Avoiding 2-binomial squares and cubes

Michaël Rao[a], Michel Rigo[b,*], Pavel Salimov[b,c,1]

[a] *CNRS, LIP, ENS Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07 - France.*
[b] *Dept of Math., University of Liège, Grande traverse 12 (B37), B-4000 Liège, Belgium.*
[c] *Sobolev Institute of Math., 4 Acad. Koptyug avenue, 630090 Novosibirsk, Russia*

Abstract

Two finite words u, v are 2-binomially equivalent if, for all words x of length at most 2, the number of occurrences of x as a (scattered) subword of u is equal to the number of occurrences of x in v. This notion is a refinement of the usual abelian equivalence. A 2-binomial square is a word uv where u and v are 2-binomially equivalent.

In this paper, considering pure morphic words, we prove that 2-binomial squares (resp. cubes) are avoidable over a 3-letter (resp. 2-letter) alphabet. The sizes of the alphabets are optimal.

Keywords: Combinatorics on words; binomial coefficient; binomial equivalence; avoidance; squarefree; cubefree.

1. Introduction

A *square* (resp. *cube*) is a non-empty word of the form xx (resp. xxx). Since the work of Thue, it is well-known that there exists an infinite square-free word over a ternary alphabet, and an infinite cubefree word over a binary alphabet [13, 14]. A main direction of research in combinatorics on words is about the avoidance of a pattern, and the size of the alphabet is a parameter of the problem.

A possible and widely studied generalization of squarefreeness is to consider an abelian framework. A non-empty word is an *abelian square* (resp. *abelian cube*) if it is of the form xy (resp. xyz) where y is a permutation of x (resp. y and z are permutations of x). Erdös raised the question whether abelian squares can be avoided by an infinite word over an alphabet of size 4 [3]. Keränen answered positively to this question, with a pure morphic word [9]. Moreover Dekking has previously obtained an infinite word over a 3-letter alphabet that avoids abelian

*Corresponding author

[1] The third author is supported by the Russian President's grant no. MK-4075.2012.1 and Russian Foundation for Basic Research grants no. 12-01-00089 and no. 11-01-00997 and by a University of Liège post-doctoral grant.

cubes, and an infinite binary word that avoids abelian 4-powers [2]. (Note that in all these results, the size of the alphabet is optimal.)

In this paper, we are dealing with another generalization of squarefreeness and cubefreeness. We consider the 2-binomial equivalence which is a refinement of the abelian equivalence, *i.e.*, if two words x and y are 2-binomially equivalent, then x is a permutation of y (but in general, the converse does not hold, see Example 1 below). This equivalence relation is defined thanks to the binomial coefficient $\binom{u}{v}$ of two words u and v which is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword). For more on these binomial coefficients, see for instance [10, Chap. 6]. Based on this classical notion, the m-binomial equivalence of two words has been recently introduced [12].

Definition 1. Let $m \in \mathbb{N} \cup \{+\infty\}$ and u, v be two words over the alphabet A. We let $A^{\leq m}$ denote the set of words of length at most m over A. We say that u and v are *m-binomially equivalent* if

$$\binom{u}{x} = \binom{v}{x}, \ \forall x \in A^{\leq m}.$$

We simply write $u \sim_m v$ if u and v are m-binomially equivalent. The word u is obtained as a permutation of the letters in v if and only if $u \sim_1 v$. In that case, we say that u and v are *abelian equivalent* and we write instead $u \sim_{\mathsf{ab}} v$. Note that if $u \sim_{k+1} v$, then $u \sim_k v$, for all $k \geq 1$.

Example 1. The four words 0101110, 0110101, 1001101 and 1010011 are 2-binomially equivalent. Let u be any of these four words. We have

$$\binom{u}{0} = 3, \ \binom{u}{1} = 4, \ \binom{u}{00} = 3, \ \binom{u}{01} = 7, \ \binom{u}{10} = 5, \ \binom{u}{11} = 6.$$

For instance, the word 0001111 is abelian equivalent to 0101110 but these two words are not 2-binomially equivalent. Let a be a letter. It is clear that $\binom{u}{aa}$ and $\binom{u}{a}$ carry the same information, *i.e.*, $\binom{u}{aa} = \binom{|u|_a}{2}$ where $|u|_a$ is the number of occurrences of a in u.

A 2-*binomial square* (resp. 2-*binomial cube*) is a non-empty word of the form xy where $x \sim_2 y$ (resp. $x \sim_2 y \sim_2 z$). For instance, the prefix of length 12 of the Thue–Morse word: 011010011001 is a 2-binomial cube. Squares are avoidable over a 3-letter alphabet and abelian squares are avoidable over a 4-letter alphabet. Since 2-binomial equivalence lies between abelian equivalence and equality, the question is to determine whether or not 2-binomial squares are avoidable over a 3-letter alphabet. We answer positively to this question in Section 2. The fixed point of the morphism $g : 0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$ avoids 2-binomial squares.

In a similar way, cubes are avoidable over a 2-letter alphabet and abelian squares are avoidable over a 3-letter alphabet. The question is to determine whether or not 2-binomial cubes are avoidable over a 2-letter alphabet. We also answer positively to this question in Section 3. The fixed point of the morphism $h : 0 \mapsto 001, 1 \mapsto 011$ avoids 2-binomial cubes.

Remark 1. The m-binomial equivalence is not the only way to refine the abelian equivalence. Recently, a notion of m-abelian equivalence has been introduced [8]. To define this equivalence, one counts the number $|u|_x$ of occurrences in u of all factors x of length up to m (it is meant factors made of consecutive letters). That is, u and v are m-abelian equivalent if $|u|_x = |v|_x$ for all $x \in A^{\leq m}$. In that context, the results on avoidance are quite different. Over a 3-letter alphabet 2-abelian squares are unavoidable: the longest ternary word which is 2-abelian squarefree has length 537 [6], and pure morphic words cannot avoid k-abelian-squares for every k [7]. On the other hand, it has been shown that there exists a 3-abelian squarefree morphic word over a 3-letter alphabet [11]. Moreover 2-abelian-cubes can be avoided over a binary alphabet by a morphic word [11].

The number of occurrences of a letter a in a word u will be denoted either by $\binom{u}{a}$ or $|u|_a$. Let $A = \{0, 1, \ldots, k\}$ be an alphabet. The *Parikh map* is an application $\Psi : A^* \to \mathbb{N}^{k+1}$ such that $\Psi(u) = (|u|_0, \ldots, |u|_k)^T$. Note that we will deal with column vectors (when multiplying a square matrix with a column vector on its right). In particular, two words are abelian equivalent if and only if they have the same Parikh vector. The mirror of the word $u = u_1 u_2 \cdots u_k$ is denoted by $\widetilde{u} = u_k \cdots u_2 u_1$.

2. Avoiding 2-binomial squares over a 3-letter alphabet

Let $A = \{0, 1, 2\}$ be a 3-letter alphabet. Let $g : A^* \to A^*$ be the morphism defined by

$$g : \left\{ \begin{array}{lll} 0 & \mapsto & 012 \\ 1 & \mapsto & 02 \\ 2 & \mapsto & 1 \end{array} \right. \quad \text{and thus, } g^2 : \left\{ \begin{array}{lll} 0 & \mapsto & 012021 \\ 1 & \mapsto & 0121 \\ 2 & \mapsto & 02. \end{array} \right.$$

It is prolongable on 0: $g(0)$ has 0 as a prefix. Hence the limit $\mathbf{x} = \lim_{n\to+\infty} g^n(0)$ is a well-defined infinite word

$$\mathbf{x} = g^\omega(0) = 012021012102012021020121 \cdots$$

which is a fixed point of g. Since the original work of Thue, this word $\mathbf{x}$ is well-known to avoid (usual) squares. It is sometimes referred to as the *ternary Thue–Morse word*. We will make use of the fact that $X = \{012, 02, 1\}$ is a prefix-code and thus an ω-code: Any finite word in X^* (resp. infinite word in X^ω) has a unique factorization as a product of elements in X. Let us make an obvious but useful observation.

Observation 1. *The factorization of* $\mathbf{x}$ *in terms of the elements in* X *permits to write* $\mathbf{x}$ *as*

$$\mathbf{x} = 0\,\alpha_1\,2\,\alpha_2\,0\,\alpha_3\,2\,\alpha_4\,0\,\alpha_5\,2\,\alpha_6\,0\cdots$$

where, for all $i \geq 1$, $\alpha_i \in \{\varepsilon, 1\}$. *That is, the image of* $\mathbf{x}$ *by the morphism* $e : 0 \mapsto 0, 1 \mapsto \varepsilon, 2 \mapsto 2$ *(which erases all the* 1*'s) is* $e(\mathbf{x}) = (02)^\omega$.

The next property is well known. For example, it comes from the fact that the image of the ternary Thue–Morse word by the morphism $0 \mapsto 011, 1 \mapsto 01, 2 \mapsto 0$ is the Thue–Morse word. However, for the sake of completeness, we give a direct proof here.

Lemma 1. *A word* u *is a factor occurring in* $\mathbf{x}$ *if and only if* $\widetilde{u}$ *is a factor occurring in* $\mathbf{x}$.

Proof. We define the morphism $\widetilde{g} : A^* \to A^*$ by considering the mirror images of the images of the letters by g,

$$\widetilde{g} : \begin{cases} 0 & \mapsto & 210 \\ 1 & \mapsto & 20 \\ 2 & \mapsto & 1 \end{cases} \quad \text{and thus, } \widetilde{g}^2 : \begin{cases} 0 & \mapsto & 120210 \\ 1 & \mapsto & 1210 \\ 2 & \mapsto & 20. \end{cases}$$

Note that $\widetilde{g}$ is not prolongable on any letter. But the morphism $\widetilde{g}^2$ is prolongable on the letter 1. We consider the infinite word

$$\mathbf{y} = (\widetilde{g}^2)^\omega(1) = 121020121012021020120202101210\cdots.$$

If $v \in A^*$ is a non-empty word ending with $a \in A$, *i.e.*, $v = ua$ for some word $u \in A^*$, we denote by va^{-1} the word obtained by removing the suffix a from v. So $va^{-1} = u$.

For every words r and s we have $r = g^2(s) \Leftrightarrow \widetilde{r} = \widetilde{g}^2(\widetilde{s})$. Obviously, u is a factor occurring in $\mathbf{x}$ if and only if $\widetilde{u}$ is a factor occurring in $\mathbf{y}$.

On the other hand, $\widetilde{g}^2$ is a cyclic shift of g^2, since $g^2(a) = 0\widetilde{g}^2(a)0^{-1}$ for every $a \in \{0, 1, 2\}$. Thus u is a factor occurring in $\mathbf{x}$ if and only if u is a factor occurring in $\mathbf{y}$. To summarize, u is a factor occurring in $\mathbf{x}$ if and only if u is a factor occurring in $\mathbf{y}$, and u is a factor occurring in $\mathbf{y}$ if and only if $\widetilde{u}$ is a factor occurring in $\mathbf{x}$. This concludes the proof.

We will be dealing with 2-binomial squares so, in particular, with abelian squares. The next lemma permit to "desubstitute", meaning that we are looking for the inverse image of a factor under the considered morphism.

Lemma 2. *Let* $u, v \in A^*$ *be two abelian equivalent non-empty words such that* uv *is a factor occurring in* $\mathbf{x}$. *There exists* $u', v' \in A^*$ *such that* $u'v'$ *is a factor of* $\mathbf{x}$, *and either:*

1. $u = g(u')$ *and* $v = g(v')$;
2. *or,* $\widetilde{u} = g(v')$ *and* $\widetilde{v} = g(u')$.

Proof. We will make an extensive use of Observation 1. Note that u and v must contain at least one 0 or one 2. Obviously $e(uv)$ is an abelian square of $(02)^\omega$, thus either $e(u) = e(v) = (02)^i$ or $e(u) = e(v) = (20)^i$ for an $i > 0$.

If $e(u) = e(v) = (02)^i$, then we have $u = a\,0 \cdots 2\,b$ and $v = c\,0 \cdots 2\,d$ with $a, b, c, d \in \{\varepsilon, 1\}$. In this case, we deduce that u and v belong to X^*. Otherwise stated, since uv is a factor of $\mathbf{x}$, there exists a factor $u'v'$ in $\mathbf{x}$ such that $g(u') = u$ and $g(v') = v$.

Otherwise we have $e(u) = e(v) = (20)^i$. Thanks to Lemma 1, $\widetilde{v}\widetilde{u}$ is a factor occurring in $\mathbf{x}$, and $e(\widetilde{u}) = e(\widetilde{v}) = (02)^i$. Thus we are reduced to the previous case, and there is a factor u', v' in $\mathbf{x}$ such that $g(u') = \widetilde{v}$ and $g(v') = \widetilde{u}$.

Let u be a word. We set

$$\lambda_u := \binom{u}{01} - \binom{u}{12}.$$

When we use the desubstitution provided by the previous lemma, the shorter factors u' and v' derived from u and v keep properties from their ancestors.

Lemma 3. *Let $u, v \in A^*$ be two abelian equivalent non-empty words such that uv is a factor occurring in $\mathbf{x}$. Let u', v' be given by Lemma 2. If $\lambda_u = \lambda_v$, then u' and v' are abelian equivalent and $\lambda_{u'} = \lambda_{v'}$.*

Proof. If we are in the second situation described by Lemma 2, then $\widetilde{v}\widetilde{u}$ is also a factor occurring in $\mathbf{x}$. Obviously $\widetilde{v}$ and $\widetilde{u}$ are also abelian equivalent, $\lambda_{\widetilde{v}} = \lambda_{\widetilde{u}}$ and the case is reduced to the first situation.

Assume now w.l.o.g. that we are in the first situation, that is, $u = g(u')$ and $v = g(v')$. First observe that we have, for all $a, b \in A$, $a \neq b$,

$$\binom{u'}{ab} = \binom{|u'|_a + |u'|_b}{2} - \binom{|u'|_a}{2} - \binom{|u'|_b}{2} - \binom{u'}{ba}. \tag{1}$$

Since $u = g(u')$, we derive that

$$\binom{u}{01} = |u'|_0 + \binom{u'}{00} + \binom{u'}{02} + \binom{u'}{12} + \binom{|u'|_0 + |u'|_1}{2} - \binom{|u'|_0}{2} - \binom{|u'|_1}{2} - \binom{u'}{01},$$

$$\begin{aligned}\binom{u}{12} &= |u'|_0 + \binom{u'}{00} + \binom{u'}{01} + \binom{|u'|_1 + |u'|_2}{2} - \binom{|u'|_1}{2} - \binom{|u'|_2}{2} - \binom{u'}{12} \\ &\quad + \binom{|u'|_0 + |u'|_2}{2} - \binom{|u'|_0}{2} - \binom{|u'|_2}{2} - \binom{u'}{02}.\end{aligned}$$

Hence

$$\lambda_u = 2\left[\binom{u'}{02} - \binom{u'}{01} + \binom{u'}{12} - \binom{|u'|_2}{2}\right] + \binom{|u'|_0 + |u'|_1}{2} - \binom{|u'|_1 + |u'|_2}{2} - \binom{|u'|_0 + |u'|_2}{2}.$$

Similar relations hold for v.

Since u' and v' occur in $\mathbf{x}$, from Observation 1, we get

$$||u'|_0 - |u'|_2| \leq 1 \text{ and } ||v'|_0 - |v'|_2| \leq 1. \tag{2}$$

Since $u \sim_{\mathsf{ab}} v$, we have $|u|_1 = |v|_1$. Hence, from the definition of g, $|u'|_0+|u'|_2 = |v'|_0 + |v'|_2$. In the same way, $|u|_2 = |v|_2$ implies that $|u'|_0 + |u'|_1 = |v'|_0 + |v'|_1$ or equivalently, $|u'|_1 - |v'|_1 = |v'|_0 - |u'|_0$. From the above relation and (2), we get

$$||v'|_0 - |u'|_0 + |u'|_2 - |v'|_2| \leq 2 \text{ and } |u'|_2 - |v'|_2 = |v'|_0 - |u'|_0.$$

Hence the difference of the following two Parikh vectors can only take three values

$$\Psi(u') - \Psi(v') \in \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

To prove that u' and v' are abelian equivalent, we will rule out the last two possibilities.

By assumption, $\lambda_u = \lambda_v$. So this relation also holds modulo 2. Hence

$$\begin{aligned} & \binom{|u'|_0 + |u'|_1}{2} - \binom{|u'|_1 + |u'|_2}{2} - \binom{|u'|_0 + |u'|_2}{2} \\ \equiv \quad & \binom{|v'|_0 + |v'|_1}{2} - \binom{|v'|_1 + |v'|_2}{2} - \binom{|v'|_0 + |v'|_2}{2} \pmod 2. \end{aligned}$$

Assume that we have

$$\Psi(u') - \Psi(v') = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, i.e., \begin{array}{rcl} |u'|_0 + |u'|_1 & = & |v'|_0 + |v'|_1, \\ |u'|_0 + |u'|_2 & = & |v'|_0 + |v'|_2, \\ |u'|_1 + |u'|_2 & = & |v'|_1 + |v'|_2 - 2. \end{array}$$

This leads to a contradiction because then

$$\binom{|u'|_1 + |u'|_2}{2} \not\equiv \binom{|v'|_1 + |v'|_2}{2} \pmod 2.$$

Indeed, it is easily seen that $\binom{4n}{2} \equiv 0 \pmod 2$, $\binom{4n+1}{2} \equiv 0 \pmod 2$, $\binom{4n+2}{2} \equiv 1 \pmod 2$ and $\binom{4n+3}{2} \equiv 1 \pmod 2$.

The case $\Psi(u') - \Psi(v') = \left(\begin{smallmatrix} -1 \\ 1 \\ 1 \end{smallmatrix}\right)$ is handled similarly. So we can assume now that $\Psi(u') = \Psi(v')$, that is, $u' \sim_{\mathsf{ab}} v'$. It remains to prove that $\lambda_{u'} = \lambda_{v'}$. By assumption $\lambda_u = \lambda_v$, and from the above formula describing λ_u (resp. λ_v) we get

$$\binom{u'}{02} - \binom{u'}{01} + \binom{u'}{12} = \binom{v'}{02} - \binom{v'}{01} + \binom{v'}{12}.$$

To conclude that $\lambda_{u'} = \lambda_{v'}$, we should simply show that $\binom{u'}{02} = \binom{v'}{02}$. But $u'v'$ is a factor occurring in $\mathbf{x}$ (from Observation 1, when discarding the 1's in $u'v'$ we just get a word made of alternating 0's and 2) and $u' \sim_{\mathsf{ab}} v'$. This concludes the proof.

Theorem 4. *The word* $\mathbf{x} = g^\omega(0) = 0120210121020120210201 21\cdots$ *avoids* 2-*binomial squares.*

Proof. Assume to the contrary that $\mathbf{x}$ contains a 2-binomial square uv where u and v are 2-binomially equivalent. In particular, u and v are abelian equivalent and moreover $\lambda_u = \lambda_v$. We can therefore apply iteratively Lemma 2 and the above lemma to words of decreasing lengths and get finally a repetition aa with $a \in A$ in $\mathbf{x}$. But $\mathbf{x}$ does not contain any such factor.

Remark 2. The fixed point of g is a 2-binomial-squarefree word, i.e., it does not contain any 2-binomial square, but g is not a 2-binomial-squarefree morphism: the image of a 2-binomial-squarefree word may contain a 2-binomial-square (*e.g.*, $g(010) = 01202012$ contains the square 2020).

3. Avoiding 2-binomial cubes over a 2-letter alphabet

Consider the morphism $h : 0 \mapsto 001$ and $h : 1 \mapsto 011$. A word is 2-binomial-cubefree if it does not contain any 2-binomial cube. In this section, we show that h is a 2-binomial-cubefree morphism: for every 2-binomial-cubefree binary word w, $h(w)$ is 2-binomial-cubefree. As a direct corollary, we get that the fixed point of h,

$$\mathbf{z} = h^\omega(0) = 001001011001001011001011011\cdots$$

avoids 2-binomial cubes.

Let u be a word over $\{0,1\}$. The *extended Parikh vector* of u is

$$\Psi_2(u) = \left(|u|_0, |u|_1, \binom{u}{00}, \binom{u}{01}, \binom{u}{10}, \binom{u}{11}\right)^T.$$

Observe that two words u and v are 2-binomially equivalent if and only if $\Psi_2(u) = \Psi_2(v)$.

Consider the matrix M_h given by

$$M_h = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 4 & 2 & 2 & 1 \\ 2 & 2 & 2 & 4 & 1 & 2 \\ 0 & 0 & 2 & 1 & 4 & 2 \\ 0 & 1 & 1 & 2 & 2 & 4 \end{pmatrix}.$$

One can check that M_h is invertible. We will make use of the following observations:

Proposition 5. *For every* $u \in \{0,1\}^*$,

$$\Psi_2(h(u)) = M_h \Psi_2(u).$$

Proposition 6. *Let $u = 1x$ and $u' = x1$ be two words over $\{0,1\}$. We have $|u|_0 = |u'|_0$, $|u|_1 = |u'|_1$,*

$$\binom{u}{00} = \binom{u'}{00}, \ \binom{u}{11} = \binom{u'}{11}, \ \binom{u'}{01} = \binom{u}{01} + |u|_0, \ \binom{u'}{10} = \binom{u}{10} - |u|_0.$$

In particular, if $1x \sim_2 1y$, then $x1 \sim_2 y1$. Similar relations hold for $0x$ and $x0$. In particular, if $x0 \sim_2 y0$, then $0x \sim_2 0y$.

Let $x, y \in \{0,1\}$. We set $\delta_{x,y} = 1$, if $x = y$; and $\delta_{x,y} = 0$, otherwise.

Lemma 7. *Let p', q' and r' be binary words, and let $a, b \in \{0,1\}$. Let $p = h(p')\,0$, $q = a\,1\,h(q')\,0\,b$ and $r = 1\,h(r')$. Then either $p \not\sim_2 q$ or $p \not\sim_2 r$.*

Proof. Assume, for the sake of contradiction, that $p \sim_2 q \sim_2 r$. Then $|p'| = |q'| + 1 = |r'| = n$. The following relations can mostly be derived from the coefficients of M_h (we also have to take into account the extra suffix 0 of p, respectively the extra prefix 1 in r):

$$\begin{aligned}
\binom{p}{01} &= 2\binom{p'}{0} + 2\binom{p'}{1} + 2\binom{p'}{00} + 4\binom{p'}{01} + \binom{p'}{10} + 2\binom{p'}{11},\\
\binom{p}{10} &= \binom{p'}{0} + 2\binom{p'}{1} + 2\binom{p'}{00} + \binom{p'}{01} + 4\binom{p'}{10} + 2\binom{p'}{11},\\
\Rightarrow \binom{p}{01} - \binom{p}{10} &= \binom{p'}{0} + 3\binom{p'}{01} - 3\binom{p'}{10};\\
\binom{r}{01} &= 2\binom{r'}{0} + 2\binom{r'}{1} + 2\binom{r'}{00} + 4\binom{r'}{01} + \binom{r'}{10} + 2\binom{r'}{11},\\
\binom{r}{10} &= 2\binom{r'}{0} + \binom{r'}{1} + 2\binom{r'}{00} + \binom{r'}{01} + 4\binom{r'}{10} + 2\binom{r'}{11},\\
\Rightarrow \binom{r}{01} - \binom{r}{10} &= \binom{r'}{1} + 3\binom{r'}{01} - 3\binom{r'}{10}.
\end{aligned}$$

We also get the following relations:

$$\begin{aligned}
\binom{q}{01} &= 2\binom{q'}{0} + 2\binom{q'}{1} + 2\binom{q'}{00} + 4\binom{q'}{01} + \binom{q'}{10} + 2\binom{q'}{11}\\
&\quad + \delta_{a,0}\left[1 + \binom{q'}{0} + 2\binom{q'}{1} + \delta_{b,1}\right] + \delta_{b,1}\left[1 + 2\binom{q'}{0} + \binom{q'}{1}\right],
\end{aligned}$$

$$\begin{aligned}
\binom{q}{10} &= 3\binom{q'}{0} + 3\binom{q'}{1} + 2\binom{q'}{00} + \binom{q'}{01} + 4\binom{q'}{10} + 2\binom{q'}{11} + 1\\
&\quad + \delta_{a,1}\left[1 + \delta_{b,0} + 2\binom{q'}{0} + \binom{q'}{1}\right] + \delta_{b,0}\left[1 + \binom{q'}{0} + 2\binom{q'}{1}\right]\\
&= (6 - 2\delta_{a,0} - \delta_{b,1})\binom{q'}{0} + (6 - \delta_{a,0} - 2\delta_{b,1})\binom{q'}{1} + 4 - 2\delta_{a,0} - 2\delta_{b,1} + \delta_{a,0}\delta_{b,1}\\
&\quad + 2\binom{q'}{00} + \binom{q'}{01} + 4\binom{q'}{10} + 2\binom{q'}{11}.
\end{aligned}$$

Where for the last equality, we have used the fact that $\delta_{a,1} = 1 - \delta_{a,0}$ and $\delta_{b,0} = 1 - \delta_{b,1}$. Finally, we obtain

$$\binom{q}{01}-\binom{q}{10} = (-4+3\delta_{a,0}+3\delta_{b,1})\left[\binom{q'}{0}+\binom{q'}{1}\right]+3\binom{q'}{01}-3\binom{q'}{10}-4+3\delta_{a,0}+3\delta_{b,1}.$$

Since $p \sim_2 q \sim_2 r$, we have $\binom{p}{10} - \binom{p}{01} = \binom{q}{10} - \binom{q}{01} = \binom{r}{10} - \binom{r}{01}$. In particular, these equalities modulo 3 give

$$\binom{p'}{0} \equiv \binom{r'}{1} \equiv 2\left[\binom{q'}{0} + \binom{q'}{1} + 1\right] \equiv 2n \pmod 3. \tag{3}$$

Now, we take into account the fact that p and r are abelian equivalent to get a contradiction. Since $p = h(p')\,0$ and $r = 1\,h(r')$, we get

$$\begin{pmatrix}|p|_0\\|p|_1\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}|p'|_0\\|p'|_1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}, \quad \begin{pmatrix}|r|_0\\|r|_1\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}|r'|_0\\|r'|_1\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}.$$

Hence, we obtain

$$\begin{pmatrix}|p|_0 - |r|_0\\|p|_1 - |r|_1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}|p'|_0 - |r'|_0\\|p'|_1 - |r'|_1\end{pmatrix} + \begin{pmatrix}1\\-1\end{pmatrix}.$$

We derive that $|p'|_0 - |r'|_0 = -1$ and $|p'|_1 - |r'|_1 = 1$. Recall that $|p'|_0 + |p'|_1 = n$. If we subtract the last two equalities, we get $|p'|_0 + |r'|_1 = n - 1$. From (3), we know that $|p'|_0 \equiv |r'|_1 \pmod 3$. Hence $2|p'|_0 \equiv n - 1 \pmod 3$ and thus

$$|p'|_0 \equiv 2n - 2 \pmod 3.$$

This contradicts the fact again given by (3) that $|p'|_0 \equiv 2n \pmod 3$.

Similarly, one get the following lemma.

Lemma 8. *Let p', q' and r' be binary words, and let $a, b \in \{0, 1\}$. Let $p = h(p')\,0\,a$, $q = 1\,h(q')\,0$ and $r = b\,1\,h(r')$. Then either $p \not\sim_2 q$ or $p \not\sim_2 r$.*

Proof. Assume, for the sake of contradiction, that $p \sim_2 q \sim_2 r$. Then $|p'| = |q'| = |r'| = n$. Taking into account the special form of p and q, we get

$$\binom{p}{01} = 2\binom{p'}{0}+2\binom{p'}{1}+2\binom{p'}{00}+4\binom{p'}{01}+\binom{p'}{10}+2\binom{p'}{11}+\delta_{a,1}\left(1+2\binom{p'}{0}+\binom{p'}{1}\right),$$

$$\binom{p}{10} = \binom{p'}{0}+2\binom{p'}{1}+2\binom{p'}{00}+\binom{p'}{01}+4\binom{p'}{10}+2\binom{p'}{11}+\delta_{a,0}\left(\binom{p'}{0}+2\binom{p'}{1}\right),$$

$$\binom{q}{01} = 2\binom{q'}{0} + 2\binom{q'}{1} + 2\binom{q'}{00} + 4\binom{q'}{01} + \binom{q'}{10} + 2\binom{q'}{11},$$

$$\binom{q}{10} = 3\binom{q'}{0} + 3\binom{q'}{1} + 2\binom{q'}{00} + \binom{q'}{01} + 4\binom{q'}{10} + 2\binom{q'}{11} + 1.$$

Hence, we get

$$\binom{p}{01} - \binom{p}{10} = -2\binom{p'}{1} + 3\binom{p'}{01} - 3\binom{p'}{10} + \delta_{a,1}\left(1 + 3\binom{p'}{0} + 3\binom{p'}{1}\right),$$

$$\binom{q}{01} - \binom{q}{10} = -\binom{q'}{0} - \binom{q'}{1} + 3\binom{q'}{01} - 3\binom{q'}{10} - 1.$$

Since, $p \sim_2 q$, the last two relations evaluated modulo 3 give

$$|p'|_1 + \delta_{a,1} \equiv 2n + 2 \pmod 3. \tag{4}$$

Similarly, the form of r gives the following relations

$$\binom{r}{01} = 2\binom{r'}{0}+2\binom{r'}{1}+2\binom{r'}{00}+4\binom{r'}{01}+\binom{r'}{10}+2\binom{r'}{11}+\delta_{b,0}\left(1+\binom{r'}{0}+2\binom{r'}{1}\right),$$

$$\binom{r}{10} = 2\binom{r'}{0}+\binom{r'}{1}+2\binom{r'}{00}+\binom{r'}{01}+4\binom{r'}{10}+2\binom{r'}{11}+\delta_{b,1}\left(2\binom{r'}{0}+\binom{r'}{1}\right),$$

$$\binom{r}{01} - \binom{r}{10} = -2\binom{r'}{0} + 3\binom{r'}{01} - 3\binom{r'}{10} + \delta_{b,0}\left(1 + 3\binom{r'}{0} + 3\binom{r'}{1}\right)$$

Since, $p \sim_2 r$, the last two relations evaluated modulo 3 give

$$|p'|_1 + \delta_{a,1} \equiv |r'|_0 + \delta_{b,0} \pmod 3. \tag{5}$$

Now, we take into account the fact that p, q and r are abelian equivalent to get a contradiction. The following two vectors are equal:

$$\begin{pmatrix}|p|_0\\|p|_1\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}|p'|_0\\|p'|_1\end{pmatrix}+\begin{pmatrix}1+\delta_{a,0}\\\delta_{a,1}\end{pmatrix}, \quad \begin{pmatrix}|r|_0\\|r|_1\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}|r'|_0\\|r'|_1\end{pmatrix}+\begin{pmatrix}\delta_{b,0}\\1+\delta_{b,1}\end{pmatrix}.$$

We derive easily that

$$|p'|_1 - |r'|_1 = 1 + \delta_{a,0} - \delta_{b,0}.$$

On the one hand, using the latter relation and (5)

$$|r'|_1 + 1 + \delta_{a,0} - \delta_{b,0} + \delta_{a,1} = |p'|_1 + \delta_{a,1} \equiv |r'|_0 + \delta_{b,0} \pmod 3$$

Replacing $|r'|_0$ by $n - |r'|_1$, we get $2|r'|_1 + 2 \equiv n + 2\delta_{b,0} \pmod 3$, or equivalently

$$|r'|_1 + 1 \equiv 2n + \delta_{b,0} \pmod 3.$$

On the other hand, using (4),

$$|r'|_1 + 1 + \delta_{a,0} - \delta_{b,0} + \delta_{a,1} = |p'|_1 + \delta_{a,1} \equiv 2n + 2 \pmod 3$$

and thus,

$$|r'|_1 \equiv 2n + \delta_{b,0} \pmod 3.$$

We get a contradiction, $2n + \delta_{b,0}$ should be congruent to both $|r'|_1$ and $|r'|_1 + 1$ modulo 3.

We are ready to prove the main theorem of this section.

Theorem 9. *Let* $h : 0 \mapsto 001, 1 \mapsto 011$*. For every* 2*-binomial-cubefree word* $w \in \{0,1\}^*$*,* $h(w)$ *is* 2*-binomial-cubefree.*

PROOF. Let w be a 2-binomial-cubefree binary word. Assume that $h(w) = z_0 \dots z_{3|w|-1}$ contains a 2-binomial cube pqr occurring in position i, *i.e.*, $p \sim_2 q \sim_2 r$ and $w = w'\, p\, q\, r\, w''$, where $|w'| = i$. We consider three cases depending on the size of p modulo 3.

As a first case, assume that $|p| = 3n$. We consider three sub-cases depending on the position i modulo 3.

1.a) Assume that $i \equiv 2 \pmod 3$. Then p, q, r have 1 as a prefix and the letter following r in $h(w)$ is the symbol $z_{i+9n} = 1$. Hence, the word $1^{-1}pqr1$ occurs in $h(w)$ in position $i+1$ and it is again a 2-binomial cube. Indeed, thanks to Proposition 6, we have $1^{-1}p1 \sim_2 1^{-1}q1 \sim_2 1^{-1}r1$. This case is thus reduced to the case where $i \equiv 0 \pmod 3$.

1.b) Assume that $i \equiv 1 \pmod 3$. Then p, q, r have 0 as a suffix and the letter preceding p in $h(w)$ is the symbol $z_{i-1} = 0$. Hence, the word $0pqr0^{-1}$ occurs in $h(w)$ in position $i - 1$ and it is also a 2-binomial cube. Thanks to Proposition 6, we have $0p0^{-1} \sim_2 0q0^{-1} \sim_2 0r0^{-1}$. Again this case is reduced to the case where $i \equiv 0 \pmod 3$.

1.c) Assume that $i \equiv 0 \pmod 3$. In this case, we can desubstitute: there exist three words p', q', r' of length n such that $h(p') = p$, $h(q') = q$, $h(r') = r$ and $p'q'r'$ is a factor occurring in w. We have $\Psi_2(p) = \Psi_2(q) = \Psi_2(r)$. By Proposition 5, and since M_h is invertible, we have $\Psi_2(p') = \Psi_2(q') = \Psi_2(r')$, meaning that w contains a 2-binomial cube $p'q'r'$.

As a second case, assume that $|p| = 3n + 1$. In this case, one of p, q and r occur in position 0 modulo 3, one in position 1 modulo 3, and one in position 2 modulo 3. Suppose w.l.o.g. that p occur in position 0 modulo 3, and q in position 1 modulo 3. Then there are three factors p', q' and r' in w, and $a, b \in \{0,1\}$ such that $p = h(p')\,0$, $q = a\,1\,h(q')\,0\,b$ and $r = 1\,h(r')$. By Lemma 7, this is impossible.

For the final case, assume that $|p| = 3n + 2$. In this case again, one of p, q and r occur in position 0 modulo 3, one in position 1 modulo 3, and one in position 2 modulo 3. Suppose w.l.o.g. that p occur in position 0 modulo 3, and q in position 1 modulo 3. Then there are three factors p', q' and r' in w, and $a, b \in \{0,1\}$ such that $p = h(p')\,0\,a$, $q = 1\,h(q')\,0$ and $r = b\,1\,h(r')$. By Lemma 8, this is impossible.

Corollary 10. *The infinite word* $\mathbf{z} = 001001011\cdots$ *fixed point of* $h : 0 \mapsto 001, 1 \mapsto 011$ *avoids* 2*-binomial cubes.*

4. Open problems

There are sufficient conditions for a morphism to be abelian or k-abelian-powerfree [1, 2, 11]. It seems more difficult to find sufficient conditions for a morphism to be k-binomial-powerfree. One can raise the following question.

Question 1. *Give a non-trivial* 2*-binomial-squarefree morphism (that is, a morphism* f *such that* $f(w)$ *is* 2*-binomial-squarefree if* w *is* 2*-binomial-squarefree).*

As shown in Section 3, the morphism $0 \mapsto 001, 1 \mapsto 011$ is 2-binomial-cubefree.

Every infinite binary word contains arbitrarily long abelian squares, while ones exist which avoid squares of period at least 3 [4]. Moreover, it is possible to construct a binary word with only 3 squares: 00, 11 and 0101 [5]. A computer experiment shows the following.

Fact 1. *It is impossible to construct an infinite binary word with only* 3 *different* 2*-binomial-squares.*

One can then ask the following questions.

Question 2. *Is there* $k \geq 3$ *such that one can construct an infinite binary word with only* 3 *different* k*-binomial-squares ?*

Question 3. *Is there* $k \geq 4$ *such that one can construct an infinite binary word with only* k *different* 2*-binomial-squares ?*

References

[1] A. Carpi, On Abelian Power-Free Morphisms, *Int. J. Algebra Comput.* **3** (1993), 151–167.

[2] F. M. Dekking, Strongly nonrepetitive sequences and progression-free sets, *J. Combin. Theory Ser. A* **27** (1979), 181-185.

[3] P. Erdös, Some unsolved problems, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** (1961), 221–254.

[4] R.C Entringer, D.E Jackson, J.A Schatz. On nonrepetitive sequences. *J. Combin. Theory Ser. A* **16**, (1974), 159–164.

[5] A. S. Fraenkel, R. J. Simpson, How Many Squares Must a Binary Sequence Contain?, *The Electronic Journal of Combinatorics*, **2**, (1995).

[6] M. Huova, J. Karhumäki, Observations and problems on k-abelian avoidability, In Combinatorial and Algorithmic Aspects of Sequence Processing (Dagstuhl Seminar 11081), (2011) 2215-2219.

[7] M. Huova, J. Karhumäki, On unavoidability of k-abelian squares in pure morphic words, *J. Integer Seq.* **16** (2013), no. 2, Article 13.2.9.

[8] J. Karhumäki, A. Saarela, L. Q. Zamboni, On a generalization of Abelian equivalence and complexity of infinite words, *J. Combin. Theory Ser. A* **120** (2013), 2189–2206.

[9] V. Keränen, Abelian squares are avoidable on 4 letters, *Lecture Notes in Comput. Sci.* **623** (1992), 41-52.

[10] M. Lothaire, *Combinatorics on Words*, Cambridge Mathematical Library, Cambridge University Press, (1997).

[11] M. Rao, On some generalizations of abelian power avoidability, preprint (2013).

[12] M. Rigo, P. Salimov, Another generalization of abelian equivalence: binomial complexity of infinite words, WORDS 2013, *Lect. Notes in Comput. Sci.* **8079**, 217–228, Springer Verlag (2013), J. Karhumäki, L. Zamboni (Eds.).

[13] A. Thue, Über unendliche Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **7** (1906), 1-22.

[14] A. Thue,Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912), 1-67.