# SUM-FREE SETS GENERATED BY THE PERIOD-*k*-FOLDING SEQUENCES AND SOME STURMIAN SEQUENCES

### JEAN-PAUL ALLOUCHE, JEFFREY SHALLIT, ZHI-XIONG WEN, WEN WU\*, AND JIE-MENG ZHANG

ABSTRACT. First, we show that the sum-free set generated by the perioddoubling sequence is not  $\kappa$ -regular for any  $\kappa \geq 2$ . Next, we introduce a generalization of the period-doubling sequence, which we call the *period-k-folding* sequences. We show that the sum-free sets generated by the period-k-folding sequences also fail to be  $\kappa$ -regular for all  $\kappa \geq 2$ . Finally, we study the sum-free sets generated by Sturmian sequences that begin with '11', and their difference sequences.

#### CONTENTS

1.	Introduction	1
2.	Preliminaries	3
3.	Sum-free sets generated by period-k-folding sequences	4
4.	Non-automaticity of $\tau_k^{\infty}(1)$	8
5.	Sum-free sets generated by certain Sturmian sequences	14
6.	Subword complexity	15
Ac	knowledgements	15
Re	ferences	15

## 1. INTRODUCTION

A set of integers S is called *sum-free* if  $S \cap (S + S) = \emptyset$ , where S + S is the set  $\{x + y \mid x, y \in S\}$ . Equivalently, S is sum-free if the equation x + y = z has no solutions  $x, y, z \in S$ . When we speak of a subset  $S \subset \mathbb{N}$ , we always arrange its elements in ascending order and treat it as an integer sequence. We write  $S = (s_n)_{n>1}$ .

One can construct an infinite sum-free set from an infinite zero-one sequence using a natural map between  $\Sigma$  and  $\mathfrak{S}$  introduced by Cameron [11], where  $\Sigma$  and  $\mathfrak{S}$ denote the set of all zero-one sequences and the set of all sum-free sets, respectively. We explain this map in Section 2.2. Calkin and Finch [8] in 1996 showed that this map, denoted by  $\theta$ , is a bijection. One might expect to be able to characterize a sum-free set in terms of its corresponding zero-one sequence **t**. However, this is not always easy, even when **t** is periodic.

An infinite sum-free set S is said to be (*ultimately*) periodic if its difference sequence  $(s_{n+1} - s_n)_{n \ge 1}$  is (ultimately) periodic. Calkin and Finch [8] showed that if a sum-free set is (ultimately) periodic, then the corresponding zero-one sequence is also (ultimately) periodic. Conversely, Cameron also asked whether sum-free sets

<sup>2010</sup> Mathematics Subject Classification. 11B85, 11B13.

Key words and phrases. sum-free sets, paper-folding sequence,  $\kappa$  automatic sequences, Sturmian sequences.

<sup>\*</sup> Wen Wu is the corresponding author.

corresponding to (ultimately) periodic zero-one sequences are (ultimately) periodic. This question is still open. With the help of a computer, Calkin and Finch [8] presented some sum-free sets, which correspond to periodic zero-one sequences, and appear to be aperiodic (the aperiodicity was checked up to  $10^7$ ). Thus far, no proof has been provided to show whether these particular sum-free sets are periodic or aperiodic. Calkin and Erdős [9] showed that a class of aperiodic sum-free sets S is incomplete, i.e.,  $\mathbb{N} \setminus (S + S)$  is an infinite set. Later, Calkin, Finch, and Flowers [10] introduced the concept of difference density, which can be used to test whether specific sets are periodic. These tests produced further evidence that certain sets are not ultimately periodic. Payne [15] studied the properties of certain sum-free sets over an additive group.

Wen, Zhang, and Wu [19] studied sum-free sets corresponding to certain zeroone automatic sequences, including the Cantor-like sequences and some substitution sequences. Those sum-free sets were proved to be 2-regular sequences, which implies that they have a simple description. In contrast, in this paper, we find that the sum-free sets corresponding to the period-doubling sequence are *not*  $\kappa$ -regular for any  $\kappa \geq 2$ .

We now summarize our results. The first result characterizes the sum-free set generated by the period-doubling sequence.

**Theorem 1.** Let  $(s_n)_{n\geq 1}$  be the sum-free set generated by the period-doubling sequence  $\mathbf{p}$  and let  $\rho_8$  be the morphism defined by sending  $0 \to 833$  and  $1 \to 86$ . Set  $d_1 = 8$  and let  $d_n = s_{n+1} - s_n$  for all  $n \geq 2$ . Then  $(d_n)_{n\geq 1} = \rho_8(\mathbf{p})$ . Furthermore,  $(d_n)_{n>1}$  is not  $\kappa$ -automatic for any  $\kappa \geq 2$ .

We introduce a general version of the period-doubling sequence. The *period-k-folding* sequence  $\mathbf{p}^{(k)}$  is the fixed point of the morphism

$$\sigma_k: 0 \to 0^k 1 \text{ and } 1 \to 0^{k+1}$$

where  $k \ge 1$  is an integer. Note that  $\mathbf{p}^{(k)}$  is (k + 1)-automatic and  $\mathbf{p}^{(1)}$  is the classical period-doubling sequence [12]. Define the morphism

$$\tau_k : \begin{cases} 1 \to 1^{k-1}2; \\ 2 \to 1^{k-1}21^{k+1} \end{cases}$$

The sum-free set generated by the period-k-folding sequence is related to the morphism  $\tau_k$  in the following way.

**Theorem 2.** Let  $S = (s_n)_{n\geq 1}$  be the sum-free set generated by  $\mathbf{p}^{(k)}$ , where  $k \geq 2$ . Then  $(s_{n+1}-s_n)_{n\geq 1} = \rho_1(\tau_k^{\infty}(1))$ , where  $\rho_1$  is the coding  $1 \to k+2$  and  $2 \to 2k+4$ .

Next result shows the non-automaticity of the sequences  $\tau_k^{\infty}(1)$ . Therefore, by Theorem 2, the sum-free set generated by period-k-folding sequence are not  $\kappa$ -regular for any  $\kappa \geq 2$ .

**Theorem 3.** The sequence  $\tau_k^{\infty}(1)$  is not  $\kappa$ -automatic for any  $\kappa \geq 2$ .

It is also interesting to investigate the sum-free sets generated by some nonautomatic sequences. For example, the famous class of non-automatic sequences: the Sturmian sequences.

**Theorem 4.** The difference sequences of the sum-free sets generated by the Sturmian sequences beginning with '11' are also Sturmian sequences.

We focus on the Sturmian sequences beginning with '11' for technical reasons. It remains unknown if similar phenomenon occurs for other Sturmian sequences. The paper is organized as follows. In Section 2, we introduce the bijection  $\theta$  and give some basic facts about the sum-free sets. In Section 3, we study the sum-free sets generated by period-k-folding sequences and prove Theorem 1 and 2. In Section 4, we prove Theorem 3, which is the non-automaticity of the sequence  $\tau_k^{\infty}(1)$ . In Section 5, sum-free sets generated by certain Sturmian sequences are investigated and Theorem 4 is proved. Finally, in Section 6, we give a conjecture about subword complexity.

### 2. Preliminaries

2.1. Notations and definitions. For a detailed discussion about the following terms, such as " $\kappa$ -automatic sequence", " $\kappa$ -regular sequence", "Sturmian sequence", and so forth, see [4, 5, 6, 14].

**Words.** An alphabet  $\mathcal{A}$  is a finite set. The elements of  $\mathcal{A}$  are called *letters*. The set of all finite words over the alphabet  $\mathcal{A}$  is  $\mathcal{A}^* := \bigcup_{\geq 0} \mathcal{A}^n$ , where  $\mathcal{A}^0 = \{\varepsilon\}$  and  $\varepsilon$  denotes the empty word. For  $w \in \mathcal{A}^*$ , let |w| denote the *length* of w. Namely, if  $w \in \mathcal{A}^n$ , then |w| = n. For two words  $w = w_1 w_2 \cdots w_{|w|}$ ,  $v = v_1 v_2 \cdots v_{|v|} \in \mathcal{A}^*$ , their concatenation is  $wv = w_1 w_2 \cdots w_{|w|} v_1 v_2 \cdots v_{|v|}$ .

**Morphism.** A morphism  $\sigma$  on  $\mathcal{A}$  is the map  $\mathcal{A} \to \mathcal{A}^*$ , which can be extended to  $\mathcal{A}^*$  satisfying  $\sigma(wv) = \sigma(w)\sigma(v)$  for all  $w, v \in \mathcal{A}^*$ . Let  $\mathcal{A}^{\omega}$  be the set of infinite sequence on  $\mathcal{A}$  and let  $\mathcal{A}^{\infty} := \mathcal{A}^{\omega} \cup \mathcal{A}^*$ . For any  $a \in \mathcal{A}$ , by  $\sigma^{\infty}(a)$  we mean the limit  $\lim_{n\to\infty} \sigma^n(a)$ , provided the limit exists. The limit is taken under the natural metric on  $\mathcal{A}^{\infty}$ .

 $\kappa$ -automatic sequences and  $\kappa$ -regular sequences. Let  $\kappa \geq 2$  be an integer. A sequence  $\mathbf{u} = (u_n)_{n\geq 0}$  over the alphabet  $\mathcal{A}$  is a  $\kappa$ -automatic sequence if and only if its  $\kappa$ -kernel  $K_{\kappa}(\mathbf{u})$  is finite, where  $K_{\kappa}(\mathbf{u}) := \{(u_{\kappa^i n+j})_{n\geq 0} \mid i \geq 0, 0 \leq j < \kappa^i\}$ . A sequence  $\mathbf{u} = (u_n)_{n\geq 0}$  taking values in  $\mathbb{Z}$  is  $\kappa$ -regular if the  $\mathbb{Z}$ -module generated by its  $\kappa$ -kernel  $K_{\kappa}(\mathbf{u})$  is finitely generated.

**Sturmian sequences.** For  $\mathbf{w} \in \mathcal{A}^{\infty}$ , its subword complexity function  $P_{\mathbf{w}} : \mathbb{N} \to \mathbb{N}$  is defined by

$$P_{\mathbf{w}}(n) = \#\{w_i w_{i+1} \cdots w_{i+n-1} \mid i \ge 0\}.$$

A sequence **w** is a *Sturmian* sequence if  $P_{\mathbf{w}}(n) = n + 1$  for all  $n \ge 1$ .

2.2. The bijection  $\theta$ . Let  $\mathbf{w} = w_1 w_2 w_3 \cdots \in \{0, 1\}^{\infty}$ . We now construct sets  $S_i, T_i, U_i$ , as follows. Define  $S_0 = T_0 = U_0 = \emptyset$ . For  $i = 1, 2, 3, \ldots$ , let  $n_i$  be the least element of  $\mathbb{N} \setminus (S_{i-1} \cup T_{i-1} \cup U_{i-1})$ . If  $w_i = 1$ , set

$$S_i = S_{i-1} \cup \{n_i\}, \quad T_i = S_i + S_i, \quad U_i = S_{i-1},$$

while if  $w_i = 0$ , set

$$S_i = S_{i-1}, \quad T_i = T_{i-1}, \quad U_i = U_{i-1} \cup \{n_i\}.$$

Let  $S = \bigcup_i S_i$ . Then, since each  $S_i$  is sum-free and  $S_i \subset S_{i+1}$ , the set S is also sum-free. We define S to be the image of **w** under  $\theta$ , i.e.,  $\theta(\mathbf{w}) = S$ . For example,

$$\begin{array}{rcl} \theta: & 111111111\cdots & \mapsto & \{1,3,5,7,9,11,13,15,\ldots\},\\ \theta: & 01010101\cdots & \mapsto & \{2,5,8,11,\ldots\}. \end{array}$$

The inverse of  $\theta$  is given as follows. Let  $S \subset \mathbb{N}$  be a sum-free set with  $\#S = \infty$ . We define the sequence  $\mathbf{v} = (v_n)_{n>1}$  over the alphabet  $\{0, 1, *\}$  by

$$v_n = \begin{cases} 1, & \text{if } n \in S; \\ *, & \text{if } n \in S + S; \\ 0, & \text{otherwise.} \end{cases}$$
(1)

Deleting all \*'s in  $\mathbf{v}$ , we obtain a zero-one sequence  $\mathbf{v}'$  and one can verify that  $\theta(\mathbf{v}') = S$ .

2.3. Basic facts. Note that 
$$S = \{i \in \mathbb{N}_{\geq 1} \mid v_i = 1\} = (s_n)_{n \geq 1}$$
 and  
 $s_{i+1} = s_i = u_i + \alpha_i + 1$  (2)

 $s_{n+1} - s_n = \mu_n + \alpha_n + 1, \tag{2}$ 

where

$$\mu_n := \#\{i \in \mathbb{N} \mid v_i = 0, \ s_n < i < s_{n+1}\}$$

and

$$\alpha_n := \#\{i \in \mathbb{N} \mid v_i = *, \ s_n < i < s_{n+1}\}.$$

The quantity  $\mu_n$  (resp.,  $\alpha_n$ ) is the number of '0's (resp., '\*'s) between the *n*-th and the (n + 1)-th occurrence of '1' in **v**. Moreover,  $\mu_n$  also counts the number of '0's between the *n*-th and the (n + 1)-th occurrence of '1' in **v**'. Let  $S' = \{i \in \mathbb{N} \mid v'_i = 1\} = (s'_n)_{n \ge 1}$ . Then

$$\mu_n = s'_{n+1} - s'_n. \tag{3}$$

## 3. Sum-free sets generated by period-k-folding sequences

Let  $k \ge 1$  be an integer. Recall that  $\sigma_k$  is the morphism  $0 \to 0^{k}1$  and  $1 \to 0^{k+1}$ and the *period-k-folding sequence* 

$$\mathbf{p}^{(k)} = (p_n)_{n \ge 0} = \sigma_k^\infty(0).$$

The sequence  $\mathbf{p}^{(k)}$  can be also defined recursively by the following recurrence relations:  $p_0 = 0$  and for all  $n \ge 0$ ,

$$p_{(k+1)n+j} = \begin{cases} 0, & \text{if } j = 0, 1, \dots, (k-1); \\ 1 - p_n, & \text{if } j = k. \end{cases}$$
(4)

Theorem 2 says that the sum-free set corresponding to  $\mathbf{p}^{(k)}$  is related to the following morphism

$$\tau_k : \begin{cases} 1 \to 1^{k-1}2; \\ 2 \to 1^{k-1}21^{k+1} \end{cases}$$

We remark that  $\tau_k^{\infty}(1)$  is the image of the period-doubling sequence under a non-uniform projection. That is, we have

$$\tau_k^{\infty}(1) = \rho_0(\sigma_k^{\infty}(0)),\tag{5}$$

where  $\rho_0$  maps  $0 \to 1^{k-1}2$  and  $1 \to 1^{k+1}$ . One can verify Eq. (5) by arguing that for all  $n \ge 1$  we have

$$\tau_k^{n+1}(1) = \rho_0(\sigma_k^n(0)). \tag{6}$$

Note that  $\tau_k^2(1) = (1^{k-1}2)^k 1^{k+1} = \rho_0(\sigma_k(0))$ , and we suppose that  $\tau_k^{m+1}(1) = \rho_0(\sigma_k^m(0))$  for all  $m \le n$ . Then

$$\begin{aligned} \tau_k^{n+2}(1) &= \tau_k^{n+1}(1^{k-1}2) \\ &= (\tau_k^{n+1}(1))^{k-1}\tau_k^{n+1}(2) \\ &= [\tau_k^{n+1}(1)]^k[\tau_k^n(1)]^{k+1} \\ &= [\rho_0(\sigma_k^n(0))]^k[\rho_0(\sigma_k^{n-1}(0))]^{k+1} \\ &= \rho_0\Big(\sigma_k^n(0^k)\sigma_k^{n-1}(0^{k+1})\Big) \\ &= \rho_0(\sigma_k^{n+1}(0)). \end{aligned}$$

So Eq. (6) follows by induction.

3.1. The blocks of zeros. Let  $\Gamma$  be the map between  $\{0,1\}^*$  and  $\mathbb{N}^*$  that measures the distance between adjacent '1's in finite binary words. More precisely, if  $w = 0^{x_0}10^{x_1}1\cdots 0^{x_n}10^{x_{n+1}}$ , then

$$\Gamma(w) = x_1 x_2 \cdots x_n$$

where  $x_i \in \mathbb{N}$  for  $i = 0, 1, \dots, n+1$ . For  $w, v \in \{0, 1\}^*$  and  $x \ge 0$ , we have

$$\Gamma(w0^x 1v) = \Gamma(w0^x 1)\Gamma(0^x 1v).$$
(7)

Let  $\rho_2$  be the coding  $1 \to k$  and  $2 \to 2k + 1$ .

**Lemma 5.** For all  $k \ge 1$ , we have  $\Gamma([\sigma_k^n(0)]^j \sigma_k(0)) = [\rho_2(\tau_k^{n-1}(k))]^j$  for all  $n, j \ge 1$ .

*Proof.* Note that  $\sigma_k(0) = 0^k 1$  is a prefix of  $\sigma_k^n(0)$  for all  $n \ge 1$ . By Eq. (7) we have  $\Gamma([\sigma_k^n(0)]^j \sigma_k(0)) = [\Gamma(\sigma_k^n(0)\sigma_k(0))]^j \tag{8}$ 

for all  $j \ge 1$ . So we only need to show that for all  $n \ge 1$ , we have

$$\Gamma(\sigma_k^n(0)\sigma_k(0)) = \rho_2(\tau_k^{n-1}(1)).$$
(9)

Obviously, Eq. (9) holds for n=1 and 2. Suppose that Eq. (9) holds for all  $m\leq n.$  We have

$$\Gamma(\sigma_k^{n+1}(0)\sigma_k(0)) = \Gamma(\sigma_k^n(0^k 1)\sigma_k(0)) 
= \Gamma([\sigma_k^n(0)]^k [\sigma_k^{n-1}(0)]^{k+1}\sigma_k(0)) 
= \Gamma([\sigma_k^n(0)]^k \sigma_k(0))\Gamma([\sigma_k^{n-1}(0)]^{k+1}\sigma_k(0)) \quad \text{(by Eq. (7))} 
= [\Gamma(\sigma_k^n(0)\sigma_k(0))]^k [\Gamma(\sigma_k^{n-1}(0)\sigma_k(0)]^{k+1} \quad \text{(by Eq. (8))} 
= \rho_2 \Big( [\tau_k^{n-1}(1)]^k [\tau_k^{n-2}(1)]^{k+1} \Big) \quad \text{(by Eq. (9))} 
= \rho_2(\tau_k^n(1)).$$

Thus Eq. (9) holds for n + 1, which completes the proof.

**Lemma 6.** For all  $k \ge 1$ ,  $(\mu_n)_{n>1} = \rho_2(\tau_k^{\infty}(1))$ .

*Proof.* Recall that  $\mu_n$  is the number of '0's between the *n*-th and the (n + 1)-th occurrence of '1' in  $\mathbf{p}^{(k)}$ . Note also that  $\sigma_k^n(0)\sigma_k(0)$  is a prefix of  $\mathbf{p}^{(k)}$  for all  $n \ge 1$ . Thus

$$(\mu_n)_{n\geq 1} = \lim_{i\to\infty} \Gamma(\sigma_k^i(0)\sigma_k(0)).$$
(10)

The result follows from Eq. (10) and Lemma 5.

3.2. The gaps for stars when  $k \geq 2$ . While we construct the sum-free set S corresponding to  $\mathbf{p}^{(k)}$ , we actually insert stars into  $\mathbf{p}^{(k)}$  and finally obtain the ternary sequence  $(v_n)_{n\geq 1}$  satisfying Eq. (1).

**Lemma 7.** For all  $k \ge 2$ ,  $(\alpha_n)_{n\ge 1} = \tau_k^{\infty}(1)$ .

*Proof.* We prove that for all  $n \ge 1$ ,

$$\begin{cases} \alpha_n = \rho_2^{-1}(\mu_n); \\ \forall x \in S_n + S_n, \ x \equiv k \pmod{k+2}, \end{cases}$$
(11)

where  $S_n = \{s_1, ..., s_n\}.$ 

Since  $0^k 10^k 1$  is a prefix of  $\mathbf{p}^{(k)}$  when  $k \ge 2$ , we have  $s_1 = k + 1$ ,  $s_2 = 2k + 3$  and  $\alpha_1 = 1 = \rho_2^{-1}(\mu_1)$ . So Eq. (11) holds for n = 1. Assume that Eq. (11) holds for all  $m \le n$ . By the inductive assumption,  $v_{s_n+s_1} = *$  and  $v_{s_n+s_1\pm i} \ne *$  for  $i = 1, \ldots, k$ . By Lemma 6, we know that  $\mu_j \in \{k, 2k + 1\}$  for all  $j \ge 1$ .

Case I. If  $\mu_{n+1} = k$ , then  $v_{s_n+i} = 0$  for  $i = 1, \ldots, k$  and  $v_{s_n+k+2} = 1$ . So  $\alpha_{n+1} = 1 = \rho_2^{-1}(\mu_{n+1})$  and  $s_{n+1} = s_n + k + 2$ . Therefore, in this case, Eq. (11) holds for n + 1.

Case II. If  $\mu_{n+1} = 2k+1$ , by the inductive assumption,  $v_{s_n+s_1} = *$ ,  $v_{s_n+s_2} = *$ and  $v_{s_n+s_1\pm i} = 0$  for  $i = 1, \ldots, k$ . So  $v_{s_n+2k+4} = 1$  and  $s_{n+1} = s_n + 2k + 4$ . It follows that  $\alpha_{n+1} = 2 = \rho_2^{-1}(\mu_{n+1})$  and  $S_{n+1} = S_n \cup \{s_n + 2k + 4\}$ , which implies that Eq. (11) holds for n+1 in this case.

By induction, we see Eq. (11) holds for all  $n \ge 1$  and this completes the proof.  $\Box$ 

Remark 8. The stars occur periodically in  $(v_n)_{n\geq 1}$ . Actually,  $v_n = *$  if and only if  $n \equiv k \pmod{k+2}$ .

3.3. The gaps for stars when k = 1. In this case, we will show that the stars occur periodically in  $(v_n)_{n>1}$  with period 6. The initial values of  $(v_n)_{n>1}$  are

													13
$v_n$	0	1	0	*	0	0	1	0	*	1	0	*	1

**Lemma 9.** Set  $S_n := \{i \mid v_i = 1, 1 \le i \le 14n + 13\}$  and  $\tilde{S}_n := \{x \pmod{14} \mid x \in S_n + S_n, x > 13\}$  for  $n \ge 0$ . Then for all  $n \ge 1$  we have

(a) for  $0 \le j \le 13$ ,  $v_{14n+j} = *$  if and only if  $j \in \mathcal{I}_* := \{0, 1, 3, 6, 9, 12\};$ 

(b)  $v_{14n+4} = 0$ ,  $\{14n+7, 14n+13\} \subset S_n \text{ and } \tilde{S}_n = \mathcal{I}_*$ .

*Proof.* From the initial values  $v_1, \ldots, v_{13}$ , we know that  $S_0 := \{2, 7, 10, 13\} \subset S$ . Since

 $S_0 + S_0 = \{4, 9, 12, 14, 15, 17, 20, 23, 26\},\$ 

we see that (a) holds for n = 1. Recall that by Lemma 6, when k = 1, the sequence  $(\mu_i)_{i\geq 1}$  is the fixed point of the morphism sending  $3 \to 311$  and  $1 \to 3$ . Therefore,  $\mu_4 = \mu_5 = 3$  yields that  $S_1 = S_0 \cup \{21, 27\}$ , which implies that (b) holds for n = 1. Now assume that (a) and (b) hold for  $m \leq n$ . We shall show the validity of them for n + 1. Using the inductive hypothesis (b) for  $n, \tilde{S}_n = \mathcal{I}_*$  and

$$S_0 + S_n \supset S_0 + \{14n + 7, 14n + 13\} = \{14n + j \mid j = 9, 14, 15, 17, 20, 23, 26\},\$$

which imply that  $v_{14(n+1)+j} = *$  if and only if  $j \in \mathcal{I}_*$ . So (a) holds for n+1. Applying (a) for n and n+1, we have the following table

j	0	1	2	3	4	5	6	7	8	9	10	11	12	13
$v_{14n+j}$	*	*		*			*	1		*			*	1
$v_{14(n+1)+j}$	*	*		*			*			*			*	

Suppose  $v_{14n+7}$  is the  $\ell$ -th '1' in  $\mathbf{p}^{(1)}$ . There are two cases  $\mu_{\ell} = 1$  and  $\mu_{\ell} = 3$ .

- If  $\mu_{\ell} = 1$ , then  $\mu_{\ell+1}$  must be 1 since  $v_{14n+13} = 1$ . Further,  $\mu_{\ell+2}$  must be 3 since '111' is not a factor of  $(\mu_n)_{n\geq 1}$ , which implies that  $v_{14(n+1)+7} = 1$  and  $v_{14(n+1)+4} = 0$ . If  $\mu_{\ell+3} = 3$ , then  $v_{14(n+1)+13} = 1$ . If  $\mu_{\ell+3} = 1$ , then  $\mu_{\ell+4} = 1$ . We also have  $v_{14(n+1)+13} = 1$ .
- If  $\mu_{\ell} = 3$ , then either  $\mu_{\ell+1} = 3$  or  $\mu_{\ell+1} = \mu_{\ell+2} = 1$ . In both cases, we have  $v_{14(n+1)+7} = 1$ . When  $\mu_{\ell+1} = 3$ , either  $\mu_{\ell+2} = 3$  or  $\mu_{\ell+2} = \mu_{\ell+3} = 1$ . In both cases,  $v_{14(n+1)+13} = 1$ . When  $\mu_{\ell+1} = \mu_{\ell+2} = 1$ , we have  $v_{14(n+1)+4} = 1$  and  $\mu_{\ell+3} = 3$ . Note that  $\mu_{\ell+3} = 3$  indicates  $v_{14(n+1)+11} = 0$ . However,  $v_7 = 1$  and  $v_{14(n+1)+4} = 1$  yields  $v_{14(n+1)+11} = *$  since 14(n+1) + 11 = [14(n+1) + 4] + 7. This contradiction implies that  $\mu_{\ell+1} = \mu_{\ell+2} = 1$  cannot happen.

In the above two cases, we have

$$S_{n+1} \subset S_n \cup \{14(n+1) + 7, 14(n+1) + 10, 14(n+1) + 13\},\$$

which together with the inductive hypothesis  $\tilde{S}_n = \mathcal{I}_*$ , gives  $\tilde{S}_{n+1} = \mathcal{I}_*$ . So (b) also holds for n+1. This completes the proof.

By Lemma 9, we are able to characterize  $(\alpha_n)_{n>1}$  through  $(\mu_n)_{n>1}$ .

**Lemma 10.** Let  $\alpha'_1 = 4$  and  $\alpha'_n = \alpha_n$  for all  $n \ge 2$ . Then

$$(\alpha'_n)_{n\geq 1} = \rho_4(\sigma_1^\infty(0))$$

where  $\rho_4: 0 \to 411, 1 \to 42 \text{ and } \sigma_1: 0 \to 01, 1 \to 00.$ 

*Proof.* Note that when k = 1, if we replace  $\rho_0$  by  $\tau_1 \circ \rho_0 : 0 \to 211, 1 \to 22$ , then Eq. (5) still holds. Recall that  $\rho_2$  is the coding  $1 \to 1, 2 \to 3$  when k = 1. Set  $\rho_3 := \rho_2 \circ \tau_1 \circ \rho_0$  which maps  $0 \to 311$  and  $1 \to 33$ . Applying Eq. (5), one can decompose  $(\mu_n)_{n\geq 1}$  into a sequence over the alphabet  $\{311, 33\}$  as follows

$$(\mu_n)_{n \ge 1} = \rho_2(\tau_1^{\infty}(1))$$
 (by Lemma 6)  
=  $\rho_2(\tau_1(\tau_1^{\infty}(1)))$   
=  $(\rho_2 \circ \tau_1 \circ \rho_0)(\sigma_1^{\infty}(1))$  (by Eq. (5))  
=  $\rho_3(\sigma_1^{\infty}(1)).$ 

From Lemma 9, we have the distribution of  $(v_n)_{n\geq 1}$  as follows: for  $n\geq 1$ ,

j	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
$v_{14n+i}$	1	*	*	0	*	0	0	*	1	0	*	$\diamond$	0	*	1

where  $\diamond \in \{0, 1\}$ . Suppose  $v_{14n-1}$  is the  $\ell$ -th '1' in  $\mathbf{p}^{(1)}$ . Then we have  $\mu_{\ell} = 3$ , which implies  $\alpha_{\ell} = 4$ . If  $\mu_{\ell+1} = 1$ , then  $\mu_{\ell+2} = 1$  and  $\alpha_{\ell+1} = \alpha_{\ell+2} = 1$ . If  $\mu_{\ell+1} = 3$ , then  $\alpha_{\ell+1} = 2$ . Thus if we treat  $(\mu_n)_{n\geq 4}$  as a sequence on  $\{311, 33\}$ , then  $(\alpha_n)_{n\geq 4}$ is a sequence on  $\{411, 42\}$  by projecting  $311 \rightarrow 411$ ,  $33 \rightarrow 42$ . This proves the lemma.

3.4. **Proof of Theorem 1 and 2.** For readers' convenience, we restate our Theorem 1 and 2 here.

**Theorem 1.** Let  $S = (s_n)_{n \ge 1}$  be the sum-free set generated by  $\mathbf{p}^{(1)}$ . Set  $d_1 = 8$ and  $d_n = s_{n+1} - s_n$  for all  $n \ge 2$ . Then

$$\mathbf{d} := (d_n)_{n \ge 1} = \rho_8(\sigma_1^\infty(0)), \tag{12}$$

where  $\rho_8 : 0 \to 833, 1 \to 86$  and  $\sigma_1 : 0 \to 01, 1 \to 00$ . Moreover, **d** is not  $\kappa$ -automatic sequence for all  $\kappa \geq 2$ .

Proof. The formula (12) follows from Eq. (2), Lemma 6 and Lemma 10.

Let  $\psi$  be the coding  $8 \to 8$ ,  $3 \to 3$ , and  $6 \to 8$ . Then  $\psi(\mathbf{d})$  is the fixed point of the morphism  $8 \to 833$  and  $3 \to 8$ . Allouche, Allouche, and Shallit [2] in 2006 showed that this sequence is not  $\kappa$ -automatic sequence for all  $\kappa \ge 2$ . So  $\mathbf{d}$  is also not an automatic sequence.

**Theorem 2.** Let  $S = (s_n)_{n \ge 1}$  be the sum-free set generated by  $\mathbf{p}^{(k)}$  where  $k \ge 2$ . Then

$$(s_{n+1} - s_n)_{n \ge 1} = \rho_1(\tau_k^{\infty}(1)),$$

where  $\rho_1$  is the coding  $1 \rightarrow k+2$  and  $2 \rightarrow 2k+4$ .

*Proof.* The result follows from Eq. (2), Lemma 6 and Lemma 7.

## 4. Non-automaticity of $\tau_k^{\infty}(1)$

Here we prove that  $\tau_k^{\infty}(1)$  is not an automatic sequence.

**Theorem 3.**  $\tau_k^{\infty}(1)$  is not a  $\kappa$ -automatic sequence for any  $\kappa \geq 2$ .

In what follows  $\tau_k$  is the morphism defined over the alphabet  $\{1,2\}$  by  $\tau_k(1) = 1^{k-1}2$ ,  $\tau_k(2) = 1^{k-1}21^{k+1}$ . The iterative fixed point of  $\tau_k$  is

$$\tau_k^{\infty}(1) = \underbrace{1^{k-1} \ 2 \ \cdots \ 1^{k-1} \ 2}_{(k-1) \text{ occurrences}} 1^{k-1} \ 2 \ 1^{k+1} \cdots$$

Letting  $\sigma_1^{\infty}(1)$  denote the iterative fixed point of the morphism  $\sigma_1$  defined by  $\sigma_1(1) = 121$ ,  $\sigma_1(2) = 12221$ , it is not very difficult to prove that  $\sigma_1^{\infty}(1) = 1\tau_1^{\infty}(1)$ . It was proved in [17] and written down in [2] that the sequence  $\sigma_1^{\infty}(1)$  is not 2automatic. Using methods similar to those in [2] for other sequences, it can be proved, using a deep result of F. Durand [13], that, for  $\kappa \ge 2$ , the sequence  $\sigma_1^{\infty}(1)$  is not  $\kappa$ -automatic either: this was actually done explicitly in [16]. Hence  $\tau_1^{\infty}(1)$  is not  $\kappa$ -automatic either, for any  $\kappa \ge 2$ .

Here we will prove, inspired by the method in [17, 2], that the iterative fixed point of  $\tau_k$  is not  $\kappa$ -automatic for any  $\kappa \geq 2$ . First we show, thanks to Durand's theorem [13], that it suffices to prove that  $\tau_k^{\infty}(1)$  is not (k + 1)-automatic.

**Lemma 11.** If the sequence  $\tau_k^{\infty}(1)$  were  $\kappa$ -automatic for some  $\kappa \geq 2$ , then it would be (k + 1)-automatic.

Proof. The transition matrix of the morphism  $\tau_k$  is the matrix  $M_k = \begin{pmatrix} k-1 & 2k \\ 1 & 1 \end{pmatrix}$ whose dominant eigenvalue is (k + 1). Hence the sequence  $\tau_k^{\infty}(1)$  is (k + 1)-substitutive. Thus, if it were  $\kappa$ -automatic for some  $\kappa \geq 2$ , then it would either be ultimately periodic (hence in particular (k + 1)-automatic), or the integers (k + 1)and  $\kappa$  would be multiplicatively dependent (see [13, Theorem 1]). If (k + 1) and  $\kappa$  are multiplicatively dependent, then there exist two nonzero integers a and bsuch that  $(k + 1)^a = \kappa^b$ . Thus the sequence  $\tau_k^{\infty}(1)$  is  $(k + 1)^a$ -automatic, hence (k + 1)-automatic.

To complete the proof of the non-automaticity of  $\tau_k^{\infty}(1)$ , we are thus going to prove that  $\tau_k^{\infty}(1)$  is not (k+1)-automatic. We begin with some lemmas.

**Lemma 12.** Define the sequence of integers  $(W_k(n))_{n\geq 0}$  by

$$W_k(n) := \frac{(k+1)^n - (-1)^n}{k+2}.$$

Then we have the following properties.

(i)  $W_k(n+1) = (k+1)W_k(n) + (-1)^n$ . (ii)  $W_k(n+2) = kW_k(n+1) + (k+1)W_k(n)$ . (iii)  $W_k(n+1) + W_k(n) = (k+1)^n$ . (iv)  $k\left(\sum_{1 \le \ell \le n} W_k(\ell)\right) = \begin{cases} W_k(n+1), & \text{if } n \text{ is odd}; \\ W_k(n+1) - 1, & \text{if } n \text{ is even.} \end{cases}$ (v)  $W_k(n) = \sum_{1 \le j \le n} (-1)^{j+1} (k+1)^{n-j}$ .

(vi) 
$$W_k(n) = \begin{cases} (k+1)^{n-1} - k \left( \sum_{1 \le j \le \frac{n-1}{2}} (k+1)^{n-2j-1} \right), & \text{if } n \text{ is odd;} \\ (k+1)^{n-1} - k \left( \sum_{1 \le j \le \frac{n-1}{2}} (k+1)^{n-2j-1} \right) - 1, & \text{if } n \ne 0 \text{ is even;} \end{cases}$$
  
(vii) The length of  $\tau_k^n(1)$  is equal to  $W_k(n+1)$ .

*Proof.* Assertions (i), (ii), (iii) and (iv) are straightforward consequences of the definition of  $W_k(n)$ . Assertion (v) is proved by induction on n using (i). Assertion (vi) is proved by calculating the sum  $\sum_{1 \le j \le \frac{n-1}{2}} (k+1)^{n-2j-1}$ .

Finally, to prove (vii), we let  $\ell_k(n)$  and  $m_k(n)$  denote the lengths of the words  $\tau_k^n(1)$  and  $\tau_k^n(2)$ . We clearly have from the definition of  $\tau_k$  that  $\ell_k(0) = m_k(0) = 1$ , and, for  $n \ge 0$ ,

$$\ell_k(n+1) = (k-1)\ell_k(n) + m_k(n)$$
  
$$m_k(n+1) = 2k\ell_k(n) + m_k(n).$$

Define  $\ell'_k$  and  $m'_k$  by  $\ell'_k(n) := W_k(n+1)$  and  $m'_k(n) := W_k(n+2) - (k-1)W_k(n+1)$ . Since  $\ell'_k$  and  $m'_k$  have the same initial values and satisfy the same recurrence (use (ii)) as  $\ell_k$  and  $m_k$ , we have  $\ell_k = \ell'_k$  and  $m_k = m'_k$ .

Remark 13. The sequence  $(W_1(n))_{n\geq 0} = 0 \ 1 \ 1 \ 3 \ 5 \ 11 \ 21 \ 43 \ \cdots$  is the Jacobsthal sequence (sequence <u>A001045</u> in the On-Line Encyclopedia of Integer Sequences (OEIS) [18]). The sequence  $(W_2(n))_{n\geq 0} = 0 \ 1 \ 2 \ 7 \ 20 \ 61 \ \cdots$  is sequence <u>A015518</u> in the OEIS. The sequences  $(W_3(n))_{n\geq 0}, (W_4(n))_{n\geq 0}, \ldots$ , up to  $(W_9(n))_{n\geq 0}$  are, respectively, the sequences <u>A015521</u>, <u>A015531</u>, <u>A015540</u>, <u>A015552</u>, <u>A015565</u>, <u>A015577</u>, <u>A015585</u> in the OEIS. The number  $W_k(n)$  counts in particular the number of walks of length n between two distinct vertices of the complete graph  $K_n$ . Also see Proposition 18 below.

Now we introduce a numeration system associated with  $\tau_k$  (where, as previously,  $k \geq 0$ ). Two propositions about this numeration system and its relation to the sequence  $\tau_k^{\infty}(1)$  will prove useful for obtaining that  $\tau_k^{\infty}(1)$  is not (k+1)-automatic.

**Definition 14.** Let r be a positive integer. Let  $x_1, x_2, \ldots, x_r$  be nonnegative integers. We let  $[x_rx_{r-1}\cdots x_1]_W$  denote the integer  $\sum_{1\leq j\leq r} x_jW_k(j)$ . We say that  $[x_rx_{r-1}\cdots x_1]_W$  is a *valid W-expansion* of the integer  $\sum_{1\leq j\leq r} x_jW_k(j)$  if all the  $x_j$ 's belong to [0, k], with  $x_r \neq 0$ , and if the word  $x_rx_{r-1}\cdots x_1$  ends with an even number (possibly equal to 0) of k's.

**Proposition 15.** Every nonzero integer admits a unique valid W-expansion.

Proof. First we show that every nonzero integer admits a valid W-expansion, by proving by induction on t that, for all  $n \in [1, W_k(t))$ , n admits a valid W-expansion  $n = [x_r x_{r-1} \cdots x_1]_W$  with r < t, and  $[x_r x_{r-1} \cdots x_1]_W$  ends with an even number (possibly equal to zero) of k's. This is true for t = 2, since  $W_k(2) = k$  and we have  $n = [n]_W$  for all  $n \in [1, k)$ . Suppose that the property holds for some t, and let n be an integer belonging to  $[W_k(t), W_k(t+1))$ . Since  $W_k(t) = [10^{t-1}]_W$  we can suppose that n belongs to  $(W_k(t), W_k(t+1))$ . Using Assertion (i) of Lemma 12, we have  $W_k(t) < n < W_k(t+1) \le (k+1)W_k(t) + 1$ . Hence  $W_k(t) < n \le (k+1)W_k(t)$ . Thus, if  $\alpha$  is the integer such that  $\alpha W_k(t) < n \le (\alpha + 1)W_k(t)$ , we have  $\alpha < k + 1$ and  $\alpha + 1 > 1$ . Hence  $1 \le \alpha \le k$ . Define  $m := n - \alpha W_k(t)$ . Then m belongs to  $(0, W_k(t)]$ . By the induction hypothesis for  $m \ne W_k(t)$  and directly for  $m = W_k(t)$ , *m* can be represented as  $[x_r x_{r-1} \cdots x_1]_W$ , with r < t and  $[x_r x_{r-1} \cdots x_1]_W$  ends with an even number (possibly equal to zero) of k's. Then

$$n = m + \alpha W_k(t) = \begin{bmatrix} \alpha & \underbrace{0 \cdots 0}_{(t-1-r) \text{ terms}} x_r x_{r-1} \cdots x_1 \end{bmatrix}_W.$$

This yields a valid W-expansion of n, except possibly if r = t-1,  $\alpha = k$ , all the  $x_j$ 's are equal to k, and t is odd. But then  $n = k(W_k(t) + W_k(t-1) + \cdots + W_k(1))$ , which is equal to  $W_k(t+1)$  (Assertion (iv) of Lemma 12): but we assumed  $n < W_k(t+1)$ .

To prove uniqueness of the valid W-expansion of every integer, it suffices to prove that the number of words  $x_r x_{r-1} \cdots x_1$  with  $x_i \in [0,k]$ ,  $x_r \neq 0$  and such that  $x_r x_{r-1} \cdots x_1$  ends with an even number (possibly 0) of k's is equal to the number of integers in the interval  $[W_k(r), W_k(r+1))$ , i.e.,  $W_k(r+1) - W_k(r)$ . To count this number of words, we note that it is the difference between the number of all words of length r over the alphabet  $\{0, 1, \ldots, k\}$  beginning with  $\alpha \in [1, k]$ (i.e.,  $k(k+1)^{r-1}$ ) and the number of all such words ending with  $\beta k$ , or  $\beta kkk$ , or  $\beta kkkkk \cdots$ , where  $\beta$  is any letter different from k, except possibly if r is odd and all the letters in  $x_r x_{r-1} \cdots x_1$  are equal to k. We thus obtain  $k(k+1)^{r-1} - (k^2(k+1)^{r-3} + k^2(k+1)^{r-5} + \cdots) - \eta_r$ , where  $\eta_r$  is equal to 0 if r is even and to 1 if r is odd, which (see Assertion (vi) of Lemma 12) is equal to  $W_k(r) + 1$  if r is even, and to  $W_k(r) - 1$  if r is odd, thus to  $W_k(r) + (-1)^r$ . And this last quantity is equal to  $W_k(r+1) - W_k(r)$  from Assertion (i) in Lemma 12.

**Proposition 16.** Let  $\tau_k^{\infty}(1) := (t_k(n))_{n\geq 0} = t_k(0)t_k(1)\cdots \in \{1,2\}^{\mathbb{N}}$ . Then  $t_k(n) = 2$  if and only if the valid W-expansion of (n+1) ends with an odd number of 0's (or equivalently if and only if the valid W-expansion of (n+1) has the form  $n = \sum_{2\ell+2 \leq j \leq r} x_j W_k(j)$ , for some  $\ell \geq 0$ , and  $x_{2\ell+2} \neq 0$ ).

*Proof.* First we note that

$$\tau_k^{m+2}(1) = (\tau_k^{m+1}(1))^k (\tau_k^m(1))^{k+1}, \text{ for all } m \ge 0.$$
(13)

For m = 0 we have

$$\tau_k^2(1) = (1^{k-1}2)^{k-1}(1^{k-1}21^{k+1}) = (1^{k-1}2)^k 1^{k+1} = (\tau_k^1(1))^k (\tau_k^0(1))^{k+1}$$

It then suffices to apply  $\tau_k^m$  to this equality. In other words, Eq. (13) means that  $\tau_k^{m+1}(1)$  is the concatenation of k blocks equal to  $\tau_k^{m+1}(1)$  and of (k+1) blocks equal to  $\tau_k^m(1)$  (of respective lengths  $W_k(m+2)$  and  $W_k(m+1)$  from Assertion (vii) of Lemma 12):

$$\tau_k^{m+2}(1) = \underbrace{\tau_k^{m+1}(1) \cdots \tau_k^{m+1}(1)}_{k \text{ blocks of } length \ W_k(m+2)} \underbrace{\tau_k^m(1) \cdots \tau_k^m(1)}_{length \ W_k(m+1)}.$$

We first prove that if  $n = W_k(m)$  for some integer m, then  $t_k(n) = 2$  if and only if the valid W-expansion of n + 1 ends with an odd number of 0's.

- if  $n = W_k(0) = 0$ , then  $t_k(0) = 1$  except for k = 1 since  $t_0(0) = 2$ . The valid W-expansion of 0 + 1 = 1 is  $[1]_W$  if  $k \ge 2$  and  $[10]_W$  if k = 1 (since  $[1]_W$  is not valid in this case);

- if n belongs to  $\{W_k(1), W_k(2)\} = \{1, k\}$ , we have that n+1 belongs to  $\{2, k+1\}$ . Hence either k = 1, thus n + 1 = 2 whose valid W-expansion is  $[11]_W$ ; or k = 2and  $n + 1 \in \{2, 3\}$ , thus  $2 = [10]_W$  and  $3 = [11]_W$ ; or  $k \ge 3$ , thus  $2 = [2]_W$  and  $k+1 = [11]_W$ . So, for  $n \in \{W_k(1), W_k(2)\} = \{1, k\}$ , the valid W-expansion of n+1ends with an odd number of 0's if and only if n + 1 = k = 2. On the other hand  $t_k(1) = 2$  if and only if k = 2, and  $t_k(k) = 1$  for all  $k \ge 1$ .

11

- if  $n = W_k(m)$  for some  $m \ge 3$ , then  $\tau_k^{m-1}(2)$  is followed by 1, so that  $t_k(n) = 1$ , and  $n + 1 = W_k(m) + 1$  ends with 1, thus with an even number of 0's

Now we prove by induction on  $m \ge 1$  that, for all  $n \in [0, W_k(m)]$ ,  $t_k(n) = 2$  if and only if the valid W-expansion of n+1 ends with an odd number of 0's. Note that, from what precedes, it suffices to prove the claim for  $n \in [0, W_k(m))$ :

- For m = 1, hence  $W_k(1) = 1$ , the set of relevant n is empty.

- For m = 2, hence  $W_k(2) = k$ , the set of relevant n that do not already satisfy  $n < W_k(1)$  is  $\{1, 2, \ldots, k-1\}$ ; thus n+1 belongs to  $\{2, \ldots, k\}$  and its valid W-expansion either belongs to  $\{[10]_W\}$  if k = 2, or it belongs to  $\{[2]_W, \ldots, [k - k]_W\}$  $1_{W}$ ,  $[10]_{W}$  if k > 2. In both cases there is only one such W-expansion ending with an odd number of 0's, namely  $[10]_W = 2$ , giving n = k - 1. Since the prefix of length  $W_k(1) = k$  of  $\tau_k^{\infty}(1)$  is  $1^{k-1}2$ , we are done with the case m = 2.

- Now suppose that our claim holds for m+2 (hence also for m and m+1 since  $W_k$  is increasing) for some m > 0. We want to prove it for m + 3. Looking at the decomposition into blocks in Eq. (13), we see that n belongs to one of (2k+1)blocks. If  $n < W_k(m+3)$  belongs to one of the first k blocks of length  $W_k(m+2)$ that compose  $\tau_k^{m+2}(1)$ , then there exists  $j \in [1, k-1]$  such that n belongs to  $[jW_k(m+2), (j+1)W_k(m+2) - 1]$ . Thus  $n = jW_k(m+2) + \ell$ , where  $\ell$  belongs to  $[0, W_k(m+2) - 1]$ . Furthermore  $t_k(n) = 2$  if and only if  $t_k(\ell) = 2$ . We distinguish between the case where  $\ell + 1 = W_k(m+2)$  and the case where  $\ell + 1 < W_k(m+2)$ . If  $\ell + 1 = W_k(m+2)$ , the valid W-expansion of  $\ell + 1$  is  $[10^{m+1}]_W$  if k > 1 or m > 0, and  $[11]_W$  if k = 1 and m = 0. The valid W-expansion of  $n + 1 = (j + 1)W_k(m + 2)$ then ends with the same number of 0's as the valid W-expansion of  $\ell + 1$ . If  $\ell + 1 < W_k(m+2)$ , then by the induction hypothesis  $\ell + 1 = [x_r x_{r-1} \cdots x_1]_W$  with r < m+2, and the valid W-expansion  $[x_r x_{r-1} \cdots x_1]_W$  ends with an odd number of 0's if and only if  $t_k(\ell) = 2$ . The valid W-expansion of (n+1) is clearly equal  $0 \cdots 0$  $x_r x_{r-1} \cdots x_1 |_W$ : it ends with an odd number of 0's if and to [j]m + 1 - r terms

only if this is also the case for  $[x_r x_{r-1} \cdots x_1]_W$  (except possibly if n = 0, but this case has already been dealt with).

- It remains to study what happens when n belongs to the last (k+1) blocks in Eq. (13). The proof is tedious and works in exactly the same way, so that we omit it. 

Now we prove one last lemma before our non-automaticity theorem.

**Lemma 17.** For  $\ell, r, n \geq 1$  define the integers  $b_k(l, n)$  and  $c_k(\ell, r, r)$  by

$$b_k(\ell, n) = [(10^{2\ell-1})^n]_W$$
  
=  $\sum_{1 \le j \le n} \frac{(k+1)^{2j\ell} - 1}{k+2} = \frac{(k+1)^{2\ell(n+1)} - (k+1)^{2\ell}}{(k+2)((k+1)^{2\ell} - 1)} - \frac{n}{k+2}$   
 $c_k(\ell, r, n) = b_k(\ell, n) - b_k(\ell, r) - (k+1)W_k(2\ell).$ 

Then the following properties hold.

(i) For all  $n \ge 1$  we have  $b_k(\ell, n) = (k+1)^{2\ell} b_k(\ell, n-1) + n \frac{(k+1)^{2\ell}-1}{k+2}$ . (ii) For all  $\ell \ge 1$  we have  $\frac{(k+1)^{2\ell}-1}{k+2} \mid b_k(\ell, n)$ . (iii) For all  $\ell, n \ge 1$ , we have  $t_k(b_k(\ell, n) - 1) = 2$ .

- (iv) For all  $\ell, r \geq 1$ , and n sufficiently large, we have  $t_k(c(\ell, r, n) 1) = 1$ .
- (v) For all  $\ell \geq 1$ , for all  $c \geq 0$ , and for all  $i \in [0, (k+1)^{2c})$ , there exist infinitely many  $r \ge 0$  such that  $b(\ell, r) \equiv i \pmod{(k+1)^{2c}}$ .

*Proof.* Claims (i) and (ii) are clear. To prove Claim (iii), note the valid W-expansion  $b(\ell, n) = [(1 \ 0^{2\ell-1})^n]_W$  and use Proposition 16. For Claim (iv), we first write (note that some W-expansions below are not valid W-expansions)

$$\begin{aligned} c_k(\ell, r, n) &= [(1 \ 0^{2\ell-1})^n]_W - [(1 \ 0^{2\ell-1})^r]_W - (k+1)[1 \ 0^{2\ell-1}]_W \\ &= [(1 \ 0^{2\ell-1})^{n-r} \ 0^{2\ell r}]_W - (k+1)[1 \ 0^{2\ell-1}]_W \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 1 \ 0^{2\ell-1+2\ell r}]_W - (k+1)[1 \ 0^{2\ell-1}]_W \quad (\text{using Lemma 12 (iv)}) \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell-1+2\ell r}]_W - (k \ 1)[1 \ 0^{2\ell-1}]_W \quad (\text{using Lemma 12 (iv)}) \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell r-1} \ 0 \ k^{2\ell-1}]_W - [1 \ 0^{2\ell-1}]_W \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell r-1} \ 0 \ k^{2\ell-1}]_W - [0 \ k^{2\ell-1}]_W \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell r-1} \ 0 \ k^{2\ell-1}]_W - [0 \ k^{2\ell-1}]_W \quad (\text{using Lemma 12 (iv)}) \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell r-1} \ 0 \ k^{2\ell-1}]_W - [0 \ k^{2\ell-1}]_W \quad (\text{using Lemma 12 (iv)}) \\ &= [(1 \ 0^{2\ell-1})^{n-r-1} \ 0 \ k^{2\ell r-1} \ 0^{2\ell}]_W. \end{aligned}$$

Since this last W-expansion is valid, Proposition 16 yields that  $t(c_k(\ell, r, n) - 1) = 1$ .

Finally to prove (v), we note that both (k+2) and  $(k+1)^{2\ell+1}$  are prime to k+1. Thus  $b(k,r) \equiv i \pmod{(k+1)^{2c}}$  holds if and only if

$$r((k+1)^{2\ell}-1) - (k+1)^{2\ell(r+1)} + (k+1)^{2\ell} \equiv -i(k+2)((k+1)^{2\ell}-1) \pmod{(k+1)^{2\ell}}.$$
  
This holds for r sufficiently large and congruent to  $-i(k+2) - (k+1)^{2\ell}((k+1)^{2\ell}-1)^{-1} \pmod{(k+1)^{2\ell}}.$ 

Now we are ready for the non-automaticity theorem (Theorem 3).

Proof of Theorem 3. As proved in Lemma 11, it suffices to show that  $\tau_k^{\infty}(1)$  is not (k+1)-automatic. Recall that this is equivalent to saying that its (k+1)-kernel is not finite, where the (k+1)-kernel of the sequence  $\tau_k^{\infty}(1) = (t_k(n))_{n\geq 0}$  is the set of subsequences

$$\left\{ \left( t_k ((k+1)^a n + b) \right)_{n \ge 0} \mid a \ge 0, b \in [0, (k+1)^a - 1] \right\}.$$

Since  $t_k((k+1)^{2c}n-1) = t_k((k+1)^{2c}(n-1) + (k+1)^{2c} - 1)$  for  $n \ge 1$ , it suffices to prove that, for all integers c, c' with  $0 \le c < c'$ , the sequences  $(t_k((k+1)^{2c}n-1))_{n\ge 0}$ and  $(t_k((k+1)^{2c'}n-1))_{n\ge 0}$  are distinct. Let  $\ell = c'-c$ . From Lemma 17 (v) applied to  $i \equiv (k - (k+1)^{2\ell+1})(k+2)^{-1} \pmod{(k+1)^{2c}}$ , there exist infinitely many r such that  $b_k(\ell, r) \equiv (k - (k+1)^{2\ell+1})(k+2)^{-1} \pmod{(k+1)^{2c}}$ . Let  $m = k+1+\frac{b_k(\ell,r)(k+2)}{(k+1)^{2\ell}-1}$ . This is an integer by Lemma 17 (ii) and

$$m-1 = k + \frac{b_k(\ell, r)(k+2)}{(k+1)^{2\ell} - 1} \equiv -\frac{(k+1)^{2\ell}}{(k+1)^{2\ell} - 1} \pmod{(k+1)^{2c}}.$$

But (see Lemma 16) the expression of

$$b_k(\ell, m-1) = \frac{(k+1)^{2\ell m} - (k+1)^{2\ell}}{(k+2)((k+1)^{2\ell} - 1)} - \frac{m-1}{k+2}$$

shows that, provided  $\ell m \geq c$  (which holds if r is sufficiently large), we have

$$b_k(\ell, m-1) \equiv -\frac{(k+1)^{2\ell}}{(k+2)((k+1)^{2\ell}-1)} - \frac{m-1}{k+2} \pmod{(k+1)^{2c}}.$$

Thus  $b_k(\ell, m-1) \equiv 0 \pmod{(k+1)^{2c}}$ . Let  $j = b_k(\ell, m-1)/(k+1)^{2c}$ . Using Lemma 17 (i) we have

$$(k+1)^{2\ell}b_k(\ell,m-1) = b_k(\ell,m) - \frac{(k+1)^{2\ell}-1}{k+2}m$$

$$= b_k(\ell, m) - \frac{(k+1)^{2\ell} - 1}{k+2} \left( k + 1 + \frac{b_k(\ell, r)(k+2)}{(k+1)^{2\ell} - 1} \right)$$
$$= b_k(\ell, m) - (k+1) \frac{(k+1)^{2\ell} - 1}{k+2} - b_k(\ell, r)$$
$$= b_k(\ell, m) - b_k(\ell, r) - (k+1) W_k(2\ell)$$
$$= c_k(\ell, r, m).$$

By Lemma 17 (iii) and (iv), we have

 $t_k(b_k(m-1)-1) = 2$  and  $t_k((k+1)^{2\ell}b_k(\ell,m-1)-1) = t_k(c_k(\ell,r,m)-1) = 1.$ Thus

$$t_k((k+1)^{2c}j-1) = t_k(b_k(\ell, m-1) - 1) = 2,$$

while

$$t_k((k+1)^{2c'}j-1) = t_k((k+1)^{2c'-2c}b_k(\ell,m-1)-1)$$
  
=  $t_k((k+1)^{2\ell}b_k(\ell,m-1)-1)$   
=  $t_k(c_k(\ell,r,m)-1) = 1,$ 

which shows that the sequences  $(t_k((k+1)^{2c}n)-1))_{n\geq 0}$  and  $(t_k((k+1)^{2c'}n)-1))_{n\geq 0}$  are distinct.

As recalled above, the sequence  $1 \tau_1^{\infty}(1) = 121122\cdots$  is also the iterative fixed point of a morphism, namely  $1 \tau_1^{\infty}(1) = \sigma_1^{\infty}(1)$  where  $\sigma_1$  is defined by  $\sigma_1(1) = 121$ ,  $\sigma_1(2) = 12221$ . One can ask whether a similar property holds for  $1 \tau_k^{\infty}(1)$ . The following proposition answers this question.

**Proposition 18.** For  $k \ge 1$ , define the morphism  $\sigma_k$  by

$$\sigma_k(1) = 1 \ (1^{k-1} \ 2)^k \ 1^k, \ \sigma_k(2) = 1 \ (1^{k-1} \ 2)^{2k+1} \ 1^k.$$

Then  $\sigma_k$  has an iterative fixed point that satisfies

$$\sigma_k^{\infty}(1) = 1 \ \tau_k^{\infty}(1).$$

*Proof.* First we note that

 $(1^{k-1} 2)^k 1^k \sigma_k(1) = \tau_k^2(1) (1^{k-1} 2)^k 1^k \text{ and } (1^{k-1} 2)^k 1^k \sigma_k(2) = \tau_k^2(2) (1^{k-1} 2)^k 1^k$ so that for words  $z \in \{1, 2\}^*$  we have

$$\sigma_k(1z) = \sigma_k(1)\sigma_k(z) = 1 \ (1^{k-1} \ 2)^k \ 1^k \ \sigma_k(z) = 1 \ \tau_k^2(z) \ (1^{k-1} \ 2)^k \ 1^k.$$

Applying this equality to  $z = z_k(m)$  the prefix of length m of  $\tau_k^{\infty}(1)$ , and letting m tend to infinity, we obtain

$$\sigma_k(1 \ \tau_k^{\infty}(1)) = 1 \ \tau_k^2(\tau_k^{\infty}(1)) = 1 \ \tau_k^{\infty}(1)$$

Hence 1  $\tau_k^{\infty}(1)$  is a fixed point of  $\sigma_k$ , thus the iterative fixed point of  $\sigma_k$ .

4.1. **Miscellanea.** The sequence  $\sigma_1^{\infty}(1)$  appears in several places in the literature: it was used by Brlek [7] for determining the block-complexity of the Thue-Morse sequence. We have already cited [2] where it is related to an Indian *kolam*. It also occurs in [1] in relation to a piecewise affine map. Finally we want to point out that the sequence of integers  $(W_k(n))_{n\geq 1}$  actually occurred (under another name) in [3]; the following result that relates the sequence  $\tau_k^{\infty}(1)$ , the sequence  $(W_k(n))_{n\geq 0}$ , and a "something-free" set, is proven there:

13

**Theorem 19** ([3]). Let  $\sigma_k^{\infty}(1) = (s_k(n))_{n\geq 0}$ . Define  $c_k(n) = \sum_{0\leq j\leq n} s_k(n)$ . Then the sequence  $\mathbf{c} = (c_k(n))_{n\geq 0}$  has the property that n belongs to  $\mathbf{c}$  if and only if (k+1)n does not belong to  $\mathbf{c}$ . Furthermore it admits the following generating function

$$\sum_{n \ge 0} c_k(n) x^n = \frac{1}{1-x} \prod_{j \ge 1} \frac{1 - x^{(k+1)W_k(j)}}{1 - x^{W_k(j)}} \cdot$$

## 5. Sum-free sets generated by certain Sturmian sequences

Let  $\mathbf{t} = (t_n)_{n\geq 0}$  be a sequence on  $\{0,1\}$  and let  $S = (s_n)_{n\geq 1}$  be the sum-free set corresponding to  $\mathbf{t}$ . Recall that  $\mathbf{v} = (v_n)_{n\geq 1}$  is the sequence defined by Eq. (1) according to S.

**Lemma 20.** If  $t_0 = t_1 = 1$  and '00' does not occur in **t**, then for all  $n \ge 1$ , we have  $v_n = *$  if and only if n is even.

*Proof.* Note that  $t_0 = t_1 = 1$ . By the construction of  $(v_n)_{n>1}$ , we have

When  $t_2 = 1$ , then  $v_5 = 1$  and  $v_8 = v_{10} = *$ . We have  $s_1 = 1$ ,  $s_2 = 3$  and  $s_3 = 5$ . When  $t_2 = 0$ , since '00' does not occur in  $(t_n)_{n\geq 0}$ , we have  $t_3 = 1$ . Hence  $v_5 = 0$ ,  $v_7 = 1$ , and  $v_8 = v_{10} = *$ . We have  $s_1 = 1$ ,  $s_2 = 3$  and  $s_3 = 7$ . So the result holds for all  $n \leq 6$ . Assume that the result holds for all m < n. We prove it for m = n.

Case 1: n = 2k. If  $v_{2k-1} = 1$ , then  $v_{2k} = *$  since  $v_1 = 1$ . If  $v_{2k-1} = 0$ , then  $v_{2k-2} = *$  by the inductive hypothesis and  $v_{2k-3} = 1$  since '00' does not occur. Note that  $v_3 = 1$ , we also have  $v_{2k} = *$ .

Case 2: n = 2k + 1. By the inductive hypothesis, for all  $\ell \leq 2k$ , if  $v_{\ell} = 1$ , then  $\ell$  is odd. If  $v_{2k+1} = *$ , then there exist  $i, j \leq 2k$  such that  $v_i = v_j = 1$  and 2k+1 = i+j. This contradicts to the fact that both i and j are odd. So  $v_{2k+1} \neq *$ .

Now we see the result is valid for n and our lemma follows from induction.  $\Box$ 

Remark 21. Under the assumption of Lemma 20, we see that  $(s_n)_{n\geq 1}$  are odd numbers.

Now we shall discuss the sum-free set  $S = (s_n)_{n\geq 1}$  generated by a Sturmian sequence **t** with  $t_0 = t_1 = 1$ . Note that '00' cannot occur in **t** (namely, a Sturmian sequence has  $\ell + 1$  factors of length  $\ell$ ; since it is not periodic, it always contain '01' and '10', thus if '11' is a factor, the Sturmian sequence has exactly 3 factors of length 2, i.e., '00', '01' and '10'). Let

$$(d_n)_{n\geq 1} := (s_{n+1} - s_n)_{n\geq 1}$$

be the difference sequence of S. It is interesting to see that the difference sequence is still a Sturmian sequence. We restate our Theorem 4 as follow.

**Theorem 4.** If **t** is a Sturmian sequence with  $t_0 = t_1 = 1$ , then the sequence  $(d_n)_{n>1}$  is a Sturmian sequence.

*Proof.* Note that  $\mu_n$  is the number of zeros between the *n*-th and the (n + 1)-th occurrences of '1' in **t**. Write  $\mathbf{u} := (\mu_n)_{n \ge 1}$ . Since **t** is a Sturmian sequence in which '00' does not occur (see above), we have  $\mu_n \in \{0, 1\}$ . Moreover,

$$\mathbf{t} = \varphi(\mathbf{u})$$

where  $\varphi$  is the morphism  $0 \to 1$  and  $1 \to 10$ . By [14, Corollary 2.3.3], we know that **u** is also a Sturmian sequence.

From Lemma 20, we obtain that for all  $n \ge 1$ ,

$$\alpha_n = \begin{cases} 1, & \text{if } \mu_n = 0, \\ 2, & \text{if } \mu_n = 1. \end{cases}$$

Then by Eq. (2), for all  $n \ge 1$ ,

$$d_n = \mu_n + \alpha_n + 1 = 2(\mu_n + 1) \in \{2, 4\}.$$

This implies that the difference sequence  $(d_n)_{n\geq 1}$  is the image of  $(\mu_n)_{n\geq 1}$  under the coding  $0 \to 2$  and  $1 \to 4$ . So  $(d_n)_{n\geq 1}$  is a Sturmian sequence.

## 6. Subword complexity

We close with a conjecture about the subword complexity of the infinite fixed points of the morphisms  $\tau_k$ . The subword complexity is the function counting the number of distinct factors of length n.

**Conjecture 22.** Let  $(f_n)_{n\geq 1}$  be the subword complexity of  $\tau_k^{\infty}(1)$ , and define  $d_n = f_{n+1} - f_n$  for  $n \geq 1$ . Then  $d_1 d_2 d_3 \cdots = 1^{a_1} 2^{a_2} 1^{a_3} 2^{a_4} \cdots$ , where

$$a_{1} = k - 1;$$
  

$$a_{2} = k;$$
  

$$a_{2n} = a_{2n-1} + a_{2n-2}, \quad n \ge 2;$$
  

$$a_{2n+1} = ka_{2n} + k(-1)^{n}, \quad n \ge 1$$

For example, consider the case k = 3. Then

$$(f_n)_{n\geq 1} = (2, 3, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 18, 20, 22, \ldots)$$
  

$$(d_n)_{n\geq 1} = (1, 1, 2, 2, 2, 1, 1, 1, 1, 1, 1, 2, \ldots),$$
  

$$(a_n)_{n\geq 1} = (2, 3, 6, 9, 30, 39, 114, 153, \ldots).$$

### Acknowledgements

The research of Z.-X. Wen was partially supported by NSFC (Nos. 11431007, 11871295). The research of W. Wu was partially supported by China Scholarship Council (File No. 201906155024) and Guangdong Natural Science Foundation (Nos. 2018B0303110005, 2018A030313971).

#### References

- S. Akiyama, H. Brunotte, A. Pethő, and W. Steiner, Periodicity of certain piecewise affine planar maps, Tsukuba J. Math. 32 (2008), 197–251.
- [2] G. Allouche, J.-P. Allouche, and J. Shallit, Kolam indiens, dessins sur le sable aux îles Vanuatu, courbe de Sierpiński et morphismes de monoïde, Ann. Inst. Fourier, Grenoble, 56 (2006), 2115–2130.
- [3] J.-P. Allouche, A. Arnold, J. Berstel, S. Brlek, W. Jockusch, S. Plouffe, and B. E. Sagan, A relative of the Thue-Morse sequence, Discrete Math., 139 (1995), 455–461.
- [4] J.-P. Allouche and J. Shallit, The ring of k-regular sequences, Theoret. Comput. Sci., 98 (1992), 163–197.
- [5] J.-P. Allouche and J. Shallit, The ring of k-regular sequences, II, Theoret. Comput. Sci., 307 (2003), 3–29.
- [6] J.-P. Allouche and J. Shallit, Automatic sequences. Theory, applications, generalizations, Cambridge University Press, Cambridge, 2003.
- [7] S. Brlek, Enumeration of factors in the Thue-Morse word, Discrete Appl. Math., 24 (1989), 83–96.
- [8] N. J. Calkin and S. R. Finch, Conditions on periodicity for sum-free sets, Exp. Math., 5 (2) (1996), 131–137.
- [9] N. J. Calkin and P. Erdős, On a class of aperiodic sum-free sets, Math. Proc. Camb. Phil. Soc., 120 (1996), 1–5.

- [10] N. J. Calkin, S. R. Finch, and T. B. Flowers, Difference density and aperiodic sum-free sets, INTEGERS: Electronic J. Comb. Number Theory, 5 (2) (2005), Paper A03.
- [11] P. J. Cameron, Cyclic automorphisms of a countable graph and random sum-free sets, Graphs and Combin., 1 (1985), 129–135.
- [12] D. Damanik, Local symmetries in the period-doubling sequence, Discrete Appl. Math., 100 (2000), 115–121.
- [13] F. Durand, Cobham's theorem for substitutions, J. Eur. Math. Soc., 13 (2011), 1799–1814.
- [14] M. Lothaire, Algebraic combinatorics on words, Cambridge University Press, 2002.
- [15] G. Payne, A solution to a problem of Cameron on sum-free complete sets, J. Combin. Theory, Series A, 70 (1995), 305–312.
- [16] N. Rampersad and M. Stipulanti, The formal inverse of the period-doubling sequence, J. Integer Seq. 21 (2018), Paper 18.9.1.
- [17] J. Shallit, The fixed point of  $1 \rightarrow 121$ ,  $1 \rightarrow 12221$  is not 2-automatic, Unpublished manuscript, dated September 10, 1992.
- [18] N. J. A. Sloane et al., On-Line Encyclopedia of Integer Sequences, electronically available at https://oeis.org.
- [19] Z.-X. Wen, J.-M. Zhang, and W. Wu, On the regular sum-free sets, Eur. J. Combin., 49 (2015), 42–56.

(J.-P. Allouche) CNRS, IMJ-PRG, Sorbonne Université, 4 Place Jussieu, F-75252 Paris Cedex 05, France

 $E\text{-}mail\ address:\ \texttt{jean-paul.allouche@imj-prg.fr}$ 

(J. Shallit) School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada

E-mail address: shallit@uwaterloo.ca

(Z.-X. Wen) School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, 430074, P.R. China.

E-mail address: zhi-xiong.wen@hust.edu.cn

(W. Wu) School of Mathematics, South China University of Technology, Guangzhou 510641, P.R. China.

Department of Mathematics and Statistics, P.O. Box 68 (Pietari Kalmin katu 5), FI-00014 University of Helsinki, Finland

 $\textit{E-mail address}, \ \texttt{Corresponding author: wwwen@scut.edu.cn}$ 

(J.-M. Zhang) School of Science, Wuhan Institute of Technology, Wuhan 430205, P.R. China.

 $E\text{-}mail \ address: \texttt{zhangjiemeng@wit.edu.cn}$