HAL
open science

Second order balance property on Christoffel words

Lama Tarsissi, Laurent Vuillon

To cite this version:

Lama Tarsissi, Laurent Vuillon. Second order balance property on Christoffel words. Daniel Slamanig; Elias Tsigaridas; Zafeirakis Zafeirakopoulos. Mathematical Aspects of Computer and Information Sciences. 8th International Conference, MACIS 2019, Gebze, Turkey, November 13–15, 2019, Revised Selected Papers, 11989, Springer, pp.295-312, 2020, Lecture Notes in Computer Science, 978-3-030-43119-8. 10.1007/978-3-030-43120-4_23 . hal-02433984v2

HAL Id: hal-02433984

https://hal.science/hal-02433984v2

Submitted on 26 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Second order balance property on Christoffel words

Lama Tarsissi[1] ⋆ and Laurent Vuillon[2]

[1] LAMA, Univ Gustave Eiffel, UPEM, Univ Paris Est Creteil, CNRS, F-77447 Marne-la-Vallée, France
lama.tarsissi@esiee.fr

[2] LAMA, Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, 73000 Chambéry, France
laurent.vuillon@univ-smb.fr

Abstract. In this paper we study the *balance matrix* that gives the order of balance of any binary word. In addition, we define for Christoffel words a new matrix called *second order balance matrix*. This matrix gives more information on the balance property of a word that codes the number of occurrences of the letter 1 in successive blocks of the same length for the studied Christoffel word. By taking the maximum of the Second order balance matrix we define the second order of balance and we are able to order the Christoffel words according to these values. Our construction uses extensively the continued fraction associated with the slope of each Christoffel word, and we prove a recursive formula based on fine properties of the Stern-Brocot tree to construct second order matrices. Finally, we show that an infinite path on the Stern-Brocot tree, which minimizes the second order of balance is given by a path associated with the Fibonacci word.

Index terms— Balance property. Second order balance property. Christoffel words. Stern-Brocot tree. Continued fractions

1 Introduction

Balanced words appear in many developments of combinatorics on words and the balance property is considered as a fine tool to investigate the structure of words [17, 21]. As a typical example of infinite balanced words, Sturmian words could be constructed equivalently by discretizations of irrational slope lines in a square grid [11, 30], by billiard words in a square [2, 24] or by coding of irrational rotations on a unit circle with a partition in two intervals [11]. The finite balanced words are given by discretizations of rational slope lines in a square grid and have been studied in particular by Christoffel [9]. Interestingly, finite and infinite balanced words show up in specific optimization problems [1, 26, 30] and for example optimal schedules for job-shop problems with two tasks are exactly given by balanced words [14, 19, 30]. Furthermore, particular solutions of job-shop problems with k tasks sharing the same ressource [1] are given by finite or infinite balanced words on a $k-$letters alphabet where the balanced property is checked on each letter of the alphabet[1, 23, 29]. More precisely, for $k = 2$ the solutions of the job-shop problem is coded by an infinite word which is either a periodic balanced word or an aperiodic balanced word [14, 19]. The situation gets more complicated for $k > 2$, which leads to the famous Fraenkel's conjecture[12, 13]. It is restated in combinatorics on words terms: An infinite word on a $k-$letters alphabet balanced on each letter of the alphabet and with all letters frequencies pairwise distinct is given by an infinite periodic word constructed on an unique period word FR_k (up to a permutation of letter and

⋆ This work was partly funded by the French Programme d'Investissements d'Avenir (LabEx Bézout, ANR-10-LABX-58) and ANR-15-CE40-0006.

circular permutation) by the recursive formula $FR_k = FR_{k-1}kFR_{k-1}$ with $FR_3 = 1213121$. Many researchers have worked on the general problem of infinite balanced words on an alphabet with one letter [16] or equivalently to cover integers by Beatty sequences [13, 27]. The conjecture is proved for $k = 3$ by Morikawa [20], for $k = 4$ by Altman, Gaujal and Hordijk [27], for $k = 5$ and $k = 6$ by Tijdeman [29] and for $k = 7$ by Barat et Varju [15] and the conjecture is still open for $k > 7$. Indeed, in order to investigate new discrete tools that allow us to deeply understand the structure of balanced words, we propose a second order balance property for Christoffel words that gives a refinement for the balance property. In fact, we define for Christoffel words a new matrix called *second order balance matrix* which gives information on the balance property of a word that codes the number of 1's of successive blocks of same length in the studied Christoffel word. Thus we investigate balance property on successive blocks instead of balance property on letters for second order balance property. The main idea, to go further in the resolution of the Fraenkel's conjecture, is to consider synchronization of the blocks instead of synchronization of letters and this is why we introduce the notion of second order of balance.

In Section 2, we recall some properties of the Christoffel words. In Section 3, we define the balance matrix which gives information on the number of occurrences of a given letter in all factors of a given binary circular word. This balance matrix gives us the order of balance, for binary words. Afterwards, in Section 4, we introduce the *second order balance matrix* of a given Christoffel word by computing the balance matrix for the rows of the associated balance matrix. We show in Section 5 that this matrix has many symmetries and is constructed by using properties of continued fractions and the Stern-Brocot tree. We present in Section 6 a recursive construction for the second order balance matrix by considering the properties of the continued fraction expansion for the slope of each Christoffel word. In Section 7, we show that an infinite path on the Stern-Brocot tree, which minimizes the second order of balance, is given by a path associated with the Fibonacci word. Section 8 is left for the perspectives of this work.

2 Notation and Christoffel words

Let A be an alphabet of cardinality m, the word w is the concatenation of letters of this alphabet and we write $w \in A^*$, where A^* represents the set of all the words formed by the alphabet A. We denote by $n = |w|$ the length of the word and by $|w|_a$ the number of occurrences of the letter a in the word w. The notation $w[i \dots j]$ refers to the factor of the word w from position i to position j. The notation w^ω represents: $w^\omega = ww \cdots w \cdots$ and named "circular word associated with w". By convention, $w^0 = \epsilon$ and a word w is said primitive if it is not the power of a nonempty word. Two words w and w' are *conjugate* of order k if and only if there exist u, v such that $|u| = k$ *with* $w = uv$ *and* $w' = vu$ and we denote: $w \equiv_k w'$. When the exact value of k is not relevant, we simply write $w \equiv w'$ and we say that the two words are conjugate. A positive integer p is a *period* of w if $w[i] = w[i + p]$; for all $1 \leq i \leq |w| - p$. Given a word $w = aw'$ where a is a letter, we note $a^{-1}w = w'$ that is the removal of the letter a at the beginning of w. If w ends with letter a, then the notation wa^{-1} is defined accordingly. Let $A = \{a_0, a_1, \dots, a_{m-1}\}$ be an alphabet, we let $\overline{\cdot}$ be the anti-morphism such that: $\overline{a_0} = a_{m-1}$, $\overline{a_1} = a_{m-2} \dots \overline{a_i} = a_{m-1-i}$. A word $w \in \{0, 1\}^*$ is $k-$balanced if and only if for all factors u, v of w, we get: $|u| = |v| \Longrightarrow ||u|_1 - |v|_1|| \leq k$. The word w is called balanced if $k = 1$.

Christoffel words [9] have many equivalent definitions and characterizations. The following geometrical definition is taken from [7] (see [5] for a self-contained survey). The lower Christoffel path of slope $\frac{a}{b}$, where a and b are relatively prime, is the path from $(0, 0)$ to (b, a) in the integer lattice $\mathbb{Z} \times \mathbb{Z}$ that satisfies the following conditions:

1. The path lies below the line segment that begins at the origin and ends at (b, a).
2. The region enclosed by the path and the line segment contains no other points of $\mathbb{Z} \times \mathbb{Z}$ besides those of the path.

We encode the lower Christoffel path (or simply Christoffel word) by means of a word in the alphabet A using 0 (resp. 1) for any unit horizontal (vertical) step. We get the *Christoffel word* of slope $\frac{a}{b}$ denoted: $C(\frac{a}{b})$, see Figure 1. Equivalently, the Christoffel word $w = C\left(\frac{a}{b}\right)$ is obtained by calculating the elements of the sequence $(r_i)_{0 \leq i \leq n}$, where $n = a + b$ as follows: $r_i = ia \mod n$. Each letter, $w[i]$, $\forall\ 1 \leq i \leq n$, of the word w and length n is obtained by computing: $w[i] = \begin{cases} 0 & \text{if } r_{i-1} < r_i, \\ 1 & \text{otherwise.} \end{cases}$

Example 1. Let $(a, b) = (3, 5)$, the sequence $(r_i)_{0 \leq i \leq 8} = (0, 3, 6, 1, 4, 7, 2, 5, 0)$ defines the Christoffel word $C\left(\frac{3}{5}\right) = 00100101$.

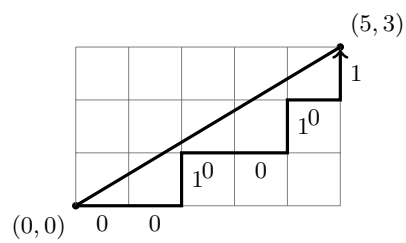

Fig. 1: Illustration of the geometrical definition of Christoffel words. The Christoffel path goes from (0,0) to (5,3) and $C\left(\frac{3}{5}\right) = 00100101$.

3 Balance matrix

In this section, we introduce a new matrix used to obtain the order of balance for any binary word in an explicit way. The ith row of the matrix M, $M[i]$, is seen as a word where each entry of the matrix is a letter. Given a word $w \in A^*$ of length n, we let S_w be the $n \times n$ matrix defined by $S_w[i, j] = w[j] + \ldots w[i + j]$ over the circular word w. By definition, we have that w is δ-balanced if

$$\delta = \max_i \left(\max(S_w[i]) - \min(S_w[i])\right).$$

The *balance matrix* B_w, is defined from S_w by subtracting the minimum value on each row,

$$B_w[i, j] = S_w[i, j] - \min(S_w[i]).$$

Obviously, we have that w is $(\max B_w)$-balanced.

It is clear that by construction, if w is not a sequence of $1's$, then the first row of B_w is equal to w. If $k < |w|$ is a period of w^ω, again by construction we have that the row $B_w[k]$ contains only zeros. The converse is also true, an integer k such that the row $B_w[k]$ contains only zeros is a period of w^ω. Consequently, a row of zeros is called a *period row*. Note that $n = |w|$ is always a period of

w^ω and therefore we define the matrix B_w as a $(n-1) \times n$ matrix since the n-th row would not add any information.

For the sake of simplicity, when working with the Christoffel word $C\left(\frac{a}{b}\right)$, the balance matrix $B_{C\left(\frac{a}{b}\right)}$ is simply denoted $B_{\frac{a}{b}}$.

Example 2. Let us consider the rational number $\frac{5}{2}$ and $n = 7$. By writing the Christoffel word $w = C(\frac{5}{2}) = 0110111$, the elements of the balance matrix B_w and S_w are the following:

$$S_w = \begin{pmatrix} 0&1&1&0&1&1&1 \\ 1&2&1&1&2&2&1 \\ 2&2&2&2&3&2&2 \\ 2&3&3&3&3&3&3 \\ 3&4&4&3&4&4&3 \\ 4&5&4&4&5&4&4 \\ 5&5&5&5&5&5&5 \end{pmatrix}, \quad B_{\frac{5}{2}} = \begin{pmatrix} 0&1&1&0&1&1&1 \\ 0&1&0&0&1&1&0 \\ 0&0&0&0&1&0&0 \\ 0&1&1&1&1&1&1 \\ 0&1&1&0&1&1&0 \\ 0&1&0&0&1&0&0 \end{pmatrix}.$$

3.1 Properties of the balance matrix

In this section, we present some properties of the matrix B_w, where w is a binary word of length $|w| = n$, allowing us to restrict the work to the upper half of the balance matrix since the lower part will be obtained by symmetry.

Definition 1. *Two words w and w' are* complementary *if $w \equiv_k \overline{w'}$ for some k. Moreover, a word is said to be* autocomplementary *if it is complementary to itself.*

Example 3. The word $w = 0110110010$ is an autocomplementary word since we have: $w \equiv_5 \overline{w}$. While $w = 0000100$ and $w' = 0111111$ are such that $w \equiv_4 \overline{w'}$.

We show some basic combinatorial properties of the matrix S_w that will be used in a further proof. All the proofs of this section can be found in the Appendix.

Property 1. For any binary word w of length n, and $1 \leq i \leq n$, the matrix S_w satisfies:

1. $\max S_w[i] + \min S_w[n-i] = |w|_1$,
2. $\min S_w[i] + \max S_w[n-i] = |w|_1$.

In particular, if w is a $1-$balanced Christoffel word, we have:

1. $\min S_w[i] + \min S_w[n-i] = |w|_1 - 1$,
2. $\max S_w[i] + \max S_w[n-i] = |w|_1 + 1$,

By applying the results of Property 1 and the definition of S_w, we can notice that the lower half of B_w is deduced from its upper half part, as we can see in Property 2.

Property 2. The balance matrix is such that for all $1 \leq i < n$, $B_w[i] \equiv_{n-i} \overline{B_w[n-i]}$.

If $B_{\frac{a}{b}}$ has a middle row, then this row is autocomplementary as we can seein Example 4.

Corollary 1. *If n is an even number, $B_{\frac{a}{b}}[\frac{n}{2}]$ is an autocomplementary row.*

Example 4. Let us consider the rational number 3/7 with $n = 10$. The balance matrix $B_{\frac{3}{7}}$ shows that $B_{\frac{3}{7}}[5]$ is an autocomplementary row.

$$B_{\frac{3}{7}} = \begin{pmatrix} 0\,0\,0\,1\,0\,0\,1\,0\,0\,1 \\ 0\,0\,1\,1\,0\,1\,1\,0\,1\,1 \\ 0\,1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,0\,1\,0\,0\,1\,0\,0\,0 \\ 0\,0\,1\,1\,0\,1\,1\,0\,0\,1 \\ 0\,1\,1\,1\,1\,1\,1\,0\,1\,1 \\ 0\,0\,0\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,1\,0\,0\,1 \\ 0\,1\,1\,1\,0\,1\,1\,0\,1\,1 \end{pmatrix}.$$

Remark 1. In the case where $a = 1$, the balance matrix is such that: $B_{\frac{1}{b}}[i] = 0^{(n-i)}1^i$.

3.2 Construction of the balance matrix for Christoffel words

In this section, we are interested in giving a direct construction for the balance matrix of a Christoffel word, by determining for each row of B_w, the positions of the letter 1. For that, we start by defining the set of positions for the letter 1 in $C(\frac{a}{b})$.

Definition 2. *Let w be a Christoffel word of slope a/b. The set of* decreasing positions *of w, denoted $D(a, b)$, is the set of the positions of the occurrences of the letter 1 in w. More formally, $D(a, b) = \{1 \leq i \leq n \mid w[i] = 1\}$.*

The following theorem is a reformulation of Paquin and Reutenauer's result([22], Corollary 3.2), that describes the set of decreasing positions of a Christoffel word.

Theorem 1. *Let α be such that $\alpha a \equiv -1 \mod n$, then the set $D(a, b)$ is :*

$$D(a, b) = \{(i\alpha \bmod n) + 1 \mid i = 1 \ldots a\}.$$

Example 5. Let us consider the rational number $a/b = 3/5$ with $n = 8$, $\alpha = 5$ and $w = 00100101$. The set of occurrences of the letter 1 is $D(3, 5) = \{(5i \bmod 8) + 1 \mid i = 1 \ldots 3\} = \{3, 6, 8\}$.

Using Theorem 1, we can conclude that $B_{\frac{a}{b}}$ and $B_{\frac{b}{a}}$ are not equal but complementary as Lemma 1 shows. The proof can be found in the Appendix.

Lemma 1. *The balance matrices $B_{\frac{a}{b}}$ and $B_{\frac{b}{a}}$ are conjugate in the sense that:*

$$B_{\frac{a}{b}}[i] \equiv_\alpha \overline{B_{\frac{b}{a}}[i]} \ \forall\ 1 \leq i < n; \ \textit{where } \alpha a \equiv -1 \bmod n.$$

For this part, w is a Christoffel word and we give a recursive construction of B_w by identifying for each row i, the set of positions of the $1'$s. This set is denoted by D_i, where for each row i in B_w, $D_i = \{j \mid B_w[i, j] = 1\}$, with $n = |w|$ and $1 \leq j \leq n$. For any set of integers S and any integer k, we denote $S + k = \{a + k \mid a \in S\}$.

Theorem 2. *If $a < b$ then the sets $(D_i)_{1 \leq i \leq n-1}$ are recursively obtained as follows: $D_1 = D(a, b)$ and for each i from 2 to $n - 1$:*

$$D_i = \begin{cases} D_{i-1} \cup (D_1 - (i-1) \bmod n) & \textit{if } i \notin D_1 \\ D_{i-1} \cap (D_1 - (i-1) \bmod n) & \textit{if } i \in D_1 \end{cases} \tag{1}$$

The proof can be found in the Appendix.

Example 6. Let us consider the rational number 2/7 *mod* 9. By calculating the set of decreasing values for each row of $B_{\frac{2}{7}}$, we get the following matrix:

$$
\begin{aligned}
D_1 &= \{5,9\}\\
D_2 &= D_1 \cup \{4,8\} = \{4,5,8,9\}\\
D_3 &= D_2 \cup \{3,7\} = \{3,4,5,7,8,9\}\\
D_4 &= D_3 \cup \{2,6\} = \{2,3,4,5,6,7,8,9\}\\
D_5 &= D_4 \cap \{1,5\} = \{5\}\\
D_6 &= D_5 \cup \{4,9\} = \{4,5,9\}\\
D_7 &= D_6 \cup \{3,8\} = \{3,4,5,8,9\}\\
D_8 &= D_7 \cup \{2,7\} = \{2,3,4,5,7,8,9\}\\
D_9 &= D_8 \cap \{1,6\} = \{\}.
\end{aligned}
\quad \Rightarrow \quad
B_{\frac{2}{7}} = \begin{pmatrix}
0&0&0&0&1&0&0&0&1\\
0&0&0&1&1&0&0&1&1\\
0&0&1&1&1&0&1&1&1\\
0&1&1&1&1&1&1&1&1\\
0&0&0&0&1&0&0&0&0\\
0&0&0&1&1&0&0&0&1\\
0&0&1&1&1&0&0&1&1\\
0&1&1&1&1&0&1&1&1
\end{pmatrix}.
$$

From this construction, we can get a relation between the number of occurrences of 1 in each row of the balance matrix of Christoffel words and the numerator of the slope related to this word.

Lemma 2. *For the balance matrix $B_{\frac{a}{b}}$, we have: $|D_i| = i.a \bmod n$, where $|D_i|$ is the cardinal of the set D_i.*

Note: The proof can be found in the Appendix. This Lemma confirms that the period row is made only of zeros since $|D_n| = n.a \bmod n = 0$, hence we have no occurrences for the letter 1 in this row.

4 Second order balance matrix

Let w^ω be a 1-balanced circular word associated with w; by computing the balanced property on each row of B_w, we get a refinement of the balanced property for w. For any factor v of length j, we can find k_j or $k_j + 1$ occurrences of the letter 1. The second order balance is the repartition of these blocks in a binary balanced word. This second order balance is computed via a matrix called the *second order balance matrix*. In other words, we are studying the balance of each row of B_w.
For a pair of integers i, j, where $1 \leq i, j \leq |w| - 1$, we consider the word $B_w[i]$ and we list all its factors of length j. Among these factors, we choose p, a factor that maximizes the number of occurrences of the letter 1 and q, a factor that minimizes it. The entry $U_w[i,j]$ is given by $|p|_1 - |q|_1$. Equivalently, if $L_\ell(w)$ is the restriction of the language of w^ω to words of length ℓ, then:

$$U_w[i,j] = \max_{v \in L_j(B_w[i])} |v|_1 - \min_{v \in L_j(B_w[i])} |v|_1.$$

In other words, we can define the second order balance matrix U_w by:

Definition 3. *Let w be a word such that w^ω is 1-balanced or, equivalently, that B_w is a binary matrix. The* second order balance matrix *$U_w = (u_{ij})_{1 \leq i,j \leq n-1}$ where $U_w[i,j] = \max(B_{(B_w[i])}[j])$.*

Definition 4. *The* second order of balance *of a circular 1-balanced word w is $\delta^2(w) = \max(U_w)$.*

Once again, in order to lighten the notation, when working with the Christoffel word $C\left(\frac{a}{b}\right)$, the second order balance matrix $U_{C\left(\frac{a}{b}\right)}$ is simply denoted $U_{\frac{a}{b}}$.
For the rest of the paper, we let w be a Christoffel word of slope a/b and length $n = a+b$, such that: $w = C(\frac{a}{b})$. The second order balance matrix of a Christoffel word of slope $\frac{a}{b}$, $U_{\frac{a}{b}}$, is of dimension $(n-1) \times (n-1)$.

Example 7. Let us consider the rational number $\frac{a}{b} = \frac{3}{7}$ with $n = 10$. The balance matrix $B_{\frac{3}{7}}$ was calculated previously and $B_{\frac{3}{7}}[5] = [0, 0, 1, 1, 0, 1, 1, 0, 0, 1]$. By computing the balance matrix for this word and taking the blocks of length 5 we get the 5th row of $S_{B_{\frac{3}{7}}}[5]$ where the difference between the maximum and the minimum values of each row of $S_{B_{\frac{3}{7}}}[5]$ determines the entries of $U_{\frac{3}{7}}[5]$. Hence, with these two blocks $\overline{00|11011|001}$, we obtain the element $U_{\frac{3}{7}}[5, 5] = 4 - 1 = 3$ as we can see in the following second order balance matrix of $\frac{3}{7}$, where we also get $\delta^2(C(\frac{3}{7})) = 3$.

$$U_{\frac{3}{7}} = \begin{pmatrix} 1&1&1&1&1&1&1&1&1 \\ 1&2&1&1&2&1&1&2&1 \\ 1&1&1&1&1&1&1&1&1 \\ 1&1&1&2&2&2&1&1&1 \\ 1&2&1&2&\color{red}{3}&2&1&2&1 \\ 1&1&1&2&2&2&1&1&1 \\ 1&1&1&1&1&1&1&1&1 \\ 1&2&1&1&2&1&1&2&1 \\ 1&1&1&1&1&1&1&1&1 \end{pmatrix}.$$

Properties of the matrix $U_{\frac{a}{b}}$

Now we give some properties of the second order balance matrix in order to show that $U_{\frac{a}{b}} = U_{\frac{b}{a}}$. Hence, we can restrict our study to the irreducible fractions a/b with $a < b$. But before that, we prove the three symmetries that appear in this matrix. From Section 3.1, we have that the rows of the upper half of $B_{\frac{a}{b}}$ are complementary to the rows of its lower half, which induces the symmetries in the matrix $U_{\frac{a}{b}}$. More precisely, the second order balance matrix $U_{\frac{a}{b}}$ of dimension $(n-1)\times(n-1)$, has horizontal, vertical and diagonal symmetries. The axis of symmetry are at position $\frac{n}{2}$ or between $\frac{n-1}{2}$ and $\frac{n+1}{2}$ depending on the parity of n.

Property 3. For any position (i, j), $U_{\frac{a}{b}}[i, j] = U_{\frac{a}{b}}[n - i, j] = U_{\frac{a}{b}}[i, n - j] = U_{\frac{a}{b}}[n - i, n - j]$.

Proof. By Lemma 2, we have that for any $1 \leq i \leq n - 1$, $B_w[i] \equiv_{n-i} \overline{B_w[n - i]}$ where $n = a + b$, Consequently :

$$\begin{aligned} U_w[i, j] &= \max(B_{\left(B_{\frac{a}{b}}[i]\right)}[j]) && \text{(def.)} \\ &= \max(B_{\left(B_{\frac{a}{b}}[n-i]\right)}[j]) && \text{(autocomplementarity of } B_{\frac{a}{b}}) \\ &= \max(B_{\left(B_{\frac{a}{b}}[n-i]\right)}[n - j]) && \text{(autocomplementarity of } B_{B_{\frac{a}{b}}[n-i]}) \\ &= \max(B_{\left(B_{\frac{a}{b}}[i]\right)}[n - j]). \end{aligned}$$

Moreover, $U_{\frac{a}{b}}$ has an extra diagonal symmetry;

Property 4. For the Christoffel word of slope a/b, we have $(U_{\frac{a}{b}})^T = U_{\frac{a}{b}}$.

Proof. Let $w = C(\frac{a}{b})$ and $i < j$, by definition: $U_w[i, j] = \max(B_{(B_w[i])}[j])$; we recall that: $B_w[i] = S_w[i] - \min(S_w[i])$ where $S_w[i, j] = w[i] + \ldots w[i + j] = |w[i, j]|_1$. Therefore $B_{B_w[i]}[j] = S_{B_w[i]}[j] - \min(S_{B_w[i]}[j])$ with $S_{B_w[i]}[j] = |w[j, \ldots, j + i - 1](j, i)|_1 = |w[j, \ldots, j + i - 1]|_1$. While for $U[j, i] = \max(B_{(B_w[j])}[i])$ we have: $B_w[j] = S_w[j] - \min(S_w[j])$. Therefore $\forall\, i, j \in \{1, \ldots, n-1\}$, $B_{B_w[j]}[i] = S_{B_w[j]}[i] - \min(S_{B_w[j]}[i])$ with $S_{S_w[j]}[i] = |w[i, \ldots, j + i - 1][i, j]|_1 = |w[j, \ldots, j + i - 1]|_1$ and $U_w[i, j] = \max(B_{(B_w[i])}[j]) = \max(B_{(B_w[j])}[i]) = U[j, i]$.

After those two properties, we are able to prove that $U_{\frac{a}{b}} = U_{\frac{b}{a}}$.

5 More about Christoffel words

Let w be a Christoffel word of length at least 2, the standard factorization is obtained by writting $w = (w_1, w_2)$ in a unique way, where w_1, w_2 are two Christoffel words by [7]. The *Christoffel tree* is an infinite tree whose vertices are all the standard factorizations of Christoffel words (see [5], Section 3.2). It uses the fact that given a standard factorization (w_1, w_2), the pairs (w_1, w_1w_2) and (w_1w_2, w_2) are also standard factorizations. Let ϕ_0, ϕ_1 be the two functions from $A^* \times A^*$ into itself defined by: $\phi_0(w_1, w_2) = (w_1, w_1w_2); \phi_1(w_1, w_2) = (w_1w_2, w_2)$. We have that any Christoffel word can be obtained in a unique way by iteration of these two functions on $(0, 1)$. Consequently, the Christoffel tree is defined as follows: the root is $(0, 1)$, the Christoffel word of slope 1. Then each node (w_1, w_2) has two sons: $\phi_0(w_1, w_2)$ on the left and $\phi_1(w_1, w_2)$ on the right. See Figure 2 for an illustration.

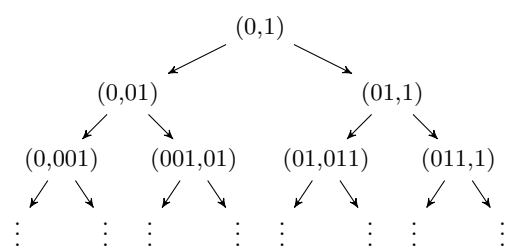

Fig. 2: The first levels of the Christoffel tree.

We introduce and study the directive sequence of Christoffel words.

Definition 5. *Let $w = C(\frac{a}{b})$ be a non-trivial Christoffel word, the* directive sequence *of w, denoted $\Delta(\frac{a}{b})$, is the word $\Delta(\frac{a}{b}) = i_1 i_2 \cdots i_n \in A^*$ such that $w = (\phi_{i_n} \circ \cdots \circ \phi_{i_2} \circ \phi_{i_1})(0, 1)$.*

Note that a directive sequence $\Delta(\frac{a}{b}) = i_1 \cdots i_n$ describes the path from the root of the Christoffel tree to the Christoffel word $C(\frac{a}{b})$ as follows : at step k, if $i_k = 0$ then go left, otherwise, if $i_k = 1$ then go right.

5.1 Stern-Brocot tree and continued fractions

In this section, we introduce the *Stern-Brocot tree* that contains all the reduced fractions $\frac{a}{b}$. It was first introduced by a German mathematician Moritz Abraham Stern and a French clockmaker Achille Brocot in the 19th century [8]. In order to construct recursively the Stern-Brocot tree, we need to introduce the *mediant* of two fractions $\frac{a}{b}$ and $\frac{c}{d}$, that is $\frac{a}{b} \oplus \frac{c}{d} = \frac{a+c}{b+d}$. In addition to that, we have to define the recursive sequence s_i, that is obtained from s_{i-1}, by completing with the mediant of each two consecutive fractions in s_{i-1}, where s_0 is given by: $s_0 = \left(\frac{0}{1}, \frac{1}{0}\right)$.

Note that, $\frac{1}{0}$, is considered as a normal fraction: $s_0 = \left(\frac{0}{1}, \frac{1}{0}\right)$, $s_1 = \left(\frac{0}{1}, \frac{1}{1}, \frac{1}{0}\right)$, $s_2 = \left(\frac{0}{1}, \frac{1}{2}, \frac{1}{1}, \frac{2}{1}, \frac{1}{0}\right)$, $s_3 = \left(\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{3}{1}, \frac{1}{0}\right)$.

The mediants added in each new step to the sequence s_i are the fractions that appear on the ith level of the Stern-Brocot tree. For example, on the third level, we have the fractions: $\frac{1}{3}, \frac{2}{3}, \frac{3}{2}, \frac{3}{1}$ that are extracted from the sequence s_3. We call consecutive fractions, two fractions that belong to the same set s_i and are next to each other, like $\frac{1}{1}$, and $\frac{2}{1}$ in s_2. Some properties about these consecutive fractions will be given in the next section.
In order to simplify the notation, we denote $\frac{a}{b} \oplus \frac{c}{d}$ by $\frac{a}{b}\frac{c}{d}$ and $\frac{a}{b} \oplus \ldots \oplus \frac{a}{b}$ repeated p times by $(\frac{a}{b})^p$. The Christoffel tree is isomorphic to the Stern-Brocot tree where each vertex of the Christoffel tree of the form (u,v) is associated to the fraction $\frac{|uv|_1}{|uv|_0}$, see Figure 3.

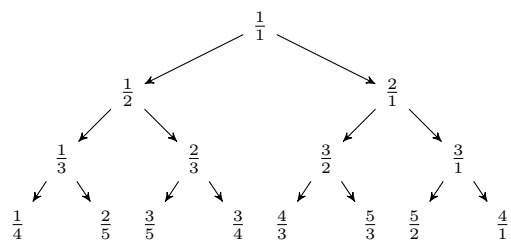

Fig. 3: The first levels of the Stern-Brocot tree.

The continued fraction of a rational number $\frac{a}{b} \geq 0$ is the sequence of integers $\frac{a}{b} = [a_0, \ldots, a_z]$, with $a_0 \geq 0; a_i \geq 1$ for $1 \leq i \leq z$ and if $z \geq 2$ then $a_z \geq 2$.

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + a_z}}.$$

In the following part, we will explain how to pass from the continued fraction of a rational number a/b to the Christoffel word of slope a/b. For that, we introduce the following theorem by Henry J.S in 1876. This theorem gives an additional characterization for the Christoffel words. In fact, he showed that the Christoffel word can be obtained by a recursive expression using the partial denominators of the rational number.

Theorem 3. *[28] Let $\frac{a}{b} = [a_0, a_1 \ldots, a_z]$, a word $w = 0u1$ is a Christoffel word of slope $\frac{a}{b}$ if and only if $u01$ or $u10$ is equal to s_{n+1}, where s_{n+1} is defined recursively by: $s_{-1} = 0, s_0 = 1$ and $s_{n+1} = s_n^{a_n} s_{n-1}$ for all $n \geq 0$.*

Moreover, in 1987, Berstel introduced, in a report for LITP, the following relation between the partial denominators of a rational number and two particular matrices based on the Theorem 2.1 in George Raney's paper [25] in 1973, that was publuished in 1990 [3]. Using the two matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and for the particular case of the rational number a/b where $a < b$, we have:

Proposition 1. *[3] For $a < b$, we associate for the rational number $\frac{a}{b} = [0, a_1, a_2, \ldots, a_n]$, the following matrix: $M(\frac{a}{b}) = A^{a_1-1}B^{a_2}A^{a_3}\ldots C^{a_{n-1}}D^{a_n-1}$, where C and D are respectively A, B or B, A depending on the parity of n. We have: $M(\frac{a}{b}).\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}b-a\\a\end{pmatrix}$.*

Furthermore, in 1993, Borel and Laubie, followed by de Luca in 1997 then Berthé, de Luca and Reutenauer in 2008, gave the following theorem, where they linked these two particular matrices to the Christoffel word of slope a/b. Let the function Pal be the iterative palindromic closure function defined recursively in [4] by $Pal(ua) = (Pal(u)a)^{(+)}$ and $Pal(\epsilon) = \epsilon$, where a is a letter and u a word and $(w)^+$ is the palindromic closure of w i.e. $(w)^+ = ww'$ with ww' is the shortest palindrome having w as a prefix.

Theorem 4. *[4, 6, 7] Let w be a Christoffel word of slope a/b, there exists a unique word v such that $w = 0Pal(v)1$, if (w_1, w_2) is the standard factorization of w then we define the multiplicative monoid morphism $\mu : \{0,1\}^* \longrightarrow SL_2(\mathbb{Z})$ such that: $\mu(0) = \begin{pmatrix}1&1\\0&1\end{pmatrix} = A$ and $\mu(1) = \begin{pmatrix}1&0\\1&1\end{pmatrix} = B$, where: $\mu(v) = \begin{pmatrix}|w_1|_0 & |w_2|_0\\|w_1|_1 & |w_2|_1\end{pmatrix}$.*

In [3, 5, 18, 25] we can find results allowing us to write, in an explicit way, the relation between the continued fraction of a rational number and its directive sequence as we can see in the following theorem (see[4]), where its proof can be found in the Appendix.

Theorem 5. *Let $\frac{a}{b} = [a_0, \ldots, a_z]$, we have: $\Delta(\frac{a}{b}) = 1^{a_0}0^{a_1}1^{a_2}...p^{a_z-1}$ where $p \in \{0,1\}$. The Christoffel word of slope a/b is written: $C\left(\frac{a}{b}\right) = 0w'1$, where w' is a palindrome and $w' = Pal(\Delta(\frac{a}{b}))$.*

6 Recursive construction of the second order balance matrix

Due to the isomorphism and the recursive construction of the Stern-Brocot tree and the Christoffel tree, we can conclude that there must exist a recursive construction for the second order balance matrix. In order to show and to prove this recursivity, we let $U_{\frac{a}{b}}$ where $a/b = [a_0, \ldots, a_z] = [a_0, a_1, \ldots, a_z - 1, 1]$ be the second order balance matrix of the Christoffel word $C\left(\frac{a}{b}\right)$. We introduce the following terminology for some specific rational numbers on the Stern-Brocot tree and that will be used for the rest of the paper. See Figure 4 for an illustration.

Definition 6. *Given $\frac{a}{b} = [a_0, \ldots, a_z]$,*

1. *the* top branch fraction *of $\frac{a}{b}$, denoted* $\mathrm{TBF}(\frac{a}{b})$, *is the fraction* $[a_0, \ldots, a_{z-1} + 1]$,
2. *the* first reduced fraction *of $\frac{a}{b}$, denoted* $\mathrm{FRF}(\frac{a}{b})$, *is the fraction* $[a_0, \ldots, a_{z-1}]$,
3. *the* first extended fraction *of $\frac{a}{b}$, denoted* $\mathrm{FEF}(\frac{a}{b})$, *is the fraction* $[a_0, a_1, \ldots, a_z + 1]$,
4. *the* first deviation fraction *of $\frac{a}{b}$, denoted* $\mathrm{FDF}(\frac{a}{b})$, *is the fraction* $[a_0, a_1, \ldots, a_z - 1, 2]$,
5. *the* first parallel fraction *of $\frac{a}{b}$, denoted* $\mathrm{FPF}(\frac{a}{b})$, *is either* $[a_0, a_1, \ldots, a_{z-1} - 1, 2]$ *if* $a_{z-1} \neq 1$ *or* $[a_0, a_1, \ldots, a_{z-2} + 2]$ *if* $a_{z-1} = 1$.
6. *the* second unidirectional father *of $\frac{a}{b}$, denoted* $\mathrm{SUF}(\frac{a}{b})$, *is either* $[a_0, a_1, \ldots, a_z - 2]$ *if* $a_z > 2$ *or* $[a_0, a_1, \ldots, a_{z-2}]$ *if* $a_z = 2$.

Note that $\mathrm{SUF}(\frac{a}{b})$ is not defined for fractions $\frac{1}{1}$, $\frac{1}{2}$ and $\frac{2}{1}$. Using Theorem 5 we can get the directive sequence of each of these fractions.

Example 8. Let $\frac{a}{b} = \frac{3}{5} = [0, 1, 1, 2]$, from Definition 6, we get: $\mathrm{TBF}(3/5) = 2/3$, $\mathrm{FEF}(3/5) = 4/7$, $\mathrm{FDF}(3/5) = 5/8$, $\mathrm{FPF}(3/5) = 1/3$ *and* $\mathrm{SUF}(3/5) = 1/1$. See Figure 4 for the positions of theses fractions in the Stern-Brocot tree.

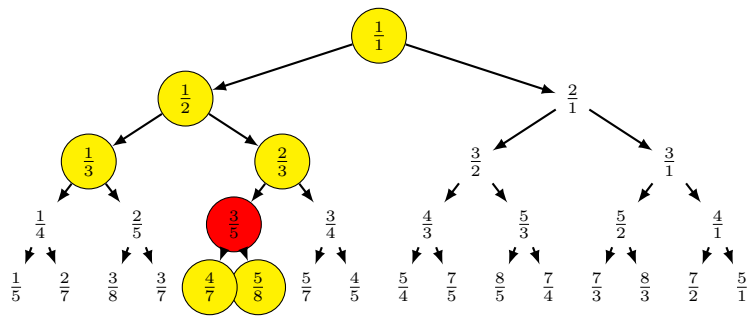

Fig. 4: Illustration of the specific rational numbers related to $3/5 = [0, 1, 1, 2]$ on the Stern-Brocot tree. We have: $\mathrm{TBF}(3/5) = 2/3$, $\mathrm{FRF}(3/5) = 1/2$, $\mathrm{FEF}(3/5) = 4/7$, $\mathrm{FDF}(3/5) = 5/8$, $\mathrm{FPF}(3/5) = 1/3$ and $\mathrm{SUF}(3/5) = 1/1$.

General form of the second order balance matrix

To construct $U_{\frac{a}{b}}$, we start by placing 4 rows of separation that divide the matrix into 9 blocks. Due to the symmetries proved in Properties 3 and 4, it is sufficient to know three of these blocks to deduce the others. These blocks are denoted α, β *and* γ and are represented in the matrix as follows :

$$U_{\frac{a}{b}} = \left(\begin{array}{c|c|c} \alpha & \cdot & \cdot \\ \hline \gamma & \beta & \cdot \\ \hline \cdot & \cdot & \cdot \end{array}\right),$$

In the following section, we show that the blocks α, β, γ are described by the second order balance matrices of simpler fractions. More precisely, α is deduced from $U_{\mathrm{TBF}(\frac{a}{b})}$ while β is given by adding one to each entry of $U_{\mathrm{SUF}(\frac{a}{b})}$. Finally, the construction of γ depends on the position of $\frac{a}{b}$ in the Stern-Brocot tree, where the fraction is, relatively to its father, either a deviation (first deviation fraction) or an extension (first extended fraction).

6.1 The construction of $U_{\frac{a}{b}}$

In this section, we start to explain how we can obtain the recursive construction of the second order balance matrix of the Christoffel word of slope $a/b = [a_0, \ldots, a_z]$. Figure 6 displays the rational numbers needed for the construction. Since $U_{\frac{a}{b}} = U_{\frac{b}{a}}$, we reduce the work to the first half of the Stern-Brocot tree, that contains all the irreducible fractions with $a \leq b$. Given $\frac{a}{b} = [a_0, \ldots, a_z]$, we consider separately the cases $z < 2$ and $z \geq 2$.

The trivial cases: $z \in \{0, 1\}$

For $z = 0$, we have $U_{\frac{0}{1}} = U_{\frac{1}{0}} = [\ ]$ and $U_{\frac{1}{1}} = [1]$. For $z = 1$, we have $\frac{a}{b} = [0, a_1]$ which implies that $a = 1$ and $b = a_1$. In this case, Remark 1, states that $B_w[i] = 0^{b+1-i}1^i$ and from Definition 3, we have: $U_w[i, j] = \max(B_{B_w[i]}[j])$. Hence, we consider the first quarter of the matrix U_w which is sufficient from properties 3 and 4. Due to the diagonal symmetry, we will consider only $j \leq i$ where $i, j \leq \frac{n}{2}$ if n is even or $i, j \leq \frac{n-1}{2}$ if n is odd.

Proposition 2. *If $a = 1$, then for all $j \leq i$ and $i, j \leq \frac{n}{2}$ we have: $U_{\frac{a}{b}}[i, j] = j$.*

Proof. Since $B_w[i] = 0^{b+1-i}1^i$ and $U_w[i, j] = \max(B_{B_w[i]}[j])$, from Property 6 and with $j < i$, we have: $B_{B_w[i]}[j] = 0^{b-i-j+2}1 \cdots j^{i-j+2} \cdots 1$. Therefore, $U_w[i, j] = \max(B_{B_w[i]}[j]) = j$.

We can observe that this matrix can also be constructed using a recursive form where the first (resp. last) row and the first (resp. last) column are all 1's and in the middle we have the matrix of $\mathrm{SUF}(\frac{1}{b})$ where its elements are increased by 1 (see Figure 5).

$$U_{\frac{1}{b}} = \left(\begin{array}{c|ccc|c} 1 & 1 & & .. 1 & 1 \\ \hline 1 & & & & 1 \\ \vdots & & U_{[0,a_1-2]} + 1 & & \vdots \\ 1 & & & & 1 \\ \hline 1 & 1 & & .. 1 & 1 \end{array}\right).$$

Fig. 5: The general form of the matrix $U_{\frac{1}{b}}$.

The general case : $z \geq 2$

Now we assume that $z \geq 2$, in order to lighten the presentation we define the following fractions, let: $\frac{u}{v} = \mathrm{TBF}(\frac{a}{b})$, $\frac{x}{y} = \mathrm{FRF}(\frac{a}{b})$, $\frac{c}{d} = \mathrm{FDF}(\frac{u}{v})$, $\frac{e}{f} = \mathrm{FPF}(\frac{c}{d})$, $\frac{g}{h} = \mathrm{SUF}(\frac{u}{v})$, $\frac{p}{q} = \mathrm{TBF}(\frac{x}{y})$, $\frac{s}{t} = \mathrm{FRF}(\frac{u}{v})$. See Figure 6 for an illustration of their relative positions on the Stern-Brocot tree.

Lemma 3. *The Christoffel words of slope $\frac{a}{b}, \frac{g}{h}$ and $\frac{e}{f}$ can be written as follows:*
$C(\frac{a}{b}) = \left(C(\frac{x}{y})\right)^{a_z - 1} C(\frac{u}{v}) = \left(C(\frac{x}{y})\right)^{a_z} C(\frac{s}{t})$, $C(\frac{x}{y}) = C(\frac{g}{h})C(\frac{s}{t})$ *and* $C(\frac{e}{f}) = C(\frac{g}{h})C(\frac{x}{y})$.

Proof. This is a consequence of Theorem 5. The iteration of the application of functions ϕ_0 and ϕ_1 provides the equalities.

Separation rows For the rest of this section, the fractions mentioned in Figure 6 are used to prove the construction of $U_{\frac{a}{b}}$. In the following part we prove that separation rows allow the decomposition of $U_{\frac{a}{b}}$ into 9 blocks. In Definition 7, we give a characteristic for the consecutive fractions of each sequence s_i, used to construct the Stern-Brocot tree.

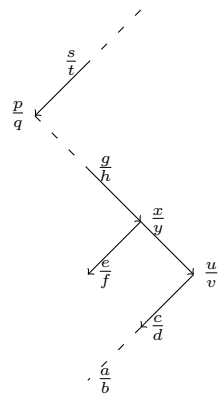

Fig. 6: Position of the fractions $\frac{u}{v}$, $\frac{x}{y}$, $\frac{c}{d}$, $\frac{e}{f}$, $\frac{g}{h}$, $\frac{p}{q}$ and $\frac{s}{t}$ relatively to $\frac{a}{b}$, for the case where $a_{z-1} \geq 2$ and z odd. By definition, we have : $\frac{a}{b} = [a_0 \ldots, a_z]$, $\frac{u}{v} = [a_0 \ldots, a_{z-1}+1]$, $\frac{x}{y} = [a_0, \ldots, a_{z-1}]$, $\frac{p}{q} = [a_0, \ldots, a_{z-2}+1]$, $\frac{s}{t} = [a_0, \ldots, a_{z-2}]$, $\frac{e}{f} = [a_0, \ldots, a_{z-1} - 1, 2]$, $\frac{c}{d} = [a_0, \ldots, a_{z-1}, 2]$, $\frac{g}{h} = [a_0, \ldots, a_{z-1} - 1]$. With respect to the Farey addition we have: $\frac{u}{v} = \frac{x}{y}\frac{s}{t}$ and $\frac{c}{d} = \frac{x}{y}\frac{u}{v}$.

Definition 7. *Let $\frac{a}{b}$ and $\frac{a'}{b'}$ be two consecutive fractions of a certain sequence s_i, $\forall i > 0$. They respect the following relation: $a'b - ab' = 1$.*

Property 5. Let $\frac{a}{b}$ and $\frac{a'}{b'}$ be two consecutive fractions in the Stern-Brocot tree, we have:

$$(a' + b').a = a'.(a + b) - 1.$$

The proof of this property is obtained by an arithmetic calculation based on Definition 7.
In the following lemma, we prove that in each $U_{\frac{a}{b}}$, we have at least 2 rows and columns full of $1'$s.

Lemma 4. *Let $k = x + y$, for the rational number $\frac{a}{b}$ with $n = a + b$, we have:*

$$U[i, k] = U[i, n - k] = U[k, j] = U[n - k, j] = 1, \ \forall \ i, j \in \{1, \ldots, n - 1\}.$$

Proof. From Property 5, we have $(x + y)a = x(a + b) - 1$, using the consecutive fractions a/b and x/y from the Stern-Brocot tree, we get: $ka = xn - 1$ hence $ka \equiv -1 \bmod n$. From Lemma 2, we have $|D_i| = ia \bmod n$, hence $|D_k| = n - 1$ and $|D_{n-k}| = 1$ since these two rows are complementary. Therefore, the words $B_{\frac{a}{b}}[k]$ and $B_{\frac{a}{b}}[n - k]$ are 1-balanced; hence $U[k, j] = U[n - k, j] = 1, \ \forall \ j \in \{1, \ldots, n - 1\}$. Using the diagonal symmetry of $U_{\frac{a}{b}}$, we get: $U[i, k] = U[i, n - k] = 1, \ \forall \ i \in \{1, \ldots, n - 1\}$.

Remark 2. From Lemma 4, we get that the row (respectively column) k and $n - k$ contain only values of $1'$s. Therefore, we define the separation rows to be between the rows (resp. columns) $(k; k + 1)$ and between the rows (resp. columns) $(n - k - 1; n - k)$ which divide $U_{\frac{a}{b}}$ into 9 blocks. See Figure 7.

Let $\frac{a}{b} = \frac{5}{8}$, the 9 blocks are represented as follows:

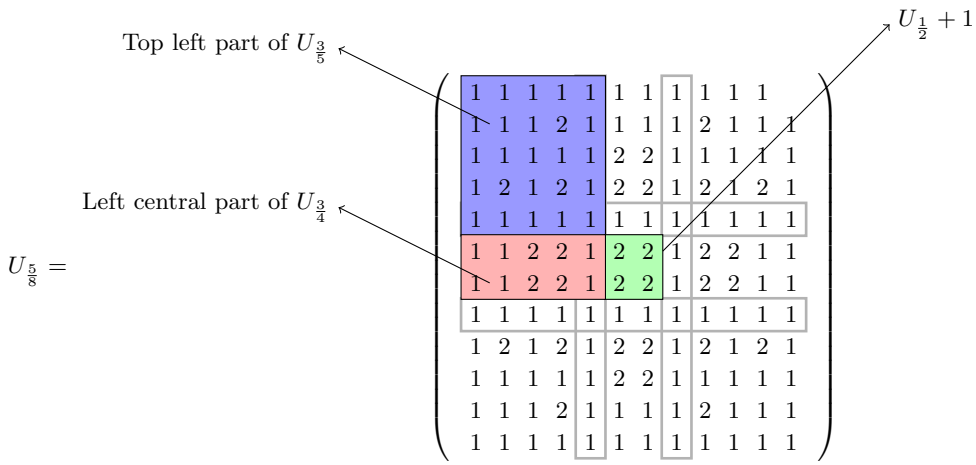

Fig. 7: The decomposition of the matrix $U_{\frac{5}{8}}$ into 9 blocks. We have $\mathrm{TBF}(\frac{5}{8}) = \frac{3}{5}$, $\mathrm{SUF}(\frac{5}{8}) = \frac{1}{2}$, $\mathrm{FPF}(\frac{5}{8}) = \frac{3}{4}$, $\frac{x}{y} = \frac{2}{3}$ and $k = 5$.

Now we prove that each block α, β and γ of $U_{\frac{a}{b}}$ can be decomposed in some smaller second order balanced matrices. In fact, to construct the matrix $U_{\frac{a}{b}}$ we first place the rows of separation. This shows that $U_{\frac{a}{b}}$ is composed of 9 blocks where the three blocks $\alpha_{k\times k}$, $\beta_{(n-2k-1)\times(n-2k-1)}$ and $\gamma_{(n-2k-1)\times k}$ are constructed while the others are obtained by symmetry.
In the following part, we construct respectively each of the blocks α, γ and β to get $U_{\frac{a}{b}}$.

α-block: Recall that $z \geq 2$ and we let for the rest of this section z to be an odd number where the even case is obtained in a similar way. Let $\frac{s}{t}$, $\frac{g}{h}$ and $\frac{x}{y}$ be the fractions defined at the beginning of Section 6.1. The α-block is formed of k rows and columns where $k = x + y = |\,\mathrm{FRF}(\frac{a}{b})|$.

Lemma 5. *The α-block of $U_{\frac{a}{b}}$ is exactly the first $k-1$ rows and columns of the matrix $U_{\frac{u}{v}}$ where $\frac{u}{v} = \mathrm{TBF}(\frac{a}{b})$.*

Proof. At the beginning, we show that the first $k-1$ rows of $B_{\frac{a}{b}}$ are exactly the concatenation of $a_z - 1$ times of the matrix $B_{\frac{x}{y}}$ and the $k-1$ first rows of the matrix $B_{\frac{u}{v}}$. Then we show that the set of factors of these rows of length shorter than k is exactly the same for the rows of $B_{\frac{u}{v}}$.

By Lemma 3 and since z is odd, we have: $C(\frac{a}{b}) = C(\frac{x}{y})^{a_z-1}C(\frac{u}{v})$ and $C(\frac{u}{v}) = C(\frac{x}{y})C(\frac{s}{t})$. Therefore, the k first columns of $B_{\frac{a}{b}}$ are given by $B_{\frac{x}{y}}$. Indeed, this part of $B_{\frac{a}{b}}$ is constructed due to factors of lengths less or equal than k where only the first k positions are considered. This means that all these factors are included in the prefix of length $2k$ of $C\left(\frac{a}{b}\right)$ which is exactly $C(\frac{x}{y})^2$. The same argument holds for each $B_{\frac{a}{b}}[i,j]$, for $1 \leq i \leq k-1$ and $1 \leq j \leq n-(u+v)$, which forms the $a_z - 1$ blocks of $B_{\frac{x}{y}}$. Note that to obtain the last block of $B_{\frac{x}{y}}$, we use the fact that $C\left(\frac{u}{v}\right)$ has $C(\frac{x}{y})$ as a prefix. For the last block, we observe that the entries of $B_{\frac{a}{b}}[i,j]$ for $1 \leq i \leq k-1$ and $n-(u+v)+1 \leq j \leq n$, are exactly the same as in $B_{\frac{u}{v}}$ since, again, we only consider factors

up to length k and both words $C\left(\frac{a}{b}\right)$ and $C\left(\frac{u}{v}\right)$ start with $C(\frac{x}{y})$ of length k. Hence we have: $B_{\frac{a}{b}}[i] = B_{\frac{x}{y}}[i]^{a_z-1} \cdot B_{\frac{u}{v}}[i]$, for all $1 \leq i \leq k-1$. Finally, we can write that $L_j(C(\frac{x}{y})^\omega) \subseteq L_j\left(C(\frac{u}{v})^\omega\right)$, for $j \leq k-1$ since $C(\frac{x}{y})$ ends with $C\left(\frac{s}{t}\right)$ and $C\left(\frac{u}{v}\right) = C(\frac{x}{y})C\left(\frac{s}{t}\right)$. Which implies

$$L_j\left(B^\omega_{\frac{a}{b}}[i]\right) = L_j\left(B^\omega_{\frac{u}{v}}[i]\right), \ \forall\ 1 \leq i \leq k-1 \ \text{ and } 1 \leq j \leq k-1.$$

γ-block: The second block of $U_{\frac{a}{b}}$ is the γ-block of dimension $(n-2k-1) \times k$. This block is located between the $k+1$ and the $(n-k-1)^{th}$ rows and bounded by the row of separation at the kth column. This part of the second order balance matrix admits a recursive form. If we consider the concatenation of matrices as a vertical stacking, then $\gamma(\frac{a}{b})$ is given by

$$\gamma\left(\frac{a}{b}\right) = \begin{cases} \gamma\left(\frac{c}{d}\right)\left[\gamma\left(\frac{u}{v}\right)\gamma\left(\frac{c}{d}\right)\right]^{a_z-2} & \text{if } a_z \geq 3 \\ \\ \gamma\left(\frac{e}{f}\right) & \text{if } a_z = 2 \end{cases} \tag{2}$$

We mentioned that we are considering the case where $z \geq 2$ and with z an odd value. In fact, the parity of z determines if a/b is a right or left child. Furthermore, it helps us to know the value of α, where α is the value needed to obtain $D(a,b)$.

Lemma 6. *Let $\frac{a}{b} = [a_0, \ldots, a_z]$, and $\mathrm{FRF}(\frac{a}{b}) = \frac{x}{y}$ with $x+y=k$.*
The set $D(a,b) = \{i\alpha + 1 \mod n | 1 \leq i \leq a\}$ where $n = a+b$ has:

$$\begin{cases} \alpha = k, & if\ z\ is\ even \\ \alpha = n-k, & if\ z\ is\ odd. \end{cases}$$

The proof can be found in the Appendix.

In order to prove the construction of the $\gamma-$block, we need to introduce some additional properties of the balance matrix $B_{\frac{a}{b}}$ used in the technical proof of the $\gamma-$block.

Property 6. Let $w = 0^x 1^y$ such that $x, y > 0$. For $i \leq x$ and $i \leq y$, we have: $B_w[i] = 0^{x-i+1}12\cdots(i-2)(i-1)\cdot i^{y-i+1}\cdot(i-1)(i-2)\cdots 21$.

Proof. Let $w = 0^x 1^y$, for $i = 2$, $S_w[i,p] = |w^\omega[p, p+i-1]|_1$ for all $1 \leq p \leq x+y$, therefore $S_w[2] = 0^{x-1}12^{y-1}1 = 0^{x-2+1}12^{y-2+1}1$ and since $x-1 > 0$ then $\min(S_w[2]) = 0$ and $B_w[2] = S_w[2]$. Suppose this is true for $i \leq x$, hence $B_w[i] = 0^{x-i+1}12\ldots i^{y-i+1}\ldots 21$ prove it true for $i+1$ such that $i+1 \leq x$. We have $S_w[i+1,p] = 0$ for all $1 \leq p \leq x-i$, for $p = x-i+1$, we have $S_w[i+1,p] = 1$ and for $x-i+1 \leq p \leq y-x-i$, we get $S_w[i+1,p] = i+1$. The values decrease until $p = x+y$ and we have $S_w[i+1, x+y] = 1$. Since $i+1 \leq x$, then $x-i \geq 1$ and $\min S_w[i+1] = 0$, therefore $B_w[i] = 0^{x-i+1}12\ldots i^{y-i+1}\ldots 21$.

Lemma 7. *Let $\frac{a}{b} = [a_0, \ldots, a_z]$ and $\frac{x}{y} = \mathrm{FRF}(\frac{a}{b})$ such that $|C(\frac{x}{y})| = x+y = k$.*

$$\begin{cases} B_{\frac{a}{b}}[k] = 0^\alpha 10^{n-\alpha-1}\ ; & B_{\frac{a}{b}}[2k] = 0^{n-2k}10^{k-1}10^{k-1} if\ z\ is\ odd, \\ B_{\frac{a}{b}}[n-k] = 0^\alpha 10^{n-\alpha-1}\ ; & B_{\frac{a}{b}}[n-2k] = 0^{n-2k}10^{k-1}10^{k-1} if\ z\ is\ even. \end{cases} \tag{3}$$

Proof. We prove the case where z is odd and the second case is obtained in a similar way. We let p_1, p_2 be two palindromes such that: $C(\frac{s}{t}) = 0p_11$, $C(\frac{g}{h}) = 0p_21$ where $C(\frac{x}{y}) = C(\frac{g}{h})C(\frac{s}{t})$; $C(\frac{a}{b}) = (C(\frac{x}{y}))^{a_z-1}C(\frac{u}{v})$ and $C(\frac{u}{v}) = C(\frac{x}{y})C(\frac{s}{t})$. The central word of $C(\frac{x}{y})$ is a palindrome then: $p_210p_1 = p_101p_2$. By the definition of the balance matrix, we have $\forall\ 1 \leq i \leq n-1$ and $\forall\ 1 \leq j \leq n$ $B_{\frac{a}{b}}[k,j] = S_{\frac{a}{b}}[k,j] - \min(S_{\frac{a}{b}}[k]) = |C(\frac{a}{b})[j \ldots k+j-1]|_1 - \min(S_{\frac{a}{b}}[k])$. Summing up all this we get:

$$C(\frac{a}{b}) = C(\frac{x}{y})^{a_z-1} \cdot C(\frac{x}{y})C(\frac{s}{t}) = C(\frac{x}{y})^{a_z-1} \cdot 0p_210p_110p_11 = C(\frac{x}{y})^{a_z-1} \cdot 0p_101p_210p_11.$$

We know that $k = |p_1| + |p_2| + 4$ and $|C(\frac{x}{y})|_1 = |p_1|_1 + |p_2|_1 + 2 = x$. Therefore all the elements $u \in L_k(C(\frac{a}{b})^\omega)$ have $|u|_1 = |p_1|_1 + |p_2|_1 + 2$ except $u' = 1p_210p_11$ which is of length k and at position $n-k$ but has $|u'|_1 = x+1$. Hence by subtracting by $\min(S_{\frac{a}{b}}[k])$, we get $B_{\frac{a}{b}}[k] = 0^{n-k}10^{k-1}$. By lemma 6, we know that since z is odd, $\alpha = n-k$, which ends the proof.
Same reasoning is used to determine the row $B_{\frac{a}{b}}[2k]$, where $2k = 2|C(\frac{x}{y})| = 2|p_1| + 2|p_2| + 8$.

The following lemma states the recursive construction of the $\gamma-$block in a formal way.

Lemma 8. *The $\gamma-$block is obtained depending on the position of the fraction in the Stern-Brocot tree.*

- *If $a_z \geq 3$ then $\gamma\left(\frac{a}{b}\right)$ is obtained by stacking vertically $\gamma(\frac{c}{d})$ over $a_z - 2$ copies of $\gamma(\frac{u}{v})$ and $\gamma(\frac{c}{d})$.*
- *Otherwise, if $a_z = 2$, then $\gamma\left(\frac{a}{b}\right)$ is given by the extension of $\gamma(\frac{e}{f})$ to the k-th column of $U_{\frac{e}{f}}$.*

where $\frac{u}{v} = \mathrm{TBF}(\frac{a}{b})$; $\frac{c}{d} = \mathrm{FDF}(\frac{u}{v})$; $\frac{e}{f} = \mathrm{FPF}(\frac{a}{b})$.

The proof can be found in the Appendix.

***β*-block** The $\beta-$block is the center of $U_{\frac{a}{b}}$, the last block needed to complete the construction of the second order balance matrix. This block is of dimension $(n-2k-1) \times (n-2k-1)$ and located between the rows (respectively columns) k and $n-k$.

Lemma 9. *The $\beta-$block of $U_{\frac{a}{b}}$ is exactly the second order balance matrix of the fraction $\frac{\rho}{\theta} =$ $\mathrm{SUF}(\frac{a}{b})$ in the Stern-Brocot tree where its elements are increased by 1.*

The proof can be found in the Appendix.

7 Perspectives

In the introduction, we stated that there were connections between the second order of balance and the Fraenkel conjecture, and we mentioned the importance of the second order of balance to construct balanced words on a $k-$letters alphabet. Hence, we have to show some hints on how to use the second order balance matrix U_w to have new information on the synchronisation of words. Of course the Fraenkel conjecture is far from being solved because of its own complexity nevertheless we have built a new tool to synchronize balanced words by considering synchronization of blocs instead of letters. The following example is based on the word $FR_3 = 1213121$ on a three letter alphabet, which is obtained from the synchronization of the 3 Christoffel words of slope 1/6 associated with the word $C_3 = ***3***$, of slope 2/5 associated with the word $C_2 = *2***2*$

$$U_{\frac{1}{6}} = \begin{pmatrix} 1&1&1&1&1&1 \\ 1&2&2&2&2&1 \\ 1&2&3&3&2&1 \\ 1&2&3&3&2&1 \\ 1&2&2&2&2&1 \\ 1&1&1&1&1&1 \end{pmatrix}, U_{\frac{2}{5}} = \begin{pmatrix} 1&1&1&1&1&1 \\ 1&2&1&1&2&1 \\ 1&1&1&1&1&1 \\ 1&1&1&1&1&1 \\ 1&2&1&1&2&1 \\ 1&1&1&1&1&1 \end{pmatrix} \text{ and } U_{\frac{4}{3}} = U_{\frac{3}{4}} = \begin{pmatrix} 1&1&1&1&1&1 \\ 1&1&1&1&1&1 \\ 1&1&2&2&1&1 \\ 1&1&2&2&1&1 \\ 1&1&1&1&1&1 \\ 1&1&1&1&1&1 \end{pmatrix}.$$

Fig. 8: The three matrices $U_{\frac{1}{6}}$, $U_{\frac{2}{5}}$ and $U_{\frac{3}{4}}$ used for the synchronization of the triplet $(1, 2, 4)$ mod 7.

and of slope 3/4 associated with the word $C_1 = 1 * 1 * 1 * 1$. Remark that in each position we have only one value and the other symbols are stars. We are able to read this synchronization on rows of matrices in Figure 8:

If we name $M_{\frac{a}{b}}$ the set of indices of the rows with a maximal values in $U_{\frac{a}{b}}$ and $m_{\frac{a}{b}}$ the set of indices of the rows with minimal values in $U_{\frac{a}{b}}$ (which are non previously chosen rows), we could extract the following set of indices: $M_{\frac{1}{6}} = \{3, 4\}$ and $m_{\frac{1}{6}} = \{1, 6\}$; $M_{\frac{2}{5}} = \{2, 5\}$ and $m_{\frac{2}{5}} = \{3, 4\}$; $M_{\frac{3}{4}} = \{3, 4\}$ and $m_{\frac{3}{4}} = \{2, 5\}$. We can observe that starting with the elements of the set $m_{\frac{3}{4}}$, if we double the values and take modulo 7, we get the elements of $m_{\frac{2}{5}}$. The same note can be also given for the sets $m_{\frac{2}{5}}$ and $m_{\frac{1}{6}}$. $2.2 = 4$; $5.2 = 10 \equiv_7 3$, where $\{3, 4\} = m_{\frac{2}{5}}$. $4.2 = 8 \equiv_7 1$; $3.2 = 6$, where $\{1, 6\} = m_{\frac{1}{6}}$. Except for the rows 1 and 6, we can see the synchronization between the sets $m_{\frac{a}{b}}$ and $M_{\frac{a}{b}}$, more precisely, if a row appears as a minimum for the set m it appears as a maximum for the set M, which allows us to read all information for the synchronization in the second order balance matrix. This will be a keypoint for our next research, in order to study the synchronization of Christoffel words over a $k-$letters alphabet and to try to tackle the Fraenkel's conjecture.

In addition, many research problems are still open in this study. For example, how to extend the second order balance matrix construction for a binary word that are not Christoffel words?

We could also investigate the Stern-Brocot tree in order to find the structure of infinite paths that minimize the second order balance and find a combinatorial construction of these paths.

To end the perspectives, we could notice that considering the balance matrix is equivalent to compute balance property of a given word. This computation appears in another context in order to compute Parikh vectors that are $k-$dimensional vectors defined for a finite word w on a $k-$letters alphabet. The ith coordinate of this vector gives the number of occurrences of the letter i in the word w. The notion of relative Parikh vectors is introduced by Turek, in [**?**,**?**], in order to understand the structure of words. Nevertheless, the computation of the second order balance matrix using Parikh vectors is not straightforward and will be considered in a future article.

References

1. Altman, E., Gaujal, B., Hordijk, A.: Balanced sequences and optimal routing. Journal of the ACM **47**(4), 752–775 (2000)
2. Baryshnikov, Y.: Complexity of trajectories in rectangular billiards. Communications in mathematical physics **174**(1), 43–56 (1995)
3. Berstel, J.: Tracé de droites, fractions continues et morphismes itérés. In: Mots, pp. 298–309. Lang. Raison. Calc., Hermès (1990), incollection
4. Berstel, J., De Luca, A.: Sturmian words, Lyndon words and trees. Theoretical Computer Science **178**(1), 171–203 (1997)
5. Berstel, J., Lauve, A., Reutenauer, C., Saliola, F.: Combinatorics on words: Christoffel words and repetition in words **27**, 147 (2008)
6. Berthé, V., De Luca, A., Reutenauer, C.: On an involution of Christoffel words and sturmian morphisms. European Journal of Combinatorics **29**(2), 535–553 (2008)
7. Borel, J.P., Laubie, F.: Quelques mots sur la droite projective réelle. Journal de théorie des nombres de Bordeaux **5**(1), 23–51 (1993)
8. Brocot, A.: Calcul des rouages par approximation. Revue chronométrique. Journal des horlogers, scientifique et pratique, 3 (1861)
9. Christoffel, E.B.: Observatio arithmetica. Annali di Matematica Pura ed Applicata (1867-1897) **6**(1), 148–152 (1873)
10. Cicalese, F., Lipták, Z., Rossi, M.: On infinite prefix normal words. SOFSEM 2019 pp. 122–135 (2019)

11. Fogg, N.P.: Substitutions in Dynamics, Arithmetics and Combinatorics. Springer Berlin Heidelberg, Berlin, Heidelberg (2002)
12. Fraenkel, A.S.: The bracket function and complementary sets of integers. Canadian Journal of Mathematics **21**, 6–27 (1969)
13. Fraenkel, A.S., Levitt, J., Shimshoni, M.: Characterization of the set of values f (n)=[na], n= 1, 2,.... Discrete Mathematics **2**(4), 335–345 (1972)
14. Gaujal, B.: Optimal allocation sequences of two processes sharing a resource. Discrete Event Dynamic Systems **7**(4), 327–354 (1997)
15. Gaujal, B., Varjú, P.P.: Partitioning the positive integers into seven beatty sequences. Indag. Math. N.S. **12**, 149–161 (2003)
16. Hubert, P.: Suites équilibrées. Theoretical Computer Science **242**(1), 91–108 (2000)
17. Lothaire, M.: Algebraic combinatorics on words, Encyclopedia of Mathematics and its Applications, vol. 90. Cambridge University Press, Cambridge (2002)
18. de Luca, A., Mignosi, F.: Some combinatorial properties of sturmian words. Theoretical Computer Science **136**(2), 361–385 (1994)
19. Mairesse, J., Vuillon, L.: Asymptotic behavior in a heap model with two pieces. Theoretical Computer Science **270**(1), 525–560 (2002)
20. Morikawa, R.: On eventually covering families generated by the bracket function iv. Natural Science **25**, 1–8 (1985)
21. Morse, M., Hedlund, G.A.: Symbolic dynamics II. Sturmian trajectories. American Journal of Mathematics **62**(1), 1–42 (1940)
22. Paquin, G., Reutenauer, C.: On the superimposition of Christoffel words. Theoretical Computer Science **412**(4), 402–418 (2011)
23. Paquin, G., Vuillon, L.: A characterization of balanced episturmian sequences. Electronic Journal of Combinatorics **14** (2007)
24. Provençal, X., Vuillon, L.: Discrete segments of Z^3 constructed by synchronization of words. Discrete Applied Mathematics (2014)
25. Raney, G.N.: On continued fractions and finite automata. Mathematische Annalen **206**(4), 265–283 (1973)
26. Sidorov, N.: Optimizing properties of balanced words pp. 240–246 (2011)
27. Skolem, T.: On certain distributions of integers in pairs with given differences. Mathematica Scandinavica **5**, 57–68 (1957)
28. Smith, H.J.S.: Note on continued fractions. Messenger Math **6**, 1–14 (1876)
29. Tijdeman, R.: Fraenkel's conjecture for six sequences. Discrete Mathematics **222**(1), 223–234 (2000)
30. Vuillon, L.: Balanced words. Bulletin of the Belgian Mathematical Society-Simon Stevin **10**(5), 787–805 (2003)