

# Infinite Self-Shuffling Words

Émilie Charlier<sup>a</sup>, Teturo Kamae<sup>b</sup>, Svetlana Puzynina<sup>c,d,1</sup>, Luca Q. Zamboni<sup>c,e,2</sup>

<sup>a</sup>*Département de Mathématique, Université de Liège, Belgium*

<sup>b</sup>*Advanced Mathematical Institute, Osaka City University, Japan*

<sup>c</sup>*FUNDIM, University of Turku, Finland*

<sup>d</sup>*Sobolev Institute of Mathematics, Novosibirsk Russia*

<sup>e</sup>*Université de Lyon, Université Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 boulevard du 11 novembre 1918, F69622 Villeurbanne Cedex, France*

---

## Abstract

In this paper we introduce and study a new property of infinite words: An infinite word  $x \in \mathbb{A}^{\mathbb{N}}$ , with values in a finite set  $\mathbb{A}$ , is said to be *k-self-shuffling* ( $k \geq 2$ ) if  $x$  admits factorizations:  $x = \prod_{i=0}^{\infty} U_i^{(1)} \dots U_i^{(k)} = \prod_{i=0}^{\infty} U_i^{(1)} = \dots = \prod_{i=0}^{\infty} U_i^{(k)}$ . In other words, there exists a shuffle of  $k$ -copies of  $x$  which produces  $x$ . We are particularly interested in the case  $k = 2$ , in which case we say  $x$  is self-shuffling. This property of infinite words is shown to be independent of the complexity of the word as measured by the number of distinct factors of each length. Examples exist from bounded to full complexity. It is also an intrinsic property of the word and not of its language (set of factors). For instance, every aperiodic uniformly recurrent word contains a non self-shuffling word in its shift orbit closure. While the property of being self-shuffling is a relatively strong condition, many important words arising in the area of symbolic dynamics are verified to be self-shuffling. They include for instance the Thue-Morse word  $t = t_0 t_1 t_2 \dots \{0, 1\}^{\mathbb{N}}$  where  $t_n$  is the sum modulo 2 of the digits in the binary expansion of  $n$ . As another example we show that all Sturmian words of slope  $\alpha \in \mathbb{R} \setminus \mathbb{Q}$  and intercept  $0 < \rho < 1$  are self-shuffling (while those of intercept  $\rho = 0, 1$  are not). Our characterization of self-shuffling Sturmian words can be interpreted arithmetically in terms of a dynamical embedding and defines an arithmetic process we call the *stepping stone model*. One important feature of self-shuffling words stems from its morphic invariance: The morphic image of a self-shuffling word is self-shuffling. This provides a useful tool for showing that one word is not the morphic image of another. In addition to its morphic invariance, this new notion has other unexpected applications particularly in the area of substitutive dynamical systems. For example, as a consequence of our characterization of self-shuffling Sturmian words, we recover a number

---

*Email addresses:* [echarlier@ulg.ac.be](mailto:echarlier@ulg.ac.be) (Émilie Charlier), [kamae@apost.plala.or.jp](mailto:kamae@apost.plala.or.jp) (Teturo Kamae), [svepuz@utu.fi](mailto:svepuz@utu.fi) (Svetlana Puzynina), [lupastis@gmail.com](mailto:lupastis@gmail.com) (Luca Q. Zamboni)

<sup>1</sup>Partially supported by the Academy of Finland under grant 251371, by Russian Foundation of Basic Research (grants 12-01-00448 and 12-01-00089).

<sup>2</sup>Partially supported by a FiDiPro grant (137991) from the Academy of Finland and by ANR grant *SUBTILE*.

theoretic result, originally due to Yasutomi, on a classification of pure morphic Sturmian words in the orbit of the characteristic.

*Keywords:* Word shuffling, Sturmian words, Lyndon words, morphic words, Thue-Morse word.

*2000 MSC:* 68R15

## 1. Introduction

Let  $\mathbb{A}$  be a finite non-empty set. We denote by  $\mathbb{A}^*$  (resp.  $\mathbb{A}^{\mathbb{N}}$ ) the set of all finite (resp. infinite) words  $u = x_0x_1x_2 \cdots$  with  $x_i \in \mathbb{A}$ .

Given  $k$  finite words  $x^{(1)}, x^{(2)}, \dots, x^{(k)} \in \mathbb{A}^*$  we let  $\mathcal{S}(x^{(1)}, x^{(2)}, \dots, x^{(k)}) \subseteq \mathbb{A}^*$  denote the collection of all words  $z$  for which there exists a factorization

$$z = \prod_{i=0}^n U_i^{(1)} U_i^{(2)} \cdots U_i^{(k)}$$

with each  $U_i^{(j)} \in \mathbb{A}^*$  and with  $x^{(j)} = \prod_{i=0}^n U_i^{(j)}$  for each  $1 \leq j \leq k$ . Intuitively,  $z$  may be obtained as a *shuffle* of the words  $x^{(1)}, x^{(2)}, \dots, x^{(k)}$ . For instance, it is readily checked that  $011100110 \in \mathcal{S}(0010, 101, 11)$ . Analogously, given  $k$  infinite words  $x^{(1)}, x^{(2)}, \dots, x^{(k)} \in \mathbb{A}^{\mathbb{N}}$  we define  $\mathcal{S}(x^{(1)}, x^{(2)}, \dots, x^{(k)}) \subseteq \mathbb{A}^{\mathbb{N}}$  to be the collection of all infinite words  $z$  for which there exists a factorization

$$z = \prod_{i=0}^{\infty} U_i^{(1)} U_i^{(2)} \cdots U_i^{(k)}$$

with each  $U_i^{(j)} \in \mathbb{A}^*$  and with  $x^{(j)} = \prod_{i=0}^{\infty} U_i^{(j)}$  for each  $1 \leq j \leq k$ .

Finite word shuffles were extensively studied in [9]. Given  $x \in \mathbb{A}^*$ , it is generally a difficult problem to determine whether there exists  $y \in \mathbb{A}^*$  such that  $x \in \mathcal{S}(y, y)$  (see Open Problem 4 in [9]). The problem has recently been shown to be NP-complete for sufficiently large alphabets [4, 14]. However, in the context of infinite words, this question is essentially trivial: In fact, it is readily verified that if  $x \in \mathbb{A}^{\mathbb{N}}$  and each symbol  $a \in \mathbb{A}$  occurring in  $x$  occurs an infinite number of times in  $x$ , then there exists at least one (and typically infinitely many)  $y \in \mathbb{A}^{\mathbb{N}}$  with  $x \in \mathcal{S}(y, y)$ . Instead, in the framework of infinite words, a far more delicate question is the following:

**Question 1.** *Given  $x \in \mathbb{A}^{\mathbb{N}}$ , does there exist an integer  $k \geq 2$  such that  $x \in \mathcal{S}(\underbrace{x, x, \dots, x}_k)$ ?*

If such a  $k$  exists, we say  $x$  is *k-self-shuffling*. In case  $k = 2$ , we say  $x$  is *self-shuffling*. It is not difficult to see that every self-shuffling word is  $k$ -self-shuffling for each  $k \geq 2$ .

If  $x \in \mathbb{A}^{\mathbb{N}}$  is  $k$ -self-shuffling, then there exists at least one word  $s \in \{1, 2, \dots, k\}^{\mathbb{N}}$  (called the *steering word*) which defines the shuffle. Typically a  $k$ -self-shuffling word  $x$  can be shuffled in more than one way so as to reproduce itself, i.e., may define more than one steering word. In contrast,

every word  $s \in \{1, 2, \dots, k\}^{\mathbb{N}}$  is a steering word for some  $k$ -self-shuffling word. Moreover, if  $s$  begins in a block of the form  $a^r b$  with  $a$  and  $b$  distinct symbols in  $\{1, 2, \dots, k\}$ , then  $s$  is the steering word of a unique (up to word isomorphism) non-constant  $k$ -self-shuffling word  $x(s)$  on an alphabet of size  $r$ . In general, the relationship between properties of  $s$  and  $x(s)$  is a mystery. For instance, it may be that  $s$  is uniformly recurrent and  $x(s)$  not. In this paper, we are primarily interested in  $k$ -self-shuffling words, and less importance is placed on the corresponding steering words. In fact, we mainly focus on self-shuffling words, although many of the results presented here extend to general  $k$ . Thus  $x \in \mathbb{A}^{\mathbb{N}}$  is self-shuffling if and only if  $x$  admits factorizations

$$x = \prod_{i=1}^{\infty} U_i V_i = \prod_{i=1}^{\infty} U_i = \prod_{i=1}^{\infty} V_i$$

with  $U_i, V_i \in \mathbb{A}^+$ .

The simplest class of self-shuffling words consists of all (purely) periodic words  $x = u^\omega$ . We note that if  $x \in \mathbb{A}^{\mathbb{N}}$  is self-shuffling, then every letter  $a \in \mathbb{A}$  occurring in  $x$  must occur an infinite number of times. Thus for instance, the ultimately periodic word  $01^\omega$  is not self-shuffling. On the opposite extreme, we show the existence of self-shuffling words having full complexity. Thus, the property of being self-shuffling is largely independent of the usual (subword) "complexity" of an infinite word as measured by the number of blocks of each given length. Moreover it is also an intrinsic feature of the word and not of its language (or set of factors). For instance, the *Fibonacci word*

$$x = 0100101001001010010100 \dots,$$

defined as the fixed point of the substitution  $0 \mapsto 01, 1 \mapsto 0$ , is self-shuffling while  $0x$  and  $1x$  are not (see §2). More generally, we will show that given any aperiodic uniformly recurrent word  $x \in \mathbb{A}^{\mathbb{N}}$ , the subshift generated by  $x$  always contains at least one point which is not self-shuffling.

While the property of being self-shuffling is quite strong, many important words arising in symbolic dynamics turn out to be self-shuffling. This includes the famous *Thue-Morse word*

$$\mathbf{T} = 0110100110010110100101100110100110010110 \dots$$

whose origins go back to the beginning of the last century with the works of the Norwegian mathematician Axel Thue [16]. The  $n$ th entry  $t_n$  of  $\mathbf{T}$  is defined as the sum modulo 2 of the digits in the binary expansion of  $n$ . The Thue-Morse word is linked to many different areas of mathematics: from discrete mathematics to number theory to differential geometry (see for example [1, 2]). While much is already known on the combinatorial properties of the Thue-Morse word, proving that  $\mathbf{T}$  is self-shuffling is less straightforward than expected.

Sturmian words constitute another important class of aperiodic self-shuffling words. Sturmian words are infinite words having exactly  $n + 1$  factors of length  $n$  for each  $n \geq 1$ . Their origins can be traced back to the astronomer J. Bernoulli III in 1772. Sturmian words arise naturally in different areas of mathematics including combinatorics, algebra, number theory, ergodic theory, dynamical

systems and differential equations. They are also of great importance in theoretical physics (as basic examples of 1-dimensional quasicrystals) and in theoretical computer science where they are used in computer graphics as digital approximation of straight lines. Sturmian words are regarded as the most basic non-ultimately periodic infinite words. Perhaps the most famous and well studied Sturmian word is the Fibonacci word defined above. In the 1940's, Hedlund and Morse [12] showed that each Sturmian word is the symbolic coding of the orbit of a point  $x$  (called the *intercept*) on the unit circle under a rotation by an irrational angle  $\alpha$  (called the *slope*), where the circle is partitioned into two complementary intervals, one of length  $\alpha$  and the other of length  $1 - \alpha$ . Conversely each such coding defines a Sturmian word. It is well known that the dynamical/ergodic properties of the system, as well as the combinatorial properties of the associated Sturmian word, hinge on the arithmetical/Diophantine qualities of the angle  $\alpha$  given by its continued fraction expansion. As in the case of the Fibonacci word, we show that for every irrational number  $\alpha$ , all (uncountably many) Sturmian words of slope  $\alpha$  and intercept  $\rho$  are self-shuffling except for the two Sturmian words corresponding to  $\rho = 0$ .

In this paper, we derive a number of necessary (and in some cases sufficient) conditions for a word to be self-shuffling. For instance, if a word  $x$  is self-shuffling, then  $x$  begins in only finitely many Abelian border-free words. As an application of this we show that the well-known *paper-folding word* is not self-shuffling. Infinite *Lyndon words* [15] constitute another class of words which are shown not to be self-shuffling. A word  $x \in \mathbb{A}^{\mathbb{N}}$  is said to be Lyndon if there exists an order on  $\mathbb{A}$  with respect to which  $x$  is lexicographically smaller than each of its tails. We prove that if  $x$  is Lyndon, then any  $z \in \mathcal{S}(x, x)$  is lexicographically smaller than  $x$ , from which it follows immediately that  $x$  is not self-shuffling. While this may appear rather intuitive, our proof of this fact is both long and delicate.

An important feature of the self-shuffling property stems from its invariance under the action of a morphism: The morphic image of a self-shuffling word is again self-shuffling. Many important classes of words (e.g., Sturmian words, pure morphic words, and Toeplitz words) are not preserved by the action of an arbitrary morphism. This invariance provides a useful tool for showing that one word is not the morphic image of another. For instance, the paper-folding word is not the morphic image of any self-shuffling word. However this application requires knowing a priori whether a given word is or is not self-shuffling. In general, to show that a word is self-shuffling, one must actually exhibit a shuffle. Self-shuffling words have other unexpected applications particularly in the study of substitutive dynamical systems. For instance, as an almost immediate consequence of our characterization of self-shuffling Sturmian words, we recover a result, originally proved by Yasutomi via number theoretic methods, which gives a characterization of pure morphic Sturmian words in the orbit of the characteristic.

The paper is organized as follows: In §2 we establish some general properties of  $k$ -self-shuffling words and along the way give various examples and non-examples. Here we also consider self-shuffling words which are fixed points of primitive substitutions. Under some additional assumptions on the substitution, we deduce the self shuffling property for other points in the shift orbit of  $x$ . For instance, if  $\tau$  is a primitive substitution having a unique periodic point  $x$ , then if  $x$  is

self-shuffling then the same is true of each shift of  $x$ . In §3 we establish the self-shuffling of the Thue-Morse word by explicitly constructing a shuffle. Our proof makes use of different morphisms associated with the Thue-Morse word. In §4 we prove that Lyndon words are not self-shuffling. In §5 we obtain a characterization of self-shuffling Sturmian words and derive various applications including Yasutomi's result mentioned above. In §6 we give an arithmetic interpretation of our characterization of self-shuffling Sturmian words in terms of a dynamical embedding of an infinite graph into the dynamical system corresponding to a circle rotation. In this framework we describe an arithmetic process we call the *stepping stone model* which may be of independent interest in the theory of Diophantine approximations. We end the paper with a few open questions.

A preliminary and incomplete version of this paper has been reported at ICALP 2103 conference [6].

## 2. General properties

In this section we develop some basic properties of  $k$ -self-shuffling words. Let  $\mathbb{A}$  be a finite non-empty set. We denote by  $\mathbb{A}^*$  the set of all finite words  $u = x_1x_2 \dots x_n$  with  $x_i \in \mathbb{A}$ . The quantity  $n$  is called the length of  $u$  and is denoted  $|u|$ . For a letter  $a \in \mathbb{A}$ , let  $|u|_a$  denote the number of occurrences of  $a$  in  $u$ . The empty word, denoted  $\varepsilon$ , is the unique element in  $\mathbb{A}^*$  with  $|\varepsilon| = 0$ . We set  $\mathbb{A}^+ = \mathbb{A} - \{\varepsilon\}$ . We denote by  $\mathbb{A}^{\mathbb{N}}$  the set of all one-sided infinite words  $x = x_0x_1x_2 \dots$  with  $x_i \in \mathbb{A}$ . Given  $x = x_0x_1x_2 \dots \in \mathbb{A}^{\mathbb{N}}$  and a finite or infinite subset  $N = \{N_0 < N_1 < N_2 < \dots\} \subseteq \mathbb{N}$ , we put  $x[N] = x_{N_0}x_{N_1}x_{N_2} \dots \in \mathbb{A}^{\mathbb{N}}$ .

**Definition 2.1.** Let  $x \in \mathbb{A}^{\mathbb{N}}$  and  $k \in \{2, 3, \dots\}$ . We say  $x$  is  $k$ -self-shuffling if  $x$  satisfies any one of the following two equivalent conditions:

- $x \in \mathcal{S}(\underbrace{x, x, \dots, x}_k)$ .
- There exists a  $k$ -element partition of  $\mathbb{N}$  into infinite subsets  $N^1, N^2, \dots, N^k$  with  $x[N^i] = x$  for each  $i = 1, \dots, k$ .

In case  $x$  is 2-self-shuffling we say simply that  $x$  is *self-shuffling*. It is evident that if  $x$  is self-shuffling, then  $x$  is  $k$ -self-shuffling for each  $k \geq 2$ . Later we give an example of a 3-self-shuffling word which is not self-shuffling.

If  $x$  is  $k$ -self-shuffling, then there exists a word  $s$  on a  $k$ -letter alphabet which defines or steers the shuffle. We call such a word a *steering word* for the shuffle. More precisely, if  $x \in \mathbb{A}^{\mathbb{N}}$  is  $k$ -self-shuffling, then there exists a  $k$ -element partition of  $\mathbb{N}$  into infinite subsets  $N^1, N^2, \dots, N^k$  with  $x[N^i] = x$  for each  $i = 1, \dots, k$ . The corresponding steering word  $s = s_0s_1s_2 \dots \in \{1, 2, \dots, k\}^{\mathbb{N}}$  is then defined by  $s_n = j \Leftrightarrow n \in N^j$ . In general, a  $k$ -self-shuffling word defines many different steering words, i.e., the shuffle is not unique. We begin with a few simple examples:

**Fibonacci word:** The Fibonacci infinite word

$$x = 0100101001001010010100\dots$$

is defined as the fixed point of the substitution  $\varphi$  given by  $0 \mapsto 01, 1 \mapsto 0$ . It is readily verified that  $\varphi^2(a) = \varphi(a)a$  for each  $a \in \{0, 1\}$ . Whence, writing  $x = x_0x_1x_2\dots$  with each  $x_i \in \{0, 1\}$  we obtain

$$\begin{aligned} x &= x_0x_1x_2\dots = \varphi(x_0)\varphi(x_1)\varphi(x_2)\dots = \varphi^2(x_0)\varphi^2(x_1)\varphi^2(x_2)\dots = \\ &\varphi(x_0)x_0\varphi(x_1)x_1\varphi(x_2)x_2\dots \end{aligned}$$

which shows that  $x$  is self-shuffling. It is readily verified that the corresponding steering word  $s = 00101001001010010100\dots$  is equal to the second shift of  $x$ .

**Period-doubling word:** The period-doubling word

$$x = 01000101010001000100010101\dots$$

is defined as the fixed point of the substitution  $\sigma$  given by  $0 \mapsto 01, 1 \mapsto 00$ . The period doubling word is also an example of a Toeplitz word (see [5]). It is readily verified that  $x$  admits factorizations  $x = \prod_{i=0}^{\infty} U_iV_i = \prod_{i=0}^{\infty} U_i = \prod_{i=0}^{\infty} V_i$  with  $U_0 = 0100, V_0 = 01$ , and  $U_i = \sigma^{i+1}(1), V_i = \sigma^i(1)$  for  $i \geq 1$ . (It suffices to show that each of the above products is fixed by  $\sigma$ ). Thus the period-doubling word is self-shuffling.

The self-shuffling property appears largely independent of the usual subword complexity of an infinite word as measured by the number of factors of each given length. In fact, on one extreme are the purely periodic infinite words all of which are easily seen to be self-shuffling. The following example illustrates the existence of a self-shuffling word having full complexity:

**A recurrent binary self-shuffling word with full complexity:** For each positive integer  $n$ , let  $z_n$  denote the concatenation of all words of length  $n$  in increasing lexicographic order. For example,  $z_2 = 00011011$ . For  $i \geq 0$  put

$$v_i = \begin{cases} z_n, & \text{if } i = n2^{n-1} \text{ for some } n, \\ 0^i1^i, & \text{otherwise,} \end{cases}$$

and define

$$x = \prod_{i=0}^{\infty} X_i = 010100110^3011^30^4010^21^2011^4\dots,$$

where  $X_0 = X_1 = 01, X_2 = 0011$ , and for  $i \geq 3, X_i = 0^i y_{i-2} 1^i$ , where  $y_{i-2} = y_{i-3} v_{i-2} y_{i-3}$ , and  $y_0 = \varepsilon$ . We note that  $x$  is recurrent (i.e., each prefix occurs twice) and has full complexity (since it contains  $z_n$  as a factor for every  $n$ ).

To show that the word  $x$  is self-shuffling, we first show that  $X_{i+1} \in \mathcal{S}(X_i, X_i)$ . Take  $N_i = \{0, \dots, i-1, i+1, \dots, 2^i - i, 2^i - i + v_{i-1}|_1, 2^{i+1} - i - 1\}$ , where  $u|_1$  denotes the positions  $j$  of a word  $u$  in which the  $j$ -th letter  $u_j$  of  $u$  is equal to 1. Then it is straightforward to see that  $X_i = X_{i+1}[N_i] = X_{i+1}[\{1, \dots, 2^{i+1}\} \setminus N_i]$ . The self-shuffle of  $x$  is built in a natural way concatenating shuffles of  $X_i$  starting with  $U_0 = V_0 = 01$ , so that  $X_0 \cdots X_{i+1} \in \mathcal{S}(X_0 \cdots X_i, X_0 \cdots X_i)$ .

The property of being self-shuffling is quite strong. Nevertheless, every infinite word  $s = s_0s_1s_2 \dots \in \{1, 2, \dots, k\}^{\mathbb{N}}$  is a steering word for some  $k$ -self-shuffling word. To see this, we define  $\ell: \mathbb{N} \rightarrow \mathbb{N}$  by  $\ell(n) = |s_0s_1 \cdots s_n|_{s_n} - 1$ . Let  $\sim$  denote the equivalence relation on  $\mathbb{N}$  generated by  $n \sim \ell(n)$ . Then  $\sim$  partitions  $\mathbb{N}$  into  $r$ -many equivalence classes where  $r$  is defined by the condition that  $a^rb$  is a prefix of  $s$  for distinct symbols  $a, b \in \{1, 2, \dots, k\}$ . Let  $x(s) = x_0x_1x_2 \dots$  be the infinite word over the alphabet  $\{a_1, \dots, a_r\}$  defined by: for all  $n \in \mathbb{N}$  and  $i \in \{1, \dots, r\}$ ,  $x_n = a_i$  if and only if  $n \sim i$ . For each  $j \in \{1, \dots, k\}$ , let  $t^j: \mathbb{N} \rightarrow \mathbb{N}$  be defined by: for all  $n \in \mathbb{N}$ ,  $\ell(t^j(n)) = n$  and  $s_{t^j(n)} = j$ . This defines a new partition of  $\mathbb{N}$  into  $k$  classes  $N_1, \dots, N_k$ : for each  $j \in \{1, \dots, k\}$ ,  $N_j = \{t^j(0) < t^j(1) < \dots\}$ . Then, for all  $n \in \mathbb{N}$  and each  $i \in \{1, \dots, r\}$ , we have

$$x_{t^j(n)} = a_i \Leftrightarrow t^j(n) \sim i \Leftrightarrow \ell(t^j(n)) \sim i \Leftrightarrow n \sim i \Leftrightarrow x_n \sim i.$$

Hence for each  $j \in \{1, \dots, k\}$ ,  $x(s)[N_j] = x(s)$ . This shows that  $x(s)$  is  $k$ -self shuffling and that  $s$  steers the described shuffle of  $x(s)$ .

We illustrate this with an example: Suppose  $s \in \{1, 2, 3\}^{\mathbb{N}}$  begins in  $s = 1111231223123 \dots$  then  $(\ell(n))_{n \geq 0}$  begins in  $0, 1, 2, 3, 0, 0, 4, 1, 2, 1, 5, 3, 2, \dots$ . This defines an equivalence relation with 4 classes given by:  $\{0, 4, 5, 6, 10, \dots\}$ ,  $\{1, 7, 9, \dots\}$ ,  $\{2, 8, 12, \dots\}$  and  $\{3, 11, \dots\}$  which in turn defines the 3-self-shuffling word  $x(s) = abcdaaabcbadc \dots \{a, b, c, d\}^{\mathbb{N}}$  having  $s$  as a steering word. It follows from Proposition 2.2 that any coding of  $x$  is also 3-self-shuffling. It follows that every binary word  $s \in \{1, 2\}^{\mathbb{N}}$  beginning in  $a^2b$  with  $\{a, b\} = \{1, 2\}$  determines as above a unique binary self-shuffling word  $x$ . In general there is no evident relationship between the words  $s$  and  $x$ . For instance,  $s$  may be uniformly recurrent while  $x$  need not be.

The next two propositions show the invariance of self-shuffling words with respect to the action of a morphism:

**Proposition 2.2.** *Let  $\mathbb{A}$  and  $\mathbb{B}$  be finite non-empty sets and  $\tau: \mathbb{A} \rightarrow \mathbb{B}^*$  a morphism. If  $x \in \mathbb{A}^{\mathbb{N}}$  is  $k$ -self-shuffling, then so is  $\tau(x) \in \mathbb{B}^{\mathbb{N}}$ .*

*Proof.* Let  $x \in \mathbb{A}^{\mathbb{N}}$  be  $k$ -self shuffling. This implies the existence of factorizations

$$x = \prod_{i=0}^{\infty} U_i^{(1)} \cdots U_i^{(k)} = \prod_{i=0}^{\infty} U_i^{(1)} = \cdots = \prod_{i=0}^{\infty} U_i^{(k)}.$$

Whence

$$\tau(x) = \prod_{i=0}^{\infty} \tau(U_i^{(1)} \cdots U_i^{(k)}) = \prod_{i=0}^{\infty} \tau(U_i^{(1)}) \cdots \tau(U_i^{(k)}) = \prod_{i=0}^{\infty} \tau(U_i^{(1)}) = \cdots = \prod_{i=0}^{\infty} \tau(U_i^{(k)})$$

as required.  $\square$

For instance, consider the fixed point  $x$  of the substitution  $a \mapsto abb, b \mapsto a$ . It is readily checked that  $x$  is the morphic image of the period doubling word under the morphism  $0 \mapsto a, 1 \mapsto bb$ . Since the period doubling word is  $k$ -self-shuffling for each  $k \geq 2$ , the same is true of  $x$ .

Following [8], if  $x \in \mathbb{A}^{\mathbb{N}}$  is uniformly recurrent, and  $u$  a non-empty prefix of  $x$ , then one defines the derived sequence  $\mathcal{D}_u(x)$  by coding  $x$  as a concatenation of first returns to  $u$ . Remind that a word  $v$  is called a *first return* to a factor  $u$  of  $x$  if  $vu$  is a factor of  $x$ , the word  $u$  is a prefix of  $vu$  and  $vu$  does not contain other occurrences of  $u$  than suffix and prefix. The following is an immediate consequence of Proposition 2.2 since in fact  $x$  is the morphic image of  $\mathcal{D}_u(x)$ :

**Corollary 2.3.** *Let  $x \in \mathbb{A}^{\mathbb{N}}$  be a uniformly recurrent word and  $u$  a non-empty prefix of  $x$ . If the derived word  $\mathcal{D}_u(x)$  is  $k$ -self-shuffling, then  $x$  is  $k$ -self-shuffling.*

The notation  $w = v^{-r}u$  means  $u = v^r w$ .

**Proposition 2.4.** *Let  $\tau : \mathbb{A} \rightarrow \mathbb{A}^*$  be a morphism, and  $x \in \mathbb{A}^{\mathbb{N}}$  be a fixed point of  $\tau$ .*

1. *Let  $u$  be a prefix of  $x$  and  $l$  be a positive integer such that  $\tau^l(a)$  begins in  $u$  for each  $a \in \mathbb{A}$ . Then if  $x$  is  $k$ -self-shuffling, then so is  $u^{-1}x$ .*
2. *Let  $u \in \mathbb{A}^*$ , and let  $l$  be a positive integer such that  $\tau^l(a)$  ends in  $u$  for each  $a \in \mathbb{A}$ . Then if  $x$  is  $k$ -self-shuffling, then so is  $ux$ .*

*Proof.* We prove only item (1) since the proof of (2) is essentially identical. Suppose

$$x = \prod_{i=0}^{\infty} U_i^{(1)} \cdots U_i^{(k)} = \prod_{i=0}^{\infty} U_i^{(1)} = \cdots = \prod_{i=0}^{\infty} U_i^{(k)}$$

with each  $U_i^{(j)} \in \mathbb{A}^+$ . By assumption, for each  $i \geq 1$  and  $j = 1, 2, \dots, k$ , we can write  $\tau^l(U_i^{(j)}) = uV_i^{(j)}$  with each  $V_i^{(j)} \in \mathbb{A}^*$ . Put  $X_i^{(j)} = V_i^{(j)}u$ . Then since

$$x = \tau^l(x) = \prod_{i=0}^{\infty} \tau^l(U_i^{(1)} \cdots U_i^{(k)}) = \prod_{i=0}^{\infty} \tau^l(U_i^{(1)}) \cdots \tau^l(U_i^{(k)}) = \prod_{i=0}^{\infty} \tau^l(U_i^{(1)}) = \cdots = \prod_{i=0}^{\infty} \tau^l(U_i^{(k)}),$$

we deduce that

$$u^{-1}x = \prod_{i=0}^{\infty} X_i^{(1)} \cdots X_i^{(k)} = \prod_{i=0}^{\infty} X_i^{(1)} = \cdots = \prod_{i=0}^{\infty} X_i^{(k)}.$$

$\square$

The following is an immediate consequence of Proposition 2.4:

**Corollary 2.5.** *Let  $\tau : \mathbb{A} \rightarrow \mathbb{A}^*$  be a primitive substitution, and  $a \in \mathbb{A}$ . Suppose  $\tau(b)$  begins (respectively ends) in  $a$  for each letter  $b \in \mathbb{A}$ . Suppose further that the fixed point  $\tau^\infty(a)$  is  $k$ -self-shuffling. Then every right shift (respectively left shift) of  $\tau^\infty(a)$  is  $k$ -self-shuffling.*

**Remark 2.6.** Since the Fibonacci word is self-shuffling and is fixed by the primitive substitution  $0 \mapsto 01, 1 \mapsto 0$ , it follows from Corollary 2.5 that every tail of the Fibonacci word is self-shuffling.

There are a number of necessary conditions that a self-shuffling word must satisfy, which may be used to deduce that a given word is not self-shuffling. For instance:

**Definition 2.7.** The *shuffling delay* of a self-shuffling word  $x$  is the length of the shortest prefix  $u$  of  $x$  such that  $(ua)^{-1}x \in \mathcal{S}(u^{-1}x, a^{-1}x)$  where  $a$  is the letter following the prefix  $u$  in  $x$ . In other words, the shuffling delay is the length of the shortest prefix of  $x$  after which one can actually start self-shuffling  $x$ .

**Proposition 2.8.** *If  $x \in \mathbb{A}^{\mathbb{N}}$  is self-shuffling, then for each positive integer  $N$  there exists a positive integer  $M$  such that every prefix  $u$  of  $x$  with  $|u| \geq M$  has an Abelian border  $v$  with  $|u|/2 \geq |v| \geq N$ . Moreover, every prefix of  $x$  longer than the shuffling delay is Abelian bordered. In particular,  $x$  must begin in only a finite number of Abelian border-free words.*

*Proof.* Suppose to the contrary that there exist factorizations  $x = \prod_{i=0}^{\infty} U_i V_i = \prod_{i=0}^{\infty} U_i = \prod_{i=0}^{\infty} V_i$  with  $U_i, V_i \in \mathbb{A}^+$ , and there exists  $N$  such that for every  $M$  there exists a prefix  $u$  of  $x$  with  $|u| \geq M$  which has no Abelian borders of length between  $N$  and  $|u|/2$ . Take  $M = |\prod_{i=0}^N U_i V_i|$  and a prefix  $u$  satisfying these conditions. Then there exist non-empty proper prefixes  $U'$  and  $V'$  of  $u$  such that  $u \in \mathcal{S}(U', V')$  with  $|U'|, |V'| > N$ . Without loss of generality we may assume  $|V'| \leq |u|/2$ . Writing  $u = U'U''$  it follows that  $U''$  and  $V'$  are Abelian equivalent. This contradicts that  $u$  has no Abelian borders of length between  $N$  and  $|u|/2$ .

For the second part of the statement, just observe that every prefix of  $x$  longer than the shuffling delay is the shuffle of two proper prefixes, and hence is Abelian bordered.  $\square$

We illustrate some applications of Proposition 2.8:

**Fibonacci word (revisited):** We already saw that the Fibonacci word  $x = 0100101001\dots$  is self-shuffling. In contrast, the word  $y = 0x$  is not self-shuffling. It is well known that  $y$  begins in infinitely many prefixes of the form  $0B1$  with  $B$  a palindrome. It is clear that  $0B1$  is Abelian border-free. It follows from Proposition 2.8 that  $y$  is not self-shuffling.

**Paper-folding word:** The paper-folding word

$$x = 00100110001101100010\dots$$

is a Toeplitz word generated by the pattern  $u = 0?1?$  (see, e.g., [5]). It is readily verified that  $x$  begins in arbitrarily long Abelian border-free words and hence by Proposition 2.8 is not self-shuffling.

More precisely, the prefixes  $u_j$  of  $x$  of length  $n_j = 2^j - 1$  are Abelian border-free. Indeed, it is verified that for each  $k < n_j$ , we have  $|\text{pref}_k(u_j)|_0 > k/2$  while  $|\text{suff}_k(u_j)|_0 \leq k/2$ . Here  $\text{pref}_k(u)$  (resp.,  $\text{suff}_k(u)$ ) denotes the prefix (resp., suffix) of length  $k$  of a word  $u$ .

**A 3-self-shuffling word which is not self-shuffling:** Let  $y$  denote the fixed point of the substitution  $\sigma : 0 \mapsto 0001$  and  $1 \mapsto 0101$ , and put

$$x = 0^{-2}y = 0100010001010100010001000101010001000100010101000101010001010100 \dots$$

Then for each prefix  $u_j$  of  $x$  of length  $4^j - 2$ , the longest Abelian border of  $u_j$  of length less than or equal to  $(4^j - 2)/2$  has length 2. Hence  $x$  is not self-shuffling (see Proposition 2.8). The 3-shuffle is given by the following:

$$\begin{array}{llll} U_0 = 0100, & U_1 = 01, & \dots, & U_{4i+2} = \varepsilon, & U_{4i+3} = \sigma^{i+1}(0100), \\ & & & U_{4i+4} = \sigma(0), & U_{4i+5} = (\sigma(0))^{-1}\sigma^{i+1}(01), \\ V_0 = 0100, & V_1 = 01, & \dots, & V_{4i+2} = (\sigma(0))^{-1}\sigma^{i+1}(0), & V_{4i+3} = \sigma(0), \\ & & & V_{4i+4} = (\sigma(0))^{-1}\sigma^{i+1}(01)\sigma(0), & V_{4i+5} = \varepsilon, \\ W_0 = 01, & W_1 = (\sigma(0))^2, & \dots, & W_{4i+2} = \varepsilon, & W_{4i+3} = (\sigma(0))^{-1}\sigma^{i+1}(01), \\ & & & W_{4i+4} = \varepsilon, & W_{4i+5} = \sigma^{i+2}(0)\sigma(0). \end{array}$$

It is then verified that

$$x = \prod_{i=0}^{\infty} U_i V_i W_i = \prod_{i=0}^{\infty} U_i = \prod_{i=0}^{\infty} V_i = \prod_{i=0}^{\infty} W_i,$$

from which it follows that  $x$  is 3-self-shuffling.

An extension of the abelian borders argument (Proposition 2.8) gives both a necessary and sufficient condition for self-shuffling in terms of Abelian borders (which is however difficult to check in practice). For  $u \in \mathbb{A}^*$  let  $\Psi(u)$  denote the *Parikh vector* of  $u$ , i.e.,  $\Psi(u) = (|u|_a)_{a \in \mathbb{A}}$ .

**Definition 2.9.** For  $x \in \mathbb{A}^{\mathbb{N}}$  and integer  $k \geq 2$ , define a directed graph  $G_x^k = (V_x^k, E_x^k)$  with the vertex set

$$V_x^k = \{(i_1, \dots, i_k) \in \mathbb{N}^k : \sum_{j=1}^k \Psi(\text{pref}_{i_j}(x)) = \Psi(\text{pref}_{i_1+\dots+i_k}(x))\},$$

and the edge set

$$E_x^k = \{((i_1, \dots, i_k), (i'_1, \dots, i'_k)) \in V_x^k \times V_x^k : i_j \leq i'_j \text{ for } j = 1, \dots, k \text{ and } \sum_{j=1}^k i_j + 1 = \sum_{j=1}^k i'_j\}.$$

We say that  $G_x^k$  connects  $\vec{0}$  to  $\vec{\infty}$  if there exists an infinite path  $v^0 v^1 v^2 \dots$  in  $G_x^k$  such that  $v^0 = (0, \dots, 0)$  and  $v_j^n \rightarrow \infty$  as  $n \rightarrow \infty$  for any  $j = 1, \dots, k$ , where  $v^n = (v_1^n, \dots, v_k^n) \in V_x^k$ .

**Theorem 2.10.** *An infinite sequence  $x \in \mathbb{A}^{\mathbb{N}}$  is  $k$ -self-shuffling if and only if the graph  $G_x^k$  connects  $\vec{0}$  to  $\vec{\infty}$ .*

*Proof.* If  $x = x_0x_1\cdots$  with the  $x_i \in \mathbb{A}$  is  $k$ -self-shuffling, then there exists a partition of  $\mathbb{N}$  into infinite subsets  $N^j \subset \mathbb{N}$  for  $j = 1, \dots, k$  such that  $x[N^j] = x$ . Let  $N^j = \{N_0^j < N_1^j < N_2^j < \dots\}$  for every  $j \in \{1, \dots, k\}$ . For any  $j \in \{1, \dots, k\}$  and  $n \in \mathbb{N}$ , let  $t_j(n)$  be the integer  $m$  such that  $N_{m-1}^j < n \leq N_m^j$ . Then  $\Psi(\text{pref}_n(x)) = \sum_{j=1}^k \Psi(\text{pref}_{t_j(n)}(x[N^j]))$ . Since  $x[N^j] = x$  for  $j = 1, \dots, k$ , we have

$$\sum_{j=1}^k \Psi(\text{pref}_{t_j(n)}(x)) = \Psi(\text{pref}_n(x)).$$

Moreover  $t_1(n) + \dots + t_k(n) = n$ . Therefore  $v^n := (t_1(n), \dots, t_k(n)) \in V_x^k$  for all  $n \in \mathbb{N}$ . Since

$$t_j(n+1) = \begin{cases} t_j(n) + 1 & \text{if } n \in N^j \\ t_j(n) & \text{else,} \end{cases}$$

we obtain  $(v^n, v^{n+1}) \in E_x^k$ . Since this holds for any  $n \in \mathbb{N}$  and each  $N^j$  is an infinite set,  $G_x^k$  connects  $\vec{0}$  to  $\vec{\infty}$ .

Conversely, assume that  $G_x^k$  connects  $\vec{0}$  to  $\vec{\infty}$ . Let  $v^0v^1v^2\dots$  be an infinite path in  $G_x^k$  such that  $v^0 = (0, \dots, 0)$  and  $v_j^n \rightarrow \infty$  as  $n \rightarrow \infty$  for any  $j = 1, \dots, k$ , where  $v^n = (v_1^n, \dots, v_k^n) \in V_x^k$ . Define

$$N^j = \{n \in \mathbb{N} : v_j^{n+1} > v_j^n\} \text{ for } j = 1, \dots, k.$$

Then the sets  $N^j$  give a partition of  $\mathbb{N}$ . Let  $N^j = \{N_0^j < N_1^j < N_2^j < \dots\}$  for  $j = 1, \dots, k$ . For any  $j = 1, \dots, k$  and  $m \in \mathbb{N}$ , let  $n = N_m^j$ . Since  $(v^n, v^{n+1}) \in E_x^k$ , we have

$$\Psi(\text{pref}_{n+1}(x)) - \Psi(\text{pref}_n(x)) = \Psi(\text{pref}_{m+1}(x)) - \Psi(\text{pref}_m(x)).$$

Hence  $(x[N^j])_m = x_n = x_m$ . Thus, the sets  $N^j$  satisfy Definition 2.1, so we have a  $k$ -self-shuffle.  $\square$

### 3. The Thue-Morse word is self-shuffling

Theorem 2.10 gives a constructive necessary and sufficient condition for self-shuffling since a path to infinity defines a self-shuffle. While the sufficient conditions in the previous section can be applied to show that certain words are not self-shuffling, in general to prove that a word is self-shuffling, one must actually explicitly exhibit a shuffle.

**Theorem 3.1.** *The Thue-Morse word  $\mathbf{T} = 011010011001\dots$  fixed by the substitution  $\tau$  mapping  $0 \mapsto 01$  and  $1 \mapsto 10$  is self-shuffling.*

*Proof.* For  $u \in \{0, 1\}^*$  we denote by  $\bar{u}$  the word obtained from  $u$  by exchanging 0s and 1s. Let  $\sigma : \{1, 2, 3, 4\} \rightarrow \{1, 2, 3, 4\}^*$  be the substitution defined by

$$\sigma(1) = 12, \quad \sigma(2) = 31, \quad \sigma(3) = 34, \quad \sigma(4) = 13.$$

Set  $u = 01101$  and  $v = 001$ ; note that  $uv$  is a prefix of  $\mathbf{T}$ . Also define morphisms  $g, h : \{1, 2, 3, 4\} \rightarrow \{0, 1\}^*$  by

$$g(1) = v\bar{u}, \quad g(2) = \bar{v}\bar{u}, \quad g(3) = \bar{v}u, \quad g(4) = vu$$

and

$$h(1) = uv, \quad h(2) = \bar{u}\bar{v}, \quad h(3) = \bar{u}\bar{v}, \quad h(4) = uv$$

We will make use of the following lemmas:

**Lemma 3.2.**  $g(\sigma(a)) \in \mathcal{S}(g(a), h(a))$  for each  $a \in \{1, 2, 3, 4\}$ . In particular  $ug(\sigma(1)) \in \mathcal{S}(ug(1), h(1))$ .

*Proof.* For  $a = 1$  we note that

$$g(\sigma(1)) = g(12) = v\bar{u}\bar{v}\bar{u} = 0011001011010010.$$

Factoring  $0011001011010010 = 0 \cdot 011 \cdot 0 \cdot 010 \cdot 11 \cdot 01 \cdot 0010$  we obtain

$$g(\sigma(1)) \in \mathcal{S}(00110010, 01101001) = \mathcal{S}(v\bar{u}, uv) = \mathcal{S}(g(1), h(1)).$$

Similarly, for  $a = 2$  we have

$$g(\sigma(2)) = g(31) = \bar{v}u\bar{v}\bar{u} = 1100110100110010.$$

Factoring  $1100110100110010 = 1 \cdot 100 \cdot 1 \cdot 1 \cdot 010 \cdot 0110 \cdot 010$  we obtain

$$g(\sigma(2)) \in \mathcal{S}(11010010, 10010110) = \mathcal{S}(\bar{v}\bar{u}, \bar{u}\bar{v}) = \mathcal{S}(g(2), h(2)).$$

Exchanging 0s and 1s in the previous two shuffles yields

$$g(\sigma(3)) = g(34) = \bar{v}u\bar{v}u \in \mathcal{S}(\bar{v}u, \bar{u}\bar{v}) = \mathcal{S}(g(3), h(3))$$

and

$$g(\sigma(4)) = g(13) = v\bar{u}\bar{v}\bar{u} \in \mathcal{S}(vu, uv) = \mathcal{S}(g(4), h(4)).$$

□

It is readily verified that

**Lemma 3.3.**  $h(\sigma(a)) = \tau(h(a))$  for each  $a \in \{1, 2, 3, 4\}$ .

Let  $w = w_0w_1w_2w_3 \dots$  with  $w_i \in \{1, 2, 3, 4\}$  denote the fixed point of  $\sigma$  beginning in 1. As a consequence of the previous lemma we deduce that

**Lemma 3.4.**  $\mathbf{T} = h(w)$ .

*Proof.* In fact  $\tau(h(w)) = h(\sigma(w)) = h(w)$  from which it follows that  $h(w)$  is one of the two fixed points of  $\tau$ . Since  $h(w)$  begins in  $h(1)$  which in turn begins in 0, it follows that  $\mathbf{T} = h(w)$ .  $\square$

**Lemma 3.5.**  $\mathbf{T} = ug(w)$ .

*Proof.* It is readily verified that:

$$ug(1) = h(1)\bar{u}$$

$$\bar{u}g(2) = h(2)\bar{u}$$

$$\bar{u}g(3) = h(3)u$$

$$ug(4) = h(4)u.$$

Moreover, each occurrence of  $g(1)$  and  $g(4)$  in  $ug(w)$  is preceded by  $u$  while each occurrence of  $g(2)$  and  $g(3)$  in  $ug(w)$  is preceded by  $\bar{u}$ . It follows that  $ug(w) = h(w)$  which by the preceding lemma equals  $\mathbf{T}$ .  $\square$

Set

$$A_0 = ug(\sigma(w_0)) \quad \text{and} \quad A_i = g(\sigma(w_i)), \text{ for } i \geq 1$$

$$B_0 = ug(w_0) \quad \text{and} \quad B_i = g(w_i), \text{ for } i \geq 1$$

and

$$C_i = h(w_i), \text{ for } i \geq 0.$$

It follows from Lemma 3.4 and Lemma 3.5 that

$$\mathbf{T} = \prod_{i=0}^{\infty} A_i = \prod_{i=0}^{\infty} B_i = \prod_{i=0}^{\infty} C_i$$

and it follows from Lemma 3.2 that  $A_i \in \mathcal{S}(B_i, C_i)$  for each  $i \geq 0$ . Hence  $\mathbf{T} \in \mathcal{S}(\mathbf{T}, \mathbf{T})$  as required.  $\square$

#### 4. Infinite Lyndon words are not self-shuffling

**Definition 4.1.** Let  $y \in \mathbb{A}^{\mathbb{N}}$ . We say  $y$  is Lyndon if there exists an order  $\leq$  on  $\mathbb{A}$  with respect to which  $y$  is lexicographically smaller than all its proper suffixes (or tails).

Recall that if  $(\mathbb{A}, \leq)$  is a linearly ordered set, then  $\leq$  induces the *lexicographic order*, denoted  $\leq_{\text{lex}}$ , on  $\mathbb{A}^+$  and  $\mathbb{A}^{\mathbb{N}}$  defined as follows: If  $u, v \in \mathbb{A}^+$  (or  $\mathbb{A}^{\mathbb{N}}$ ) we write  $u \leq_{\text{lex}} v$  if either  $u = v$  or if  $u$  is lexicographically smaller than  $v$ . In the latter case we write  $u <_{\text{lex}} v$ . Thus  $y \in \mathbb{A}^{\mathbb{N}}$  is Lyndon if and only if there exists an order  $\leq$  on  $\mathbb{A}$  with respect to which  $y <_{\text{lex}} y'$  for all proper suffixes  $y'$  of  $y$ . Note that the property of being Lyndon is an intrinsic property of a word in the sense that if  $y$  is Lyndon and  $z$  is word isomorphic to  $y$ , then  $z$  is Lyndon. For instance  $10^\omega$  and  $01^\omega$ , are each Lyndon. Similarly, any suffix of  $x = 101001000100001 \cdots$  beginning in 1 is Lyndon. Note that in this last example, all Lyndon words in the shift orbit closure of  $x$  begin in 1. It follows from Definition 4.1 that a Lyndon word is never purely periodic (although it may be ultimately periodic) as in the first two examples above. We also note that if  $y$  is Lyndon with respect to  $\leq$ , then every prefix  $u$  of  $y$  is *minimal* in  $y$  with respect to  $\leq$ , i.e.,  $u \leq_{\text{lex}} v$  for all factors  $v$  of  $y$  with  $|v| = |u|$ .

**Theorem 4.2.** *Let  $y \in \mathbb{A}^{\mathbb{N}}$  be Lyndon relative to some order  $\leq$  on  $\mathbb{A}$  and let  $z \in \mathbb{A}^{\mathbb{N}}$  be any point in the shift orbit closure of  $y$ . Then for each  $w \in \mathcal{S}(y, z)$ , we have  $w <_{\text{lex}} z$ . In particular, taking  $z = y$ , it follows that  $y$  is not self-shuffling.*

As an immediate consequence we have

**Corollary 4.3.** *Every aperiodic uniformly recurrent word  $x$  contains a point  $y$  in its shift orbit closure which is not self-shuffling.*

*Proof.* Fix any linear order  $\leq$  on  $\mathbb{A}$ . For each  $n \geq 1$  let  $u_n$  denote the minimal factor of  $x$  of length  $n$ . Then for each  $n$  we have that  $u_n$  is a prefix of  $u_{n+1}$ . Let  $y$  denote the limit of the sequence  $(u_n)_{n \geq 1}$ . Then since  $x$  is both aperiodic and uniformly recurrent it follows that  $y$  is Lyndon.  $\square$

In order to prove Theorem 4.2 we will make use of the following four lemmas. In each of the following lemmas, assume  $(\mathbb{A}, \leq)$  is a linearly ordered set and  $x \in \mathbb{A}^{\mathbb{N}}$ .

**Lemma 4.4.** *Let  $u = u_1 u_2 \cdots u_n$  be a factor of  $x$  with each  $u_i$  minimal in  $x$ . Then  $u$  is minimal in  $x$ .*

*Proof.* It is readily verified by induction on  $k$  that each prefix  $u_1 u_2 \cdots u_k$  is minimal in  $x$ .  $\square$

**Lemma 4.5.** *Let  $u$  be a minimal factor of  $x$  and let  $v$  be the longest unbordered prefix of  $u$ . Then  $v$  is a period of  $u$ , i.e.,  $u$  is a prefix of  $v^n$  for some positive integer  $n$ .*

*Proof.* The result of the lemma is clear in case  $v = u$ . So suppose  $|v| < |u|$  and let us write  $u = v u'$ . Since every prefix of  $u$  of length longer than  $|v|$  is bordered, and since  $v$  itself is unbordered, it follows that  $u'$  is a product  $v_k v_{k-1} \cdots v_1$  where each  $v_i$  is a prefix of  $u$ . In fact, since  $u$  is bordered, there exists a border  $v_1$  of  $u$ . Moreover since  $v$  is unbordered, the suffix  $v_1$  of  $u$  does not overlap  $v$ , and hence  $v_1$  is a suffix of  $u'$ . If  $v_1 = u'$  we are done, otherwise the prefix  $u v_1^{-1}$  of  $u$  admits a border  $v_2$ . Again since  $v$  is unbordered, the suffix  $v_2$  of  $u v_1^{-1}$  does not overlap  $v$  and hence  $v_2 v_1$  is

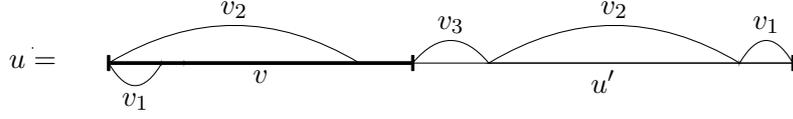


Figure 1:  $u'$  as a product of prefixes of  $u$ .

a suffix of  $u'$ . (See Figure 1). Continuing in this way we can write  $u'$  as a product of prefixes of  $u$  whence  $u'$  is a product of minimal factors of  $x$ . It follows from Lemma 4.4 that  $u'$  is a minimal factor of  $x$  and hence  $u'$  is both a prefix and a suffix of  $u$ . If  $|u'| \leq |v|$  then  $u'$  is a prefix of  $v$  and hence  $u$  is a prefix of  $v^2$ . Otherwise, if  $|u'| > |v|$  then the occurrences of  $u'$  at the beginning and end of  $u$  overlap; whence  $v$  is a period of  $u$  (see Figure 2).  $\square$

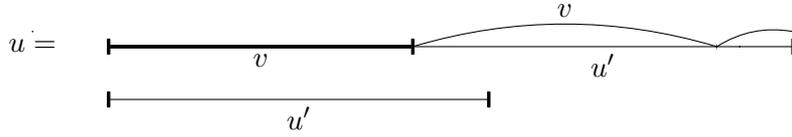


Figure 2:  $u$  has period  $v$ .

**Lemma 4.6.** *Let  $u$  and  $v$  be factors of  $x$  with  $u$  minimal. Then either  $uv <_{\text{lex}} v$  or else  $v$  is minimal.*

*Proof.* Suppose  $\neg(uv <_{\text{lex}} v)$ . Since  $u$  is minimal,  $v$  is a prefix of  $uv$ . If  $|v| \leq |u|$  then  $v$  is a prefix of  $u$  and hence  $v$  is minimal. If  $|v| > |u|$  then the prefix  $v$  of  $uv$  overlaps the suffix  $v$  of  $uv$  whence  $u$  is a period of  $uv$  and hence  $u$  is also a period of  $v$ . Thus by Lemma 4.4 we deduce that  $v$  is minimal.  $\square$

Given two finite non-empty words  $u$  and  $v$  we say that  $s \in \mathcal{S}(u, v)$  is a *proper shuffle* of  $u$  and  $v$  if there exists a positive integer  $k$  and factorings

$$u = \prod_{i=0}^k U_i \quad \text{and} \quad v = \prod_{i=0}^k V_i$$

and either

$$s = \prod_{i=0}^k U_i V_i$$

with each of  $U_0, V_0$  and  $U_1$  non-empty, or

$$s = \prod_{i=0}^k V_i U_i$$

with each of  $V_0, U_0$  and  $V_1$  non-empty. In other words,  $s$  is a proper shuffle of  $u$  and  $v$  means that  $s$  is obtained as a non-trivial shuffle of  $u$  and  $v$ . We let  $\mathcal{S}^*(u, v)$  denote the set of all proper shuffles of  $u$  and  $v$ .

Let  $\mathcal{C}(x)$  denote the set of all factors  $v$  of  $x$  with the property that no suffix of  $v$  (including  $v$  itself) is minimal in  $x$ .

**Lemma 4.7.** *Let  $u$  and  $v$  be factors of  $x$  with  $u$  minimal in  $x$  and  $v \in \mathcal{C}(x)$ . Let  $s \in \mathcal{S}^*(u, v)$ . Then  $s <_{\text{lex}} v$ .*

*Proof.* Let 0 denote the minimal element of  $\mathbb{A}$  with respect to the linear order  $\leq$ . We proceed by induction on  $|u| + |v|$ . Since  $s$  is assumed to be a proper shuffle of  $u$  and  $v$ , it follows that  $|u| + |v| \geq 3$ . For the base case  $|u| + |v| = 3$  we have either  $u = 0a$  and  $v = b$  with  $b \neq 0$  or  $u = 0$  and  $v = cd$  with  $d \neq 0$ . In the first case  $s = 0ba <_{\text{lex}} b$  and in the second case  $s = c0d <_{\text{lex}} cd$ .

Let  $N \geq 4$ . By induction hypothesis we suppose the result of the proposition is true whenever  $|u| + |v| < N$ . Now suppose  $u$  and  $v$  are factors of  $x$  with  $u$  minimal in  $x$ ,  $v \in \mathcal{C}(x)$  and  $|u| + |v| = N$ . Let

$$u = \prod_{i=0}^k U_i \quad \text{and} \quad v = \prod_{i=0}^k V_i$$

be factorings of  $u$  and  $v$  with each  $U_i$  and  $V_i$  non-empty except for possibly  $U_k$  or  $V_k$ .

**Case 1.** We consider first the case in which  $v$  dishes out the initial segment  $V_0$  of  $s$ , i.e.,

$$s = \prod_{i=0}^k V_i U_i.$$

Set  $v' = V_1 \cdots V_k$  so that  $v = V_0 v'$  (see Figure 3). Since  $s$  is assumed to be a proper shuffle of  $u$  and  $v$ , it follows that  $v'$  is non-empty and  $v' \in \mathcal{C}(x)$ . Let us write  $s = V_0 s'$ . We have two possibilities: either  $s' = uv'$  or  $s' \in \mathcal{S}^*(u, v')$ . If  $s' = uv'$ , then, because  $v'$  is not minimal, we have  $s' = uv' <_{\text{lex}} v'$  by Lemma 4.6. If  $s' \in \mathcal{S}^*(u, v')$ , then by induction hypothesis  $s' <_{\text{lex}} v'$ . Then, in both cases, we have  $s = V_0 s' <_{\text{lex}} V_0 v' = v$ .

**Case 2.** We may now suppose that  $u$  dishes out the initial segment  $U_0$  of  $s$ , i.e.,

$$s = \prod_{i=0}^k U_i V_i.$$

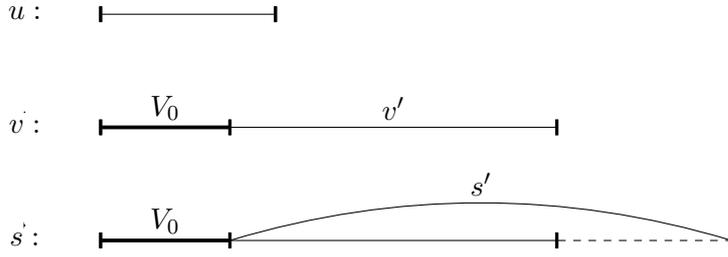


Figure 3: Case 1:  $v$  dishes out the initial segment of  $s$ .

Let  $Ub$  denote the shortest non-minimal prefix of  $v$  with  $U \in \mathbb{A}^*$  and  $b \in \mathbb{A}$ . Then  $Ub$  belongs to  $\mathcal{C}(x)$  for otherwise  $Ub$  would be the concatenation of two minimal factors, and hence itself minimal by Lemma 4.4. Since  $U_0$  is a prefix of  $s$  and  $U_0$  is minimal, we deduce that either  $U_0$  is a prefix of  $v$  (and hence of  $U$ ) or  $s <_{\text{lex}} v$ . Thus we may suppose that  $U_0$  is a prefix of  $U$  so that  $|U| \geq |U_0|$ . We consider two sub-cases according to the length of  $u$ .

**Case 2.1.** Suppose  $|u| > |U|$ , i.e.,  $U$  is a proper prefix of  $u$ . In this case let  $s'$  denote the prefix of  $s$  of length  $|Ub|$  (see Figure 4). Then  $s'$  is the prefix of a proper shuffle  $z$  of  $U$  and  $Ub$ . Remark

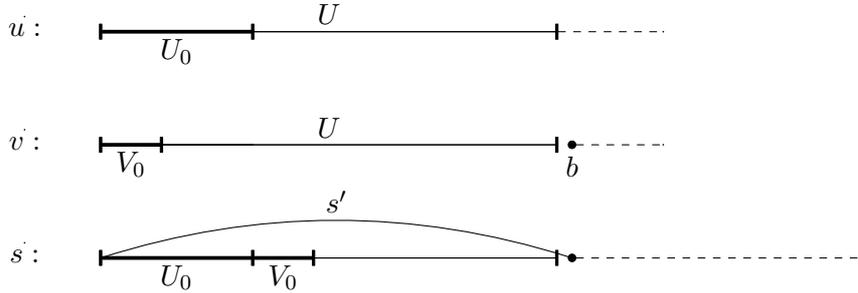


Figure 4: Case 2.1:  $|u| > |U|$ .

that here we mean a prefix of some proper shuffle of  $U$  and  $Ub$ , not necessarily the same shuffle we have in  $s$ . We have  $|U| + |Ub| < |u| + |v|$ ,  $U$  is minimal and  $Ub \in \mathcal{C}(x)$ . Hence by induction hypothesis  $z <_{\text{lex}} Ub$ , and hence  $s' <_{\text{lex}} Ub$  as  $s'$  is the prefix of length  $|Ub|$  of  $z$ . Thus  $s <_{\text{lex}} v$ .

**Case 2.2.** Suppose  $|u| \leq |U|$ , i.e.,  $u$  is a prefix of  $U$ . Let  $r$  denote the longest unbordered prefix of  $u$ . Suppose first that  $|r| > |U_0|$ , then we can write  $r = U_0r'$ . Let  $s'$  be such that  $|s'| = |r'|$  and

$U_0s'$  is a prefix of  $s$  (see Figure 5). Then  $s'$  is a prefix of a proper shuffle  $z$  of  $r$  (which is minimal)

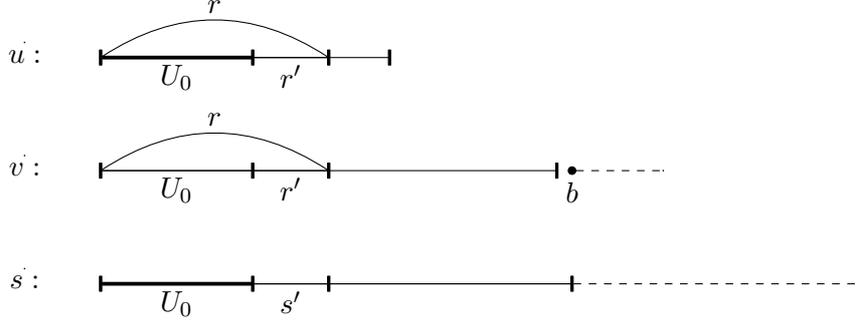


Figure 5: Case 2.2:  $u$  is a prefix of  $U$  and  $|r| > |U_0|$ .

and  $r'$  (which is in  $\mathcal{C}(x)$  since  $r$  is unbordered). By induction hypothesis  $z <_{\text{lex}} r'$ . This implies  $s' <_{\text{lex}} r'$  as  $s'$  is the prefix of length  $|r'|$  of  $z$ . Thus  $s <_{\text{lex}} v$ .

Thus we can assume that  $|r| \leq |U_0|$ . In this case let  $v'$  be such that  $v = rv'$  and  $u'$  such that  $u = ru'$ . By Lemma 4.5,  $r$  is a period of  $u$  and hence  $u'$  is also a prefix of  $u$  and hence minimal. Let  $s'$  be such that  $rs'$  is a prefix of  $s$  of length  $|Ub|$  (see Figure 6). Then  $s'$  is a prefix of a proper shuffle

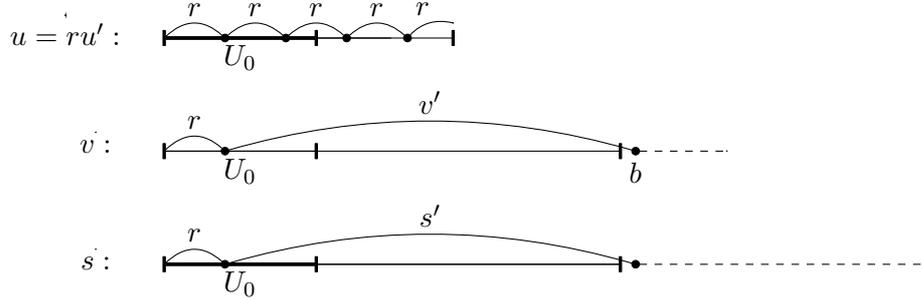


Figure 6: Case 2.2:  $u$  is a prefix of  $U$  and  $|r| \leq |U_0|$ .

$z$  of  $u'$  and  $Ub$  and hence by induction hypothesis  $z <_{\text{lex}} Ub$ . As  $Ub = rv' <_{\text{lex}} v'$  by Lemma 4.6 we obtain  $z <_{\text{lex}} v'$ . This implies  $s' <_{\text{lex}} v'$  as  $s'$  is the prefix of length  $|v'|$  of  $z$ . So, we get  $s <_{\text{lex}} v$  as required. This concludes the proof of Lemma 4.7.  $\square$

*Proof of Theorem 4.2.* Suppose  $y \in \mathbb{A}^{\mathbb{N}}$  is Lyndon relative to some order  $\leq$  on  $\mathbb{A}$ . We will apply the previous lemmas to  $x = y$ . Recall that every prefix of  $y$  is minimal relative to  $\leq_{\text{lex}}$ . We first

consider the case where  $z = y$ , i.e, when  $w \in \mathcal{S}(y, y)$ . Set

$$w = \prod_{i=0}^{\infty} (U_i V_i)$$

where

$$y = \prod_{i=0}^{\infty} U_i = \prod_{i=0}^{\infty} V_i$$

with each  $U_i$  and  $V_i$  non-empty. Since  $y$  is Lyndon, and hence in particular not purely periodic, it follows from Lemma 4.5 that  $y$  contains arbitrarily long unbordered prefixes. Let  $v$  be an unbordered prefix of  $y$  with  $|v| > |U_0|$ . Writing  $v = U_0 v'$ , since  $v$  is unbordered and  $y$  is Lyndon,  $v' \in \mathcal{C}(y)$ . Let  $s$  be such that  $U_0 s$  is a prefix of  $w$  and  $|s| = |v'|$ . Then  $s$  is a prefix of a proper shuffle  $x$  of  $v'$  and a prefix of  $y$ . By Lemma 4.7 we deduce that  $x <_{\text{lex}} v'$ . Then  $s <_{\text{lex}} v'$  as  $s$  is the prefix of length  $|v'|$  of  $x$ . Whence  $U_0 s <_{\text{lex}} U_0 v' = v$  and hence  $w <_{\text{lex}} y$ .

Next suppose  $z \neq y$ . Let

$$w = \prod_{i=0}^{\infty} (U_i V_i)$$

where

$$y = \prod_{i=0}^{\infty} U_i \quad \text{and} \quad z = \prod_{i=0}^{\infty} V_i$$

with each  $U_i$  and  $V_i$  non-empty except for possibly  $U_0$ . Suppose first that  $U_0$  is non-empty, that is to say  $y$  dishes out the initial segment of  $w$ . Let  $r$  be the shortest non-minimal prefix of  $z$ . Then by Lemma 4.4,  $r \in \mathcal{C}(y)$ . Let  $t$  denote the prefix of  $w$  of length  $|r|$ . If  $|r| \leq |U_0|$  then  $t <_{\text{lex}} r$ . Suppose that  $|r| > |U_0|$ . Let  $u$  be the prefix of length  $|r|$  of  $y$ . Then  $t$  is a prefix of a proper shuffle  $x$  of  $u$  and  $r$ . By Lemma 4.7,  $x <_{\text{lex}} r$ , whence  $t <_{\text{lex}} r$ . So in both cases  $w <_{\text{lex}} z$ . Finally suppose that  $U_0$  is empty, so that  $z$  dishes out the initial segment of  $w$ . Let  $z'$  be a tail of  $z$  so that  $z = V_0 z'$ . Writing  $w = V_0 w'$ , it follows that  $w'$  is a shuffle of  $z'$  and  $y$  in which  $y$  dishes out the initial segment of  $z'$ . If  $z' \neq y$  then we are done by the preceding case, i.e,  $w' <_{\text{lex}} z'$  whence  $w <_{\text{lex}} z$ . If  $z' = y$ , then as we saw in the beginning of the proof, we again have  $w' <_{\text{lex}} y = z'$  whence  $w <_{\text{lex}} z$ .  $\square$

**Remark 4.8.** Let  $x = 11010011001011010010110\dots$  denote the first shift of the Thue-Morse infinite word. It is easily checked that  $x$  is Lyndon and hence is not self-shuffling; yet it can be verified that  $x$  begins in only a finite number of Abelian border-free words.

## 5. A characterization of self-shuffling Sturmian words

Sturmian words admit various types of characterizations of geometric and combinatorial nature, e.g., they can be defined via balance, complexity, morphisms, etc. (see Chapter 2 in [11]). In [12],

Morse and Hedlund showed that each Sturmian word may be realized geometrically by an irrational rotation on the circle. More precisely, every Sturmian word  $x$  is obtained by coding the symbolic orbit of a point  $\rho(x)$  on the circle (of circumference one) under a rotation by an irrational angle  $\alpha$  where the circle is partitioned into two complementary intervals, one of length  $\alpha$  (labeled 1) and the other of length  $1 - \alpha$  (labeled 0) (see Fig. 7). And conversely each such coding gives rise to a Sturmian word. The irrational  $\alpha$  is called the *slope* and the point  $\rho(x)$  is called the *intercept* of the Sturmian word  $x$ . A Sturmian word  $x$  of slope  $\alpha$  with  $\rho(x) = \alpha$  is called a *characteristic Sturmian word*. It is well known that every prefix  $u$  of a characteristic Sturmian word is *left special*, i.e., both  $0u$  and  $1u$  are factors of  $x$  [11]. Thus if  $x$  is a characteristic Sturmian word of slope  $\alpha$ , then both  $0x$  and  $1x$  are Sturmian words of slope  $\alpha$  and  $\rho(0x) = \rho(1x) = 0$ . The fact that  $\rho$  is not one-to-one stems from the ambiguity of the coding of the boundary points 0 and  $1 - \alpha$ .

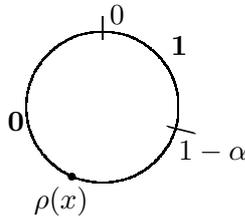


Figure 7: Geometric picture of a Sturmian word of slope  $\alpha$ .

**Theorem 5.1.** *Let  $S$ ,  $M$  and  $L$  be Sturmian words of the same slope  $\alpha$ ,  $0 < \alpha < 1$ , satisfying  $S \leq_{\text{lex}} M \leq_{\text{lex}} L$ . Then  $M \in \mathcal{S}(S, L)$  if and only if the following conditions hold: If  $\rho(M) = \rho(S)$  (respectively,  $\rho(M) = \rho(L)$ ), then  $\rho(L) \neq 0$  (respectively  $\rho(S) \neq 0$ ).*

In particular (taking  $S = M = L$ ), we obtain the following characterization of self-shuffling Sturmian words:

**Corollary 5.2.** *A Sturmian word  $x \in \{0, 1\}^{\mathbb{N}}$  is self-shuffling if and only if  $\rho(x) \neq 0$ , or equivalently,  $x$  is not of the form  $aC$  where  $a \in \{0, 1\}$  and  $C$  is a characteristic Sturmian word.*

As another immediate consequence we have:

**Corollary 5.3.** *Let  $C \in \{0, 1\}^{\mathbb{N}}$  be a characteristic Sturmian word beginning in  $a^n b$  with  $\{a, b\} = \{0, 1\}$ . Then every point  $x$  in the shift orbit closure of  $C$  beginning in  $a^n$  belongs to  $\mathcal{S}(C, C)$ .*

*Proof.* Let  $X$  denote the shift orbit closure of  $C$ . In case  $x = C$  the result follows from Corollary 5.2. Next suppose  $x \neq C$ . Suppose  $bx \in X$ . Then by Theorem 5.1 it follows that  $bx \in \mathcal{S}(bC, C)$  whence  $x \in \mathcal{S}(C, C)$ . Next suppose  $ax \in X$ . Then again by Theorem 5.1 we have  $ax \in \mathcal{S}(aC, C)$ . Since  $ax$  begins in  $a^{n+1}$ , then for any shuffle of  $aC$  and  $C$  which produces  $ax$ , one of the initial  $n + 1$  leading  $a$ 's in  $ax$  must come from the leading  $a$  in  $aC$ . Thus it can always be arranged that the first  $a$  in  $ax$  comes from the leading  $a$  in  $aC$ . Thus again we have that  $x \in \mathcal{S}(C, C)$ .  $\square$

*Proof of Theorem 5.1.* We begin by showing that the conditions stated in Theorem 5.1 are in fact necessary for  $M \in \mathcal{S}(S, L)$ . To see this, suppose  $\rho(M) = \rho(S)$  and  $\rho(L) = 0$  (the other symmetric condition works analogously). This implies that  $L \in \{0x, 1x\}$  where  $x$  is the characteristic Sturmian word of slope  $\alpha$ . If  $L = 0x$ , then as  $0x$  is minimal in the Sturmian subshift of slope  $\alpha$ , it follows that  $S = M = L$ . Whence by Proposition 2.8,  $M \notin \mathcal{S}(M, M) = \mathcal{S}(S, L)$ . If  $L = 1x$ , we consider the lexicographic order induced by  $0 > 1$ . Then  $L \leq_{\text{lex}} M \leq_{\text{lex}} S$  and moreover  $L$  is minimal. Since  $\rho(M) = \rho(S)$  we have that either case i)  $M = S$  or case ii)  $S = 0x$  and  $M = 1x$  or case iii) there exists  $u \in \{0, 1\}^*$  such that  $S = u01x$  and  $M = u10x$  where in each case  $x$  denotes the characteristic Sturmian word of slope  $\alpha$ . In case i), using Theorem 4.2 we deduce that each element of  $\mathcal{S}(S, L)$  is lexicographically smaller than  $S$  and hence since  $M = S$  we have  $M \notin \mathcal{S}(S, L)$ . In case ii), if  $M \in \mathcal{S}(S, L)$ , then  $x \in \mathcal{S}(x, 0x)$  which contradicts Theorem 4.2. Finally, in case iii), suppose to the contrary that  $M \in \mathcal{S}(S, L)$ . Then since  $u0$  is not a prefix of  $M$ , it follows that there exists a non-empty prefix  $v$  of  $L$  and a prefix  $w$  of  $M$  such that  $|w| = |u0| + |v|$  and  $M' \in \mathcal{S}(L', 1x)$  where  $M'$  and  $L'$  are defined by  $M = wM'$  and  $L = vL'$ . But this implies  $M' = L'$ , whence  $L' \in \mathcal{S}(L', 1x)$  which again contradicts Theorem 4.2.

We next show that the conditions stated in Theorem 5.1 are sufficient. Without loss of generality we can assume  $0 < \alpha < 1/2$ .

Our proof explicitly describes an algorithm for shuffling  $S$  and  $L$  so as to produce  $M$ . It is formulated in terms of the circle rotation description of Sturmian words. Geometrically speaking, points  $\rho(S)$  and  $\rho(L)$  will take turns following the trajectory of  $\rho(M)$  so that the respective codings agree; as one follows the other waits its turn (remains neutral). The algorithm specifies this following rule depending on the relative positions of the trajectories of all three points and is broken down into several cases. The proof can be summarized by the directed graph in Fig. 8 in which each state  $n$  corresponds to “case  $n$ ” in the proof.

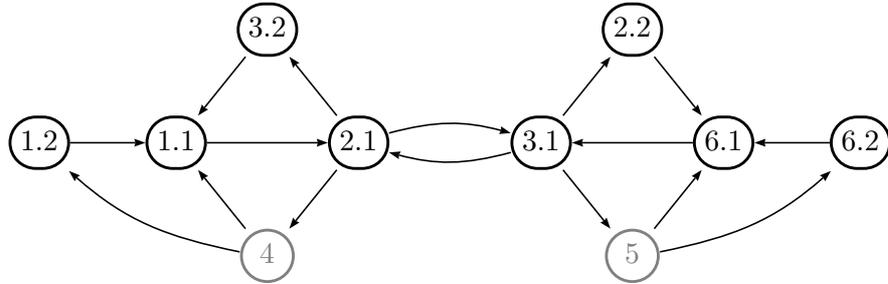


Figure 8: Graphical depiction of the proof of Theorem 5.1.

**The states:** We denote by  $s$ ,  $m$ , and  $\ell$  the current tail of the words  $S$ ,  $M$ , and  $L$ . They are initialized as

$$s := S, \ell := L, \text{ and } m := M.$$

While  $m$  is always a tail of  $M$ , the letters  $s$  and  $\ell$  may be tails of  $S$  or  $L$ , depending on which is the current lexicographically largest<sup>3</sup>. Each state, or case in the proof, is described by a figure depicting the relative positions of  $\rho(s)$ ,  $\rho(m)$  and  $\rho(\ell)$ , which for the sake of simplicity, are actually labeled  $s, m$  and  $\ell$  respectively. If  $x \in \{s, m, \ell\}$  is depicted inside the interval  $(0, 1 - \alpha)$  (resp.  $(1 - \alpha, 1)$ ), then this implies that the first letter of the coding of  $x$  is 0 (respectively 1). Moreover the endpoints of the partition interval are regarded as both open and closed. For example, even if  $s$  and  $m$  are both depicted in the interval  $(0, 1 - \alpha)$ , it could be that  $\rho(s) = 0$  and  $\rho(m) = 1 - \alpha$ . In the same way, even if  $s$  and  $m$  are depicted in distinct intervals of the circle partition, it could be that  $\rho(s) = \rho(m)$ . In addition to their relative positions on the circle, each state lists a set of relations which the variables  $s, m$  and  $\ell$  must satisfy. These conditions are written to the right of the circle picture and are described in terms of the following predicates:

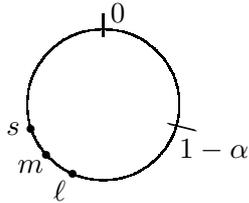
$$\begin{aligned} C(s, m, \ell) &\equiv [\rho(m) = \rho(s) \text{ and } \rho(\ell) = 0] \text{ or } [\rho(m) = \rho(\ell) \text{ and } \rho(s) = 0]; \\ P_1(s, m, \ell) &\equiv \rho(s) = \alpha \text{ and } \rho(m) = 0 \text{ and } \rho(\ell) = 1 - \alpha; \\ P_2(s, m, \ell) &\equiv [(\rho(\ell) - \rho(m)) \bmod 1 = \alpha \text{ and } \rho(s) = 1 - \alpha] \text{ or } \rho(m) = 0. \end{aligned}$$

All states, except those labeled 4 and 5, can be taken to be initial states.

**The edges:** Each directed edge corresponds to a precise set of instructions which specify which of  $s$  or  $\ell$  is neutral, which of  $s$  or  $\ell$  follows  $m$  and for how long, and in the end a possible relabeling of the variables  $s$  and  $\ell$ . In each case the outcome leads to a new case in which there is a switch in the follower. In other words, if there is an edge from case  $i$  to case  $j$  in the graph, then either the instructions for case  $i$  and case  $j$  specify different followers (as is the case for cases 1.1 and 2.1) in which case the passage from  $i$  to  $j$  leaves the labeling of  $s$  and  $\ell$  unchanged, or the instructions for case  $i$  and case  $j$  specify the same follower (as is the case for cases 1.2 and 1.1) in which case the passage from  $i$  to  $j$  exchanges the labeling of  $s$  and  $\ell$ . The proof of Theorem 5.1 amounts to showing that for each state  $n$  in the graph, the specified instructions will take  $n$  to an adjacent state in the graph.

What follows is a complete listing of all ten cases with their respective set of instructions.

**Case 1.1:**



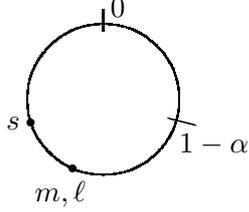
$$s \leq_{\text{lex}} m <_{\text{lex}} \ell; \neg C(s, m, \ell)$$

---

<sup>3</sup>The choice of the letter  $s, m$ , and  $\ell$  is intended to refer to *small*, *medium*, and *large* respectively.

*Instruction:*  $s$  is neutral and  $\ell$  follows  $m$  until they lie in different elements of the circle partition. No relabeling of  $s$  and  $\ell$ .

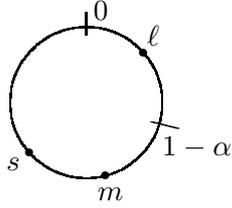
**Case 1.2:**



$$s \leq_{\text{lex}} m = \ell; \neg C(s, m, \ell)$$

*Instruction:*  $s$  is neutral and  $\ell$  follows  $m$  until  $0 < \rho(m) = \rho(\ell) < \rho(s)$ . We note that this is always possible since  $\rho(s) \neq 0$  and the set  $\{(\rho(m) + n\alpha) \bmod 1 : n \in \mathbb{N}\}$  is dense in the unit circle. Exchange the labels  $s \leftrightarrow \ell$ .

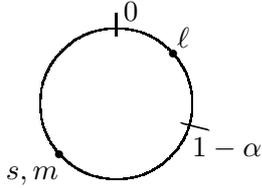
**Case 2.1:**



$$s <_{\text{lex}} m; \neg C(s, m, \ell)$$

*Instruction:*  $\ell$  is neutral and  $s$  follows  $m$  until they lie in different elements of the circle partition. No relabeling of  $s$  and  $\ell$ . Three cases are possible according to the relative position of  $m$  and  $\ell$  in the partition  $(1 - \alpha, 1)$ .

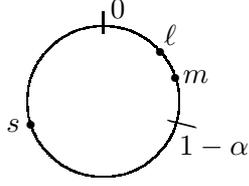
**Case 2.2:**



$$s = m; \neg C(s, m, \ell)$$

*Instruction:*  $\ell$  is neutral and  $s$  follows  $m$  until  $\rho(m) = \rho(s) > \rho(\ell)$ . This is possible because  $\rho(\ell) \neq 0$ . Exchange the labels  $s \leftrightarrow \ell$ .

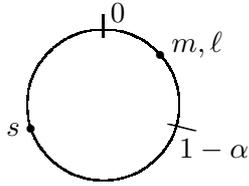
**Case 3.1:**



$$m <_{\text{lex}} \ell; \neg C(s, m, \ell)$$

*Instruction:*  $s$  is neutral and  $\ell$  follows  $m$  until they lie in different elements of the circle partition. No relabeling of  $s$  and  $\ell$ . Three cases are possible according to the relative position of  $m$  and  $s$  in the partition  $(0, 1 - \alpha)$ .

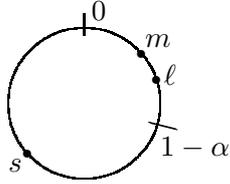
**Case 3.2:**



$$m = \ell; \neg C(s, m, \ell)$$

*Instruction:*  $s$  is neutral and  $\ell$  follows  $m$  until  $0 < \rho(m) = \rho(\ell) < \rho(s)$ . This is possible because  $\rho(s) \neq 0$ . Exchange the labels  $s \leftrightarrow \ell$ .

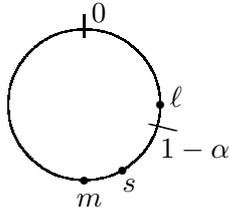
**Case 4:**



$$m >_{\text{lex}} \ell; \rho(s) \geq \alpha; \neg P_1(s, m, \ell)$$

*Instruction:*  $s$  is neutral and  $\ell$  follows  $m$  for just one rotation by  $\alpha$ . Exchange the labels  $s \leftrightarrow \ell$ . Because  $\rho(s) \geq \alpha$ , we have either  $m <_{\text{lex}} s$  or  $m = s$ .

**Case 5:**

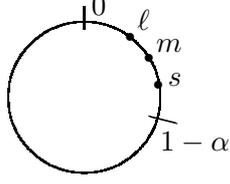


$$m <_{\text{lex}} s; (\rho(\ell) - \rho(m)) \bmod 1 \leq \alpha; \neg P_2(s, m, \ell)$$

*Instruction:*  $\ell$  is neutral and  $s$  follows  $m$  for just one rotation by  $\alpha$ . Exchange the labels  $s \leftrightarrow \ell$ .

Because  $(\rho(\ell) - \rho(m)) \bmod 1 \leq \alpha$ , we have either  $m >_{\text{lex}} \ell$  or  $m = \ell$ .

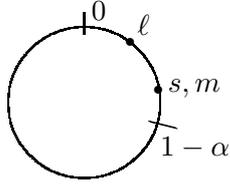
**Case 6.1:**



$$s <_{\text{lex}} m \leq_{\text{lex}} \ell; \neg C(s, m, \ell)$$

*Instruction:*  $\ell$  is neutral and  $s$  follows  $m$  until they lie in different elements of the circle partition. No relabeling of  $s$  and  $\ell$ .

**Case 6.2:**



$$s = m \leq_{\text{lex}} \ell; \neg C(s, m, \ell)$$

*Instruction:*  $\ell$  is neutral and  $s$  follows  $m$  until  $\rho(m) = \rho(s) > \rho(\ell)$ . This is possible because  $\rho(\ell) \neq 0$ . Exchange the labels  $s \leftrightarrow \ell$ .

Here we verify four of the ten cases in the proof of Theorem 5.1. The verifications of all cases are similar to one another.

**Verification of Case 1.1:** To see that Case 1.1 leads to Case 2.1, let  $m'$  and  $\ell'$  denote the positions of  $m$  and  $\ell$  respectively, the first time they lie in different elements of the circle partition. Then clearly  $0 \leq \rho(s) < \rho(m') \leq 1 - \alpha \leq \rho(\ell')$  as required. It remains to show that after the rotation  $\neg C(s, m', \ell')$  holds. Suppose to the contrary that  $C(s, m', \ell')$  holds. Because  $\rho(m') > \rho(s)$ , we must have  $\rho(m') = \rho(\ell')$  and  $\rho(s) = 0$ . But this implies  $\rho(m) = \rho(\ell)$  and  $\rho(s) = 0$ , which is impossible since we had assumed  $\neg C(s, m, \ell)$ .

**Verification of Case 1.2:** To see that Case 1.2 leads to Case 1.1, let  $m'$  and  $\ell'$  denote the positions of  $m$  and  $\ell$  respectively, the first time  $0 < \rho(m') = \rho(\ell') < \rho(s)$ . Then clearly after exchanging the labels  $s \leftrightarrow \ell$  the points  $s, m$ , and  $\ell$  are situated as specified in Case 1.1. It remains to show that  $\neg C(\ell', m', s)$ . Suppose to the contrary that  $C(\ell', m', s)$  holds. Because  $m' = \ell'$ , we must have  $\rho(m') = \rho(\ell')$  and  $\rho(s) = 0$ . But this implies  $\rho(m) = \rho(\ell)$  and  $\rho(s) = 0$ , which is impossible since we had assumed  $\neg C(s, m, \ell)$ .

**Verification of Case 2.1:** Let  $s'$  and  $m'$  denote the positions of  $s$  and  $m$  respectively, the first time they lie in different elements of the circle partition. Note that since we have assumed  $\alpha < 1/2$ , it follows that  $\rho(s') \geq \alpha$  for otherwise  $m$  and  $s$  would have already differed earlier. Three cases are possible:  $m' <_{\text{lex}} \ell$ ,  $m' = \ell$  or  $m' >_{\text{lex}} \ell$ . We show that this leads to cases 3.1, 3.2 and 4 respectively. Assume first that  $m' \leq_{\text{lex}} \ell$ . To show that this leads to Case 3.1 or Case 3.2, we must verify that  $\neg C(s', m', \ell)$ . However we have  $\alpha \leq \rho(s') \leq 1 - \alpha$ , and hence  $\rho(s') \neq 0$ . If  $\rho(m') = \rho(s')$ , then  $\rho(m) = \rho(s)$ , and hence  $\rho(\ell) \neq 0$  since we had assumed  $\neg C(s, m, \ell)$ . Next we suppose that  $m' > \ell$ . To show this result in Case 4, we must show that

$$\neg P_1(s', m', \ell).$$

Assume to the contrary that  $P_1(s', m', \ell)$ , that is, that  $\rho(s') = \alpha$ ,  $\rho(m') = 0$  and  $\rho(\ell) = 1 - \alpha$ . This implies  $m = 0m'$  and  $s = 0s'$ , and hence  $\rho(m) = 1 - \alpha = \rho(\ell)$  and  $\rho(s) = 0$ , which is impossible since we had assumed  $\neg C(s, m, \ell)$ .

**Verification of Case 4:** Let  $\ell'$  and  $m'$  denote the positions of  $\ell$  and  $m$  after rotation by  $\alpha$ . Because  $\rho(s) \geq \alpha$ , we have either  $m' <_{\text{lex}} s$  or  $m' = s$ . We will show that this leads to cases 1.1 and 1.2 respectively. In view of the label exchange  $s \leftrightarrow \ell$ , the relative positions of the three points is as required. It remains to check in both cases that  $\neg C(\ell', m', s)$  holds. Since  $\rho(s) \geq \alpha$ , it follows that  $\rho(s) \neq 0$ . If  $\rho(m') = \rho(s)$ , then we actually have  $\rho(m') = \rho(s) = \alpha$ . This implies  $\rho(m) = 0$ . Because we had assumed  $\neg P_1(s, m, \ell)$ , we obtain  $\rho(\ell) \neq 1 - \alpha$ , and hence  $\rho(\ell') \neq 0$  as required.  $\square$

We know by Proposition 2.8 that the shuffling delay is necessarily longer than the longest Abelian border-free prefix of  $x$ . We will show that, in the case of self shuffling Sturmian words, we can actually start the shuffle right after the longest Abelian border-free prefix.

**Lemma 5.4.** *Given a Sturmian word  $x$  of slope  $\alpha < 1/2$  and beginning in 0 (resp. in 1), and  $u$  a non-empty prefix of  $x$ , the following are equivalent:*

- (a)  $u$  is the longest Abelian border free prefix of  $x$ ;
- (b)  $u$  is the longest border free prefix of  $x$ ;
- (c)  $|u|$  is the least positive integer  $n$  such that  $T^n(x) <_{\text{lex}} x$  (resp.  $T^n(x) >_{\text{lex}} x$ ).

*Proof.* Border-free factors of a Sturmian word are of the form  $0B1$  or  $1B0$  where  $B$  is a palindrome. Consequently, a Sturmian factor is border-free if and only if it is Abelian border-free. Hence (a)  $\Leftrightarrow$  (b).

(a)  $\Leftrightarrow$  (c): Let us denote by  $D$  the shuffling delay of  $x$ , by  $L$  the length of the longest (Abelian) border-free prefix of  $x$ , and by  $N$  (resp.  $M$ ) the least positive integer  $n$  such that  $T^n(x) <_{\text{lex}} x$  (resp.  $T^n(x) >_{\text{lex}} x$ ). We know that  $D \geq L$  by Proposition 2.8. From the proof of Theorem 5.1, we also know that  $D \leq N$  (resp.  $D \leq M$ ) if  $x$  begins in 0 (resp. in 1). Suppose  $u$  is the longest



and

$$10C = 10 \prod_{k \geq 1} w_k = 10^{k_1(1)} \prod_{i \geq 2} w_{k_1(i)} = 1 \prod_{i \geq 1} w_{k_0(i)}. \quad (2)$$

*Proof.* Note that we have

$$w_{k_0(i)} = w_i = 0 \quad \text{for } 1 \leq i < k_1(1) \quad \text{and} \quad w_{k_1(1)} = 10^{k_1(1)-1}.$$

This implies  $01w_1 = 01w_{k_0(1)} = 010 \in \mathcal{S}(01, 0)$  and

$$10 \prod_{k=1}^{k_1(1)} w_k = 10^{k_1(1)} w_{k_1(1)} \in \mathcal{S}(10^{k_1(1)}, 1 \prod_{i=1}^{k_1(1)-1} w_{k_0(i)}) = \mathcal{S}(w_{k_1(1)}0, 1 \prod_{i=1}^{k_1(1)-1} w_i).$$

Therefore it is sufficient to prove the equalities (1) and (2). In both cases, these equalities follow from the following observation due to Risley and the fourth author [13]: For  $a \in \{0, 1\}$  and  $i \geq 1$  such that  $k_a(i) > k_1(1)$ , we have

$$w_{k_a(i)} = (\Phi(k_a(i-1) - 1))^{-1} \Phi(k_a(i) - 1). \quad (3)$$

In other words, this means that, if the letter  $a_k$  is equal to 0 (resp. to 1), the  $k$ -th iteration  $\Phi(k)$  is obtained from  $\Phi(k-1)$  by concatenating to  $\Phi(k-1)$  its suffix starting from the last  $\hat{0}$  (resp. from the last  $\hat{1}$ ).  $\square$

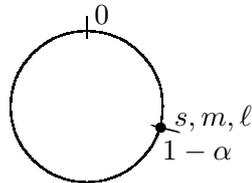
For the words built on the directive sequence  $\mathbf{d}$  we obtain:

$$01C = (01)(\hat{0})(\underbrace{\hat{0}}_{w_2})(\underbrace{\hat{1}00}_{w_3})(\underbrace{\hat{0}\hat{1}00}_{w_4})(\underbrace{\hat{1}000100\hat{1}000100}_{w_5 w_6})(\underbrace{\hat{0}\hat{1}0010001001000100}_{w_7})(\hat{1}000\dots)$$

and

$$10C = (\underbrace{100\hat{0}}_{w_3 0})(\underbrace{\hat{1}00\hat{0}100}_{1 w_1 w_2 w_4})(\underbrace{\hat{1}000100\hat{1}000100}_{w_5 w_6})(\underbrace{\hat{0}\hat{1}0010001001000100}_{w_7})(\hat{1}000\dots)$$

**Remark 5.7.** It turns out that this shuffle is the same as the one described by the general algorithm for shuffling Sturmian words described in the proof of Theorem 5.1. We will show this fact in the case of  $01C$ . The other case can be handled similarly. Because we have assumed that the directive sequence starts with 0, we know that  $0 < \alpha < 1/2$ . We start in Case 1.2 with  $s = m = \ell = 01C$ :



$$s \leq_{\text{lex}} m = \ell; \neg C(s, m, \ell)$$

According to our general algorithm,  $\ell$  follows  $m$  until  $0 < \rho(m) = \rho(\ell) < \rho(s)$ . In this case, this means  $\ell = m = C$ . We exchange the labels  $s \leftrightarrow \ell$ , hence  $s = m = C$  and  $\ell = 01C$ . The first copy of  $01C$  has output  $01$ . We are now in Case 1.1 and so,  $\ell$  follows  $m$  as long as possible. When it stops, we arrive in Case 2.1. At this point, depending on the directive sequence, the second copy has dished out  $0 \prod_{i=1}^j w_{k_1(i)}$  for some  $j \geq 0$ . For the next step,  $s$  follows  $m$  as long as possible, and the first copy dishes out  $\prod_{i=2}^j w_{k_0(i)}$  for some  $j \geq 2$ . When it stops, in principle, we arrive in Case 3.1, Case 3.2 or Case 4. Let us show that, in the case we are concerned with, we necessarily arrive in Case 3.1. Clearly, we cannot arrive in Case 3.2 because we started with  $s = m = \ell$  and the intercepts of  $m$  and  $\ell$  (resp.  $m$  and  $s$ ) cannot coincide more than once while the algorithm is performed. The fact that we actually arrive in Case 3.1 follows from (3):  $m$  and  $\ell$  coincide until  $m$  sees  $\hat{0}$  and  $\ell$  sees  $\hat{1}$ , meaning that  $m <_{\text{lex}} \ell$ . Using the same kind of arguments, we can show that from Case 3.1, we necessarily arrive in Case 2.1. Then, following the general algorithm, we simply alternate between Case 2.1 and Case 3.1. At each step, this corresponds to dishing out either a product of  $w_{k_0(i)}$  or a product of  $w_{k_1(i)}$ .

We can also exhibit an explicit self-shuffle of the characteristic Sturmian words. This shuffle is described in Proposition 5.9. We will need the following auxiliary lemma. For a self-shuffling word  $x = x_0x_1x_2\dots$  we say that letters  $x_i$  and  $x_j$  are congruent modulo its self-shuffle defined by  $N^1$  and  $N^2$  (see Definition 2.1), if  $i, j \in N^1$  or  $i, j \in N^2$ .

**Lemma 5.8.** *Let  $x \in \{0, 1\}^{\mathbb{N}}$  be of the form  $\prod_{i=1}^{\infty} (0^{k_i}1)$  with  $k_i \in \mathbb{N}$ . Suppose that for each  $n \geq 1$*

$$\sum_{i=1}^n k_i \leq \min\left\{ \sum_{i=n+1}^{2n} k_i, \sum_{i=n+2}^{2n+1} k_i \right\}$$

and for each  $n \geq 2$

$$\sum_{i=1}^n k_i \geq \max\left\{ \sum_{i=n+1}^{2n-1} k_i, \sum_{i=n+2}^{2n} k_i \right\}.$$

Then  $x$  is self-shuffling, and moreover there exists a self-shuffle of  $x$  such that no two consecutive 1's in  $x$  are congruent modulo this shuffle.

*Proof.* For each  $n \geq 1$ , define

$$u_n^1 = \begin{cases} k_1, & \text{if } n \leq 2 \\ \sum_{i=1}^{n-1} k_i - \sum_{i=n+1}^{2n-2} k_i, & \text{if } n \geq 3 \end{cases}, \quad v_n^1 = \sum_{i=n+1}^{2n} k_i - \sum_{i=1}^n k_i$$

and

$$u_n^2 = \begin{cases} k_1, & \text{if } n = 1 \\ \sum_{i=1}^n k_i - \sum_{i=n+1}^{2n-1} k_i, & \text{if } n \geq 2 \end{cases}, \quad v_n^2 = \sum_{i=n+2}^{2n+1} k_i - \sum_{i=1}^n k_i.$$

Then, for each  $n \geq 1$ , we have

$$u_n^1 \geq 0, v_n^1 \geq 0, u_n^2 \geq 0, v_n^2 \geq 0$$

by our assumption. Moreover, it is easy to see that

$$x = \prod_{n=1}^{\infty} (0^{u_n^1} 10^{v_n^1}) = \prod_{n=1}^{\infty} (0^{u_n^2} 10^{v_n^2}) = \prod_{n=1}^{\infty} (0^{u_n^1} 10^{v_n^1})(0^{u_n^2} 10^{v_n^2}).$$

Thus, this self-shuffle of  $x$  holds in a way that no two consecutive 1's in  $x$  are congruent modulo this shuffle.  $\square$

**Proposition 5.9.** *Let  $x = \prod_{i=1}^{\infty} (0^{k_i} 1)$  be a characteristic Sturmian word beginning in 0. Then  $x$  satisfies each of the inequalities of the previous lemma and hence is self-shuffling.*

*Proof.* We shall verify only the first inequality as the second is proved analogously. Let  $x = \prod_{i=1}^{\infty} (0^{k_i} 1)$  be a characteristic Sturmian word. We begin by observing that if  $u$  is a prefix of  $x$  ending in 0, then  $u$  is *rich* in 0, i.e., there exists a factor  $v$  of  $x$  with  $|u| = |v|$  such that  $|u|_0 = |v|_0 + 1$ . In fact we can take  $v = 1u0^{-1}$ . Similarly if  $u$  ends in 1 then  $u$  is *poor* in 0, i.e., rich in 1. Fix  $n \geq 1$  and consider the prefix  $X = \prod_{i=1}^n (0^{k_i} 1)$ . Then  $X$  is poor in 0. Set

$$Y = \prod_{i=n+1}^{2n} (0^{k_i} 1) \quad \text{and} \quad Z = \prod_{i=n+2}^{2n+1} (0^{k_i} 1).$$

We claim that

$$|X| \leq \min\{|Y|, |Z|\}$$

from which it follows that

$$\sum_{i=1}^n k_i = |X|_0 \leq \min\{|Y|_0, |Z|_0\} = \min\left\{ \sum_{i=n+1}^{2n} k_i, \sum_{i=n+2}^{2n+1} k_i \right\}.$$

In fact, suppose to the contrary that  $|X| > \min\{|Y|, |Z|\}$ ; note that  $0X1^{-1}$ ,  $1Y$  and  $1Z$  are each factors of  $w$  and  $|0X1^{-1}| \geq \min\{|1Y|, |1Z|\}$ . But  $|0X1^{-1}|_1 = n - 1$  while  $|1Y|_1 = |1Z|_1 = n + 1$  contradicting that  $x$  is balanced.  $\square$

As an almost immediate application of Corollary 5.2 we recover the following result originally proved by Yasutomi in [17] and later reproved by Berthé, Ei, Ito and Rao in [3] and independently by Fagnot in [10]. We say an infinite word is *pure morphic* if it is a fixed point of some morphism different from the identity.

**Theorem 5.10** (Yasutomi [17]). *Let  $x \in \{0, 1\}^{\mathbb{N}}$  be a characteristic Sturmian word. If  $y$  is a pure morphic word in the orbit of  $x$ , then  $y \in \{x, 0x, 1x, 01x, 10x\}$ .*

*Proof.* We begin with some preliminary observations. Let  $\Omega(x)$  denote the set of all left and right infinite words  $y$  such that  $\mathcal{F}(x) = \mathcal{F}(y)$  where  $\mathcal{F}(x)$  and  $\mathcal{F}(y)$  denote the set of all factors of  $x$  and  $y$  respectively. If  $y \in \Omega(x)$  is a right infinite word, and  $0y, 1y \in \Omega(x)$ , then  $y = x$ . This is because every prefix of  $y$  is a left special factor and hence also a prefix of the characteristic word  $x$ . Similarly if  $y$  is a left infinite word and  $y0, y1 \in \Omega(x)$ , then  $y$  is equal to the reversal of  $x$ . If  $\tau$  is a morphism fixing some point  $y \in \Omega(x)$ , then  $\tau(z) \in \Omega(x)$  for all  $z \in \Omega(x)$ .

Suppose to the contrary that  $\tau \neq id$  is a substitution fixing a proper tail  $y$  of  $x$ . Then  $y$  is self-shuffling by Corollary 5.2. Put  $x = uy$  with  $u \in \{0, 1\}^+$ . Using the characterization of Sturmian morphisms (see Theorem 2.3.7 & Lemma 2.3.13 in [11]) we deduce that  $\tau$  must be primitive. Thus we can assume that  $|\tau(a)| > 1$  for each  $a \in \{0, 1\}$ . If  $\tau(0)$  and  $\tau(1)$  end in distinct letters, then as both  $0\tau(x), 1\tau(x) \in \Omega(x)$ , it follows that  $\tau(x) = x$ . Since also  $\tau(y) = y$  and  $|\tau(u)| > |u|$ , it follows that  $y$  is a proper tail of itself, a contradiction since  $x$  is aperiodic. Thus  $\tau(0)$  and  $\tau(1)$  must end in the same letter. Whence by Corollary 2.5 it follows that every left extension of  $y$  is self-shuffling, which is again a contradiction since  $0x$  and  $1x$  are not self-shuffling.

Next suppose  $\tau \neq id$  is a substitution fixing a point  $y = uabx \in \Omega(x)$  where  $u \in \{0, 1\}^+$  and  $\{a, b\} = \{0, 1\}$ . Again we can suppose  $\tau$  is primitive and  $|\tau(0)| > 1$  and  $|\tau(1)| > 1$ . If  $\tau(0)$  and  $\tau(1)$  begin in distinct letters, then  $\tau(\tilde{x})0, \tau(\tilde{x})1 \in \Omega(x)$  where  $\tilde{x}$  denotes the reverse of  $x$ . Thus  $\tau(\tilde{x}) = \tilde{x}$ . Thus for each prefix  $v$  of  $abx$  we have  $\tau(\tilde{x}v) = \tilde{x}\tau(v)$  whence  $\tau(v)$  is also a prefix of  $abx$ . Hence  $\tau(abx) = abx$ . As before this implies that  $abx$  is a proper tail of itself which is a contradiction. Thus  $\tau(0)$  and  $\tau(1)$  begin in the same letter. Whence by Corollary 2.5 it follows that every tail of  $y$  is self-shuffling, which is again a contradiction since  $0x$  and  $1x$  are not self-shuffling.  $\square$

**Remark 5.11.** In the case of the Fibonacci infinite word  $x$ , each of  $\{x, 0x, 1x, 01x, 10x\}$  is pure morphic. For a general characteristic word  $x$ , since every point in the orbit of  $x$  except for  $0x$  and  $1x$  is self-shuffling, it follows that if  $\tau$  is a morphism fixing  $x$  (respectively  $01x$  or  $10x$ ), then  $\tau(0)$  and  $\tau(1)$  must end (respectively begin) in distinct letters.

## 6. Dynamical embedding and the stepping stone model

Let  $\mathbb{A}$  be a finite set with at least 2 elements. For  $k = 2$  and  $z \in \mathbb{A}^{\mathbb{N}}$ , let  $G_z^k$  be the directed graph defined in Definition 2.9. We denote this  $G_z^k$  by  $G_z = (V_z, E_z)$ .

We can sometimes embed the graph  $G_z$  into a dynamical system nicely in the following sense.

**Definition 6.1.** Let  $X$  be a compact metric space and  $R$  be a continuous mapping from  $X$  onto itself. Let  $x_0 \in X$  and  $K$  be a Borel subset of  $X^2$ . We say that the quadruple  $(X, R, x_0, K)$  is a *dynamical embedding* of the graph  $G_z$  if  $(i, j) \in V_z$  if and only if  $(R^i x_0, R^j x_0) \in K$ .

**Definition 6.2.** Let  $(X, R, x_0, K)$  be a dynamical embedding of the graph  $G_z$ . The minimum subset  $D$  of  $X^2$  satisfying

1.  $D \supset X^2 \setminus K$ , and

2.  $(x, y) \in D$  if  $(Rx, y) \in D$  and  $(x, Ry) \in D$

is called the *dead set*. Let

$$T = \{(x, y) \in K : \text{exactly one of } (Rx, y) \text{ or } (x, Ry) \text{ is in } D\}$$

and  $F = K \setminus (D \cup T)$ . We call  $T$  the *deterministic set* and  $F$  the *free set*.

**Definition 6.3.** If  $0 < \alpha < 1$  is an irrational number and  $0 \leq \rho < 1$  is any real number, we let  $z(\alpha, \rho) = z_0 z_1 z_2 \dots \in \{0, 1\}^{\mathbb{N}}$  denote the Sturmian word defined by

$$z_n = \lfloor (n+1)\alpha + \rho \rfloor - \lfloor n\alpha + \rho \rfloor.$$

We also define  $z(\alpha, 1)$  to be the limit of  $z(\alpha, \rho)$  as  $\rho \rightarrow 1$ . That is,  $z(\alpha, 1)_0 = 1$  and  $z(\alpha, 1)_n = z(\alpha, 0)_n$  for any  $n \geq 1$ .

Remark that here  $\alpha$  is the slope of the Sturmian word  $z(\alpha, \rho)$  and if  $\rho < 1$ , then  $\rho$  is its intercept. In Section 5, we noted that the intercept  $\rho(x)$  of a Sturmian word  $x$  is not one-to-one. So the intercept of  $z(\alpha, 1)$  is 0 and not 1. This will not be confusing in any way in the following.

For  $x', x'' \in \mathbb{R}$ , we will use the following notation:

$$1_{x' \geq x''} = \begin{cases} 1, & \text{if } x' \geq x'', \\ 0, & \text{if } x' < x''. \end{cases}$$

**Theorem 6.4.** Let  $z = z(\alpha, \rho)$ . Then the graph  $G_z$  has a dynamical embedding  $(\mathbb{T}, R_\alpha, 0, K)$ , where  $\mathbb{T} = \mathbb{R}/\mathbb{Z} = [0, 1)$ ,  $R_\alpha$  is the rotation by  $\alpha$  on  $\mathbb{T}$ , and

$$K = \{(x, y) \in [0, 1)^2 : 1_{x \geq 1-\rho} + 1_{y \geq 1-\rho} = \lfloor x + y + \rho \rfloor\}.$$

The set  $K$  is illustrated on Figure 9.

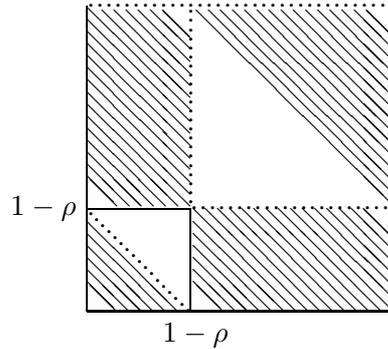


Figure 9: The set  $K$  of Theorem 6.4.

*Proof.* Let  $G_z = (V_z, E_z)$ . Let  $z(\alpha, \rho) = z_0 z_1 z_2 \dots$  with each  $z_i \in \{0, 1\}$ . If  $0 \leq \rho < 1$  then

$$|z_0 z_1 \cdots z_{n-1}|_1 = \sum_{i=0}^{n-1} z_i = \sum_{i=0}^{n-1} ([ (i+1)\alpha + \rho ] - [ i\alpha + \rho ]) = [ n\alpha + \rho ].$$

The same holds for  $\rho = 1$  since in this case,

$$|z_0 z_1 \cdots z_{n-1}|_1 = \sum_{i=0}^{n-1} z_i = 1 + \sum_{i=1}^{n-1} ([ (i+1)\alpha ] - [ i\alpha ]) = 1 + [ n\alpha ] = [ n\alpha + 1 ].$$

Hence it holds that

$$\begin{aligned} (i, j) \in V_z &\Leftrightarrow [ (i+j)\alpha + \rho ] = [ i\alpha + \rho ] + [ j\alpha + \rho ] \\ &\Leftrightarrow \{ (i+j)\alpha + \rho \} = \{ i\alpha + \rho \} + \{ j\alpha + \rho \} - \rho \end{aligned}$$

where  $\{x\}$  means  $x - \lfloor x \rfloor$ . Since

$$\{ (i+j)\alpha + \rho \} = \{ \{i\alpha\} + \{j\alpha\} + \rho \} = \{i\alpha\} + \{j\alpha\} + \rho - [ \{i\alpha\} + \{j\alpha\} + \rho ]$$

and, for each  $k \in \{i, j\}$ ,

$$\{ k\alpha + \rho \} = \{ \{k\alpha\} + \rho \} = \{k\alpha\} + \rho - 1_{\{k\alpha\} \geq 1-\rho},$$

we obtain that

$$\begin{aligned} (i, j) \in V_z &\Leftrightarrow 1_{\{i\alpha\} \geq 1-\rho} + 1_{\{j\alpha\} \geq 1-\rho} = [ \{i\alpha\} + \{j\alpha\} + \rho ] \\ &\Leftrightarrow (\{i\alpha\}, \{j\alpha\}) \in K \\ &\Leftrightarrow (R_\alpha^i 0, R_\alpha^j 0) \in K. \end{aligned}$$

This implies that the quadruple  $(\mathbb{T}, R_\alpha, 0, K)$  is a dynamical embedding of  $G_z$ . □

**Theorem 6.5.** *Let  $z = z(\alpha, \rho)$ . Then  $z$  is self-shuffling if and only if there exists a sequence  $(i_n, j_n) \in \mathbb{N}^2$  such that*

1.  $i_n \leq i_{n+1}, j_n \leq j_{n+1}, i_n + j_n = n$  for any  $n \in \mathbb{N}$ ,
2.  $\lim_{n \rightarrow +\infty} i_n = \lim_{n \rightarrow +\infty} j_n = \infty$ , and
3.  $1_{\{i_n\alpha\} \geq 1-\rho} + 1_{\{j_n\alpha\} \geq 1-\rho} = [ \{i_n\alpha\} + \{j_n\alpha\} + \rho ]$  for any  $n \in \mathbb{N}$ .

*Proof.* Clear from Theorems 2.10 and 6.4. □

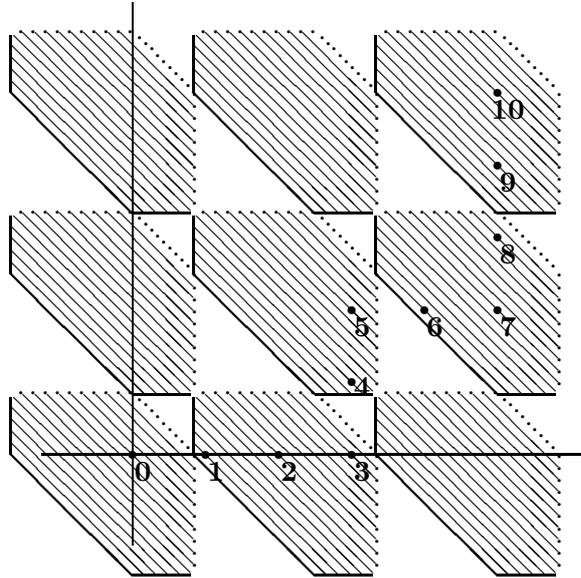


Figure 10: The stepping stone for  $0 < \rho < 1$  and a stepping stone path:  $(0,0) \rightarrow (\alpha,0) \rightarrow (2\alpha,0) \rightarrow (3\alpha,0) \rightarrow (3\alpha,\alpha) \rightarrow (3\alpha,2\alpha) \rightarrow (4\alpha,2\alpha) \rightarrow (5\alpha,2\alpha) \rightarrow (5\alpha,3\alpha) \rightarrow (5\alpha,4\alpha) \rightarrow (5\alpha,5\alpha) \rightarrow \dots$

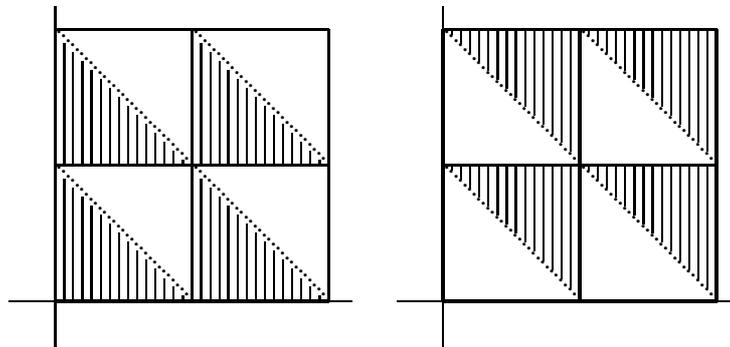


Figure 11: The stepping stones for  $\rho = 0$  (left) and  $\rho = 1$  (right).

Put  $i\alpha = x$ ,  $j\alpha = y$  and consider the condition  $(\{x\}, \{y\}) \in K$ , where  $K$  is defined as in Theorem 6.4. That is,  $(x, y)$  is in  $K + \mathbb{Z}^2$  (see Figures 10 and 11).

We call  $K + \mathbb{Z}^2$  the *stepping stone* for  $\rho$ . Hence,  $z$  is self-shuffling if and only if there is a sequence  $(i_n\alpha, j_n\alpha) \in K + \mathbb{Z}^2$  satisfying the conditions (1) and (2) of Theorem 6.5. We call it a *stepping stone path* with respect to  $(\alpha, \rho)$ . Therefore, by Theorem 5.1, we have the following corollary.

**Corollary 6.6.** *There exists a stepping stone path with respect to  $(\alpha, \rho)$  if  $0 < \rho < 1$ .*

Now we give another proof of the fact that Sturmian words of the form  $aC$ , where  $C$  is a characteristic Sturmian word, are not self-shuffling.

**Theorem 6.7.** *Let  $C$  be a characteristic Sturmian word. Both  $0C$  and  $1C$  have no stepping stone path, and hence, are not self-shuffling.*

*Proof.* We have  $0C = z(\alpha, 0)$  and  $1C = z(\alpha, 1)$ . Let  $p_k/q_k$ , with  $k \geq 1$ , be the convergents of  $\alpha$ .

Assume first that  $\rho = 0$  and  $z = z(\alpha, 0)$ . By Theorem 6.4,  $(i, j) \in V_z$  if and only if  $\{i\alpha\} + \{j\alpha\} < 1$  since in this case,  $1_{\{i\alpha\} \geq 1-\rho} = 1_{\{j\alpha\} \geq 1-\rho} = 0$ . Since  $\{q_{2k+1}\alpha\}$  is so close to 1, we do not have  $\{i\alpha\} + \{j\alpha\} < 1$  for any  $(i, j) \in \mathbb{N}^2$  such that either  $i = q_{2k+1}$  and  $1 \leq j \leq q_{2k+1}$  or  $j = q_{2k+1}$  and  $1 \leq i \leq q_{2k+1}$ . Therefore, for each  $k \geq 1$ , no point in

$$\{q_{2k+1}\} \times \{1, 2, \dots, q_{2k+1}\} \cup \{1, 2, \dots, q_{2k+1}\} \times \{q_{2k+1}\}$$

belongs to the vertex set  $V_z$ . Hence, there is no path in  $G_z$  connecting  $\vec{0}$  to  $\vec{\infty}$ . Thus,  $z(\alpha, 0)$  is not self-shuffling.

We have a similar proof for  $\rho = 1$  and  $z = z(\alpha, 1)$ . By Theorem 6.4,  $(i, j) \in V_z$  if and only if  $\{i\alpha\} + \{j\alpha\} \geq 1$  since in this case,  $1_{\{i\alpha\} \geq 1-\rho} = 1_{\{j\alpha\} \geq 1-\rho} = 1$ . Since  $\{q_{2k+2}\alpha\}$  is so close to 0, we have  $\{i\alpha\} + \{j\alpha\} < 1$  if either  $i = q_{2k+2}$  and  $1 \leq j \leq q_{2k+2}$  or  $j = q_{2k+2}$  and  $1 \leq i \leq q_{2k+2}$ . Hence, there is no path in  $G_z$  connecting  $\vec{0}$  to  $\vec{\infty}$ . Thus,  $z(\alpha, 1)$  is not self-shuffling.  $\square$

Let  $z = z(\alpha, \rho)$ . Let  $(\mathbb{T}, R_\alpha, 0, K)$  be the dynamical embedding of the graph  $G_z$  in Theorem 6.5. Let  $z = z(\alpha, \rho)$ . Let  $(\mathbb{T}, R_\alpha, 0, K)$  be the dynamical embedding of the graph  $G_z$  in Theorem 6.5. We determine the dead set, the deterministic set and the free set in the easy case where  $(1 - \rho)/2 < \alpha < \min\{\rho, 1 - \rho\}$ . In fact, let

$$\begin{aligned} D_1 &= \{(x, y) \in [0, 1]^2 : (R_\alpha x, y) \in \mathbb{T}^2 \setminus K\}, \quad \text{and} \\ D_2 &= \{(x, y) \in [0, 1]^2 : (x, R_\alpha y) \in \mathbb{T}^2 \setminus K\}. \end{aligned}$$

Then,  $D_1 \cap D_2 \cap K$  is obtained as in Figure 12.

That is,

$$D_1 \cap D_2 \cap K = \{(x, y) \in [0, 1 - \alpha - \rho]^2 : x + y \geq 1 - \alpha - \rho\}.$$

It is easily verified that:

1. If  $(R_\alpha x, y) \in D_1 \cap D_2 \cap K$ , then  $(x, R_\alpha y) \notin (D_1 \cap D_2 \cap K) \cup (\mathbb{T}^2 \setminus K)$ .

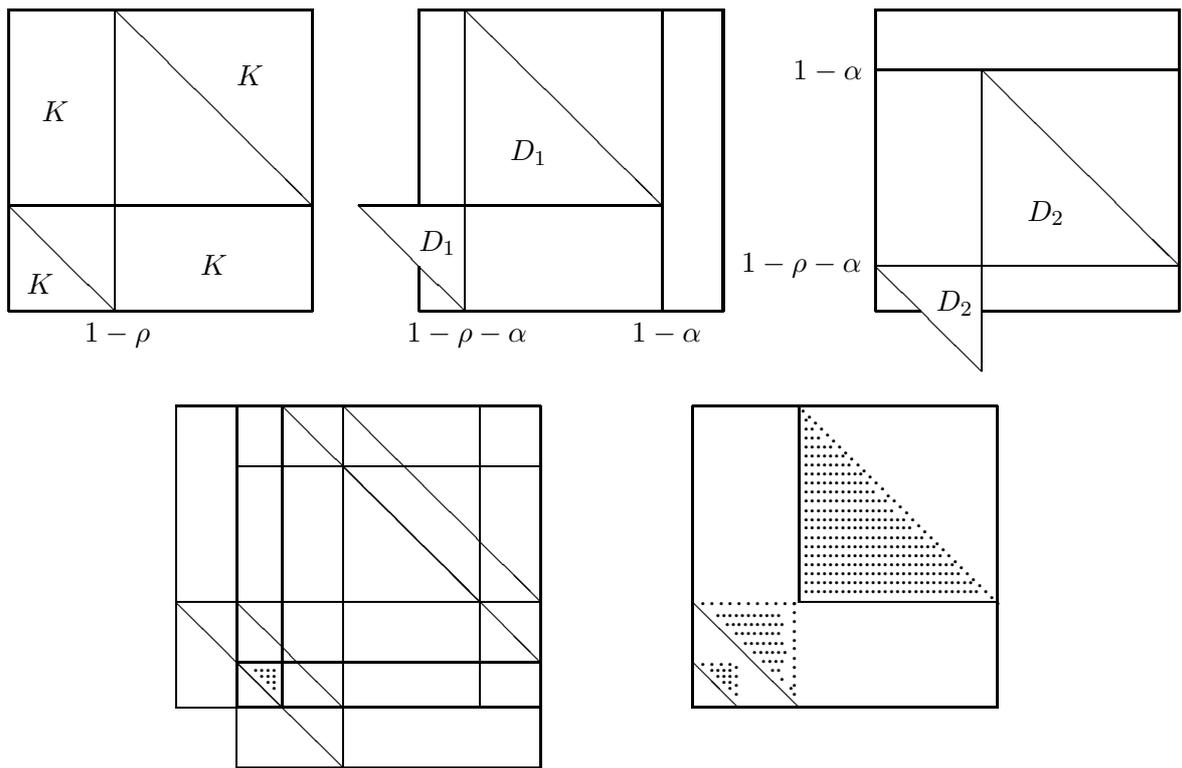


Figure 12: The sets  $K$ ,  $D_1$ ,  $D_2$  (above) and their intersection (below left),  $D$  (below right).

2. If  $(x, R_\alpha y) \in D_1 \cap D_2 \cap K$ , then  $(R_\alpha x, y) \notin (D_1 \cap D_2 \cap K) \cup (\mathbb{T}^2 \setminus K)$ .

Hence,  $D = (D_1 \cap D_2 \cap K) \cup (\mathbb{T}^2 \setminus K)$  is the dead set.

Let

$$T_1 = \{(x, y) \in \mathbb{T}^2 \setminus D : (x, R_\alpha y) \in D\} \text{ and } T_2 = \{(x, y) \in \mathbb{T}^2 \setminus D : (R_\alpha x, y) \in D\}.$$

Then,

$$\begin{aligned} T_1 = \{ & (x, y) \in (1 - \rho - \alpha, 1 - \rho] \times [0, 1 - \rho - \alpha) : x + y < 1 - \rho \} \\ & \cup \{(x, y) \in [1 - \rho, 1) \times [1 - \rho - \alpha, 1 - \rho) : x + y < 2 - \rho - \alpha \} \\ & \cup \{(x, y) \in [0, 1 - \rho) \times [1 - \rho, 1) : x + y \geq 2 - 2\alpha - \rho \} \end{aligned}$$

holds. If  $(x, y) \in T_1$ , then  $(R_\alpha x, y) \in \mathbb{T}^2 \setminus D$  since otherwise,  $(x, y) \in D$  by the definition of  $D$ . Hence,  $(R_\alpha x, y) \in \mathbb{T}^2 \setminus D$  and  $(x, R_\alpha y) \in D$  holds if  $(x, y) \in T_1$ . The same things hold for  $T_2$  in the symmetrical sense.

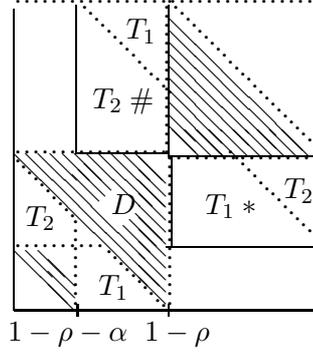


Figure 13: The dead set  $D$  together with  $T_1 \cup T_2$ .

Let  $T = T_1 \cup T_2$  and  $F = K \setminus (D \cup T)$ . By the definition, it holds that  $(R_\alpha x, y) \notin D$  and  $(x, R_\alpha y) \notin D$  for any  $(x, y) \in F$ . For each  $i \in \{1, 2\}$ , define a mapping  $\tilde{R}_{\alpha, i}: F \rightarrow F$ . For  $(x, y) \in \mathbb{T}^2$ , we denote

$$R_{\alpha, i}(x, y) = \begin{cases} (R_\alpha x, y), & \text{if } i = 1 \\ (x, R_\alpha y), & \text{if } i = 2. \end{cases}$$

Take an arbitrary  $(x_0, y_0) \in F$ . Let  $(x_1, y_1) = R_{\alpha, i}(x_0, y_0)$ . If  $(x_1, y_1) \in F$ , then let  $\tilde{R}_{\alpha, i}(x_0, y_0) = (x_1, y_1)$ . If  $(x_1, y_1) \notin F$ , then either  $(x_1, y_1) \in T_1$  or  $(x_1, y_1) \in T_2$ . Let  $(x_2, y_2) = R_{\alpha, 1}(x_1, y_1)$  in the former case, and let  $(x_2, y_2) = R_{\alpha, 2}(x_1, y_1)$  in the latter case. If  $(x_2, y_2) \in F$ , then let  $\tilde{R}_{\alpha, i}(x_0, y_0) = (x_2, y_2)$ . Repeat this procedure until we get  $(x_n, y_n) \in F$ . Then, we define  $\tilde{R}_{\alpha, i}(x_0, y_0) = (x_n, y_n)$ . If  $(x_n, y_n) \notin F$  for any  $n \geq 1$ , then  $\tilde{R}_{\alpha, i}(x_0, y_0)$  is not defined, which never happens in our case. This is easily seen from Figure 13. That is, if  $(x, y) \in T_1$ , then  $R_{\alpha, 1}(x, y) \in F$  except for the

case when  $(x, y)$  is in the \*-marked region. If  $(x, y)$  is in the \*-marked region, then take the first  $n > 0$  such that  $R_{\alpha,1}^n(x, y) \notin T_1$ . Then, either  $R_{\alpha,1}^n(x, y) \in F$  or  $R_{\alpha,1}^n(x, y) \in T_2$ . In the latter case,  $R_{\alpha,1}^n(x, y)$  is not in the #-marked region so that  $R_{\alpha,2}R_{\alpha,1}^n(x, y) \in F$ . The same for the case  $(x, y) \in T_2$ . Therefore, the infinite paths starting  $(0, 0)$  in  $G_z$  correspond bijectively to the infinite sequences of mappings applied to  $(0, 0)$

$$\cdots \tilde{R}_{\alpha,i_3} \tilde{R}_{\alpha,i_2} \tilde{R}_{\alpha,i_1}(0, 0)$$

with  $i_1, i_2, \dots \in \{1, 2\}$ . Note that both of  $\tilde{R}_{\alpha,1}$  and  $\tilde{R}_{\alpha,2}$  are domain exchange transformations on  $F$ .

## 7. Open questions

Typically a self-shuffling word can be shuffled in more than one way, i.e., it defines several different steering words. One may ask:

**Question 7.1.** *Does there exist a self-shuffling word admitting a unique steering word, i.e., which can be self-shuffled to produce itself in one and only one way?*

We saw that every aperiodic uniformly recurrent word contains an element in its shift orbit closure which is not self-shuffling.

**Question 7.2.** *Does there exist a word  $x \in \mathbb{A}^{\mathbb{N}}$  for which no element of its shift orbit closure is self-shuffling?*

## References

- [1] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence. In: Sequences and their applications, Proceedings of SETA'98, C. Ding, T. Hellesteth and H. Niederreiter (Eds.) (1999), Springer Verlag, 1–16.
- [2] J.-P. Allouche, J. Shallit, Automatic sequences. Theory, applications, generalizations. Cambridge University Press, 2003.
- [3] V. Berthé, H. Ei, S. Ito, H. Rao, On substitution invariant Sturmian words: an application of Rauzy fractals. Theor. Inform. Appl. 41 (2007), pp. 329–349.
- [4] S. Buss, M. Soltys, *Unshuffling a square is NP-hard*, arXiv:1211.7161
- [5] J. Cassaigne, J. Karhumäki, Toeplitz Words, Generalized Periodicity and Periodically Iterated Morphisms. European J. Combin. 18 (1997), pp. 497–510.
- [6] É. Charlier, T. Kamae, S. Puzynina, L. Q. Zamboni *Self-shuffling words*. ICALP 2013, Part II, LNCS 7966 (2013), p. 113–124.

- [7] A. de Luca, Sturmian words: structure, combinatorics, and their arithmetics, *Theoret. Comput. Sci.* 183 (1997), pp. 45–82.
- [8] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math.* 179 (1998), pp. 89–101.
- [9] D. Henshall, N. Rampersad, J. Shallit, Shuffling and unshuffling. *Bull. EATCS*, 107 (2012), pp. 131–142.
- [10] I. Fagnot, A little more about morphic Sturmian words, *Theor. Inform. Appl.* 40 (2006), pp. 511–518.
- [11] M. Lothaire, *Algebraic Combinatorics On Words*, vol. 90 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, U.K., 2002.
- [12] M. Morse, G.A. Hedlund, Symbolic dynamics II: Sturmian sequences. *Amer. J. Math.* 62 (1940), pp. 1–42.
- [13] R. Risley, L.Q. Zamboni, A generalization of Sturmian sequences: combinatorial structure and transcendence. *Acta Arith.*, 95 (2000), pp. 167–184.
- [14] R. Rizzi and S. Vialette, *On recognizing words that are squares for the shuffle product*, LNCS 7913 (CSR 2013): 235–245.
- [15] R. Siromoney, L. Mathew, V.R. Dare, K.G. Subramanian, Infinite Lyndon words, *Inform. Process. Lett.* 50 (1994) 101–104
- [16] A. Thue, Über unendliche Zeichenreihen, *Norske Vid. Selsk. Skr. I Math-Nat. Kl.* 7 (1906), pp. 1–22.
- [17] S.-I. Yasutomi, On sturmian sequences which are invariant under some substitutions. In: *Number theory and its applications*, Proceedings of the conference held at the RIMS, Kyoto, Japan, November 10–14, 1997, edited by Kanemitsu, Shigeru et al. Kluwer Acad. Publ. Dordrecht (1999) 347–373.