

## CHRISTOFFEL MATRICES AND STURMIAN DETERMINANTS

CHRISTOPHE REUTENAUER<sup>1</sup>  AND JEFFREY SHALLIT<sup>2,\*</sup> 

**Abstract.** We discuss certain matrices associated with Christoffel words, and show that they have a group structure. We compute their determinants and show a relationship with the Zolotareff symbol from number theory. We show that the  $n \times n$  determinants constructed from the factors of length  $n$  of a Sturmian sequence form a perfectly clustering word on three letters.

**Mathematics Subject Classification.** 68R15, 11A15, 05E16.

Received November 1, 2024. Accepted January 25, 2026.

### 1. INTRODUCTION

Christoffel words and their cyclic conjugates, the standard words, are a finitary version of Sturmian sequences. The Burrows–Wheeler matrix of such a word, that is, the square tableau whose rows are all the conjugates of such a word, lexicographically sorted, is called a Christoffel matrix when the letters of the word are elements of a field. By construction, Christoffel matrices have two distinct entries, typically but not necessarily 0 and 1.

A remarkable property of the Christoffel matrices of order  $n$  over a field  $\mathbb{K}$  is that they form a commutative subgroup of  $\text{GL}_n(\mathbb{K})$  (Thm. 3.1).

In order to prove this result, and to give the structure of this group, we note that the Christoffel matrices have the vector  $(1, \dots, 1)$  as a left eigenvector, since the row- and column- sums of each matrix in the group are constant. They have as supplementary subspace the space spanned by the differences  $e_{i-1} - e_i$  ( $e_i$  the canonical basis), and the restriction to this subspace belongs to a finite commutative group. The group is isomorphic to the Cartesian product of the square of the group of invertible elements of the ground field by the group of invertible elements of the ring  $\mathbb{Z}/n\mathbb{Z}$ . This allows us to compute the inverse of a Christoffel matrix; see Corollary 4.1. Another consequence is that the determinant of such a matrix is equal to  $\pm$  its row-sum, multiplied by the difference of its two distinct entries raised to the power  $n$  (Cor. 4.2). The sign is the Zolotareff (Zolotarev) symbol, a generalization of the Jacobi symbol.

The next task of this article is to compute the determinants obtained from Sturmian sequences. One takes such a sequence, say  $g$ , over the alphabet  $\{0, 1\}$  and forms the determinant whose rows are  $n$  factors (contiguous subwords) of this sequence. Recall that such a sequence has  $n + 1$  distinct factors of length  $n$ , so one obtains  $n + 1$  determinants, which are the components of some vector  $V_n$  of length  $n + 1$ . It turns out that this vector,

---

*Keywords and phrases:* Combinatorics on words, Burrows–Wheeler matrix, Christoffel word, perfectly clustering word, determinant, Sturmian sequence, group structure, Zolotareff symbol, determinant, Fibonacci number.

<sup>1</sup> Département de mathématiques, Université du Québec à Montréal, Canada.

<sup>2</sup> School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

\* Corresponding author: [reutenauer.christophe@uqam.ca](mailto:reutenauer.christophe@uqam.ca)

viewed as a word, is a *perfectly clustering word* over a three-letter alphabet, which may be computed precisely (Thm. 8.1).

Recall that perfectly clustering words on  $\ell$  letters are equivalently defined as *word-encodings* of *symmetric discrete exchanges of  $\ell$  intervals*, as follows from the work of Ferenczi and Zamboni [1]. Perfectly clustering words were studied by Simpson and Puglisi [2] and by Restivo and Rosone [3]. When  $\ell = 2$ , by a result of Mantaci, Restivo and Sciortino [4], these words are the Christoffel words and their conjugates, including the *standard words* of Aldo de Luca. One property of perfectly clustering words that we use is that such a word has a unique palindromic factorization (product of two palindromes), a result due to [2], and proved previously for Christoffel words by Chuan [5]; this property is preserved by conjugation.

We show that for arbitrarily large values of  $n$  (more precisely, if  $n = |w| - 1$ , where  $w$  is a Christoffel word that is a factor of  $s$ ), the word  $V_n$  is a Christoffel word that may be computed using our result on Burrows–Wheeler matrices: it is the last row of the inverse of the Burrows–Wheeler matrix of  $w$ . Then we describe how to pass from  $V_n$  to  $V_{n-1}$ , up to a sign. This is done in two ways. One of them is to use the palindromic factorization of  $V_n$ : write a product of words  $V_n = XxyY$  with  $Xx, yY$  palindromes, over the alphabet  $\{x, y\}$ . Then  $V_{n-1}$  is obtained by merging and adding  $x$  and  $y$  (Cor. 8.2).

In Section 13, Theorem 8.1 is given an interpretation in terms of continued fractions; the Sturmian determinants are expressed using the semi-convergents and convergents of the slope of the Sturmian sequence.

In Section 14, we apply our results to compute the determinants associated with the Fibonacci sequence (aka the Fibonacci word), which are shown to be Fibonacci numbers (up to a sign). The sign is precisely computed in Section 14.1; it amounts to determining the Zolotareff symbol  $\left(\frac{F_{m-2}}{F_m}\right)_{\mathbb{Z}}$ ; that is, the sign of the permutation of  $\mathbb{Z}/F_m\mathbb{Z}$  given by multiplication by  $F_{m-2}$  modulo  $F_m$ . See Proposition 14.1.

**Convention:** We consider words whose letters are elements of a field, and therefore identify words with vectors over this field. We index rows and columns of matrices starting from 0; likewise for vectors and words.

## Part 1. Group of Christoffel matrices

### 2. BURROWS–WHEELER MATRIX OF A CHRISTOFFEL WORD

We consider words over an alphabet  $A$ . Let  $w = a_0 \cdots a_{n-1}$  be such a word, of length  $n$ , so that  $a_i \in A$ . We assume that  $w$  is *primitive*, that is, not the power of another word. Then  $w$  has  $n$  distinct rotations called *conjugates*, of the form  $a_i \cdots a_n a_0 \cdots a_{i-1}$  for  $i = 0, \dots, n-1$ . For example, the conjugates of the French word **rai** are **{rai, air, ira}**.

The length of a word  $w$  is written  $|w|$ . The number of occurrences of a letter  $a$  in  $w$  is written  $|w|_a$ , and called the *a-length*.

We consider  $n \times n$  matrices. We index their rows and columns by  $\{0, 1, \dots, n-1\}$ , and we identify this set naturally with  $\mathbb{Z}/n\mathbb{Z}$ .

The *Burrows–Wheeler matrix* (or BW-matrix for short), due to [6], of a primitive word  $w = a_0 \cdots a_{n-1}$  over a totally ordered alphabet, is the  $n \times n$  matrix  $(a_{ij})$ , where the words  $a_{i,0} \cdots a_{i,n-1}$  for  $i = 0, \dots, n-1$  are the distinct conjugates of  $w$ , in *decreasing* lexicographical order<sup>1</sup>. See Figure 1.

Among the many equivalent definitions of Christoffel words (see, for example, the books [7, 8]), we take the following, which actually is a characterization due to Mantaci, Restivo and Sciortino [4]:  $w$  is a *lower (resp., upper) Christoffel word* if it is a word over a totally ordered two-letter alphabet  $\{a < b\}$  such that

- (a)  $|w|_a$  and  $|w|_b$  are relatively prime;
- (b)  $w$  is the lowermost row (resp., uppermost row) of its BW-matrix; and
- (c) the last column of this matrix is (not necessarily strictly) increasing, from the first row to the last.

<sup>1</sup>In contrast with what is commonly used in the literature, where one takes the *increasing* order.

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

FIGURE 1. Burrows–Wheeler matrix of the Christoffel word 0001001.

Of course, we naturally identify rows and columns with the corresponding word. Note that the two previous Christoffel words, the first row and the last row, are conjugate. See Figure 1. It is known that  $|w_a|$  and  $|w_b|$ , together with the order  $a < b$ , completely determine each of these two words. Note that in each row, and each column, there are  $|w_a|$  (resp.,  $|w_b|$ ) occurrences of letter  $a$  (resp.,  $b$ ).

### 3. GROUP OF CHRISTOFFEL MATRICES

Let  $\mathbb{K}$  be a field. We identify words over  $\mathbb{K}$ , in a natural way, with row vectors. For instance, the word  $1^n$  is identified with the vector of length  $n$  each component of which is equal to 1.

We say that a matrix is a *Christoffel matrix* if it is the BW matrix of some Christoffel word whose alphabet is a two-element subset of  $\mathbb{K}$ , where this subset is equipped with some total order. We let  $M_n(a, b, r)$  denote the Christoffel matrix of length  $n$  associated with the lower and upper Christoffel word over the two-letter alphabet  $\{a, b\} \subset \mathbb{K}$ , with the order  $a < b$ , with  $b$ -length  $r$ . For example,  $M_7(0, 1, 2)$  is the matrix represented in Figure 1.

We now fix  $n$  and let  $G_n$  denote the group of permutations of  $\mathbb{Z}/n\mathbb{Z}$  (where the latter is identified with  $[n] = \{0, 1, \dots, n-1\}$ ) of the form  $\omega_r : x \mapsto rx$ , for some  $r$  relatively prime with  $n$ . This group is isomorphic to the group of invertible elements of the ring  $\mathbb{Z}/n\mathbb{Z}$ , under the isomorphism  $r \mapsto \omega_r$ .

**Theorem 3.1.** *Suppose that the characteristic of  $\mathbb{K}$  is  $> n$ . The set of all matrices  $M_n(a, b, r)$ , where  $a, b \in \mathbb{K}$ ,  $a \neq b$ ,  $r \in [n]$ ,  $r, n$  relatively prime, and  $(n-r)a + rb \neq 0$ , forms a commutative subgroup  $\mathcal{G}$  of  $\mathrm{GL}_n(\mathbb{K})$ . The mapping  $\mathcal{G} \rightarrow \mathbb{K}^* \times \mathbb{K}^* \times G_n$ ,  $M_n(a, b, r) \mapsto ((n-r)a + rb, b-a, \omega_r)$ , is an isomorphism.*

Note that  $n-r$  is the number of  $a$ 's and  $r$  the number of  $b$ 's in the Christoffel word corresponding to the BW matrix  $M_n(a, b, r)$ . Thus  $(n-r)a + rb$  is the common value of the row-sums, and also of the column-sums, of this matrix.

We need the following lemma.

**Lemma 3.2.** *Let  $M$  be the BW matrix of the Christoffel word  $w$  of length  $n$  over the alphabet  $\{a < b\}$ , with  $|w|_b = r$ . Then the word obtained by reading down the first column of  $M$  is  $b^r a^{n-r}$ , and for  $i = 2, \dots, n$ , the  $i$ -th column is obtained from the  $(i-1)$ -th column by cyclically shifting the latter downwards by  $r$  steps.*

See, for example, Figure 1, where  $r = 2$ .

*Proof.* We adapt the definition of lower Christoffel words to upper Christoffel words (see [7, 8]). Let  $q = n-r$ . Then the upper Christoffel word  $w = a_0 \cdots a_{n-1}$  with  $|w|_b = r$  is defined by

$$a_j = \begin{cases} b, & \text{if } qj \bmod n \in \{0, \dots, r-1\}; \\ a, & \text{if } qj \bmod n \in \{r, \dots, n-1\}. \end{cases}$$

The conjugates of  $w$  are, in decreasing lexicographical order, the words

$$w_i = a_{i,0} \cdots a_{i,n-1}, i = 0, \dots, n-1,$$

with

$$a_{i,j} = \begin{cases} b, & \text{if } (i + qj) \bmod n \in \{0, \dots, r-1\}; \\ a, & \text{if } (i + qj) \bmod n \in \{r, \dots, n-1\}. \end{cases}$$

Thus the BW matrix  $M$  of  $w$  is the matrix  $(a_{ij})_{0 \leq i, j \leq n-1}$ . In particular,  $w_0$  is an upper Christoffel word, and the corresponding lower Christoffel word is  $w_{n-1}$ .

The formulas show that the first column of  $M$ , viewed as a word, is  $b^r a^{n-r}$ . Moreover, each column is a conjugate of this word.

Let  $j \in \{1, \dots, n-1\}$ . Since  $i + qj \equiv 0 \pmod{n}$  is equivalent to  $i \equiv rj \pmod{n}$ , the  $r$  occurrences of  $b$  in column  $j$  are located in rows  $rj, rj+1, \dots, rj+r-1$ , viewed modulo  $n$ . These occurrences are cyclically consecutive, which implies the lemma.  $\square$

We deduce a well-known and remarkable property of the BW matrix of a Christoffel word; namely, how one passes from each row to the next one.

In what follows, we let  $r^*$  denote the unique inverse in  $\{1, \dots, n-1\}$  of  $r$  modulo  $n$ , for  $0 < r < n$  and  $r, n$  relatively prime.

**Corollary 3.3.** *Suppose that  $n \geq 2$  and  $1 \leq i < n$ . Then the  $(i-1)$ -th and  $i$ -th rows of  $M$  are equal, except for the elements in the submatrix*

$$\begin{bmatrix} a_{i-1, j-1} & a_{i-1, j} \\ a_{i, j-1} & a_{i, j} \end{bmatrix} = \begin{bmatrix} b & a \\ a & b \end{bmatrix},$$

where  $j = ir^* \bmod n$ .

See Figure 1, where  $r = 2$ ,  $r^* = 4$ , and  $n = 7$ .

*Proof.* By inspection of the matrix, using Lemma 3.2, we see that two consecutive rows differ only by a submatrix as indicated (see Fig. 1). It remains to determine its exact location. Let  $i-1$  and  $i$  be the indices of the two rows. Then the right lower entry of this  $2 \times 2$  submatrix is in position  $i, j$  say, and in this position there is a  $b$  that is the first  $b$  of the cyclic run of  $b$ 's in column  $j$ .

In the proof of Lemma 3.2, we have seen that in column  $j$ , the first  $b$  in the cyclic run of  $b$ 's is in row  $jr \bmod n$ . Hence, modulo  $n$ , we have  $i \equiv jr$ , and thus  $j \equiv ir^*$ .  $\square$

*Proof of Theorem 3.1.* Let square matrices act on the right on row vectors of length  $n$ . Recall that the row-sums of  $M = M_n(a, b, r)$  are equal to its column-sums, and to  $(n-r)a + rb$ . Since the column-sums of  $M$  are constant, the row vector  $1^n$  is an eigenvector of this matrix, with eigenvalue  $(n-r)a + rb$ . Since the characteristic is  $> n$ , the line spanned by this vector has as supplementary subspace the subspace spanned by the  $n-1$  vectors  $v_i = e_{i-1} - e_i$ ,  $i = 1, \dots, n-1$ , where  $e_0, \dots, e_{n-1}$  is the canonical basis of the space of row vectors.

Now  $v_i M = (e_{i-1} - e_i)M = L_{i-1} - L_i$ , where  $L_k$  is the  $k$ -th row of  $M$ . It follows from Corollary 3.3 that  $L_{i-1} - L_i = (b-a)(e_{j-1} - e_j)$  where  $j \equiv ir^* \pmod{n}$ . Hence  $v_i M = (b-a)v_{ir^* \bmod n}$ .

Recall that the *direct sum* of two square matrices  $F$  and  $G$  of dimension  $f \times f$  and  $g \times g$ , respectively, is the square matrix of dimension  $(f+g) \times (f+g)$  with  $F$  in the upper left,  $G$  in the lower right, and zeros elsewhere.

In the basis  $v_1, \dots, v_{n-1}, 1^n$ , the matrix of the endomorphism defined by right multiplication by  $M$  is a direct sum of two matrices: the product of  $b-a$  by a permutation matrix  $N$  of order  $n-1$ , and the  $1 \times 1$  matrix  $(n-r)a + rb$ . The matrix  $N$  has a 1 in row  $i$  and column  $ir^* \bmod n$  for  $i = 1, \dots, n-1$ , and 0's elsewhere;

equivalently, the 1's are in row  $jr \bmod n$  and column  $j$ . In other words,  $N$  is the matrix of the permutation  $x \mapsto rx$  of  $(\mathbb{Z}/n\mathbb{Z})^*$ . It follows that the mapping of the theorem is a group homomorphism.

Let  $(c, d, \omega)$  be in the group  $\mathbb{K}^* \times \mathbb{K}^* \times G_n$ . Looking for a preimage  $M_n(a, b, r)$ , we must have  $c = (n-r)a + rb$ ,  $d = b - a$ ,  $\omega = \omega_r$ . The last equation determines  $r \in [n]$  uniquely, since  $r = \omega_r(1) = \omega(1)$ . Then we must have  $b = a + d$ ,  $c = (n-r)a + r(a+d) = na + rd$ . Hence, by the hypothesis on the characteristic of  $\mathbb{K}$ , we see that  $a = (1/n)(c - rd)$  and  $b = (1/n)(c - rd + nd) = (1/n)(c + (n-r)d)$ . We deduce that the mapping is an isomorphism. In particular,  $\mathcal{G}$  is commutative.  $\square$

For later use, we state the following result.

**Corollary 3.4.** *Let  $q = n - r$ . Let  $h \equiv jq \pmod{n}$  for  $1 \leq j < n$ . Then the row vectors  $(a_{h-1,1}, \dots, a_{h-1,n-j-1})$  and  $(a_{h,1}, \dots, a_{h,n-j-1})$  are equal, except for their last two entries, which satisfy  $a_{h-1,n-j-1} = b$  and  $a_{h,n-j-1} = a$ .*

*Proof.* Use Corollary 3.3, with  $i = h$ , thus  $i = jq$ , with equalities modulo  $n$ . Let  $J = ir^*$ . Then  $J = jqr^* = -j = n - j$ . It follows from Corollary 3.3 that rows  $i - 1$  and  $i$  of the matrix coincide in positions  $0, 1, \dots, J - 2$  and that their  $(J - 1)$ -th entry is  $b$  for the first and  $a$  for the second. Since  $J - 1 = n - j - 1$ , we get the result.  $\square$

#### 4. CONSEQUENCES

**Corollary 4.1.** *The inverse of  $M_n(0, 1, r)$  is  $M_n(-Q/r, 1 - Q/r, r^*)$ , where  $r^* \in [n]$  is defined by  $rr^* = 1 + Qn$ .*

*Proof.* Taking the image of both matrices under the isomorphism of Theorem 3.1, we obtain the two elements  $(r, 1, \omega_r)$  and  $((n - r^*)\frac{-Q}{r} + r^*(1 - \frac{Q}{r}), 1, \omega_{r^*})$ . Since the first component of the second triple is  $(1/r)(-nQ + r^*Q + rr^* - r^*Q) = 1/r$ , these two elements are inverses of each other, which implies the corollary.  $\square$

**Corollary 4.2.** *The determinant of  $M_n(a, b, r)$  is equal to*

$$((n - r)a + rb)(b - a)^{n-1} \operatorname{sgn}(\omega_r).$$

*Moreover,  $(n - r)a + rb$  is an eigenvalue of  $M_n(a, b, r)$ , and the other eigenvalues are products of  $b - a$  by  $\varphi(n)$ -th roots of unity.*

*Proof.* By the proof of Theorem 3.1, the matrix  $M_n(a, b, r)$  is conjugate to the direct sum of two matrices: the  $(n - 1) \times (n - 1)$  matrix  $(b - a)P$ , and the  $1 \times 1$  matrix  $((n - r)a + rb)$ , where  $P$  is the matrix of the permutation  $\pi : x \mapsto rx$  of  $(\mathbb{Z}/n\mathbb{Z})^*$ . This permutation has the same sign as the permutation  $\omega_r$  of  $(\mathbb{Z}/n\mathbb{Z})$ , since  $\omega_r$  fixes 0. Hence  $\det(P) = \operatorname{sgn}(\omega_r)$ . The assertion for the determinant follows.

Now, since the order of the group of invertible elements in the ring  $\mathbb{Z}/n\mathbb{Z}$  is  $\varphi(n)$ , we have  $\omega_r^{\varphi(n)} = 1$ . Hence  $\pi^{\varphi(n)} = 1$ , which implies that  $P^{\varphi(n)} = I_{n-1}$ . This implies the assertion about the eigenvalues.  $\square$

The sign of the permutation  $\omega_r : x \mapsto rx$  of  $\mathbb{Z}/n\mathbb{Z}$  is known as the *Zolotareff symbol*, denoted by  $(\frac{r}{n})_{\mathbb{Z}}$ . It is known that for  $n$  odd, the Zolotareff symbol is equal to the *Jacobi symbol*  $(\frac{r}{n})$ , which by definition is the product  $\prod_i (\frac{r}{p_i})$ , where  $n = \prod_i p_i$  is written as a product of primes, and where  $(\frac{r}{p})$  is the *Legendre symbol*; see [9–12]. For  $n$  even, see Lerch [13] or Barnes [14].

The Christoffel words over the alphabet  $\{0, 1\}$  are particularly meaningful; their BW matrices multiply as follows.

**Corollary 4.3.** *One has the product formula*

$$M_n(0, 1, r)M_n(0, 1, s) = M_n(Q, Q + 1, R),$$

*where  $rs = R + Qn$  (Euclidean division by  $n$ ).*

*Proof.* One has simply to apply the isomorphism of Theorem 3.1. The identity to prove is equivalent to

$$(r, 1, \omega_r)(s, 1, \omega_s) = (nQ + R, 1, \omega_R),$$

which is true. □

As an example, the matrix in Figure 1 is  $M_7(0, 1, 2)$ . Since  $2^2 = 4 + 7 \cdot 0$  and  $2^3 = 1 + 7 \cdot 1$ , its square and its cube are the matrices  $M_7(0, 1, 4)$  and  $M_7(1, 2, 1)$ , which are

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 2 \end{bmatrix},$$

the Burrows–Wheeler matrices of the lower Christoffel words 0101011 and 1111112.

## Part 2. Sturmian determinants

### 5. DETERMINANTS OF STURMIAN FACTORS

We consider Sturmian sequences (also called Sturmian words, but in this article we prefer to reserve the term “word” for finite words) over the alphabet  $\{0, 1\}$ . It is well known that each such infinite sequence has, for each natural number  $n$ , exactly  $n + 1$  factors (also called *subwords*, but for some authors, “subword” has another meaning) of length  $n$ . Taking  $n$  of these  $n + 1$  words, one obtains an  $n$  by  $n$  matrix with entries in  $\{0, 1\}$ , whose determinants we want to compute.

To be precise, we need some notation and terminology. Let  $A$  be a  $(k + 1) \times k$  matrix over a commutative ring. Write  $A = (L_0, \dots, L_k)$ , where the  $L_i$  are the rows of  $A$ . Define  $a_i = (-1)^{k-i} \det(L_0, \dots, \hat{L}_i, \dots, L_k)$ . Then we write

$$D(A) = (a_0, \dots, a_k),$$

and call  $D(A)$  the *determinantal vector* of  $A$ .

Now, let  $g$  be a Sturmian sequence, order the  $n + 1$  factors of length  $n$  in reverse lexicographic order, and form the  $(n + 1) \times n$  matrix  $G_n$  whose rows are these words, the largest being the uppermost row.

We call  $D(G_n)$  the  $n$ -th *determinantal vector* of the given Sturmian sequence, and denote it by  $V_n$ . Our aim is to compute it. It will turn out that  $V_n$ , viewed as a word, is a perfectly clustering word over an alphabet with two or three letters.

For more on Sturmian sequences, see [7, 15].

### 6. PERFECTLY CLUSTERING WORDS AND SYMMETRIC DISCRETE INTERVAL EXCHANGES

A word is *perfectly clustering* if the last column of its Burrows–Wheeler matrix is (not necessarily strictly) increasing from top to bottom. This matrix was defined in Section 2; recall that the rows of the matrix are decreasing from top to bottom.

A word  $w$  is a *Lyndon word* if for each proper factorization  $w = uv$ , the word  $w$  is smaller than  $vu$  in lexicographic order; equivalently,  $w$  of length  $n$  is a Lyndon word  $w$  if and only if it is the last row of its Burrow–Wheeler matrix, and if the latter has  $n$  rows. For more on Lyndon words, see, e.g., [7] or [10].

Let  $n \geq 1$ . Let  $(c_1, \dots, c_\ell)$  be a *composition* of  $n$ , that is, an  $\ell$ -tuple of natural numbers whose sum is  $n$  (for convenience, we allow 0's in the composition). We decompose the interval

$$[n] = \{0, 1, 2, \dots, n-1\}$$

into intervals in two ways: the intervals  $I_1, \dots, I_\ell$  (resp.,  $J_1, \dots, J_\ell$ ) are defined by the condition that they are consecutive and  $|I_j| = c_j$  (resp.,  $|J_h| = c_{\ell+1-h}$ ). Let  $S_n$  denote the group of permutations of  $[n]$ . We define the permutation  $\sigma \in S_n$  by the condition that its restriction to each  $I_h$  is the unique increasing bijection onto  $J_{\ell+1-h}$ . We call such a permutation a *symmetric discrete interval exchange*<sup>2</sup>, and we say it is *associated with the composition*  $(c_1, \dots, c_\ell)$ .

A symmetric discrete interval exchange may be equivalently defined by using *local translations*. Indeed, the permutation  $\sigma$  acts on  $[n]$  as a discrete version of an interval exchange as we now describe in the case of three intervals. Let  $\ell = 3$  and  $(c_1, c_2, c_3)$  the composition, with  $c_1 + c_2 + c_3 = n$ . Then the permutation  $\sigma$  is defined by

$$\sigma(i) = \begin{cases} i + c_2 + c_3, & \text{if } i \in I_1 = \{0, \dots, c_1 - 1\}; \\ i + c_3 - c_1, & \text{if } i \in I_2 = \{c_1, \dots, c_1 + c_2 - 1\}; \\ i - c_1 - c_2, & \text{if } i \in I_3 = \{c_1 + c_2, \dots, n - 1\}. \end{cases} \quad (6.1)$$

In particular, if  $c_2 = 0$ , the permutation  $\sigma$  is translation by  $c_3$  in  $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \dots, n-1\}$ .

As an example, consider the composition  $(2, 2, 5)$  of 9; then  $J_1 = \{0, 1, 2, 3, 4\}$ ,  $J_2 = \{5, 6\}$ ,  $J_3 = \{7, 8\}$  and

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 8 & 5 & 6 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}. \quad (6.2)$$

The second row of  $\sigma$  is obtained by concatenating the intervals  $J_3, J_2, J_1$ .

We call  $\sigma$  a *circular symmetric discrete interval exchange* if  $\sigma$  is a circular permutation; that is, it has only one cycle. The example in equation (6.2) is circular, since in cycle form we have

$$\sigma = (0, 7, 3, 6, 2, 5, 1, 8, 4).$$

Note that there is no simple criterion known for a symmetric discrete interval exchange to be circular, except for the cases  $\ell = 2$  ( $c_1$  and  $c_2$  must be relatively prime), and the case  $\ell = 3$  ( $c_1 + c_2$  and  $c_2 + c_3$  must be relatively prime, by a result of Pak and Redlich [16]<sup>3</sup>). However, there is a general construction of all circular discrete interval exchanges, for all  $\ell$ , due to Mélodie Lapointe [25].

Let  $A$  be a totally ordered alphabet of cardinality  $\ell$ . A *word encoding* of a circular symmetric discrete interval exchange  $\sigma$  is one of the words obtained by replacing each number  $i$  in one of the cycle forms of  $\sigma$  by the  $j$ -th letter in  $A$ , if  $i \in I_j$ . In the example, we have

$$I_1 = \{0, 1\}, \quad I_2 = \{2, 3\}, \quad I_3 = \{4, 5, 6, 7, 8\},$$

and we obtain the word *acbcacc*, with  $A = \{a < b < c\}$ .

Note that we have  $n$  choices for the cycle form, so we obtain  $n$  words from  $\sigma$ , which clearly form a conjugacy class. The word encoding is called *standard* if one chooses the cycle form beginning with 0, as in the example.

<sup>2</sup>The word “symmetric” refers to the fact that the intervals are exchanged according to the central symmetry of the set  $\{1, 2, \dots, \ell\}$ , that is, the mapping  $h \mapsto \ell + 1 - h$ .

<sup>3</sup>They also prove that the probability that a symmetric discrete exchange of three intervals in  $S_n$  is circular tends to  $1/\zeta(2) = 6/\pi^2$ , when  $n \rightarrow \infty$ .

**Theorem 6.1.** (Ferenczi and Zamboni [1]) *Let  $w$  be a primitive word  $w$  over a totally ordered alphabet. Then the following conditions are equivalent:*

- (i)  $w$  is a perfectly clustering Lyndon word;
- (ii)  $w$  is the standard encoding of some circular symmetric interval exchange.

Note that if one of the parts is 0 in the composition  $(c_1, c_2, c_3)$ , then the corresponding perfectly clustering Lyndon word is a lower Christoffel word. In particular, we use this when the composition is  $(p, 0, q)$ : then the word is the lower Christoffel word of slope  $q/p$  (slope is defined below).

## 7. FACTORS OF STURMIAN SEQUENCES

The *slope* of a Christoffel word  $w \in \{0 < 1\}^*$  is by definition  $|w|_1/|w|_0$ ; lower Christoffel words are in bijection with the positive rationals, *via* their slope. To also obtain the words 0 and 1 of length 1, one has to include the slopes 0/1 and 1/0, so that the bijection is now onto  $\mathbb{Q} \cup \infty$ .

Let  $g = (a_n)_{n \in \mathbb{N}}$  be a Sturmian sequence over the alphabet  $\{0, 1\}$ . The limit of the sequence  $|w|_1/|w|_0$  for  $w$  a factor of  $g$ , when the length  $w$  tends to infinity, is called the *slope* of this sequence. Note that we are not considering the quantity  $|w|_1/|w|$ , as many authors do; see Section 13 on continued fractions, for comparing both notions.

It is well known that the set of factors of a given Sturmian sequence  $g$  depends only on its slope. We describe this set as follows. Consider the infinite path, in the Stern–Brocot tree (see [17]), starting from the root, determined by the slope of  $g$ . This path is coded by a sequence  $(\lambda_n)_{n \geq 1}$  over the set  $\{\ell, r\}$ , where  $\ell$  means “left” and  $r$  means “right”. Each node in this path is a positive rational number, and corresponds to some prefix of the sequence; this number is the slope of some lower Christoffel word. Let  $w_0, w_1, w_2, \dots$  with  $w_0 = 01$  be the corresponding sequence of Christoffel words. Equivalently, these words are the nodes on the infinite path determined by the sequence  $(\lambda_n)_{n \geq 1}$  in the Christoffel tree (see [8], Fig. 14.1). Recall that a *circular factor* of a word  $u$  is a factor of some conjugate of  $u$ ; equivalently, a prefix of some conjugate of  $u$ .

Then one has the following result: *a word is a factor of the Sturmian sequence  $g$  if and only if it is a circular factor of some word  $w_n$ .*

Note that the sequence of lengths  $|w_n|$  is strictly increasing, with first element  $|w_0| = 2$ .

Each Christoffel word  $w$ , of length  $\geq 2$ , is uniquely the product of two Christoffel words (Theorem of Borel and Laubie, see [8], Thm. 2.4.1), and all three are simultaneously lower or upper. This product is called the *standard factorization* of  $w$ , and we denote it by  $w = w'w''$ .

One has the following fact: for each  $n$ ,  $w_n$  is equal to the longer of the two words  $w'_{n+1}, w''_{n+1}$  (this follows from the direct construction of the Christoffel tree, see [8] Thm. 14.1.1). In particular,  $w_n$  is a factor, hence a circular factor, of  $w_{n+1}$ .

It is known also that if a word  $u$  is a factor of  $g$ , then it is a circular factor of  $w_p$  as soon as  $|u| \leq w_p$ . This can be proved as follows: the set of circular factors of length  $p$  of  $w_p$  has cardinality  $p + 1$  (see [8], Thm. 15.3.1) and is contained in the set of factors of length  $p$  of  $g$  (the displayed result above). The latter set has cardinality  $p + 1$  (because  $g$  is a Sturmian sequence). Hence the two sets coincide, and  $u$  is in the latter set, hence in the former.

## 8. DETERMINANTS

In the next result, vectors are identified with words.

**Theorem 8.1.** *Let  $g$  be a Sturmian sequence over the alphabet  $\{0, 1\}$ , with associated sequence of lower Christoffel words  $01 = w_0, w_1, w_2, \dots$ . Let  $n \geq 2$ . The  $n$ -th determinantal vector  $V_n$  of  $g$  is a perfectly clustering word over a two- or three-letter alphabet.*

*More precisely, let  $\nu$  the unique integer such that  $|w_{\nu-1}| \leq n \leq |w_\nu| - 1$ , let  $N = |w_\nu|$ , and let  $\epsilon$  be the sign of the permutation  $x \mapsto |w_\nu|_1 x$  of  $(\mathbb{Z}/N\mathbb{Z})^*$ .*

1. If  $n = |w_\nu| - 1$ , then  $V_n$  is the lower Christoffel word associated with the composition

$$(|w''_\nu|, |w'_\nu|),$$

over the alphabet

$$\{-|w'_\nu|_1, |w''_\nu|_1\},$$

multiplied by  $\epsilon$ .

2. If  $|w_{\nu-1}| \leq n < |w_\nu| - 1$ , let  $i = |w_\nu| - 1 - n$ . Then  $V_n$  is the perfectly clustering word over the three-letter alphabet

$$\{-|w'_\nu|_1, |w''_\nu|_1 - |w'_\nu|_1, |w''_\nu|_1\},$$

associated with the composition

$$(|w''_\nu| - i, i, |w'_\nu| - i),$$

multiplied by  $\epsilon(-1)^t$ , where

$$t = \sum_{1 \leq j \leq i} (|w_\nu| - j + d_j - (jq \bmod N)),$$

with  $d_i$  equal to the number of  $j$ ,  $1 \leq j \leq i$ , such that  $jq \bmod N$  is smaller than  $iq \bmod N$ .

Define  $q = |w_\nu|_0$  and  $r = |w_\nu|_1$ ; in particular  $q + r = N$ . Then  $|w''_\nu| = q^*$ ,  $|w'_\nu| = r^*$ , the inverses in  $[N]$  of  $q$  and  $r$  modulo  $N$  (see e.g., [8], Cor. 14.1.5).

As an example, consider  $w_\nu = 01^201^301^3$ . The standard factorization is  $w_\nu = w'_\nu w''_\nu$ ,  $w'_\nu = 01^201^3$ ,  $w''_\nu = 01^3$ , of respective lengths 7, 4, and respective 1-lengths 5, 3. For  $n = |w_\nu| - 1$ , the vector  $V_{10}$  is, up to a factor of  $\pm 1$ , the lower Christoffel word associated with the composition (4, 7); the alphabet is  $\{-5, 3\}$ , and  $V_{10} = \pm(-5, 3, -5, 3, 3, -5, 3, 3, -5, 3, 3)$ .

Now consider  $n = |w_\nu| - 3$ , so that  $i = 2$ . Then  $V_8$  is, up to a factor of  $\pm 1$ , the perfectly clustering word associated with the composition  $(4 - 2, 2, 7 - 2) = (2, 2, 5)$ , over the alphabet  $\{-5, -2, 3\}$ , and  $V_8 = \pm(-5, 3, -2, 3, -2, 3, -5, 3, 3)$  (compare to the example in Sect. 6).

It is known that each perfectly clustering word has a unique factorization as the product of two palindromes (Simpson and Puglisi [2], Cor. 4.4). We call this the *palindromic factorization*.

**Corollary 8.2.** *For all  $n \geq 1$ , up to multiplication by  $\pm 1$ , the word  $V_n$  is obtained from  $V_{n+1}$  by considering its unique palindromic factorization, and adding together the last component of its first factor and the first component of its second factor.*

## 9. PROOF OF THEOREM 8.1, FIRST PART

For simplicity, write  $w = w_\nu$ . Recall that  $|w| = N$ . The BW matrix of  $w$  is  $M_N(0, 1, r)$ . Since  $n = N - 1$ , the matrix  $G_n$ , which is the matrix of circular factors of length  $N - 1$  of  $w$ , is obtained from the latter matrix by removing its last column. Let  $D(F_n) = (a_0, \dots, a_n)$ ; it follows that  $a_i$  is the cofactor of the element in the last column, row  $i$ , of  $M_n(0, 1, r)$ . Hence  $V_n = D(G_n)$  is the last row of  $M_n(0, 1, r)^{-1}$ , multiplied by the determinant of  $M_n(0, 1, r)$ . The latter is equal to  $r\epsilon$ , by Corollary 4.2. Moreover,  $M_N(0, 1, r)^{-1} = M_N(-Q/r, 1 - Q/r, r^*)$ , where  $r^* \in [N]$  is defined by  $rr^* = 1 + QN$ , by Corollary 4.1. The last row of this matrix is the lower Christoffel word associated with the composition  $(N - r^*, r^*)$  over the alphabet  $\{-Q/r, 1 - Q/r\}$ . Multiplying by  $r$ , we obtain the lower Christoffel word associated with the composition  $(N - r^*, r^*)$  over the alphabet  $\{-Q, r - Q\}$ .

Thus, to finish the first part of the proof, it is enough to show the four equalities  $|w''| = N - r^*$ ,  $|w'| = r^*$ , and  $-Q = -|w'|_1$ ,  $r - Q = |w''|_1$ .

To prove the four equalities, we use the fact that the matrix

$$\begin{pmatrix} |w'|_0 & |w''|_0 \\ |w'|_1 & |w''|_1 \end{pmatrix}$$

has determinant 1 (see, *e.g.*, [8], Prop. 14.1.2). Thus

$$|w'|_0|w''|_1 - |w'|_1|w''|_0 = 1.$$

This implies that

$$(|w'|_1 + |w''|_1)(|w'|_1 + |w'|_0) = 1 + |w'|_1(|w'|_0 + |w'|_1 + |w''|_0 + |w''|_1).$$

That is, we have  $|w|_1|w'| = |w'|_1|w| + 1$ ; in other words,  $r|w'| = |w'|_1N + 1$ . Thus  $r^* = |w'|$  and  $Q = |w'|_1$ . The two other equalities follow from  $w = w'w''$ .

## 10. SOME PRELIMINARY RESULTS

In this section we assume the hypothesis of Theorem 8.1, in the second case. Recall that  $G_n$  is the matrix whose rows are the  $n + 1$  factors of length  $n$  of  $g$ , in decreasing lexicographic order.

For two matrices  $A, B$ , we write

$$A \xrightarrow{h} B \tag{10.1}$$

if for some  $k$ ,  $A$  is a  $(k + 1) \times k$  matrix,  $B$  is a  $k \times (k - 1)$  matrix, rows  $h - 1$  and  $h$  of  $A$  are equal except for their last entry, which is 1 for row  $h - 1$  and 0 for row  $h$ , and  $B$  is obtained from  $A$  by removing row  $h$  from this matrix, and then removing column  $k - 1$ , the last column. Note that, equivalently,  $B$  is obtained by removing row  $h - 1$  and the last column. Recall that we index matrix rows and columns starting from 0.

### Proposition 10.1.

1.  $G_n$  is obtained from  $G_{N-1}$  by removing the rows  $jq \bmod n$ ,  $j = 1, \dots, i$  and by keeping only the first  $n$  columns, with  $i = N - 1 - n$ .
2. Let  $h_i = (iq \bmod N) - d_i$ , where  $d_i$  is the number of  $j, 1 \leq j \leq i$ , such that  $jq \bmod N$  is smaller than  $iq \bmod N$ . Then

$$G_{n+1} \xrightarrow{h_i} G_n.$$

*Proof.*

1. Call *square* a  $2 \times 2$  submatrix of  $M = M_N(0, 1, r)$  as the one appearing in Corollary 3.3, so that it appears in two contiguous rows and two contiguous columns, and it is the identity matrix of order 2. As indicated in the lemma, these rows are numbered  $x - 1, x$  and the columns are numbered  $y - 1, y$  with  $x \equiv ry \pmod{N}$ , for some  $x, 1 \leq x \leq N - 1$ . We claim that if a square does not intersect the first  $n$  columns, then  $x \equiv qj \pmod{N}$  for some  $j = 1, \dots, i$ . The claim will be proved below.

Consider the matrix  $G$  obtained from  $G_{N-1}$  by removing the rows  $jq \bmod n$ ,  $j = 1, \dots, i$  and by keeping only the first  $n$  columns. Since  $G_{N-1}$  is obtained from  $M$  by removing its last column,  $G$  is equivalently obtained from  $M$  by removing the rows  $jq \bmod n$ ,  $j = 1, \dots, i$  and by keeping only the first  $n$  columns. In what follows, for simplicity, we index the rows of  $G$  by their index in  $M$  (equivalently in

$G_{N-1}$ ) before removal (*e.g.*, if we remove row 1 and keep 0 and 2, the first two rows of  $G$  are indexed by 0 and 2). Since the rows of  $M$  are strictly decreasing from the upper one to the lowest, the rows of  $G$  are weakly decreasing from the upper to the lower. Suppose, in order to get a contradiction, that they are not strictly decreasing. Then for some  $x' < x$ , rows  $x'$  and  $x$  appear in  $G$ , are successive, and are equal; hence the squares located in rows  $x', x' + 1, \dots, x$  do not intersect the first  $n$  columns of  $M$ . By the claim,  $x \equiv qj \pmod{N}$  for some  $j = 1, \dots, i$ , and this contradicts the fact that  $G$  contains row  $x$ . Thus the  $N - i = n + 1$  rows of  $G$  are strictly decreasing, hence distinct; each row of  $G$  being a prefix of a row of  $M$ , it is a circular factor of  $w$ , hence the rows of  $G$  are the  $n + 1$  circular factors of  $w$ , and  $G = G_n$ .

We prove the claim: if a square (with  $x, y$  as above) does not intersect the first  $n$  columns, then  $y - 1 \geq n$ , hence  $y \geq n + 1$ . Hence  $x \equiv ry$  for some  $y = n + 1, \dots, N - 1$ . Since  $q + r = N$ , and  $n + 1 = N - i$ , by letting  $j = N - y$ , we have

$$x \equiv ry \equiv -q(N - j) \equiv qj \pmod{N},$$

with  $j = i, \dots, 1$ .

2. By Part 1, the matrix  $G_n$  is obtained from  $G_{n+1}$  by removing its row  $iq \pmod{N}$ , if the rows are indexed as they are in  $G_{N-1}$ , and then removing the last column. But to get their actual index, one has to subtract the number of rows which were removed before from  $G_{N-1}$ , which is  $d_i$ . Note that  $G_{n+1}$  without its last column has two successive equal rows, which (in  $M$ ) are indexed  $(iq \pmod{N}) - 1$  and  $iq \pmod{N}$ . By Corollary 3.3, the last elements of these rows in  $G_{n+1}$  are necessarily 1 and 0. Hence we have  $G_{n+1} \xrightarrow{h_i} G_n$ . □

For two vectors  $V, W$ , we write

$$V \xrightarrow{h} W$$

if for some  $k$  we have  $V = (a_0, \dots, a_k)$  and  $W = (a_0, \dots, a_{h-2}, a_{h-1} + a_h, a_{h+1}, \dots, a_k)$ . That is,  $W$  is obtained from  $V$  by merging and adding the components ranked  $h - 1$  and  $h$  in  $V$ .

**Proposition 10.2.** *If  $A \xrightarrow{h} B$ , with the notation of (10.1), then  $D(A) \xrightarrow{h} (-1)^{k-h} D(B)$ , with the notation  $D(A)$  of Section 5.*

*Proof.* Write  $A = (L_0, \dots, L_k)$ , where the  $L_j$ 's are the rows of  $A$ ; similarly,  $B = (H_0, \dots, H_{k-1})$ . Also let  $D(A) = (a_0, \dots, a_k)$  and  $D(B) = (b_0, \dots, b_{k-1})$ . We compute within the matrix  $A$ , omitting mention of the rows before  $L_{h-1}$  and those after  $L_h$  for notational simplicity:

$$\begin{aligned} (-1)^{k-h}(a_{h-1} + a_h) &= -\det(\dots, \hat{L}_{h-1}, L_h, \dots) + \det(\dots, L_{h-1}, \hat{L}_h, \dots) \\ &= \det(\dots, L_{h-1} - L_h, \dots) = \det(\dots, (0, \dots, 0, 1), \dots) \\ &:= e. \end{aligned}$$

This is the cofactor of the entry in row  $h - 1$ , column  $k - 1$ , of  $A$ ; its corresponding submatrix is  $B$  without its row of rank  $h - 1$ , so that  $e = (-1)^{h-1+k-1} \det(H_0, \dots, \hat{H}_{h-1}, \dots, H_{k-1}) = b_{h-1}$ . Hence  $b_{h-1} = (-1)^{k-h}(a_{h-1} + a_h)$ .

We now show that  $b_j = (-1)^{k-h} a_j$  if  $j < h - 1$ . We have

$$\begin{aligned} (-1)^{k-j} a_j &= \det(\dots, \hat{L}_j, \dots, L_{h-1}, L_h, \dots) \\ &= \det(\dots, \hat{L}_j, \dots, L_{h-1} - L_h, L_h, \dots). \end{aligned}$$

Since  $L_{h-1} - L_h$  is equal to  $(0, \dots, 0, 1)$ , we see that  $(-1)^{k-j} a_j$  is equal to the cofactor of  $A_j$  (that is,  $A$  without its row  $j$ ) in row  $h - 2$ , column  $k - 1$  ( $A$  has  $k$  columns, the last is indexed  $k - 1$ ). The corresponding minor

is obtained from  $A_j$  by removing row  $h - 2$  (which is row  $h - 1$  of  $A$ ) and the last column; this matrix is  $B$  without its row  $j$ . Hence

$$\begin{aligned} (-1)^{k-j} a_j &= (-1)^{h-2+k-1} \det(\dots, \hat{H}_j, \dots) \\ &= (-1)^{h+k-1} (-1)^{k-1-j} b_j, \end{aligned}$$

implying that  $b_j = (-1)^{k-h} a_j$ .

The verification that  $b_j = (-1)^{k-h} a_{j+1}$  for  $j > h - 1$  is similar and left to the reader.  $\square$

For two words  $v, w$ , we write

$$v \xrightarrow{h} w$$

if  $v, w$  are perfectly clustering words over the alphabet  $\{a < b < c\}$ , if  $v = a_0 \cdots a_{h-2} a_{h-1} a_h a_{h+1} \cdots a_k$ ,  $a_{h-1} = a$ ,  $a_h = c$ , and  $w = a_0 \cdots a_{h-2} b a_{h+1} \cdots a_k$ . In other words,  $w$  is obtained from  $v$  by replacing a factor  $ac$  by  $b$ , and the indicated  $a, c$  are the letters ranked  $h - 1, h$  in  $v$ .

**Proposition 10.3.** *Let  $n = |w_\nu| - 1, \dots, |w_{\nu-1}| - 1$ , and  $i = |w_\nu| - 1 - n$ . Let  $u_n$  denote the perfectly clustering word associated with the composition  $(|w'_\nu| - i, i, |w'_\nu| - i)$  over the alphabet  $\{a < b < c\}$ . Then for  $n = |w_\nu| - 2, \dots, |w_{\nu-1}| - 1$  we have*

$$u_{n+1} \xrightarrow{h_i} u_n.$$

We need some preparation. Given a permutation  $\sigma$  of  $[N]$ , and  $0 \leq k \leq N - 1$ , as in [18], we let the *cyclic restriction* of  $\sigma$  to the interval  $[k]$  be the permutation of  $[k]$  obtained by removing each number larger than  $k - 1$  from any cycle form of  $\sigma$  (this does not depend on the chosen cycle form). Note that this definition is a discrete version of *Rauzy induction* for continuous interval exchanges [19, 20].

**Lemma 10.4.**

1. *Let  $\sigma$  be the circular symmetric discrete exchange of two intervals, associated with the composition  $(\gamma, \rho)$ ,  $\gamma, \rho$  relatively prime,  $\gamma + \rho = N$ . Let  $n = N - 1, \dots, N - \min(\gamma, \rho) - 1$ , and  $i = N - 1 - n$ , so that  $i = 0, \dots, \min(\gamma, \rho)$ . Let  $\sigma_n$  denote the cyclic restriction of  $\sigma$  to  $[n + 1]$ , so that  $\sigma = \sigma_{N-1}$ . Then  $\sigma_n$  is the symmetric discrete exchange of intervals, associated with the composition  $(\gamma - i, i, \rho - i)$ .*
2. *Let  $v = a_0 \cdots a_{N-1}$  be the lower Christoffel word over the alphabet  $\{a < c\}$  associated with  $\sigma$ . Suppose that  $\gamma < \rho$ , so that  $\min(\gamma, \rho) = \gamma$ . For  $n = N - 1, \dots, N - \gamma - 1$ , let  $v_n$  denote the standard word encoding of  $\sigma_n$  over the alphabet  $\{a < b < c\}$ , so that  $v = v_{N-1}$ . Then if  $n \leq N - 2$ , the quantity  $v_n$  is obtained from  $v$  by replacing in it the factor  $a_{j\gamma^* - 1} a_{j\gamma^*}$  by  $b$ , for each  $j = 1, \dots, i$ , where  $i = N - 1 - n$ , and where  $\gamma^* \in [N]$  is the inverse of  $\gamma$  modulo  $N$ ; moreover, all these factors are equal to  $ac$ , and are not overlapping.*

*Proof.*

1. This result was proved during the proof of Lemma 2 of [18].
2. By definition,  $v$  has  $\gamma$  occurrences of the letter  $a$ , which correspond to the numbers  $0, 1, \dots, \gamma - 1$  in the standard cyclic form of  $\sigma$ , and  $\rho$  occurrences of the letter  $c$ , which correspond to the numbers  $\gamma, \gamma + 1, \dots, N - 1$  in it. Likewise, by Part (1),  $v_n$  has  $\gamma - i$  occurrences of letter  $a$ ,  $i$  occurrences of letter  $b$  and  $\rho - i$  occurrences of letter  $c$ ; they correspond, in the standard cyclic form of  $\sigma_n$ , to the numbers  $0, 1, \dots, \gamma - i - 1$  for  $a$ ,  $\gamma - i, \dots, \gamma - 1$  for  $b$ , and  $\gamma, \dots, N - i - 1$  for  $c$ .

We know that  $\sigma$  is translation by  $\rho$  in  $\mathbb{Z}/N\mathbb{Z}$ :  $\sigma(x) = x + \rho$  in  $\mathbb{Z}/N\mathbb{Z}$ , and in particular  $\sigma^e(0) = e\rho$ . Since  $\rho$  is the opposite of  $\gamma$ , we have  $\gamma^* \rho \equiv -1 \pmod{N}$ ; hence  $\sigma^{j\gamma^*}(0) \equiv N - j \pmod{N}$  and  $\sigma^{j\gamma^* - 1}(0) \equiv \gamma - j \pmod{N}$ . Now by Part (1) above, the standard cyclic form of  $\sigma_n$  is obtained by removing the numbers  $N - j, j = 1, \dots, i$  from the cyclic form of  $\sigma$ . It follows that  $v_n$  is obtained from  $v$  by first removing the

letters  $a_{j\gamma^*}$ , which are all equal to  $c$  (since for  $j = 1, \dots, i$  we have  $N - j = N - 1, \dots, N - i = \gamma + \rho - i \geq \gamma$ , since  $i \leq \gamma < \rho$ ); and then by replacing the letters  $a_{j\gamma^*-1}$  (which are all equal to  $a$  since for  $j = 1, \dots, i$ ,  $\gamma - j = \gamma - 1, \dots, \gamma - i$ ) by  $b$ .

Since these factors are all equal to  $ac$ , they cannot overlap. □

**Corollary 10.5.** *For  $i = 1, \dots, \gamma$ , let  $d_i$  denote the number of  $j \leq i$  such that  $j\gamma^* \bmod N$  is smaller than  $i\gamma^* \bmod N$ . Then one has  $v_{n+1} \xrightarrow{i\gamma^* \bmod N - d_i} v_n$  for  $n = N - 2, \dots, N - \gamma - 1$ .*

*Proof.* By Lemma 10.4, for  $i = 1, \dots, \gamma$ , one has  $v_{i-1} \xrightarrow{e_i} v_i$  for some value of  $e_i$ , which is the position of the  $c$  being removed. If one takes into account the previous removals, one obtains the indicated value of  $e_i$ . □

*Proof of Proposition 10.3.* We prove it in the case  $|w''_\nu| < |w'_\nu|$  (the other case is similar). It is known (Berstel and de Luca; see [8], Cor. 14.5.6) that  $q^* = |w''_\nu|$  and  $r^* = |w'_\nu|$ . Hence  $q^* < r^*$  (recall that  $q^*$  is the inverse of  $q$  modulo  $N$ , and that we identify  $\mathbb{Z}/N\mathbb{Z}$  and  $[N] = \{0, 1, \dots, N - 1\}$ ).

By construction,  $u_{N-1}$  (which corresponds to  $i = 0$ ) is the perfectly clustering word over the alphabet  $\{a < b < c\}$  associated with the composition  $(|w''_\nu|, 0, |w'_\nu|)$ . In other words,  $u_{N-1}$  is the lower Christoffel word over the alphabet  $\{a, c\}$  associated with the composition  $(\gamma, \sigma) = (q^*, r^*)$ . With the notation of Lemma 10.4, we have  $u_n = v_n$ , and since  $\gamma^* = q$ , Corollary 10.5 implies the proposition, with  $h_i$  defined in Proposition 10.1. □

## 11. PROOF OF THEOREM 8.1, SECOND PART, AND COROLLARY 8.2

Let  $w = w_\nu$ . Consider the substitution  $\phi$  that maps

- $a$  onto  $-|w'|_1$ ,
- $b$  onto  $|w''|_1 - |w'|_1$ ,
- $c$  onto  $|w''|_1$ .

Since  $a < b < c$  and  $-|w'|_1 < |w''|_1 - |w'|_1 < |w''|_1$ , the map  $\phi$  maps perfectly clustering words onto perfectly clustering words, and preserves the associated composition. Moreover, for all words  $u, u' \in \{a, b, c\}^*$  and  $U, U' \in \{-|w'|_1, |w''|_1 - |w'|_1, |w''|_1\}^*$ , one has

$$u' \xrightarrow{h} u, U' \xrightarrow{h} U, \phi(u') = U' \Rightarrow \phi(u) = U. \quad (11.1)$$

The number of rows of  $G_n$  is  $k_i = n + 1$ , with  $i = N - 1 - n$ . Letting  $U_n = D(G_n)$ , for  $n = N - 2, \dots, |w_{\nu-1}| - 1$  we have  $U_{n+1} \xrightarrow{h_i} (-1)^{k_i - h_i} U_n$  by Propositions 10.1 and 10.2. Moreover  $u_{n+1} \xrightarrow{h_i} u_n$  by Proposition 10.3. By the first case of the Theorem, that is  $n = N - 1$ , one has  $U_{n-1} = \epsilon \phi(u_{N-1})$ . It follows then by (11.1) that  $U_n$  is equal to  $\phi(u_n)$  multiplied by  $\epsilon$  and by  $(-1)^t$ , where  $t = \sum_{1 \leq j \leq i} (k_j - h_j)$ . This proves the theorem.

We now prove Corollary 8.2. It is known from [2] that if  $w$  is a perfectly clustering Lyndon word, then  $w$  is the smallest, and its reversal  $\tilde{w}$ , the largest, word in the conjugation class of  $w$ . If  $w = uv$  is the palindromic factorization, then clearly  $\tilde{w} = vu$ . But it is also known that the position in  $w$  of the first letter of  $v$  corresponds to the largest digit in the standard cyclic form of the symmetric discrete interval exchange associated with  $w$  ([1] Proof of Theorem 4; a direct proof is given in [21], Lem. 6). Hence the factor  $ac$  in Lemma 10.4 corresponds to the palindromic factorization of  $v$ . The corollary follows.

## 12. EXAMPLE

We take  $w_\nu = 01101110111$ , so that  $q = 3, r = 8, w'_\nu = 0110111, w''_\nu = 0111$ . One has  $w_{\nu-1} = w'_\nu$ , of length 7, and  $N = 11$  is the length of  $w_\nu$ . The value of  $n$  varies from  $N - 1 = 10$  to  $|w_{\nu-1}| - 1 = 6$ . The rows of the matrices  $G_{10}, G_9, G_8, G_7, G_6$  are, in reverse lexicographic order, the factors of respective lengths 10, 9, 8, 7, 6 of the given Sturmian sequence, and actually the circular factors of  $w_\nu$ . The entries 1 and 0 of  $G_{n+1}$  are indicated

in boldface corresponding to the definition of  $G_{n+1} \xrightarrow{h_i} G_n$ : the two rows of  $G_{n+1}$  with the boldface entries coincide up to these entries. The numbers  $h_i$  are equal to  $(iq \bmod N) - d_i$ ; the corrective term  $d_i$  takes into account the rows that are previously removed. One has, for  $i = 1, 2, 3, 4$ ,  $iq \bmod N = 3, 6, 9, 1$ , and  $d_i = 0, 1, 2, 0$ , so that  $h_i = 3, 5, 7, 1$ , and

$$G_{10} \xrightarrow{3} G_9 \xrightarrow{5} G_8 \xrightarrow{7} G_7 \xrightarrow{1} G_6.$$

$$G_{10} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{1} \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{0} \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \end{pmatrix},$$

$$G_9 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{1} \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{0} \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix},$$

$$G_8 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{1} \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{0} \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix},$$

$$G_7 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{1} \\ 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{0} \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad G_6 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

The word  $u = u_{10}$  is the lower Christoffel word over  $\{a, c\}$  associated with the composition  $(|w''_\nu|, |w'_\nu|) = (4, 7)$ . Hence  $u = acaccaccacc$ . One has

$$\begin{aligned} u_{10} &= ac\mathbf{accaccacc} \xrightarrow{3} u_9 = abc\mathbf{accacc} \\ &\xrightarrow{5} u_8 = acbc\mathbf{bcacc} \xrightarrow{7} u_7 = acbc\mathbf{bc}bc \xrightarrow{1} u_6 = bbcb\mathbf{c}bc. \end{aligned}$$

The determinantal vectors are correspondingly, up to a factor  $\pm 1$ :

$$\begin{aligned} (-5, 3, -\mathbf{5}, \mathbf{3}, 3, -5, 3, 3, -5, 3, 3) &\xrightarrow{3} (-5, 3, -2, 3, -\mathbf{5}, \mathbf{3}, 3, -5, 3, 3) \\ &\xrightarrow{5} (-5, 3, -2, 3, -2, 3, -\mathbf{5}, \mathbf{3}, 3) \xrightarrow{7} (-\mathbf{5}, \mathbf{3}, -2, 3, -2, 3, -2, 3) \\ &\xrightarrow{1} (-2, -2, 3, -2, 3, -2, 3), \end{aligned}$$

where the two components which are added are indicated in boldface.

As the reader may verify, the transformations between vectors above correspond to the palindromic factorization, as described in Corollary 8.2, and the boldface components indicate the palindromic factorization.

For the subsequent determinantal vectors, of smaller length, one has

$$\begin{aligned} (-2, -\mathbf{2}, \mathbf{3}, -2, 3, -2, 3) &\rightarrow (-2, 1, -\mathbf{2}, \mathbf{3}, -2, 3) \rightarrow (-2, 1, 1, -\mathbf{2}, \mathbf{3}) \\ &\rightarrow (-\mathbf{2}, \mathbf{1}, 1, 1) \rightarrow (-\mathbf{1}, \mathbf{1}, 1) \rightarrow (\mathbf{0}, \mathbf{1}) \rightarrow (1). \end{aligned}$$

### 13. CONTINUED FRACTIONS

Define the matrix

$$P(a) = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}.$$

Define the *continuant polynomials*  $K_n(x_1, \dots, x_n)$  by  $K_0() = 1$ ,  $K_1(x_1) = x_1$  and the recursion

$$K_n(x_1, \dots, x_n) = K_{n-1}(x_1, \dots, x_{n-1})x_n + K_{n-2}(x_1, \dots, x_{n-2})$$

if  $n \geq 2$ . It is well known (combine, e.g., [22], Eq. (4), Sect. 4.5.3) with [23], that

$$P(n_0) \cdots P(n_k) = \begin{pmatrix} K(n_0, \dots, n_k) & K(n_0, \dots, n_{k-1}) \\ K(n_1, \dots, n_k) & K(n_1, \dots, n_{k-1}) \end{pmatrix}, \quad (13.1)$$

where the index of  $K$  is omitted, since it is redundant. Continued polynomials are used to compute the numerators and denominators of finite continued fractions. One has

$$[n_0, \dots, n_k] = \frac{K(n_0, \dots, n_k)}{K(n_1, \dots, n_k)}.$$

In particular, if a Christoffel word  $w$  has slope  $[n_0, \dots, n_k]$ , then its length is

$$|w| = K(n_0, \dots, n_k) + K(n_1, \dots, n_k).$$

The following result is essentially well known. We give a proof for sake of completeness.

**Proposition 13.1.** *Let  $w$  be a lower Christoffel word of slope  $[n_0, \dots, n_m]$  over the alphabet  $\{0 < 1\}$ , with standard factorization  $w = w'w''$ . Then*

$$P(n_0) \cdots P(n_{m-1})P(n_m - 1) = \begin{cases} \begin{pmatrix} |w'|_1 & |w''|_1 \\ |w'|_0 & |w''|_0 \end{pmatrix} & \text{if } m \text{ even;} \\ \begin{pmatrix} |w''|_1 & |w'|_1 \\ |w''|_0 & |w'|_0 \end{pmatrix} & \text{if } m \text{ odd.} \end{cases}$$

*Proof.* We use the fact that the set of lower Christoffel words of length  $\geq 2$  is the smallest set containing  $01$  and closed under the two substitutions  $G, \tilde{D}$ , where  $G(0) = 0, G(1) = 01, \tilde{D}(0) = 01, \tilde{D}(1) = 1$  (see [7] or [8]). Moreover, these substitutions preserve lower Christoffel words, and hence the standard factorization.

The result is clear for  $w = 01$ , of slope  $1 = [1]$ , of standard factorization  $w = 0 \cdot 1$ , because  $P(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , and since we are here in the even case.

Suppose now that the result is true for  $u$ , of slope having the continued fraction expansion  $[n_0, \dots, n_m]$ . It is enough to prove the proposition for  $w = G(u)$  and  $w = \tilde{D}(u)$ . We carry that out for the first case; the second one is similar and somewhat easier. Thus we assume that  $w = G(u)$ . This implies that  $w' = G(u'), w'' = G(u'')$ , and that

$$|w|_0 = |u|_0 + |u|_1, |w|_1 = |u|_1,$$

and similarly for  $w', w''$ . Let  $s(w)$  denote the slope of  $w$ . Then

$$\frac{1}{s(w)} = \frac{|w|_0}{|w|_1} = \frac{|u|_0 + |u|_1}{|u|_1} = 1 + \frac{1}{s(u)} = 1 + [0, n_0, \dots, n_m].$$

If  $n_0 > 0$ , then  $1 + \frac{1}{s(u)} = [1, n_0, \dots, n_m]$  and  $s(w)$  has the continued fraction expansion  $[0, 1, n_0, \dots, n_m]$ . If  $n_0 = 0$ , then

$$\frac{1}{s(w)} = 1 + \frac{1}{s(u)} = 1 + [n_1, \dots, n_m] = [n_1 + 1, \dots, n_m].$$

Hence  $s(w)$  has the continued fraction expansion  $[0, n_1 + 1, \dots, n_m]$ . Note that in both cases, the continued fractions of the slopes of  $u$  and  $w$  have the same length parity.

Suppose that  $m$  is even (the odd case is similar). Then by induction

$$P(n_0) \cdots P(n_{m-1})P(n_m - 1) = \begin{pmatrix} |u'|_1 & |u''|_1 \\ |u'|_0 & |u''|_0 \end{pmatrix}.$$

Suppose that  $n_0 > 0$ ; noting that  $P(0)$  is the matrix of the transposition  $(1, 2)$ , we compute

$$\begin{aligned} & P(0)P(1)P(n_0) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} P(n_0) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} |u'|_1 & |u''|_1 \\ |u'|_0 & |u''|_0 \end{pmatrix} \\ &= \begin{pmatrix} |u'|_1 & |u''|_1 \\ |u'|_0 + |u'|_1 & |u''|_0 + |u''|_1 \end{pmatrix} = \begin{pmatrix} |w'|_1 & |w''|_1 \\ |w'|_0 & |w''|_0 \end{pmatrix}, \end{aligned}$$

which settles this case. Suppose that  $n_0 = 0$ ; we compute

$$\begin{aligned} & P(0)P(n_1 + 1) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ n_1 + 1 & 1 \end{pmatrix} P(n_2) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ n_1 & 1 \end{pmatrix} P(n_2) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} P(0)P(n_1)P(n_2) \cdots P(n_{m-1})P(n_m - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} |u'|_1 & |u''|_1 \\ |u'|_0 & |u''|_0 \end{pmatrix}, \end{aligned}$$

and we conclude as previously.  $\square$

Let  $g$  be a Sturmian sequence whose slope  $s$  has the infinite continued fraction expansion  $s = [n_0, n_1, n_2, \dots]$ . The slopes of the Christoffel words that are factors of  $g$  are the *semi-convergents* of  $s$ ; that is, the rationals of the form  $[n_0, \dots, n_{m-1}, h]$ , with  $m \geq 0$ , and  $1 \leq h \leq n_m$ . This follows from, *e.g.*, Section 6.7 on the Stern–Brocot tree in the book [17] of Graham, Knuth and Patashnik; also see [8] Theorem 14.2.3 and Corollary 14.2.4.

After this preparation, we may translate Theorem 8.1 in terms of continued fractions; we use the notation of this theorem. The integer  $\nu$  is defined by  $|w_{\nu-1}| \leq n \leq |w_\nu| - 1$ . Since the slope of  $w_\nu$  is a semi-convergent of  $s$ ,  $n$  is equivalently defined by the inequalities

$$m \geq 0, \quad 1 \leq h \leq n_m,$$

and

$$\begin{aligned} & K(n_0, \dots, n_{m-1}, h - 1) + K(n_1, \dots, n_{m-1}, h - 1) \leq n \\ & \leq K(n_0, \dots, n_{m-1}, h) + K(n_1, \dots, n_{m-1}, h) - 1, \end{aligned}$$

with the convention that  $h - 1$  is omitted if  $h = 1$ . Note that  $w_\nu$  is the lower Christoffel word of slope  $[n_0, \dots, n_{m-1}, h]$ . Then we have, by Proposition 13.1,

$$P(n_0) \cdots P(n_{m-1})P(h-1) = \begin{cases} \begin{pmatrix} |w'_\nu|_1 & |w''_\nu|_1 \\ |w'_\nu|_0 & |w''_\nu|_0 \end{pmatrix} & \text{if } m \text{ even;} \\ \begin{pmatrix} |w''_\nu|_1 & |w'_\nu|_1 \\ |w''_\nu|_0 & |w'_\nu|_0 \end{pmatrix} & \text{if } m \text{ odd.} \end{cases}$$

To finish, note that many authors consider, instead of the slope  $s$  as defined here, the limit of  $|w|_1/|w|$  when  $|w|$  tends to infinity,  $w$  factor of  $g$ ; we call this limit the *Slope* of  $g$ , denoted by  $S$ . These two numbers are related by

$$S = s/(1 + s), s = S/(1 - S), S^{-1} = 1 + s^{-1}.$$

It is easy to deduce that if the continued fraction of  $S$  is of the form  $[0, a_1, \dots, a_m, \dots]$ , where the  $a_i$  are positive integers, then

$$s = \begin{cases} [0, a_1 - 1, a_2, \dots, a_m, \dots], & \text{if } a_1 \geq 2; \\ [a_2, \dots, a_m, \dots], & \text{if } a_1 = 1. \end{cases}$$

## 14. FIBONACCI

The Fibonacci sequence  $f$  is the limit of the sequence of finite words  $\Phi_\nu$ , where  $\Phi_1 = 1$ ,  $\Phi_2 = 0$ , and  $\Phi_\nu = \Phi_{\nu-1}\Phi_{\nu-2}$  for  $\nu \geq 3$  (see, e.g., [15], Sect. 7.1). The length of  $\Phi_\nu$  is  $F_\nu$ , the Fibonacci number (with  $F_1 = F_2 = 1$ ). The slope of  $f$ , which is the limit of  $|\Phi_\nu|_1/|\Phi_\nu|_0$ , is  $(\sqrt{5} - 1)/2$ . This number corresponds in the Stern–Brocot tree to the infinite path, starting from the root, and indexed by  $lrlrlr \dots$  (alternatively left and right).

As explained in Section 7, the corresponding sequence  $(w_\nu)_{\nu \in \mathbb{N}}$  of lower Christoffel words is the sequence of words on the tree of the Christoffel word on the infinite path, starting from the root and indexed as above. It follows from the direct construction of the Christoffel tree (see [8], Thm. 14.1.1) that these words are defined by  $w_{-2} = 1, w_{-1} = 0$  and the recursion, for any natural number  $\nu$

$$w_\nu = \begin{cases} w_{\nu-1}w_{\nu-2}, & \text{if } \nu \text{ even;} \\ w_{\nu-2}w_{\nu-1}, & \text{if } \nu \text{ odd.} \end{cases} \quad (14.1)$$

One has  $|w_\nu|_0 = F_{\nu+2}, |w_\nu|_1 = F_{\nu+1}$ , as is easily verified.

Consider the  $n$ -th determinantal vector  $V_n$  of  $f$ . Then the integer  $\nu$  in Theorem 8.1 is defined by  $F_{\nu+2} \leq n \leq F_{\nu+3} - 1$ ; consequently  $i = F_{\nu+3} - 1 - n$ .

Since the standard factorization of  $w_\nu$  is given by (14.1), we obtain that, up to multiplication by  $\pm 1$ , the word  $V_n$  is the perfectly clustering word associated with

- the composition  $(F_{\nu+1} - i, i, F_{\nu+2} - i)$  over the alphabet  $\{-F_\nu, -F_{\nu-2}, F_{\nu-1}\}$  if  $\nu$  is even, and
- the composition  $(F_{\nu+2} - i, i, F_{\nu+1} - i)$  over the alphabet  $\{-F_{\nu-1}, F_{\nu-2}, F_\nu\}$  if  $\nu$  is odd.

In particular, if  $n = F_{\nu+3} - 1$ , the Sturmian determinants of  $f$  of order  $n$  take, up to the sign, only the two values  $F_\nu, F_{\nu-1}$ . And if  $F_{\nu+2} \leq n < F_{\nu+3} - 1$ , they take, up to the sign, only the three values  $F_\nu, F_{\nu-1}, F_{\nu-2}$ .

Moreover, since a perfectly clustering word begins with its smallest letter and ends with its largest, and recalling the definition of  $V_n$  (Sect. 5), we see that the determinant of the  $n$  lexicographically largest factors

of length  $n$  of  $f$  is  $\pm F_{\nu-1}$  if  $\nu$  is even, and  $\pm F_\nu$  if  $\nu$  is odd; and for the  $n$  smallest, it is  $\pm F_\nu$  if  $\nu$  is even, and  $\pm F_{\nu-1}$  if  $n$  is odd.

Recall that a factor  $u$  of a Sturmian sequence on  $\{0, 1\}$  is called (*right*) *special* if  $u0, u1$  are both factors of the sequence. In Proposition 10.1, the description of how one passes from  $G_{n+1}$  to  $G_n$ , and the subsequent proofs, show that the determinant of the factors of the sequence of length  $n$ , without the special factor, is the middle element of the alphabet. Hence, for the Fibonacci sequence, it is  $\pm F_{\nu-2}$ .

### 14.1. Computation of the sign

It follows from what precedes that for  $n = F_{\nu+3} - 1$ , the word  $V_n$  is the product of  $\epsilon$  by the lower Christoffel word associated with the composition  $(F_{\nu+1}, F_{\nu+2})$  over the alphabet  $\{-F_\nu, F_{\nu-1}\}$  if  $\nu$  is even, and the composition  $(F_{\nu+2}, F_{\nu+1})$  over the alphabet  $\{-F_{\nu-1}, F_\nu\}$  if  $\nu$  is odd. Here  $\epsilon$  is the sign of the permutation ‘‘multiplication by  $F_{\nu+1}$  in  $\mathbb{Z}/F_{\nu+3}\mathbb{Z}$ ’’. It is therefore of interest to compute this sign.

Recall that the *Lucas numbers*  $L_n$  are defined by the same recursion as the Fibonacci numbers, with the initial conditions  $L_0 = 2, L_1 = 1$ .

**Proposition 14.1.** *Let  $m \geq 3$  be an integer. Consider the permutation induced by multiplication by  $F_{m-2}$  in  $\mathbb{Z}/F_m\mathbb{Z}$ . The cyclic type of this permutation is*

- (a)  $1^{L_{m/2}2^{(F_m-L_{m/2})/2}}$  if  $m \equiv 0 \pmod{4}$ ;
- (b)  $1^{F_{m/2}2^{(F_m-F_{m/2})/2}}$  if  $m \equiv 2 \pmod{4}$ ;
- (c)  $1^14^{(F_m-1)/4}$  if  $m \equiv 1, 5 \pmod{6}$ ;
- (d)  $1^24^{(F_m-2)/4}$  if  $m \equiv 3 \pmod{6}$ .

Furthermore, if  $s_m$  denotes the sign of this permutation, we have  $s_m = 1$  if  $m \equiv 1, 2, 3, 4, 9, 11 \pmod{12}$  and  $s_m = -1$  if  $m \equiv 0, 5, 6, 7, 8, 10 \pmod{12}$ .

Note that  $s_n = \left(\frac{F_{m-2}}{F_m}\right)_Z$ , the Zolotareff symbol. We begin with a lemma.

**Lemma 14.2.** *For  $k \geq 0$  we have*

- (a)  $\gcd(F_{6k+1} - 1, F_{6k+3}) = 2$ .
- (b)  $\gcd(F_{6k+3} - 1, F_{6k+5}) = 1$ .
- (c)  $\gcd(F_{6k-1} - 1, F_{6k+1}) = 1$ .

*Proof.*

- (a) If  $n \equiv 0 \pmod{3}$  then both  $F_{2n+1} - 1$  and  $F_{2n+3}$  are even. To see that  $\gcd(F_{2n+1} - 1, F_{2n+3}) = 2$ , observe that

$$(F_{2n+1} - 1)(F_{2n} - 2) - F_{2n+3}(F_{2n-2} - 1) = 4$$

for all  $n$ . Furthermore, if  $n \equiv 0 \pmod{3}$ , then  $F_{2n}$  is even and  $F_{2n-2}$  is odd, so we can divide both  $F_{2n} - 2$  and  $F_{2n-2} - 1$  by 2 to get 2 as an integral linear combination of  $F_{2n+1} - 1$  and  $F_{2n+3}$ . Thus this identity holds for  $n = 6k + 1$ .

- (b) It is easy to check  $(F_{2n+1} - 1)(F_{2n+2} - 1) - F_{2n+3}(F_{2n} - 1) = 2$  for all  $n$ . If  $n \equiv 1 \pmod{3}$ , then both  $F_{2n+2}$  and  $F_{2n}$  are odd, so we can divide both  $F_{2n+2} - 1$  and  $F_{2n} - 1$  by 2 to get an integral linear combination of  $F_{2n+1} - 1$  and  $F_{2n+3}$  that equals 1. Thus this identity holds for  $n = 6k + 3$ .
- (c) It is easy to check that  $F_{2n+3}F_{2n-1} - (F_{2n+1} - 1)(F_{2n+1} + 1) = 2$  for all  $n$ . If  $n \equiv 2 \pmod{3}$  then  $F_{2n-1}$  is even and  $F_{2n+1}$  is odd, so we can divide both  $F_{2n-1}$  and  $F_{2n+1} + 1$  by 2 to get an integral linear combination of  $F_{2n+1} - 1$  and  $F_{2n+3}$  that equals 1. Thus this identity holds for  $n = 6k - 1$ .

□

*Proof of Proposition 14.1.* Let  $s_m$  denote the sign of the permutation induced by multiplication by  $F_{m-2}$  in  $\mathbb{Z}/F_m\mathbb{Z}$ .

*Case 1:  $m$  is even.*

From the classical Fibonacci number identity  $F_i^2 = F_{i-2}F_{i+2} + (-1)^i$ , valid for all  $i$ , we see that if  $i$  is even then  $F_i^2 = F_{i-2}F_{i+2} + 1$  and so  $F_i^2 \equiv 1 \pmod{F_{i+2}}$ . Setting  $i = m - 2$  shows that multiplication by  $F_{m-2}$  induces cycles of size either 1 or 2.

The number  $a$  will be in a cycle of length 1 iff  $aF_{m-2} = a \pmod{F_m}$ , or in other words  $a(F_{m-2} - 1) \equiv 0 \pmod{F_m}$ . Now there are two cases, namely (a)  $m = 4k$ , and (b)  $m = 4k + 2$ .

*Case 1a:  $m = 4k$ .*

We use the identities

$$\begin{aligned} F_{4k-2} - 1 &= L_{2k}F_{2k-2} \\ F_{4k} &= F_{2k}L_{2k}. \end{aligned}$$

Suppose  $a(F_{4k-2} - 1) \equiv 0 \pmod{F_{4k}}$ . Then

$$\begin{aligned} aL_{2k}F_{2k-2} &= a(F_{4k-2} - 1) \\ &\equiv 0 \pmod{F_{4k}} \\ &\equiv 0 \pmod{F_{2k}L_{2k}}, \end{aligned}$$

so cancelling  $L_{2k}$  gives the congruence  $aF_{2k-2} \equiv 0 \pmod{F_{2k}}$ . But  $F_{2k-2}$  is relatively prime to  $F_{2k}$ , so it must be that  $a$  is actually a multiple of  $F_{2k}$  and every such multiple works. There are  $F_{4k}/F_{2k} = L_{2k}$  such multiples in the numbers mod  $F_{4k}$ . So there are  $L_{2k}$  cycles of length 1 and  $(F_{4k} - L_{2k})/2$  cycles of length 2. The sign of the permutation is then  $s_m = (-1)^e$  where  $e \equiv (F_{4k} - L_{2k})/2 \pmod{2}$ . It is easy to see that  $e = 0$  iff  $k \equiv 0 \pmod{3}$  and  $e = 1$  otherwise. So  $s_m = 1$  if  $m \equiv 4 \pmod{12}$  and  $-1$  if  $m \equiv 0, 8 \pmod{12}$ .

*Case 1b:  $m = 4k + 2$ .*

We use the identities

$$\begin{aligned} F_{4k} - 1 &= L_{2k-1}F_{2k+1} \\ F_{4k+2} &= F_{2k+1}L_{2k+1}. \end{aligned}$$

As above,

$$\begin{aligned} aL_{2k-1}F_{2k+1} &= a(F_{4k} - 1) \\ &\equiv 0 \pmod{F_{4k+2}} \\ &\equiv 0 \pmod{F_{2k+1}L_{2k+1}}, \end{aligned}$$

and so cancelling  $F_{2k+1}$  from this congruence we see  $aL_{2k-1} \equiv 0 \pmod{L_{2k+1}}$ . But  $L_{2k-1}$  is relatively prime to  $L_{2k+1}$ , so it must be that  $a$  is a multiple of  $L_{2k+1}$ , and every such multiple works. There are  $F_{4k+2}/L_{2k+1} = F_{2k+1}$  such multiples in the numbers mod  $F_{4k+2}$ . So there are  $F_{2k+1}$  cycles of length 1 and  $(F_{4k+2} - F_{2k+1})/2$  cycles of length 2.

The sign of the permutation is then  $s_m = (-1)^e$ , where  $e \equiv (F_{4k+2} - F_{2k+1})/2 \pmod{2}$ . It is easy to see that  $e = 0$  if  $k \equiv 0 \pmod{3}$  and  $e = 1$  otherwise. So now we see  $s_m = 1$  if  $m \equiv 2 \pmod{12}$ , and  $-1$  if  $m \equiv 6, 10 \pmod{12}$ .

*Case 2:  $m$  is odd.*

We use the identity  $F_{m-2}^4 - 1 = F_{m-4}F_{m-3}F_{m-1}F_m$ , valid for all  $m$ . Then  $aF_{m-2}^4 = a \pmod{F_m}$ , so it is clear that there can only be cycles of length 1, 2, or 4.

We now show there can never be cycles of length 2. If there were, as above, we would have  $a(F_{m-2}^2 - 1) \equiv 0 \pmod{F_m}$ . But since  $m$  is odd we have  $F_{m-2}^2 - 1 = F_{m-4}F_m - 2$ , so

$$\begin{aligned} a(F_{m-2}^2 - 1) &= a(F_{m-4}F_m - 2) \\ &\equiv -2a \pmod{F_m} \\ &\equiv 0 \pmod{F_m}, \end{aligned}$$

which forces  $F_m$  to be even, which in turn forces  $m = 6k + 3$  and  $a = F_{6k+3}/2$ . Now  $F_{6k+1}$  is odd, so  $F_{6k+1} \equiv 1 \pmod{2}$ . Multiplying through by  $a$  gives  $aF_{6k+1} \equiv a \pmod{F_{6k+3}}$ , so in fact this supposed cycle of size 2 is actually of size 1.

*Case 2a:  $m = 6k + 1$ .* Except for 0, the only cycles are of length 4, and there are  $(F_{6k+1} - 1)/4$  of them.

From above it suffices to show there are no cycles of length 1. If there were we would have  $aF_{6k-1} = a \pmod{F_{6k+1}}$  so  $a(F_{6k-1} - 1) \equiv 0 \pmod{F_{6k+1}}$ . But  $\gcd(F_{6k-1} - 1, F_{6k+1}) = 1$  by Lemma 14.2 (c), which forces  $a \equiv 0$ , impossible. So  $s_m = (-1)^e$  where  $e \equiv (F_{6k+1} - 1)/4 \pmod{2}$ . It is easy to check  $s_m = 1$  if  $k \equiv 1 \pmod{2}$  and  $s_m = -1$  otherwise. So we get  $s_m = -1$  for  $m \equiv 7 \pmod{12}$  and  $s_m = 1$  for  $m \equiv 1 \pmod{12}$ .

*Case 2b:  $m = 6k + 3$ .*

We will see there are exactly two cycles of length 1; the rest therefore have to be of length 4. Suppose  $aF_{6k+1} \equiv a \pmod{F_{6k+3}}$ . Then  $a(F_{6k+1} - 1) \equiv 0 \pmod{F_{6k+3}}$ . But  $\gcd(F_{6k+1} - 1, F_{6k+3}) = 2$  by Lemma 14.2 (a), so the only possibilities are  $a = 0$  or  $a = F_{6k+3}/2$ . Thus  $s_m = (-1)^e$  where  $e \equiv (F_{6k+3} - 2)/4 \pmod{2}$ . It is now easy to check  $s_m = 1$  in all cases.

*Case 2c:  $m = 6k + 5$ .*

Except for 0, there are only cycles of length 4, and there are  $(F_{6k+5} - 1)/4$  of them. Suppose there is another cycle of length 1. Then we would have  $aF_{6k+3} = a \pmod{F_{6k+5}}$  so  $a(F_{6k+3} - 1) \equiv 0 \pmod{F_{6k+5}}$ . But by Lemma 14.2 (b), we know that  $\gcd(F_{6k+3} - 1, F_{6k+5}) = 1$ , which forces  $a \equiv 0$ , impossible.

So  $s_m = (-1)^e$  where  $e \equiv (F_{6k+5} - 1)/4 \pmod{2}$ . It is easy to check  $s_m = 1$  if  $k \equiv 1 \pmod{2}$  and  $-1$  otherwise. So we get  $s_m = -1$  for  $n \equiv 3 \pmod{12}$  and  $s_m = 1$  for  $m \equiv 9 \pmod{12}$ .

Summing up, we have

$$s_m = \begin{cases} 1, & \text{if } m \equiv 1, 2, 3, 4, 9, 11 \pmod{12}; \\ -1, & \text{if } m \equiv 0, 5, 6, 7, 8, 10 \pmod{12}. \end{cases}$$

□

## 15. POSTSCRIPT

Most of the work presented in this article was done between March and June 2024.

As we were putting the finishing touches on this paper in mid-September 2024, we were informed about the interesting new paper of Luca Zamboni [24], where he independently found the result about the group structure.

### ACKNOWLEDGMENTS

This article arose after mail exchanges with Jean-Paul Allouche. We heartily thank him for his suggestions and encouragements. We also thank the two anonymous referees for their thorough reading and suggestions.

## FUNDING

The research of Christophe Reutenauer is supported by a grant from NSERC, 2020-04685. The research of Jeffrey Shallit is supported by grants from NSERC, 2018-04118 and 2024-03725.

## CONFLICTS OF INTEREST

The authors have nothing to disclose.

## DATA AVAILABILITY STATEMENT

This article has no associated data generated.

## AUTHOR CONTRIBUTION STATEMENT

Both authors contributed to the research reported here, and both authors did the writing.

## REFERENCES

- [1] S. Ferenczi and L.Q. Zamboni, Clustering words and interval exchanges. *J. Integer Sequences* **16** (2013) Article 13.2.1.
- [2] J. Simpson and S.J. Puglisi, Words with simple Burrows–Wheeler transforms. *Electron. J. Combinatorics* **15** (2008) Paper R83.
- [3] A. Restivo and G. Rosone, Burrows–Wheeler transform and palindromic richness. *Theor. Comput. Sci.* **410** (2009) 3018–3026.
- [4] S. Mantaci, A. Restivo and M. Sciortino, Burrows–Wheeler transform and Sturmian words. *Inform. Process. Lett.* **86** (2001) 241–246.
- [5] W.-F. Chuan,  $\alpha$ -words and factors of characteristic sequences. *Discrete Math.* **177** (1997) 33–50.
- [6] M. Burrows and D.J. Wheeler, A block-sorting lossless data compression algorithm. Systems Research Center research report (1994). Available at <https://www.cs.jhu.edu/~langmea/resources/burrows-wheeler.pdf>.
- [7] M. Lothaire, Algebraic Combinatorics on Words. Cambridge University Press (2002).
- [8] C. Reutenauer, From Christoffel Words to Markoff Numbers. Oxford University Press (2019).
- [9] M. Zolotareff, Nouvelle démonstration de la loi de réciprocité quadratique. *Nouvelles Ann. Math.* **11** (1872) 354–362.
- [10] M. Riesz, Sur le lemme de Zolotareff et sur la loi de réciprocité des restes quadratiques. *Math. Scand.* **1** (1953) 159–169.
- [11] J.L. Brenner, Zolotareff’s theorem on the Legendre symbol, *Pac. J. Math.* **45** (1973) 413–414.
- [12] G. Rousseau, On the Jacobi symbol. *J. Number Theory* **48** (1994) 109–111.
- [13] M. Lerch, Sur un théorème arithmétique de Zolotarev. *Bull. Int. l’Ac. Prague* **3** (1896) 34–37.
- [14] F.W. Barnes, A permutation reciprocity law. *Ars Combin.* **29A** (1990) 155–159.
- [15] J.-P. Allouche and J. Shallit, Automatic Sequences: Theory, Applications, Generalizations, Cambridge University Press (2003).
- [16] I. Pak and A. Redlich, Long cycles in  $abc$ -permutations. *Funct. Anal. Other Math.* **2** (2008) 87–92.
- [17] R.L. Graham, D.E. Knuth and O. Patashnik, Concrete Mathematics. Addison Wesley (1994).
- [18] V. Berthé and C. Reutenauer, On the three-distance theorem, *Math. Intell.* **46** (2024) 183–188.
- [19] G. Rauzy, Échanges d’intervalles et transformations induites. *Acta Arith.* **34** (1979) 315–328.
- [20] W.A. Veech, Interval exchange transformations. *J. Analyse Math.* **33** (1978) 222–272.
- [21] M. Lapointe and C. Reutenauer, Characterizations of perfectly clustering words. *Electron. J. Combinatorics* **32** (2025) #P3.8.
- [22] D.E. Knuth, *The Art of Computer Programming*, vol. 2 of *Seminumerical Algorithms*, 3rd edn. Addison-Wesley (1998).
- [23] J.S. Frame, Continued fractions and matrices. *Am. Math. Monthly* **56** (1949) 98–103.
- [24] L.Q. Zamboni, On Christoffel words & their lexicographic array. arXiv preprint (2024). <https://arxiv.org/abs/2409.07974>.
- [25] M. Lapointe, *Combinatoire des mots: mots parfaitement amassants, triplets de Markoff et graphes chenilles*, Thèse de doctorat, Université du Québec à Montréal (2021).



**Please help to maintain this journal in open access!**

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org).

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.