Information Processing Letters 115 (2015) 543–547

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

# Infinite binary words containing repetitions of odd period

Golnaz Badkobeh [a,c,*], Maxime Crochemore [a,b]

[a] King's College London, UK
[b] Université Paris-Est, France
[c] University of Sheffield, UK

ARTICLE INFO

*Article history:*
Received 10 July 2014
Received in revised form 15 January 2015
Accepted 17 January 2015
Available online 22 January 2015
Communicated by M. Chrobak

*Keywords:*
Combinatorial problems
Combinatorics on words
Repetitions
Repetitive threshold
Word morphisms

ABSTRACT

A square is the concatenation of a nonempty word with itself. A word has period $p$ if its letters at distance $p$ match. The exponent of a nonempty word is its length divided by its smallest period. In this article, we give some new results on the trade-off between the number of squares and the number of cubes in infinite binary words whose square factors have odd periods.



## 1. Introduction

Enumerating the repetitions in infinite words is a classic problem in combinatorics on words that has been studied in depth over the last 100 years (see for example, [10,3] and references therein).

A square is the concatenation of a nonempty word with itself. Let $g(n)$ be the length of a longest binary word containing at most $n$ distinct squares. Then $g(0) = 3$ (e.g., 010), $g(1) = 7$ (e.g., 0001000) and $g(2) = 18$ (e.g., 010011000111001101).

In 1974, Entringer, Jackson, and Schatz [5] showed that there exists an infinite word with 5 distinct squares. Therefore, they proved that $g(5) = \infty$. Later, Fraenkel and Simpson [6] showed that there exists an infinite binary word that contains only three squares, 00, 11, and 0101, and thus $g(3) = \infty$. A somewhat simplified proof of this result was given by Rampersad, Shallit and Wang [9]. Later, in 2006, Harju and Nowotka [7] provided a simpler proof of this result and, finally, Badkobeh [2] presented yet another proof by exploiting two simple morphisms.

Instead of avoiding all squares, one interesting variation on the same problem is to avoid larger repetitions. Entringer, Jackson, and Schatz [5] showed that there exist infinite binary words avoiding squares of period at least three. Later works aimed at avoiding large squares, such as, for instance, Dekking [4], Rampersad et al. [9], Shallit [11], Ochem [8].

In this article, we provide study pattern avoidance from a different point of view. We analyse the possibility of avoiding repetitions of even and odd periods, and further impose a constraint on their maximal exponent. This new approach enables us to provide new and interesting results.

We show that there exists no infinite $3^+$-free binary word avoiding all squares of odd period. We also show that there exists no infinite binary word simultaneously avoiding cubes and squares of even period. Moreover, we show that there exists an infinite $3^+$-free binary word avoiding squares of even period.

The trade-off between the maximal period length and the number of repetitions follows a similar trade-off be-

* Corresponding author.
*E-mail address:* golnaz.badkobeh@gmail.com (G. Badkobeh).

http://dx.doi.org/10.1016/j.ipl.2015.01.004

tween the number of cubes and the number of distinct squares. A similar study was comprehensively carried out by the first author in [1].

The article is structured as follows. We provide some definitions in Section 2. In Section 3, we present the proof technique that will be used throughout this article. In addition, we prove that there exists no infinite $3^+$-free binary word avoiding all squares of odd period and there exists no infinite binary word simultaneously avoiding cubes and squares of even period. In Section 4, we show that in fact there exists an infinite $3^+$-free binary word avoiding squares of even period. In Section 5, we reduce the number of repetitions contained in infinite binary words without compromising the constraint on the parity of the periods of squares. We conclude that the minimal number of squares in such words is 7 when only 1 cube occurs. The number reduces to 4 when 2 cubes are allowed in the word. In Section 6, we give a summary of our results.

## 2. Preliminaries

An alphabet is any non-empty set, the members of which are called letters. A word, or a string, is a sequence of letters drawn from the alphabet. The empty word $\epsilon$ is a string of length 0 that is considered to be a word over every alphabet. The length of the word $w$, denoted by $|w|$, is the number of occurrences of letters in $w$. For example, $|\mathtt{abaca}| = 5$.

We consider the ternary alphabet $\mathtt{A} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, the binary alphabet $\mathtt{B} = \{\mathtt{0}, \mathtt{1}\}$, and the $n$-ary alphabet $\Sigma_n$ for $n > 3$.

The word $v$ is called a factor of $x$ if there exist words $u$ and $w$ such that $x = uvw$. In the case $u = \epsilon$ (resp., $w = \epsilon$), $v$ is a prefix (resp., a suffix) of $x$. A nonempty word $x$ has period $p$ if $x[i] = x[i+p]$ for all $i$ for which the equation is meaningful. The exponent of $x$ is its length divided by its smallest period.

The maximum exponent of a word $w$ is the supremum of $E(x)$, where $E(x)$ is the set of exponents of all finite factors of $x$.

A *square* is a word of the form $xx$, where $x$ is a nonempty word. *Cubes* and $k$-th *powers* are defined accordingly. A word is *overlap-free* if it does not contain any factor of the form $xyxyx$ for a non-empty $x$. In general, a word is said to be $\alpha$-free if it contains no factor of the form $u^\beta$ for any rational number $\beta \geq \alpha$. It is $\alpha^+$-free if it contains no factor of the form $u^\beta$ for any rational number $\beta > \alpha$.

A morphism is a map $h: \Sigma_n^* \to \Sigma_m^*$ such that $h(uv) = h(u)h(v)$ for all $u, v \in \Sigma_n^*$. This implies that $h(\epsilon) = \epsilon$. In addition, the morphism $h$ is completely defined by the pairs $(a, h(a))$ for $a \in \Sigma_n$. We refer to images of letters as codewords. If $h(a) = ax$ for some letter $a \in \Sigma_n$, then we say that $h$ is prolongable on $a$, and we can then iterate $h$ infinitely often to get the fixed point $h^\infty(a) := axh(x)h^2(x)h^3(x)\cdots$. For $q \geq 2$ a morphism $h$ is said to be $q$-uniform if $|h(a)| = q$ for all $a \in \Sigma_n$. A uniform morphism $h$ is *synchronising* when $h(ab) = vh(c)w$ implies that either $v = \epsilon$ and $a = c$ or $w = \epsilon$ and $b = c$, for any $a, b, c \in \Sigma_n$ and $v, w \in \Sigma_m^*$. Notice that a synchronising morphism $h$ is always injective (actually it is injective on the set $\Sigma_n$ of monoid generators). Moreover, if it is $q$-uniform then, for each factor $u$ of a word in $h(\Sigma_n^*)$ such that $|u| \geq 2q-1$, there exists a unique factorisation $u = xh(u')y$ where $u' \in \Sigma_n^*$ and $0 \leq |x|, |y| < q$.

## 3. Words containing only repetitions of odd period

Here, we study further the infinite binary words and the squares they contain. Looking at the parity of the periods of the squares reveals interesting properties.

Note that the only infinite binary words omitting 00 and 11 are $(01)^\infty$ and $(10)^\infty$, both of which contain $3^+$-powers. This proves the following proposition.

**Proposition 1.** *There exists no infinite $3^+$-free binary word avoiding all squares of odd period.*

**Proposition 2.** *There exists no infinite binary word, simultaneously avoiding cubes and squares $xx$ with $|x| = 2k$ for $k > 0$. The length of a cube-free binary word containing only squares of odd period does not exceed 23.*

**Proof.** Here, we try to build a binary word that avoids cubes and squares of even period. The following list contains all possible strings with prefix 00, avoiding cubes and squares of even period:

| | | |
|---|---|---|
| 00100100 | 00110010010 | 0011011001001100 |
| 001001100 | 0011001001100 | 0011011001001101100 |
| 00100110110010010 | 0011001001101100100 10 | 0011011001001101101 |
| 0010011011001001100 | 001100100110110010011 | 00110110011 |
| 0010011011001001101 | 00110010011011 0011 | 001101101 |
| 0010011011 0011 | 0011001001101101 | |
| 001001101101 | 00110110010010 | |

The maximum length of these words is 21. This is also true for words starting with 11. Now the only binary words avoiding 00, 11, cubes, and squares of even period are: $\{\mathtt{0}, \mathtt{1}, \mathtt{01}, \mathtt{10}, \mathtt{010}, \mathtt{101}\}$. Concatenating these two sets will not produce a word complying with the properties whose length exceeds 23. □

The remainder of this section is dedicated to demonstrating that if the constraint on the maximal exponent is relaxed so that the word may contain cubes, then avoiding squares of even period becomes possible.

The same technique is used to prove each of the theorems in this article. The technique is stated below. To demonstrate how this technique works, a step-by-step proof is given for Proposition 3, as an example.

**Proof technique.** Let $g$ be a synchronising morphism $g: \mathtt{A}^* \to \mathtt{B}^*$, and let $s$ be an infinite square-free word in $\mathtt{A}^*$. Notice that the only squares occurring in $g(s)$ also occur in the images of square-free factors of $s$ of length 3. Therefore, to study the squares contained in $g(s)$ it is enough to look at all the images of triplets in $\mathtt{A}^*$ (a triplet is a word of length 3). This set is finite and therefore it is possible to count all the squares contained in the images of the set. In order to prove the theorems presented in this article, it is sufficient to show that the given morphisms are synchronising. To demonstrate this we look at the images of all the doublets (words of length 2) in $\mathtt{A}^*$ to investigate if they comply with the definition of synchronising morphisms.