changing the side letters (e. g., $aba \leftrightarrow cbc$) if the 0's are separated by 11;
- switching the letters (e. g., $aba \leftrightarrow bab$) if the 0's are separated by 111.

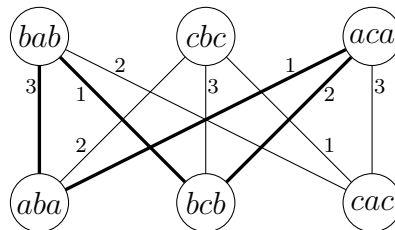


Figure 1: The graph of jumps in ternary words. Bold edges mark the closed codewalk (1213).

In order to describe square-free codewords, the complete bipartite graph $K_{3,3}$ is used. Left [right] jumps correspond to the bottom [resp., top] part of the graph. The number of 1's between two jumps equals the weight of the edge connecting the corresponding vertices. Each square-free codeword not equal to 0 corresponds to a walk represented as a sequence of edge weights (i. e., words over $\{1, 2, 3\}$) in the weighted graph shown in Fig. 1. We call such sequences *codewalks*. Note that Thue [19] proved that for any pair of vertices (x, y) , where x is from the top part and y is from the bottom part, there exists an infinite walk which does not contain x and y and corresponds to a square-free word. In order to decode a word from a codewalk uniquely, one has to keep the first two letters of this word *and the information about the leading and trailing 0's in the codeword*.

Example 2. A ternary word $abacbabcbcabcb$ has the codeword 0111011010111 and the codewalk 3213. Depending in the leading and trailing zeroes, this codewalk corresponds to three more codewords 111011010111, 1110110101110, and 01110110101110. Starting with the same letters ab , the second of these codewords decodes as $abcabacbcacbac$, which has little in common with the initial word.

Since we are interested in constructing codewalks (corresponding to square-free words), for convenience we assume that codewords corresponding to these codewalks always start with 0.

Remark 3. If a codewalk X decodes to a word W , then a suffix of X decodes, in general, to an image of the corresponding suffix of W under some permutation of the alphabet.

Due to symmetry, the sequence of weights in $K_{3,3}$ determines whether the walk is closed independently of the initial vertex. Any closed walk is a combination of simple cycles (a closed walk of length two is considered as a simple cycle also).

Remark 4. [17] There are no minimal squares with periods 5, 7, 9, 10 over a ternary alphabet. The roots of length ≥ 11 in periodic words are coded by the codewalks of length ≥ 4 .

Remark 5. [17] Any infinite codewalk of a square-free word does not contain 11, 222, 223, 322, 333 as factors.

The next lemma is crucial for constructing ternary square-free words from codewalks.

Lemma 6 ([17]). *A codewalk having (a) no factors 11, 222, 223, 322, 333, and (b) no factors of the form XYX such that $|Y| = 2$, $|X|$ is even, and (XY) is the label of a closed walk, decodes to a square-free word.*

By *square-free codewalks* we mean the codewalks decoding to square-free words. Note that square-free codewalks are ternary words with much weaker restrictions than square-freeness: squares in codewalks are permitted if their roots are not closed walks.

3 Letter patterns and codewalks

In this section we start the proof of Theorem 1. We show which of 5-letter and 6-letter patterns are unavoidable using the properties of codewalks and present an idea for

constructing ternary square-free words avoiding the remaining such patterns. Note that all words represented by a letter pattern have the same codeword.

First, consider all square-free 5-letter patterns and their codewords.

$$\begin{array}{c|c|c|c|c} x & y & x & z & x \\ 0 & 1 & 0 & & \end{array} \quad \begin{array}{c|c|c|c|c} x & y & x & z & y \\ 0 & 1 & 1 & & \end{array} \quad \begin{array}{c|c|c|c|c} x & y & z & x & y \\ 1 & 1 & 1 & & \end{array} \quad \begin{array}{c|c|c|c|c} x & y & z & x & z \\ 1 & 1 & 0 & & \end{array} \quad \begin{array}{c|c|c|c|c} x & y & z & y & x \\ 1 & 0 & 1 & & \end{array}$$

By Remark 5, letter patterns $xyxzy$ and $xyzxz$ are unavoidable since 11 is prohibited in codewalks of square-free words. The pattern $xyzyx$ is also unavoidable since 00 and 1111 in codewords correspond to squares. The other two letter patterns of length 5 are avoidable if there exist codewalks of infinite square-free words using only two of the letters $\{1, 2, 3\}$ ($\{1, 2\}$ for $xyxzy$ and $\{2, 3\}$ for $xyzxz$).

Suppose that we have constructed such codewalks and proved the avoidability of $xyxzy$ and $xyzxz$ by square-free words. Then, we need to consider only those 6-letter patterns which do not contain avoidable 5-letter patterns as factors:

$$\begin{array}{c|c|c|c|c|c} x & y & x & z & y & z \\ 0 & 1 & 1 & 0 & & \end{array} \quad \begin{array}{c|c|c|c|c|c} x & y & z & x & z & y \\ 1 & 1 & 0 & 1 & & \end{array} \quad \begin{array}{c|c|c|c|c|c} x & y & z & y & x & z \\ 1 & 0 & 1 & 1 & & \end{array}$$

By the same argument as for the 5-letter case, we conclude that the last two 6-letter patterns are unavoidable. The first one corresponds to 2 in a codewalk, hence, if there exists a codewalk of an infinite square-free word using only the letters 1 and 3 then $xyxzyz$ is an avoidable letter pattern.

Summarizing the above considerations, we want to answer the following question: *are there infinite square-free codewalks using only two of the letters $\{1, 2, 3\}$?* It is easy to see that if such codewalks exist, they contain no cubes of one letter and no squares of the other letter according to Remark 5. (If the pair is $\{2, 3\}$, then the codewalk cannot contain the factor 22 because it has no possible extensions leading to square-free words.) This property is exactly the property 4 of the Fibonacci words. What if we take the ω -word f as a codewalk?

4 Constructing square-free words from the Fibonacci words

Consider the codewalks obtained from Fibonacci words by three substitutions:

$$\sigma_{21} : a \rightarrow 2, b \rightarrow 1 \tag{1a}$$

$$\sigma_{31} : a \rightarrow 3, b \rightarrow 1 \tag{1b}$$

$$\sigma_{32} : a \rightarrow 3, b \rightarrow 2 \tag{1c}$$

We call such codewalks *Fibonacci codewalks* and denote by F_n [F] the codewalk obtained from the word f_n [resp., the ω -word f] under one of these substitutions. If we need to specify the substitution being applied we write F_n^{ij} [F^{ij}] for $\sigma_{ij}(f_n)$ [resp., $\sigma_{ij}(f)$].

Let us have a look at ternary words w^{21}, w^{31}, w^{32} decoded from Fibonacci codewalks F^{21}, F^{31}, F^{32} respectively. Suppose we always start decoding with ab .

- $$f = abaababaabaab \dots$$
- (1) $F^{21} = 2122121221221 \dots$
 $w^{21} = abacbcacbabcbacbcabcbacabcbacbcabac \dots$
- (2) $F^{31} = 3133131331331 \dots$
 $w^{31} = abacbabcabcbacbcabcbacabcbacabcbacabcbac \dots$
- (3) $F^{32} = 32332323323323232 \dots$
 $w^{32} = abacbabcbacabcbacbabcbacbcabcbacbcabcbacbcabcbac \dots$

Combining the properties of Fibonacci words with Lemma 6 we will show that Fibonacci codewalks correspond to square-free words. The following lemma and the considerations from Section 3 together imply Theorem 1.

Lemma 7. *Fibonacci codewalks decode to square-free words.*

Proof. In this proof, we show that Fibonacci codewalks F^j satisfy the conditions (a) and (b) of Lemma 6 and hence they are decoded into square-free words.

Using property 4 of Fibonacci words we can easily see that Fibonacci codewalks do not contain forbidden factors from condition (a) of Lemma 6. Then we need to check that Fibonacci codewalks do not contain factors of the form XYX , where XY labels a closed walk and $|Y| = 2$.

Let us consider periodic factors of Fibonacci ω -word. The argument in the case where the period is a Fibonacci number is quite different from the argument in the other case; so we study these cases separately.

Case 1: Periods not equal to Fibonacci numbers. Due to property 7 we know that f does not contain periodic factors with exponent greater than $(2 + \rho)/2$ whose root is not a Fibonacci word. For p big enough, $(2p - 2)/p > (2 + \rho)/2$, so we have no forbidden factors XYX , where $|XY| = p$. Let us check short periods p such that p is even, $(2p - 2)/p \leq (2 + \rho)/2$ and p is not a Fibonacci number. Such values of p are 4, 6, and 10. Factors in f of these lengths corresponding to closed codewalks are $baab$, $abaaba$, $baabaa$, $aabaab$, $abaababaab$, $baababaaba$, $aababaabab$, $ababaababa$, $babaababaa$ (for all three encodings of f into codewalks). Let us check the words XYX in each of these cases. It is obvious that $baabba$ is not a factor of f . Consider the factor $abaaba$ and its cyclic shifts. If $XY = abaaba$ then $XYX = abaabaabaa$. But this is $\phi(ababab)a$ and $ababab = \phi(aaa)$, contradicting property 4. We conclude that such a factor does not exist in f . A similar analysis is applied to the other length 6 closed codewalks.

Suppose $XY = abaababaab$. Then $abaababaab = (abaab)^2 = f_3^2$. Then $XYX = (abaab)^3aba$. This has period $5 = \Phi_3$ and length 18 which is greater than $(|f_4| + 2|f_3| - 2 = 16$, contradicting property 8). The other listed factors of length 10 are conjugates of f_3^2 , so the same observation is true for them.

Case 2: Periods equal to Fibonacci numbers. We will show that the codewalk generated by f_n is not closed for any n . Without loss of generality we will use substitution (1a).

Due to the symmetry in $K_{3,3}$, the same proof works in the two other cases. Consider a sequence $\{S_n\}$ of shortest codewalks (over $\{1, 2\}$) such that $F_n S_n$ is a closed codewalk (see Fig. 1). We want to show that all codewalks S_n are nonempty. Since $F_n S_{n-2} S_{n-1} = F_{n-1} F_{n-2} S_{n-2} S_{n-1}$ is a closed codewalk, the codewalk S_n is obtained from $S_{n-2} S_{n-1}$ by deleting all maximal cycles.

$$\begin{aligned} F_0 = 2 &\Rightarrow S_0 = 2, \\ F_1 = 21 &\Rightarrow S_1 = 12, \\ S_2 &= 212, \\ S_3 &= \cancel{122}12 = 2, \\ S_4 &= 21\cancel{22} = 21, \\ S_5 &= \cancel{22}1 = 1, \\ S_6 &= 2\cancel{11} = 2 = S_0, \\ S_7 &= 12 = S_1, \\ S_8 &= 212 = S_2, \dots \end{aligned}$$

We see that S_n is a periodic sequence with period 6 and all words S_n are nonempty, implying that the codewalk generated by any Fibonacci word is not closed. Now note that any conjugate of a closed codewalk is closed (effectively, two such closed walks coincide up to the origin). Hence we conclude that the codewalks generated by all conjugates of Fibonacci words are not closed. Note also that if $F_n S_n$ is a closed walk then $F_n F_n S_n S_n$ is closed too, and $S_n S_n$ is closed if and only if $F_n F_n$ is closed. Looking at the sequence S_n , we conclude that $S_n S_n$ is closed for $n = 6k, 2 + 6k, 3 + 6k, 5 + 6k, k \in \mathbb{N}_0$, where the lengths of the words S_n are odd. Therefore, $F_n F_n$ generates a closed walk if and only if the length of F_n is odd.

Assume that f contains a factor XYX from Lemma 6(b), where XY is a cyclic shift of $f_n f_n$ for some n . Then the word XYX has the period Φ_n and the length $< (2 + \rho)\Phi_n$ by property 6. Then the length of XYX is less than $(2|XY| - 2)$ for all n big enough. It is easy to check that $n = 5$, and so $|XY| = 2\Phi_5 = 26$, is sufficient. The smaller cases, which are $n = 2$ and $n = 3$, lead to the periods 6 and 10, considered above.

Thus, the lemma is proved. □

5 Exponents of square-free words avoiding 5 and 6-letter patterns

In this section we estimate the minimal exponents of ternary square-free words avoiding letter patterns $xyxzx$, $xyzxy$, and $xyxzyz$. By finding critical exponents of the constructed words w^{ij} and obtaining some lower bounds, we prove the following

Theorem 8. *The minimal critical exponent of a ternary square-free word avoiding a letter pattern of length ≤ 6 is:*

(1) $15/8$ for the pattern $xyxzx$; (2) $11/6$ for the pattern $xyzxy$; and (3) at most $1 + \rho/2$ for the pattern $xyxzyz$, where ρ is the golden ratio.

Remark 9. Using property 2 we obtain the next sequence of equalities:

$$\begin{aligned}
 f_n &= f_{n-1}f_{n-2} \\
 &= f_{n-2}f_{n-3}f_{n-2} \\
 &= f_{n-3}f_{n-4}f_{n-3}f_{n-3}f_{n-4} \\
 &= f_{n-3}f_{n-4}f_{n-4}f_{n-5}f_{n-4} \\
 &= f_{n-3}f_{n-4}f_{n-4}f_{n-5}f_{n-5}f_{n-6} \\
 &= f_{n-3}f_{n-4}f_{n-4}f_{n-5}f_{n-6}f_{n-7} \\
 &= f_{n-3}f_{n-4}f_{n-4}f_{n-4}f_{n-7} \dots
 \end{aligned}$$

Hence, for each $n \geq 2$ a cube $f_n f_n f_n$ occurs in f and for $n \geq 3$ this cube is a ϕ -image of a cube $f_{n-1} f_{n-1} f_{n-1}$. Note that Fibonacci words alternately end with ab and ba ; so the period $|f_n|$ cannot be extended to the left.

Lemma 10. *One has $\text{lexp}(w^{21}) = 11/6$, $\text{lexp}(w^{31}) = 1 + \rho/2$, $\text{lexp}(w^{32}) = 15/8$.*

Proof. To compute the critical exponents of w^{21}, w^{31}, w^{32} we need to consider two types of factors in F decoded into periodic words:

- (1) the factors of the form XYX , where $|XY| \geq 6$ and XY is a label of closed walk (the case $|XY| = 4$ is impossible, see the proof of Theorem 7; as we saw earlier, the maximal such factors of F have the form $F_n F_n F_n Z$, where Z is a prefix of F_n);
- (2) the factors decoded into words with periods less than 11 (see Remark 4).

For convenience, we assume that codeword corresponding to any given codewalk begins with 0 and ends with 1. For example, the codewalk $(121)^3$ corresponds to the codeword $(0101101)^3$. Then the length of the codeword always equals the length of the codewalk plus the sum of its digits. Recall that a codeword of length n is decoded to a ternary word of length $n+2$.

1. Firstly, let us estimate exponents of words generated in w^{ij} by the factors of type (1). The minimal period of the word u decoded from $F_n F_n F_n Z$, where $F_n F_n$ is a closed walk, equals the length of the codeword corresponding to $F_n F_n$, which is $p = s + l$, where s is the sum of digits in $F_n F_n$ and $l = |F_n F_n|$. The length of the word u is the sum of digits of $F_n F_n F_n Z$ plus $|F_n F_n F_n Z| + 2$.

Further, the periodic factor in w^{ij} with the period generated by $F_n F_n$ is several symbols longer than the word u decoded from $F_n F_n F_n Z$. Consider F^{21} . Assume that $F_n F_n F_n Z$ is preceded by 1. Since this 1 breaks the period $|F_n|$ (see Remark 9), this period would extend if we replace this 1 by 2. Hence, in terms of codewords, the period can be extended to the left by 011; since we have 01 instead, the period extends to the left just by one symbol in the codeword, and then, by one symbol in w^{21} . Note that exchanging the roles of 1 and 2 does not affect this result. Now consider the right extension: if the next symbol after Z in the codewalk is 1 [2], then the codeword continues with 010 [resp., 011]. Hence, the period in the codeword extends by exactly two symbols to the right. In total, the periodic factor in w^{21} with the period generated by $F_n F_n$ has the length equal to sum of digits of $F_n F_n F_n Z$ plus $|F_n F_n F_n Z| + 5$. Exactly the same argument for the other two words give the same constant five for w^{31} and the constant seven for w^{32} .

Using property 8, we obtain $|F_n F_n F_n Z| \leq \Phi_{n+1} + 2\Phi_n - 2 = \Phi_n + \Phi_{n-1} + 2\Phi_n - 2 = 3\Phi_n + \Phi_{n-1} - 2$, i.e., Z is a prefix of F_n of length $\leq \Phi_{n-1} - 2$. The number of letters a in f_n is Φ_{n-1} and the number of letters b is Φ_{n-2} . Let α and β denote images of a and b respectively under the substitution σ_{ij} .

Remembering that the last two letters of any f_n is ab or ba , we can establish an upper bound for the local exponent of a periodic word w generated in w^{ij} by $F_n F_n F_n Z$:

$$|w| \leq 3(\alpha\Phi_{n-1} + \beta\Phi_{n-2} + \Phi_n) + \alpha(\Phi_{n-2} - 1) + \beta(\Phi_{n-3} - 1) + (\Phi_{n-1} - 2) + m,$$

where $m = 5$ for $w^{21}, w^{31}, m = 7$ for w^{32} ,

$$\begin{aligned} \text{lexp}(w) &\leq \frac{3(\alpha\Phi_{n-1} + \beta\Phi_{n-2} + \Phi_n) + \alpha(\Phi_{n-2} - 1) + \beta(\Phi_{n-3} - 1) + (\Phi_{n-1} - 2) + m}{2(\alpha\Phi_{n-1} + \beta\Phi_{n-2} + \Phi_n)} \\ &= \frac{3}{2} + \frac{\alpha\Phi_{n-2} + \beta\Phi_{n-3} + \Phi_{n-1} - \alpha - \beta - 2 + m}{2(\alpha\Phi_{n-1} + \beta\Phi_{n-2} + \Phi_n)}. \end{aligned}$$

For $\alpha = 2, \beta = 1$ one has

$$\text{lexp}(w) \leq \frac{3}{2} + \frac{2\Phi_{n-2} + \Phi_{n-3} + \Phi_{n-1}}{2(2\Phi_{n-1} + \Phi_{n-2} + \Phi_n)} = \frac{3}{2} + \frac{\Phi_{n+1}}{2\Phi_{n+2}} = 1 + \frac{\Phi_{n+3}}{2\Phi_{n+2}}.$$

This bound tends to $1 + \rho/2$ as n approaches infinity and its maximum $29/16$ is reached when $n = 2$.

For $\alpha = 3, \beta = 1$:

$$\text{lexp}(w) \leq \frac{3}{2} + \frac{3\Phi_{n-2} + \Phi_{n-3} + \Phi_{n-1} - 1}{2(3\Phi_{n-1} + \Phi_{n-2} + \Phi_n)} = \frac{3}{2} + \frac{2\Phi_n - 1}{4\Phi_{n+1}} = 1 + \frac{\Phi_{n+2} - 1/2}{2\Phi_{n+1}}.$$

This bound also tends to $1 + \rho/2$ as n approaches infinity. Using the explicit formula

$$\Phi_n = \frac{1}{\sqrt{5}} \left(\rho^n + (-1)^{n+1} \rho^{-n} \right)$$

for Fibonacci numbers, it is easy to check that this bound is less than $1 + \rho/2$ for all $n \geq 2$. Thus, this value can be approached arbitrarily close, but is never reached.

For $\alpha = 3, \beta = 2$:

$$\text{lexp}(w) \leq \frac{3}{2} + \frac{3\Phi_{n-2} + 2\Phi_{n-3} + \Phi_{n-1}}{2(3\Phi_{n-1} + 2\Phi_{n-2} + \Phi_n)} = \frac{3}{2} + \frac{\Phi_{n+1} + \Phi_{n-1}}{2(\Phi_{n+2} + \Phi_n)} = 1 + \frac{\Phi_{n+3} + \Phi_{n+1}}{2(\Phi_{n+2} + \Phi_n)}.$$

Again, the limit of this bound as $n \rightarrow \infty$ is $1 + \rho/2$. The fraction behaves, up to the factor $1/2$, quite similar to the ratio between the $(n+1)$ st and n th Fibonacci numbers: converges to ρ and is alternately higher and lower than ρ . Its maximum $9/11$, giving us the upper bound $20/11$, is reached for $n = 2$.

2. Now let us check short periodic factors directly. Due to Remark 4, it is sufficient to check the periods 6 and 8 only.

1. The Fibonacci codewalk F obtained by substitution (1a) contains the factor 1221 decoded to the word $abacabcacab a$ of local exponent $11/6 > 29/16$. w^{21} has no factors with period 8, because codewalks of such factors contain 33. Hence, the local exponent of w^{21} is $11/6$.
2. The Fibonacci codewalk F obtained by substitution (1b) does not contain factors 1221 and 2332 decoded to words of local exponents greater than $1 + \rho/2$ with periods 6 and 8 respectively, hence, the local exponent of w^{31} is $1 + \rho/2$. Note that, unlike the two other cases, this value is unreachable.
3. The Fibonacci codewalk F obtained by substitution (1c) contains the factor 2332 decoded to word $abacbcabcacbcab a$ of local exponent equal to $15/8 > 20/11$. Hence, local exponent of w^{32} is $15/8$. \square

Proof of Theorem 8. (1) Every infinite ternary square-free word w avoiding letter pattern $xyxzx$ have a factor with codewalk 2332 due to Lemma 6, hence $\text{lexp}(w) \geq 15/8$, as we see in the proof of Lemma 10. The word w^{32} is the example proving that this value is precise.

(2) Replacing 2332 with 1221, w^{32} with w^{21} and $15/8$ with $11/6$ in the previous case we obtain the required result.

(3) The word w^{31} avoids $xyxzyz$ and $\text{lexp}(w^{31}) \leq 1 + \rho/2$. \square

6 Discussion

The exploration of letter pattern avoidance by ternary square-free words can be extended to larger lengths of patterns. It is easy to check that the only 7-letter pattern which does not contain avoidable letter patterns of smaller lengths as factors is $xyxzyx$. Proving its avoidance will finalize the classification of letter patterns avoidable by ternary square-free words. This pattern has codeword 11011. The codewalk of an infinite ternary word avoiding this pattern does not contain 22, 23, 32, and 33 as factors, hence, after each 2 and 3 there must be a 1. It is not clear whether such a codewalk can be constructed; but again, what we need is a binary sequence over the alphabet $\{12, 13\}$, so the approach similar to that used in this paper could work.

References

- [1] S. E. Arshon. Proof of the existence of asymmetric infinite sequences. *Mat. Sbornik* 2: 769–779, 1937, in Russian, with French abstract.
- [2] G. Badkobeh, M. Crochemore. Finite-repetition threshold for infinite ternary words. *Proc. 8th Int. Conf. WORDS 2011, EPTCS* 63: 37–43, 2011.
- [3] G. Badkobeh, M. Crochemore, M. Rao: Finite-repetition threshold for large alphabets. *RAIRO Inform. Théor. Appl.* 48: 419–430, 2014.

- [4] M. Crochemore. Sharp characterizations of squarefree morphisms. *Theor. Comp. Sci.* 18: 221–226, 1982.
- [5] J. D. Currie, K. Saari. Least periods of factors of infinite words. *RAIRO Inform. Théor. Appl.* 43: 165–178, 2009.
- [6] F. Dejean. Sur un théorème de Thue. *J. Combin. Theory. Ser. A* 13: 90–99, 1972.
- [7] J. Karhumäki. On cube-free ω -words generated by binary morphisms. *Discr. App. Math.* 5: 279–297, 1983.
- [8] M. Lothaire. *Combinatorics on Words*. Addison-Wesley, 1983.
- [9] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University Press, 2002.
- [10] R. Mercas, P. Ochem, A. Samsonov, A. Shur. Binary patterns in binary cube-free words: avoidability and growth. *RAIRO Inform. Théor. Appl.* 48: 369–389, 2014.
- [11] F. Mignosi, G. Pirillo. Repetitions in the Fibonacci infinite word. *RAIRO Inform. Théor. App.* 26: 199–204, 1992.
- [12] J. J. Pansiot. A propos d’une conjecture de F. Dejean sur les répétitions dans les mots. *Discrete Appl. Math.* 7: 297–311, 1984.
- [13] G. Pirillo. Fibonacci numbers and words. *Discrete Mathematics* 173: 197–207, 1997.
- [14] N. Rampersad, J. Shallit, M. Wang. Avoiding large squares in infinite binary words. *Theor. Comp. Sci.* 339: 19–34, 2005.
- [15] M. Rao. Last Cases of Dejean’s Conjecture. *Theor. Comp. Sci.* 412: 3010–3018, 2011.
- [16] G. Richomme, F. Wlazinski. Some results on k -power-free morphisms. *Theor. Comp. Sci.* 273: 119–142, 2002.
- [17] A. M. Shur. On ternary square-free circular words. *Electronic J. Combinatorics* 17(#R140), 2010.
- [18] A. Thue. Über unendliche Zeichenreihen. *Kra. Vidensk. Selsk. Skrifter. I. Mat.-Nat. Kl., Christiana*: 1–22, 1906.
- [19] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichentreihen. *Norske Vid. Selsk. Skr. I. Mat.-Nat. Kl., Christiana*: 1–67, 1912.