

# Repetitions in Toeplitz Words and the Thue Threshold

Antonio Boccuto<sup>®</sup> and Arturo Carpi<sup>(⊠)</sup><sup>®</sup>

Dipartimento di Matematica e Informatica, University of Perugia, via Vanvitelli 1, 06123 Perugia, Italy {antonio.boccuto,arturo.carpi}@unipg.it

**Abstract.** A (finite or infinite) word is said to be k-th power-free if it does not contain k consecutive equal blocks. A colouring of the integer lattice points in the n-dimensional Euclidean space is power-free if there exists a positive integer k such that the sequence of colours of consecutive points on any straight line is a k-th power-free word. The Thue threshold of  $\mathbb{Z}^n$  is the least number of colours t(n) allowing a power-free colouring of the integer lattice points in the n-dimensional Euclidean space.

Answering a question of Grytczuk (2008), we prove that t(2) = t(3) = 2. Moreover, we show the existence of a 2-colouring of the integer lattice points in the Euclidean plane such that the sequence of colours of consecutive points on any straight line does not contain squares of length larger than 26.

In order to obtain these results, we study repetitions in Toeplitz words. We show that the Toeplitz word generated by any sequence of primitive partial words of maximal length k is k-th power-free. Moreover, adding a suitable hypothesis on the positions of the holes in the generating sequence, we obtain that also the subwords occurring in the considered Toeplitz word according to an arithmetic progression of suitable difference, are k-th power-free words.

**Keywords:** Power-free word  $\cdot$  Toeplitz word  $\cdot$  Partial word  $\cdot$  Word with bounded square  $\cdot$  Thue threshold  $\cdot$  Arithmetic subword

# 1 Introduction

The study of repetitions in words has been one of the main fields of interest in Combinatorics on Words since its origins [6,20,21]. This subject has several applications in other fields such as Algebra, Symbolic Dynamics, Game Theory.

Let k be a positive integer. We recall that a word is said to be k-th powerfree if it does not contain k consecutive equal blocks. For instance, the word "barbarian" is not square-free, since it contains two consecutive occurrences of the block "bar", while it is cube-free. It is known that there exist cube-free infinite words over a two-letter alphabet and square-free infinite words over a three-letter alphabet [20]. In [8], one of the authors considered the following multidimensional extension of k-th power-freeness, suggested by J. Berstel. Consider a colouring of the integer lattice points in the n-dimensional Euclidean space, with finitely many colours. Such a colouring will be called k-th power-free if the sequence of colours of consecutive points on any straight line is an infinite k-th power-free word. The main result of [8] is the existence of a square-free colouring of the integer lattice points in the n-dimensional Euclidean space, for all positive integer n.

The minimum number of colours C(n) required for such a colouring is actually unknown. While the construction of [8] uses  $4^n$  colours, recently [16] it has been shown that  $C(n) \ge 9 \cdot 2^{n-2}$ , for all  $n \ge 2$ . Kao *et al.* [15] used a smaller number of colours to build k-th power-free colourings of the integer lattice points of the plane, with k > 2. In this context, an interesting notion is the Thue threshold, introduced by Grytczuk [13]. A colouring of the integer lattice points in the n-dimensional Euclidean space is power-free if it is k-th power-free for some positive k. The Thue threshold of  $\mathbb{Z}^n$  is the minimum number of colours t(n) needed for such a colouring. One of the problems proposed by Grytczuk asks for the value of t(n), at least in the case n = 2. The existence of infinite cube-free words on a binary alphabet shows that t(1) = 2. The results of [15] quoted above imply, in particular, that  $t(2) \leq 4$  and, more generally,  $t(n) \leq 2^n$ . On the other side, it is clear that  $t(n) \ge 2$  for all n. In this paper, we will show that t(2) = t(3) = 2. More precisely, we construct a 9-th power-free 2-colouring of the 3-dimensional Euclidean space and an 8-th power-free 2-colouring of the Euclidean plane. In view of these results, we conjecture that t(n) = 2 for all positive n.

Another notion very close to power-freeness is that of bounded repetition. An infinite word is said to have bounded repetition if the length of the squares occurring in it is upperbounded by a constant. An infinite word with bounded repetition over a binary alphabet has been given in [10]. Further examples can be found, for instance, in [1,11,19]. Similarly to the Thue threshold, one can introduce the bounded repetition threshold as follows. A colouring of the integer lattice points in the *n*-dimensional Euclidean space is said to have bounded repetition if the length of the squares occurring in it is upperbounded by a constant. The minimum number of colours needed for such a colouring will be called the bounded repetition threshold of  $\mathbb{Z}^n$ . We will show that the bounded repetition threshold of  $\mathbb{Z}^2$  is 2.

Now we describe some of the tools used for the construction of the power-free colourings considered above. According to [4], we call arithmetic subsequence (of difference d) of an infinite word the subsequence obtained by extracting the letters which are positioned according to an arithmetic progression of difference d. The study of infinite words whose arithmetic subsequences, for some prescribed difference, are power-free has been started in [8]. Further results on this subject can be found in [9,15]. In [8] it is shown that there exists an infinite word on a 4-letter alphabet such that all its arithmetic subsequences of odd difference are square-free. As noticed in [15] this word is related to paperfolding words. There is a large literature on paperfolding words and on the more general class of Toeplitz words (see, e.g., [1,3,14,18,19] and the references therein). Avgusti-

novich *et al.* [4] showed that the factors of arithmetic subsequences of odd difference of paperfolding words are themselves factors of paperfolding words. Since all paperfolding words are 4-th power-free [1], one obtains that arithmetic subsequences of odd difference of paperfolding words are 4-th power-free words.

In this paper we study in more details power-freeness of Toeplitz words and of their arithmetic subwords, obtaining some extension of the results above. In order to describe our results, we recall some further notion. Partial words were introduced by Berstel and Boasson [5] in order to study some extension of the Fine and Wilf Periodicity Theorem. Roughly speaking, a partial word is a word such that the letters in some positions (called holes) are unknown. Primitive partial words have been studied by Blanchet-Sadri [7]. Given a sequence of partial words  $s^{(n)}$ ,  $n \ge 1$ , one can construct an infinite word as follows: first concatenate infinitely many copies of  $s^{(1)}$ , thus obtaining an infinite word with holes; next, ordinately replace the holes by the letters of the infinite word obtained by concatenating infinitely many copies of  $s^{(2)}$ , thus obtaining a new infinite word with holes; and so on... If there are infinitely many n such that  $s^{(n)}$  does not start with a hole, then the previous construction converges to an infinite word with no hole, which is called the Toeplitz word generated by the sequence  $s^{(n)}$ . We will show that, if the words  $s^{(n)}$ ,  $n \ge 1$ , are primitive partial words of maximal length k, then the generated Toeplitz word is k-th power-free. Concerning the arithmetic subsequences of Toeplitz words, we prove a useful structural property. If the positions of the holes in the words of the generating sequence of a Toeplitz word U satisfy a suitable hypothesis and d is an integer relatively prime with their lengths, then any arithmetic subsequence of difference d of Uis a Toeplitz word and its generating sequence can be obtained by rearranging the letters of the generating sequence according to a particular rule. We notice that some results related to this one can be found in [4, 12].

As a consequence of the two previous results, we obtain that if U is a Toeplitz word generated by a sequence of primitive partial words of maximal length ksatisfying the condition on the position of the holes considered above, if q is the least common multiple of the lengths of the generating words, and if d is any integer which is not a multiple of q, then all arithmetic subsequences of Uof difference d are k-th power-free. This last result gives us a powerful tool to construct words with power-free arithmetic subsequences.

The paper is organized as follows. In the next section we recall some basic definitions and properties needed in the sequel. In Sect. 3 we establish the result on power-freeness of Toeplitz words generated by primitive partial words. Arithmetic subsequences of Toeplitz words are studied in Sect. 4. In Sect. 5 we prove that the Thue threshold of  $\mathbb{Z}^3$  is 2 and in Sect. 6 we prove that the bounded square threshold of  $\mathbb{Z}^2$  is 2. Finally, in Sect. 7 we discuss some open problems.

#### 2 Preliminaries

Let A be a finite nonempty set, or *alphabet*, and  $A^*$  be the free monoid generated by A. The elements of A are usually called *letters* and those of  $A^*$  words. The identity element of  $A^*$  is called *empty word* and denoted by  $\varepsilon$ . We set  $A^+ = A^* \setminus \{\varepsilon\}$ . A word  $w \in A^+$  can be written uniquely as a sequence of letters as  $w = w_1 w_2 \cdots w_n$ , with  $w_i \in A$ ,  $1 \leq i \leq n$ , n > 0. The integer n is called the *length* of w and denoted by |w|. All words  $w_i w_{i+1} \cdots w_j$  with  $1 \leq i \leq j \leq n$  are called *factors* of w. A *period* of the word w is any positive integer p such that  $w_i = w_{i+p}$  for all  $i = 1, 2, \ldots, n-p$ . For all  $n \geq 0$ , the set of all words of length n on the alphabet A is denoted by  $A^n$ .

Let k be a positive integer. Any word of the form  $u^k$ , with  $u \neq \epsilon$ , is called a k-th power. In particular, 2-nd and 3-rd powers are usually called squares and cubes. As is known, a word w is a k-th power if and only if w has a period p such that |w| = kp. A word is primitive if it is not a k-th power, for all  $k \geq 2$ .

An infinite word U on the alphabet A is any unending sequence of letters. For all  $n \geq 1$ , we let  $U_n$  denote the *n*-th letter of the infinite word U. Thus,  $U = U_1 U_2 \cdots U_n \cdots$ . The set of infinite words on the alphabet A is denoted by  $A^{\omega}$ . A bi-infinite word on the alphabet A is any map  $V : \mathbb{Z} \to A$ , where  $\mathbb{Z}$  denotes the semiring of relative integers. The image of any  $n \in \mathbb{Z}$  by V is denoted by  $V_n$ . The factors of an infinite (resp., bi-infinite) word U are the words  $U_i U_{i+1} \cdots U_j$ with  $1 \leq i \leq j$  (resp.,  $i, j \in \mathbb{Z}$  and  $i \leq j$ ). A period of an infinite (resp., biinfinite) word U is any positive integer p such that  $U_i = U_{i+p}$  for all  $i \geq 1$ (resp.,  $i \in \mathbb{Z}$ ). If s is a finite word, we let  $s^{\omega}$  denote the infinite word obtained by concatenating infinitely many copies of s.

A (finite or infinite or bi-infinite) word is said to be k-th power-free if none of its factors is a k-th power.

Let U be an infinite word, and i and d be two positive integers. The infinite word  $V = U_i U_{i+d} U_{i+2d} \cdots U_{i+nd} \cdots$  is called an *arithmetic subsequence* of U of *difference* d. Any factor of any arithmetic subsequence of U of difference d is said to be an *arithmetic subword* of U of *difference* d. Arithmetic subsequences and subwords of bi-infinite words can be defined similarly.

Let A be a k-letter alphabet and n be a positive integer. Any map  $\alpha : \mathbb{Z}^n \to A$ will be called a k-colouring of the lattice points of the n-dimensional space, or, briefly, a  $\mathbb{Z}^n$ -word. A bi-infinite word V is a line of  $\alpha$  if there exist integers  $j_1, \ldots, j_n, m_1, \ldots, m_n \in \mathbb{Z}$  such that  $gcd(m_1, \ldots, m_n) = 1$  and

$$V_q = \alpha(j_1 + qm_1, \dots, j_n + qm_n)$$
 for all  $q \in \mathbb{Z}$ .

This definition [8] is motivated by the fact that, as one can easily verify, the lines of  $\alpha$  are the sequences of letters corresponding to the integer lattice points of the *n*-dimensional space  $\mathbb{R}^n$  which lie on a same straight line. For instance, if  $\alpha$  is the 'chessboard coloring' of the plane, that is the 2-coloring of  $\mathbb{Z}^2$  defined by  $\alpha(x, y) = x + y \mod 2$ , then its lines are the bi-infinite words

$$\cdots 01010101 \cdots$$
,  $\cdots 00000000 \cdots$  and  $\cdots 111111111 \cdots$ 

Indeed, the first one corresponds to all straight lines with directive numbers  $(m_1, m_2)$  with  $m_1 + m_2$  odd, while the other two correspond to the straight lines with directive numbers  $(m_1, m_2)$  where both  $m_1$  and  $m_2$  are odd and, consequently,  $m_1 + m_2$  is even.

# 3 Repetitions in Toeplitz Words

Partial words were introduced by Berstel and Boasson [5]. We call partial word over the alphabet A any word over the alphabet  $A \cup \{?\}$ , where ? is a distinguished letter, not belonging to A. The occurrences of ? in a partial word are usually called *holes*. Let x and y be partial words. We say that x is *contained* in y and we write  $x \subset y$ , if the partial word y can be obtained from x replacing some hole by letters of A. A partial word x is *primitive* if there do not exist another partial word z and an integer  $n \ge 2$  such that  $x \subset z^n$ . For instance, if x = ab??ab?aand y = ab?bab?a, then one has  $x \subset y$ . The word x is not primitive, since  $x \subset (abaa)^2$ , while the word y is primitive.

The following proposition is a slight modification of a result of [5]. The proof is identical.

**Proposition 1.** Let x and y be partial words. One has  $x^k \subset y^{\ell}$  for some integers  $k, \ell$  if and only if  $x \subset z^n$  and  $z^m \subset y$  for some partial word z and integers n, m.

The following is a straightforward consequence of the previous proposition.

**Corollary 1.** Let x be a primitive partial word. If one has  $x^k \subset y^{\ell}$  for some partial word y and integers  $k, \ell$ , then  $x^m \subset y$  for some integer m.

The following proposition [7] will be useful in the sequel.

**Proposition 2.** Let x and y be partial words. If xy is primitive, then yx is primitive.

Now, we recall the notion of Toeplitz word [14]. Let U be an infinite word over the alphabet  $A \cup \{?\}$  containing infinitely many holes and s be a partial word over the alphabet  $A \cup \{?\}$ . We let  $T_s(w)$  denote the infinite word obtained by ordinately replacing in U the holes by the letters of  $s^{\omega}$ . For instance, if  $U = (aa?)^{\omega}$  and s = b?c?, then  $T_s(U) = (aabaa?aacaa?)^{\omega}$ .

Now, let  $s^{(n)}$ ,  $n \ge 1$  be a sequence of non-empty partial words over the alphabet A. We suppose that each word  $s^{(n)}$  contains at least one hole and that there are infinitely many n such that the initial letter of  $s^{(n)}$  is not a hole. We define a sequence of periodic words as follows:

$$U^{(0)} = ?^{\omega}, \quad U^{(n)} = T_{s^{(n)}}(U^{(n-1)}), \ n \ge 1.$$
 (1)

Our assumption on the initial letters of  $s^{(n)}$  ensures that the sequence  $U^{(n)}$  converges to an infinite word  $U \in A^{\omega}$ . It will be called the *Toeplitz word* generated by the sequence  $s^{(n)}$ . The following lemma, whose proof is left to the reader, will be useful in the sequel.

**Lemma 1.** Let U be the Toeplitz word generated by a sequence  $s^{(n)}$ ,  $n \ge 1$  and V be the Toeplitz word generated by the sequence  $s^{(n)}$ ,  $n \ge 2$ . Then U can be obtained by ordinately replacing in  $(s^{(1)})^{\omega}$  the holes by the letters of V.

Now, we are ready to state the main result of this section.

**Theorem 1.** Let U be the Toeplitz word generated by a sequence  $s^{(n)}$ ,  $n \ge 1$ . If all  $s^{(n)}$  are primitive partial words of maximal length k, then U is k-th power-free.

*Proof.* We suppose, by contradiction, that U contains a k-th power  $u^k$ , with  $u \in A^+$ . With no loss of generality, we assume that  $u^k$  is the shortest k-th power occurring in any Toeplitz word satisfying the statement.

The proof requires several steps. The first one consists in showing that |u| is a multiple of  $|s^{(1)}|$ .

Indeed, set  $s = s^{(1)}$ , p = |s|, q = |u|. By the construction of U, there is a factor v of  $U^{(1)} = s^{\omega}$  such that  $v \subset u^p$ . Since |v| = pq, one derives that s = xy and  $v = (yx)^q$  for some partial words x, y. Hence, in view of Proposition 2 and Corollary 1, one has  $(yx)^m \subset u$ , for a suitable integer m. Replacing the words by their lengths, one has mp = q, which proves our claim.

As a second step, we show that the word V generated by the sequence  $s^{(n)}$ ,  $n \ge 2$ , also contains a k-th power.

Let *h* be the integer such that  $u^k = U_h U_{h+1} \cdots U_{h+kq-1}$  and  $i_1, i_2, \ldots, i_t$  be the positions of the holes of  $s^{\omega}$  in the interval [h, h+kq-1]. In view of Lemma 1, the word  $w = U_{i_1} U_{i_2} \cdots U_{i_t}$  is a factor of *V*. We shall verify that *w* is a *k*-th power. Indeed, let  $\ell$  denote the number of holes in  $s^m$ . Then each factor of  $s^{\omega}$ of length q = mp contains exactly  $\ell$  holes. One derives, in particular,  $t = k\ell$ . Moreover, one has

$$i_{j+\ell} = i_j + q, \quad j = 1, 2, \dots, t - q.$$

Taking into account that q is a period of the word  $U_h U_{h+1} \cdots U_{h+kq-1} = u^k$ , one obtains  $U_{i_j} = U_{i_{j+\ell}}, j = 1, 2, \ldots, t-q$ . This proves that w has period  $\ell$ . Since  $|w| = t = k\ell$ , we conclude that  $w = z^k$ , for some word z of length  $\ell$ .

By the minimality of u, one derives  $\ell \ge q = |s^m|$ . But s can contain at most p-1 holes, so that  $\ell \le m(p-1) < q$ . This leads to a contradiction.  $\Box$ 

*Example 1.* A Toeplitz word generated by a sequence of words  $s^{(n)} \in \{0?1?, 1?0?\}$ ,  $n \geq 1$ , is called a *paperfolding word*. By Theorem 1, one derives the known fact [2] that paperfolding words are 4-th power-free.

Let  $s^{(n)} = 01$ ? for all  $n \ge 1$ . The Toeplitz word generated by the sequence  $s^{(n)}$  is cube-free. More generally, if p > 2 and  $r^{(n)} = 0^{p-2}1$ ? for all  $n \ge 1$ , then the Toeplitz word generated by the sequence  $r^{(n)}$  is p-th power-free.

#### 4 Arithmetic Subsequences of Toeplitz Words

A permutation  $\sigma$  of the set  $\{1, 2, ..., n\}$  will be called *arithmetic* if there exists an integer d such that

$$\sigma(i+1) \equiv \sigma(i) + d \pmod{n}, \quad i = 1, 2, \dots, n-1.$$

One easily verifies that, in such a case, one has  $\sigma(j) - \sigma(i) \equiv (j - i)d \pmod{n}$ , for all  $i, j \in \{1, 2, ..., n\}$ . Moreover, the composition of arithmetic permutations

is an arithmetic permutation. Thus, the set of all arithmetic permutations is a subgroup of the symmetric group on n objects.

Now, let x and y be two words of length n. We say that x and y are arithmetically conjugate and we write  $x \approx y$  if there exists an arithmetic permutation  $\sigma$  of  $\{1, 2, \ldots, n\}$  such that

$$x = x_1 x_2 \cdots x_n, \quad y = x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(n)}, \tag{2}$$

with  $x_1, x_2, \ldots, x_n$  letters. Roughly speaking, x and y are arithmetically conjugate if y is obtained by rearranging the letters of x according to an arithmetic permutation. Taking into account that by composing (or inverting) arithmetic permutations one obtains an arithmetic permutation, one easily checks that the relation  $\approx$  is an equivalence relation.

The following statement extends Proposition 2 to the case of arithmetic conjugacy. For the sake of brevity, the proof is omitted .

**Proposition 3.** Let  $x, y \in A^*$  be arithmetically conjugate partial words. Then x is primitive if and only if y is primitive.

The following lemma shows a useful application of arithmetic conjugacy to the analysis of arithmetic subsequences of periodic words.

**Lemma 2.** Let s be a word of length p and V be an arithmetic subsequence of difference d of the word  $s^{\omega}$ . If d and p are coprime, then there exists a word r such that

$$r \approx s$$
 and  $V = r^{\omega}$ .

We will study Toeplitz words with a generating sequence whose elements lie in the set

$$P = \bigcup_{0 \le i \le \ell} A^{i}?(A^{\ell}?)^* A^{\ell-i}.$$
 (3)

Arithmetic subwords of Toeplitz words of this type have been studied in [4, 12] as they are prototypes of uniformly recurrent words with linear arithmetic complexity. In this section, we will investigate in more details the relationship between a Toeplitz word with a generating sequence whose elements are in P and its arithmetic subsequences.

Let U be an infinite word over the alphabet  $A \cup \{?\}$ . The sequence of the integers i such that  $U_i = ?$ , in increasing order, will be called the *hole sequence* of U. One easily verifies that one has  $s \in P$  if and only if the hole sequence of  $s^{\omega}$  is an arithmetic progression whose difference divides  $|s^{(n)}|$ . The following lemma shows that, in the case of our interest, the hole sequences of the words approximating a Toeplitz word are, in fact, arithmetic progressions.

**Lemma 3.** For all  $n \ge 1$ , let  $s^{(n)} \in P$ . Then, for all  $n \ge 0$ , the hole sequence of the infinite word  $U^{(n)}$  defined by (1) is an arithmetic progression whose difference divides the number  $\prod_{i=1}^{n} |s^{(i)}|$ .

Now we are ready to prove the main result of this section.

**Theorem 2.** Let U be the Toeplitz word generated by a sequence of words  $s^{(n)} \in P$ ,  $n \ge 1$  and V be an arithmetic subsequence of U of difference d. If one has  $gcd(d, |s^{(n)}|) = 1$  for all  $n \ge 1$ , then V is a Toeplitz word with a generating sequence  $r^{(n)}$  such that for all  $n \ge 1$ ,  $r^{(n)} \approx s^{(n)}$ .

*Proof.* Let  $(i_m)_{m>1}$  be the arithmetic progression of difference d such that

$$V = U_{i_1} U_{i_2} \cdots U_{i_m} \cdots$$

Moreover, for all  $n \geq 0$ , let  $U^{(n)}$  be the infinite word defined by (1) and  $V^{(n)}$  be the arithmetic subsequence  $V^{(n)} = U_{i_1}^{(n)} U_{i_2}^{(n)} \cdots U_{i_m}^{(n)} \cdots$ . Fix  $n \geq 1$ , let  $(h_m)_{m\geq 1}$  be the hole sequence of  $U^{(n-1)}$  and  $(h_{j_m})_{m\geq 1}$  be the subsequence of the elements occurring also in the arithmetic progression  $(i_m)_{m\geq 1}$ . In view of Lemma 3,  $(h_m)_{m\geq 1}$  is an arithmetic progression whose difference  $\ell$  is coprime with d. From this fact, one easily derives that  $(h_{j_m})_{m\geq 1}$  is an arithmetic progression of difference  $d\ell$ . This implies that  $(j_m)_{m\geq 1}$  is an arithmetic progression of difference d.

Let W be the arithmetic subsequence of  $(s^{(n)})^{\omega}$  obtained by taking the letters of position  $j_m$ ,  $m \geq 1$ . By Lemma 2, there is a partial word  $r^{(n)}$  such that  $r^{(n)} \approx s^{(n)}$  and  $W = (r^{(n)})^{\omega}$ . Since the word  $U^{(n)}$  is obtained by ordinately replacing the holes of  $U^{(n-1)}$  by the letters of  $(s^{(n)})^{\omega}$ , the word  $V^{(n)}$  will be obtained by ordinately replacing the holes of  $V^{(n-1)}$  by the letters of W, that is  $V^{(n)} = T_{r^{(n)}}(V^{(n-1)})$ .

One easily verifies that the sequence  $V^{(n)}$  converges to V, so that V is the Toeplitz word with generating sequence  $r^{(n)}$ ,  $n \ge 1$ .

From Theorems 1 and 2 and Proposition 3, one easily derives the following

**Corollary 2.** Let U be the Toeplitz word generated by a sequence of primitive partial words  $s^{(n)} \in P$ ,  $n \ge 1$ . If  $k = \max\{|s^{(n)}|, n \ge 1\}$ , then every arithmetic subsequence of U whose difference d is coprime with  $\operatorname{lcm}\{|s^{(n)}|, n \ge 1\}$  is a k-th power-free infinite word.

*Example 2.* From the corollary above, one obtains that all arithmetic subsequences of odd difference of a paperfolding word are 4-th power-free. However, this fact has been directly proved in [15] using a result of [4].

Let U be the Toeplitz word generated by the (constant) sequence  $s^{(n)} = 01$ ?,  $n \ge 1$ . By the corollary above, all arithmetic subsequences of U whose difference is not a multiple of 3 are cube-free. More generally, if p > 2 and V is the Toeplitz word generated by the (constant) sequence  $s^{(n)} = 0^{p-2}1$ ?,  $n \ge 1$ , then all arithmetic subsequences of U whose difference is not a multiple of p are p-th power-free.

## 5 Thue Threshold

In this section we shall prove that the 3-dimensional Thue threshold is equal to 2. More precisely, we will exhibit a 2-colouring of the lattice points of the 3dimensional space such that any line is a 9-th power-free word. As a consequence, also the 2-dimensional Thue threshold is equal to 2.

We let  $\mathbb{Z}_3$  denote the field of the integers modulo 3. The following lemma gives a useful combinatorial property of the vector space  $\mathbb{Z}_3^3$ .

**Lemma 4.** There is a partition  $(A_1, A_2, A_3)$  of  $\mathbb{Z}_3^3$  such that for all  $j, m \in \mathbb{Z}_3^3$  with  $m \neq 0$ , there exist  $k \in \mathbb{Z}_3$  and  $i \in \{1, 2, 3\}$  such that

$$j + km \in A_i$$
,  $m_i \neq 0$ .

We limit ourselves to list the elements of the classes  $A_1, A_2, A_3$ , as the verification of Lemma 4 merely requires a finite but long and tedious check.

$$A_1 = \{ \boldsymbol{x} + (k, k, k) \mid \boldsymbol{x} = (0, 0, 0), (0, 1, 0), (0, 0, 1), k = 0, 1, 2 \}$$

$$A_2 = \{ (i_3 + 1, i_1, i_2) \mid (i_1, i_2, i_3) \in A_1 \}, \quad A_3 = \{ (i_3 + 1, i_1, i_2) \mid (i_1, i_2, i_3) \in A_2 \}.$$

Remark 1. The previous lemma has an interesting interpretation in terms of Galois geometry: let AG(3,3) be the 3-dimensional affine space over the 3element field  $\mathbb{Z}_3$ . Then there is a partition  $(A_1, A_2, A_3)$  of the points of AG(3,3) with the following property: for every line r there is an axis  $x_i$ , i = 1, 2, 3 such that r intersects  $A_i$  and is not orthogonal to  $x_i$ . We notice that an analogous property holds for the 2-dimensional affine space over the 2-element field  $\mathbb{Z}_2$ , taking the partition  $(A_1, A_2)$  with  $A_1 = \{(0, 0), (1, 1)\}$  and  $A_2 = \{(0, 1), (1, 0)\}$ .

In Example 2, we have shown an infinite binary word such that all arithmetic subsequences whose difference is not a multiple of 3 are cube-free. A bi-infinite word with the same property can be easily obtained via Koenig's Lemma (see, e.g., [17]). Now, we define a 2-colouring of the lattice points of the 3-dimensional space whose lines are 9-th power-free. Let  $(A_1, A_2, A_3)$  be the partition of  $\mathbb{Z}_3^3$ given by Lemma 4 and let U be a binary bi-infinite word such that all arithmetic subsequences whose difference is not a multiple of 3 are cube-free. We let  $\pi$ denote the natural projection  $\pi: \mathbb{Z}^3 \to \mathbb{Z}_3^3$  and define the map  $\alpha: \mathbb{Z}^3 \to \{0, 1\}$ as follows:

if 
$$\pi(j_1, j_2, j_3) \in A_i$$
, then  $\alpha(j_1, j_2, j_3) = U_{\lfloor j_i/3 \rfloor}$ ,

 $j_1, j_2, j_3 \in \mathbb{Z}, i \in \{1, 2, 3\}.$ 

**Proposition 4.** All lines of  $\alpha$  are 9-th power-free.

*Proof.* Actually, it suffices to verify that no line of  $\alpha$  contains a cube  $u^3$  with |u| multiple of 3. Indeed, any 9-th power  $v^9$  can be written as  $v^9 = u^3$ , with  $u = v^3$  and |u| = 3|v|.

Let V be a line of  $\alpha$ . Then, there are  $j_1, j_2, j_3, m_1, m_2, m_3 \in \mathbb{Z}$  such that  $gcd(m_1, m_2, m_3) = 1$  and

$$V_n = \alpha(j_1 + nm_1, j_2 + nm_2, j_3 + nm_3) \text{ for all } n \in \mathbb{Z}.$$
 (4)

By Lemma 4, there are  $i \in \{1, 2, 3\}$  and  $k \in \mathbb{Z}$  such that

$$\pi(j_1 + km_1, j_2 + km_2, j_3 + km_3) \in A_i, \quad m_i \not\equiv 0 \pmod{3}.$$

From (4) and the definition of  $\alpha$ , it follows that for all  $n \in \mathbb{Z}$ , one has

$$V_{k+3n} = \alpha(j_1 + km_1 + 3nm_1, j_2 + km_2 + 3nm_2, j_3 + km_3 + 3nm_3) = U_{j'+nm_i},$$

where  $j' = \lfloor (j_i + km_i)/3 \rfloor$ . This equation shows that there is a bi-infinite word W which is simultaneously an arithmetic subsequence of difference  $m_i$  of U and an arithmetic subsequence of difference 3 of V. Since  $m_i \neq 0 \pmod{3}$ , the word W is cube-free. Now suppose, by contradiction, that a cube  $u^3$  with  $|u| = 3\ell$  occurs in  $V, \ell \geq 1$ . Then the subword made by the letters of this occurrence whose position is congruent to  $k \pmod{3}$  would have the form  $v^3$  with  $|v| = \ell$  and would be a factor of W. This yields a contradiction, since W is cube-free.

Thus, no line of  $\alpha$  contains a factor of the form  $u^3$  with |u| multiple of 3. This implies, in particular, that all lines of  $\alpha$  are 9-th power-free.

As a corollary of the proposition above one obtains

**Theorem 3.** The Thue threshold of  $\mathbb{Z}^3$  is equal to 2. Thus, the Thue threshold of  $\mathbb{Z}^2$  is also equal to 2.

## 6 Avoiding Long Squares

In this section we will produce a 2-colouring of the lattice points of the plane such that any line contains only squares of length not larger than 26. In this case, the results of Sects. 3 and 4 do not seem to help, so that we need an adhoc construction based on paperfolding words. For the sake of brevity, we limit ourselves to outline the construction.

As a consequence of Koenig's Lemma, there exists a bi-infinite binary word U whose factors are all finite paperfolding words. We define the map  $\beta \colon \mathbb{Z}^2 \to \{0, 1\}$  as

$$\beta(j_1, j_2) = \begin{cases} U_{\lfloor j_1/2 \rfloor} & \text{if } j_1 + j_2 \text{ is odd,} \\ U_{\lfloor j_2/2 \rfloor} & \text{if } j_1 + j_2 \text{ is even,} \end{cases}$$
(5)

 $j_1, j_2 \in \mathbb{Z}.$ 

Finite factors of paperfolding words are usually called *finite paperfolding* words. By exploiting some known properties of paperfolding words we obtain the following description of the factors of the lines of  $\beta$ .

**Lemma 5.** Let x be a factor of even length of any line of  $\beta$ . Then one can write

$$x = b_1 c_1 b_2 c_2 \cdots b_k c_k,$$

with  $b_1, b_2, \ldots, b_k, c_1, c_2, \ldots, c_k \in \{0, 1\}, k \ge 0$ , and the words  $y = b_1 b_2 \cdots b_k$ and  $z = c_1 c_2 \cdots c_k$  are either finite paperfolding words or words of period 2. Moreover, at least one of the words y and z is a paperfolding word. We need also the following combinatorial property, which we state without proof, for the sake of brevity.

**Proposition 5.** Let  $x, y \in \{0, 1\}^*$  be words such that both xy and yx are finite paperfolding words. If |x|, |y| > 6, then |xy| is even.

Now we are ready to establish the main result of this section.

**Proposition 6.** Let  $\beta$  be the 2-colouring of the integer lattice points of the plane defined by (5). No line of  $\beta$  contains a square  $u^2$  with |u| > 13.

*Proof (outline).* Suppose that  $u^2$  is a factor of some line of  $\beta$ . Then  $x = u^2$  can be factorized as in Lemma 5. If k = |u| is even, then both y and z are squares. Since one of them is a paperfolding word and no paperfolding word of length larger than 10 is a square [1], one derives that  $|u| = k \leq 10$ .

Now suppose that k = |u| is odd, say  $k = 2\ell + 1$  and  $\ell > 6$ . Since x is a square, with some computation, one obtains

$$y = b_1 b_2 \cdots b_k = c_{\ell+1} c_{\ell+2} \cdots c_k c_1 c_2 \cdots c_\ell.$$

$$(6)$$

If y and z are both paperfolding words, then from (6) and Proposition 5 one derives that k is even, which is a contradiction. If, on the contrary, one of the words y and z has period 2, then from (6) one derives that both y and z have a factor of period 2 and length 4, that is one of the words  $0^4$ ,  $1^4$ , 0101, or 1010. Since none of them is a paperfolding word, we conclude that neither y nor z are paperfolding words, obtaining a contradiction. We conclude that  $\ell \leq 6$ , so that  $k \leq 13$ .

Remark 2. A careful analysis of the proof of the previous proposition shows that if  $u^2$  is a factor of a line of  $\beta$ , then either |u| is an odd integer not larger than 13 or  $|u| \in \{2, 6, 10\}$ . In particular, no line of  $\beta$  contains a square whose length is a multiple of 8. Since any 8-th power is also a square whose length is a multiple of 8, we conclude that all lines of  $\beta$  are 8-th power-free words.

# 7 Concluding Remarks

We have proved that the Thue threshold t(n) satisfies the equality t(1) = t(2) = t(3) = 2 and that the bounded square threshold of  $\mathbb{Z}^2$  is 2. It would be interesting to evaluate the Thue threshold and the bounded square threshold of  $\mathbb{Z}^n$  for all positive integer n. In [15] it is shown that  $t(n) \leq 2^n$ . Combining the results of this paper with the construction of [8], one could obtain the tighter bound  $t(3n) \leq 2^n$ . We conjecture that t(n) = 2 for all  $n \geq 1$ . Actually, since t(n) is a non-decreasing sequence, it would suffice to verify that t(n) = 2 for arbitrarily large n. We notice that our proof of the equality t(n) = 2 for n = 3 could be easily generalized to the case n = p for any prime p provide one is able to establish, for the affine Galois space AG(p, p), a property analogous to that of Remark 1. Moreover, we have treated the case of Toeplitz words such that the elements of the generating sequence belong to the set P, where P is as in (3). Other interesting unanswered questions are the following:

- 1. What is the least integer k such that there exists a 2-colouring of  $\mathbb{Z}^2$  (resp.,  $\mathbb{Z}^3$ ) whose lines are k-th power-free?
- 2. What is the least integer k such that there exists a 2-colouring of  $\mathbb{Z}^2$  whose lines do not contain squares of length larger than k?

We have built a 2-colouring of  $\mathbb{Z}^2$  such that all lines are 8-th power-free and do not contain squares of length larger than 26 and a 2-colouring of  $\mathbb{Z}^3$  such that all lines are 9-th power-free. However, there is no evidence that these bounds are optimal.

## References

- Allouche, J.-P.: Suites infinies à répétitions bornées, Séminaire de Théorie des Nombres de Bordeaux (1983–1984) Exposé no. 20, 11 p
- Allouche, J.-P., Bousquet-Mélou, M.: Facteurs des suites de Rudin-Shapiro généralisées. Bull. Belg. Math. Soc. 1, 145–164 (1994)
- Allouche, J.-P., Bousquet-Mélou, M.: Canonical positions for the factors in paperfolding sequences. Theoret. Comput. Sci. 129, 263–278 (1994)
- Avgustinovich, S.V., Fon-Der-Flaas, D.G., Frid, A.E.: Arithmetical complexity of infinite words. In: Ito, M., Imaoka, T. (eds.) Words, Languages & Combinatorics III, pp. 51–62. World Scientific, Singapore (2003)
- Berstel, J., Boasson, L.: Partial words and a theorem of Fine and Wilf. Theoret. Comput. Sci. 218, 135–141 (1999)
- Berstel, J., Perrin, D.: The origins of combinatorics on words. Eur. J. Comb. 28, 996–1022 (2007)
- Blanchet-Sadri, F.: Primitive partial words. Discret. Appl. Math. 148, 195–213 (2005)
- Carpi, A.: Multidimensional unrepetitive configurations. Theoret. Comput. Sci. 56, 233–241 (1988)
- 9. Currie, J., Simpson, J.: Non-repetitive tilings. Electron. J. Comb. 9, #R28 (2002)
- Etringer, R.C., Jackson, D.E., Schatz, J.A.: On nonrepetitive sequences. J. Comb. Theory (A) 16, 159–164 (1974)
- Fraenkel, A.S., Simpson, J.: How many squares can a string contain? J. Comb. Theory (A) 82, 112–120 (1998)
- Frid, A.E.: Sequences of linear arithmetical complexity. Theoret. Comput. Sci. 339, 68–87 (2005)
- Grytczuk, J.: Thue type problems for graphs, points, and numbers. Discret. Math. 308, 4419–4429 (2008)
- Jacobs, K., Keane, M.: 0-1-sequences of Toeplitz type. Z. Wahr. verw. Geb. 13(2), 123–131 (1969). https://doi.org/10.1007/BF00537017
- Kao, J.-Y., Rampersad, N., Shallit, J., Silva, M.: Words avoiding repetitions in arithmetic progressions. Theoret. Comput. Sci. **391**, 126–137 (2008)
- Kenkireth, B.G., Singh, M.: On the minimal alphabet size in multidimensional unrepetitive configurations. Discret. Appl. Math. 255, 258–266 (2019)
- Lothaire, M.: Combinatorics on Words. Encyclopaedia of Mathematics and its Applications, vol. 17. Addison-Wesley, Reading (1983)
- Mendès France, M., van der Poorten, A.J.: Arithmetic and analytic properties of paperfolding sequences. Bull. Austr. Math. Soc. 24, 123–131 (1981)

- Prodinger, H., Urbanek, F.J.: 0–1 sequences without long adjacent identical blocks. Discret. Math. 28, 277–289 (1979)
- Thue, A.: Über unendliche Zeichenreihen, Norske vid. Selsk. Skr. Mat. Nat. Kl. 7, 1–22 (1906). Reprinted in T. Nagell (ed.), Selected mathematical papers of Axel Thue, Universitetsforlaget (Oslo, 1977) pp. 139–158
- Thue, A.: Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, Norske vid. Selsk. Skr. Mat. Nat. Kl. 1, 1–67 (1912). Reprinted in T. Nagell (ed.), Selected mathematical papers of Axel Thue, Universitetsforlaget (Oslo, 1977) pp. 413–478