Characterizations of perfectly clustering words

Mélodie Lapointe[a] Christophe Reutenauer[b]

Submitted: Mar 5, 2024; Accepted: May 6, 2025; Published: Jul 4, 2025

Abstract

Perfectly clustering words are one of many possible generalizations of Christoffel words. In this article, we propose a factorization of a perfectly clustering word on a n letters alphabet into a product of $n-1$ palindromes with a letter between each of them. This factorization allows us to generalize two combinatorial characterization of Christoffel words due to Pirillo (1999) and de Luca and Mignosi (1994).

Mathematics Subject Classifications: 68R15

Contents

[a]Département de mathématiques et statistique, Université de Moncton, Moncton, Canada (melodie.lapointe@umoncton.ca).

[b]Département de mathématiques, Université du Québec à Montréal, Montréal, Canada (Reutenauer.Christophe@uqam.ca).

1 Introduction

Christoffel words and related families of binary words like standard words and central words have been extensively studied, resulting in many equivalent but seemingly unrelated definitions of these sets of words. Several generalizations have been studied in the literature, for larger alphabets, or in larger dimension [4, 8, 9, 16, 18]. However, the equivalence between the different definitions of Christoffel words does not extend in general.

In this article, we aim to study properties of perfectly clustering words, a generalisation of Christoffel words proposed in [22, 5]. A word w is perfectly clustering if its Burrows-Wheeler transform is a weakly decreasing word. In [15], Mantaci, Restivo and Sciortino showed that Christoffel words and their conjugates are exactly the binary words whose Burrows-Wheeler transform is a weakly decreasing word. Later, Ferenczi and Zamboni [5] related these words with symmetric interval exchange transformations. Moreover, they also are primitive elements of the free group [11], and product of two palindromes [22]. Thus perfecly clustering words share several properties of Christoffel words.

Our main result (Theorem 7) uses what we call the *special factorization* of a perfectly clustering Lyndon word, in order to give two characterizations of these words. This result generalizes the following characterizations of Christoffel words:

- a binary word amb is a Christoffel word if and only if the word amb is a product of two palindromes and m is also a palindrome (de Luca and Mignosi [14]),
- a binary word amb is a Christoffel word if and only if the words amb and bma are conjugates (Pirillo [20]).

The article is structured as follows. In Section 2, we recall known results and definitions. In Section 3, we introduced the special factorization and characterize perfectly clustering word with it. In Section 4, we prove several lemmas that are necessary for our characterization. In Section 5, we present the proof of it. In Section 6, we give some results on the structure of the Burrows-Wheeler matrix of a perfectly clustering word.

2 Perfectly clustering words

2.1 Words

Let $A = \{a_1, a_2, \dots, a_k\}$ be a totally ordered alphabet, where $a_1 < a_2 < \cdots < a_k$. Let A^* denote the free monoid generated by A and $F(A)$ denote the free group on A. An element in $F(A)$ which lies actually in A^* is called *positive*.

The length of $w = b_1 \cdots b_n$ (with $b_i \in A$), denoted by $|w|$, is n. The *empty word* is denoted by 1 and is the only element of length 0. The number of occurrences of a letter a in w is denoted by $|w|_a$. The *commutative image* of w is the integer vector $(|w|_{a_1}, \dots, |w|_{a_k})$. The function *Alph* is defined by $Alph(w) = \{x \in A \mid |w|_x \geqslant 1\}$.

A word w is called *primitive* if it is not the power of another word, that is: for any word z such that $w = z^n$, one has $n = 1$. The *conjugates* of a word w, as above, are the

a	p	a	r	t	m	e	n	t
a	r	t	m	e	n	t	a	p
e	n	t	a	p	a	r	t	m
m	e	n	t	a	p	a	r	t
n	t	a	p	a	r	t	m	e
p	a	r	t	m	e	n	t	a
r	t	m	e	n	t	a	p	a
t	a	p	a	r	t	m	e	n
t	m	e	n	t	a	p	a	r

Figure 1: The Burrows-Wheeler matrix of the word *apartment* sorted in lexicographic order. The last column is its Burrows-Wheeler transform

words $b_i \dots b_n b_1 \dots b_{i-1}$. In other words, two words $u, v \in A^*$ are *conjugate* if for some words $x, y \in A^*$, one has $u = xy, v = yx$. The *conjugation class* of a word is the set of its conjugates. If a word w is primitive, then it has exactly n distinct conjugates.

A *palindrome* in the free group $F(A)$ is an element fixed by $g \mapsto \tilde{g}$ the unique anti-automorphism of $F(A)$ such that $\tilde{a} = a$ for any $a \in A$. We call $\tilde{g}$ the *reversal* of g. It is well-known that a word w in A^* is conjugate to its reversal if and only if w is a product of two palindromes (see [2]).

The *lexicographic* order is an extension of the total order on A defined as follows: if $u, v \in A^*$, we have $u < v$ if either u is a proper prefix of v, or $u = rxs$ and $v = ryt$ such that $x, y \in A$, $x < y$, and $r, s, t \in A^*$. A word w is called a *Lyndon word* if it is a primitive word, and it is the minimal word in lexicographic order among its conjugates.

2.2 Perfectly clustering words

Let v be a primitive word of length n on the alphabet A. Let $v_1 < v_2 < \cdots < v_n$ be its conjugates, lexicographically ordered. Let l_i be the last letter of the word v_i for $1 \leqslant i \leqslant n$. The *Burrows-Wheeler transform* of the word v, denoted by $\mathrm{bw}(v)$, is the word $l_1 \cdots l_n$. Define the *Burrows-Wheeler matrix* of v to be the matrix, with rows corresponding bijectively to the conjugates of w, lexicographically ordered, and with row-elements being the letters of the corresponding conjugate; then $bw(v)$ is the last column of this matrix. For example, the Burrows-Wheeler transform of the word *apartment* is $\mathrm{bw}(\textit{apartment}) = \textit{tpmteaanr}$ (see Figure 1). It follows from the definition that two words u and v are conjugates if and only if $\mathrm{bw}(u) = \mathrm{bw}(v)$ (see [15, Proposition 1]).

Following [5], we say that a primitive word v is π-*clustering* if

$$\mathrm{bw}(v) = a_{\pi(1)}^{|v|_{a_{\pi(1)}}} \dots a_{\pi(k)}^{|v|_{a_{\pi(k)}}},$$

where π is a permutation on $\{1, \dots, k\}$.

For example, the word $\textit{aluminium} = a_1 a_3 a_6 a_4 a_2 a_5 a_2 a_6 a_4$ is 451623-clustering, since $\mathrm{bw}(\textit{aluminium}) = \textit{mmnauuiil}$. A word is *perfectly clustering* if the permutation π is the symmetric permutation, i.e. $\pi(i) = k - i + 1$, for all $i \in \{1, \ldots, k\}$. It was proved by Sabrina Mantaci, Antonio Restivo and Marinella Sciortino that the perfectly clustering words on a two-letter alphabet are the Christoffel words and their conjugates (see [15, Theorem 9]; see also [21, Theorem 15.2.1]). If a primitive word is perfectly clustering, then all its conjugates are. Consequently, there is no loss of generality to study perfectly clustering Lyndon words.

For each letter $\ell \in A$, we define, following [10], two automorphisms λ_ℓ and ρ_ℓ of the free group $F(A)$ by:

$$\lambda_\ell(a) = \begin{cases} a\ell^{-1}, & \text{if } a < \ell; \\ a, & \text{if } a = \ell; \\ \ell a, & \text{if } a > \ell; \end{cases} \quad \text{and} \quad \rho_\ell(a) = \begin{cases} a\ell, & \text{if } a < \ell; \\ a, & \text{if } a = \ell; \\ \ell^{-1} a, & \text{if } a > \ell; \end{cases}$$

for any $a \in A$.

The following result, due to the first author, will be crucial in our proofs.

Theorem 1. *([10, Theorem 4.29]) For each perfectly clustering Lyndon word $w \in A^*$ of length at least 3, there exists a shorter perfectly clustering Lyndon word $u \in A^*$, and an automorphism $f = \lambda_\ell$ or ρ_ℓ such that $w = f(u)$. One may also assume that either a) $Alph(w) = Alph(u) = A$, or b) $A = Alph(w) = Alph(u) \cup \ell$ and $\ell \notin Alph(u)$.*

Corollary 2. *With the notations of the previous theorem, one has, in the case $f = \rho_\ell$, $\sum_{a \in A, a<\ell} |u|_a > \sum_{a \in A, a>\ell} |u|_a$.*

Proof. Since, for any $a \in A$, ρ_ℓ leaves invariant the a-degree, except if $a = \ell$, the fact that w is longer than u implies that $|w|_\ell > |u|_\ell$, that is,

$$|w|_\ell = |u|_\ell + \sum_{a<\ell} |u|_a - \sum_{a>\ell} |u|_a > |u|_\ell$$

(by the special form of ρ_ℓ); therefore $\sum_{a<\ell} |u|_a > \sum_{a>\ell} |u|_a$. ☐

We also need

Lemma 3. *([10, Lemma 4.6]) For each word $u \in A^*$, $\lambda_\ell(u)$ (resp. $\rho_\ell(u)$) $\in A^*$ (that is, is positive) if and only if each letter $< \ell$ (resp. $> \ell$) in w is followed (resp. preceded) in w by a letter $\geqslant \ell$ (resp. $\leqslant \ell$).*

2.3 Symmetric discrete interval exchanges

We follow [5]. Let $n \geqslant 1$ be an integer. Let $(c_1, \ldots, c_k)$ be a *composition* of n, that is, a k-tuple of positive integers whose sum is n. We decompose in two ways the interval $[n] = \{1, 2, \ldots, n\}$ into intervals: the intervals $I_1, \ldots, I_k$ (resp. $J_1, \ldots, J_k$) are defined by the condition that they are consecutive and that $|I_h| = c_h$ (resp. $|J_h| = c_{k+1-h}$). Denote by S_n the group of permutations of $[n]$. We define the permutation $\sigma \in S_n$ by the condition that it sends increasingly each interval I_h onto the interval J_{k+1-h} (they have the same cardinality c_h). We call such a permutation a *symmetric discrete interval exchange*[1], and it is said to be *associated with the composition* $(c_1, \dots, c_k)$.

As an example, consider the 3-tuple $(3, 3, 4)$, a composition of 10. The intervals I_1, I_2, I_3 are $\{1, 2, 3\}$, $\{4, 5, 6\}$, $\{7, 8, 9, 10\}$ and the intervals J_1, J_2, J_3 are $\{1, 2, 3, 4\}$, $\{5, 6, 7\}$, $\{8, 9, 10\}$. The permutation σ sends increasingly $\{1, 2, 3\}$ onto $\{7, 8, 9\}$, $\{4, 5, 6\}$ onto $\{5, 6, 7\}$ and $\{7, 8, 9, 10\}$ onto $\{1, 2, 3, 4\}$ whence

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 8 & 9 & 10 & 5 & 6 & 7 & 1 & 2 & 3 & 4 \end{pmatrix}$$

and its cycle form is $(1, 8, 2, 9, 3, 10, 4, 5, 6, 7)$.

It follows from this definition that for each $i < j$ in $[n]$ such that i, j are in the same interval I_h, one has $\sigma(i) < \sigma(j)$.

Note that the restriction of σ to I_h is the translation by some integer t_h; since its image is J_{k+1-h}, one has $t_h = \max(J_{k+1-h}) - \max(I_h)$. Thus σ is completely defined by:

$$\forall h = 1, \dots, k, \forall x \in I_h,\ \sigma(x) = x + t_h. \tag{1}$$

One has $\max(I_h) = c_1 + \cdots + c_h$ and $\max(J_{k+1-h}) = c_k + \cdots + c_h$, hence $t_h = \sum_{i>h} c_i - \sum_{i<h} c_i$. We call σ a *circular* symmetric discrete interval exchange if σ is a circular permutation, that is, has only one cycle.

A *word encoding* of a circular symmetric discrete interval exchange σ is one of the words obtained by replacing in one of the cycle forms of σ each number i by the letter a_j such that $i \in I_j$. Note that if w is a word encoding of σ, one has, with the notations above, $|w|_{a_j} = c_j$ for any j.

Theorem 4. *The following conditions are equivalent, for a primitive word w on a totally ordered alphabet:*

(i) w is primitive perfectly clustering word;

(ii) w is a word encoding of some circular symmetric discrete interval exchange.

This result is due to Sébastien Ferenczi and Luca Zamboni [5, Theorem 4]. A brief history around this result is given in Section 2.4.

In our example, the symmetric discrete interval exchange associated with $(3, 3, 4)$ is circular and its word encoding on the alphabet $a < b < c$ with $a_1 = a$, $a_2 = b$ and $a_3 = c$ is *acacacbbbc* which is a primitive perfectly clustering word as its Burrows-Wheeler transform is *ccccbbbaaa*, as shown in Figure 2.

Denote by m_h the minimum of the interval I_h. For later use, we prove two lemmas, the first one is an easy observation.

Lemma 5. *For any $h = 1, \dots, k - 1$, one has $\sigma(m_h) = m_{h+1}$ if and only if $\sum_{i>h} c_i - \sum_{i \leqslant h} c_i = 0$.*

[1] The word "symmetric" refers to the fact that the intervals are exchanged according to the central symmetry of the set $\{1, 2, \dots, k\}$, that is, the mapping $h \mapsto k + 1 - h$.

a	c	a	c	a	c	b	b	b	c
a	c	a	c	b	b	b	c	a	c
a	c	b	b	b	c	a	c	a	c
b	b	b	c	a	c	a	c	a	c
b	b	c	a	c	a	c	a	c	b
b	c	a	c	a	c	a	c	b	b
c	a	c	a	c	a	c	b	b	b
c	a	c	a	c	b	b	b	c	a
c	a	c	b	b	b	c	a	c	a
c	b	b	b	c	a	c	a	c	a

Figure 2: Burrows-Wheeler matrix of the perfectly clustering Lyndon word *acacacbbbc*

Proof. By definition of the intervals I_h, one has $m_h = \sum_{i<h} c_i + 1$. Hence, $\sigma(m_h) = m_h + t_h = \sum_{i<h} c_i + 1 + \sum_{i>h} c_i - \sum_{i<h} c_i = 1 + \sum_{i>h} c_i$. This is equal to $m_{h+1} = \sum_{i \leqslant h} c_i + 1$ if and only if the equality in the lemma holds. □

Lemma 6. *([5, Proof of Theorem 4]) Let σ be a circular symmetric discrete interval exchange, with the notations above. For each cyclic representation*

$$\gamma_r = (r, \sigma(r), \cdots, \sigma^{n-1}(r))$$

of σ, denote by w_r the word obtained by replacing in γ_r each number i by a_j if $i \in I_j$. Then $w_r <_{lex} w_s \Leftrightarrow r < s$.

Proof. Suppose that $r < s$. Let $p \leqslant n$ be maximal such that the prefixes of length p of w_r and w_s coincide. Then for each $i = 1, \ldots, p$, the i-th letters of w_r and w_s are equal, hence $\sigma^{i-1}(r)$ and $\sigma^{i-1}(s)$ are in the same interval I_j. It follows recursively from the observation before (1) that $\sigma^p(r) < \sigma^p(s)$. Suppose that $p < n$; since p is maximal, the previous two numbers are not in the same interval, so that they are respectively in two intervals I_h, I_j with $h < j$ (the intervals are successive); hence the $(p+1)$-th letter of w_r (obtained by replacing $\sigma^p(r)$ according to the rule in the statement) is a_h, which is strictly smaller than a_j, which is the $(p+1)$-th letter of w_s and therefore $w_r <_{lex} w_s$. Suppose now by contradiction that $p = n$; then we have $\sigma^{i-1}(r) < \sigma^{i-1}(s)$ for any $i = 1, \ldots, n$, and this is not possible since σ is circular so that some $\sigma^{i-1}(s)$ is equal to 1. This ends the proof of the equivalence, since both orders are total. □

In the context of Theorem 4, this lemma has several consequences: if we identify the set $\mathcal{C}$ of conjugates of w, lexicographically ordered, with $[n]$, under the unique order isomorphism between the two sets, then σ is identified with the *conjugator* C, which is the mapping sending each nonempty word au onto ua ($a \in A, u \in A^*$). Moreover the interval I_s is identified with the conjugates of w beginning by a_s. And m_s, the minimum of the interval I_s, is identified with the smallest conjugate of w beginning by a_s.

2.4 Brief history of Theorem 4

Continuous interval exchanges have a long history, beginning in 1966 by Osedelets [17]. But discrete interval exchanges appeared more recently, in 2013, in the article of Ferenczi and Zamboni [5]. They prove a more general theorem that Theorem 4; the latter is obtained from theirs by taking π to be the longest permutation in S_k, that is $\pi(i) = k+1-i$. However, for a two-letter alphabet, the result was given implicitly in [15, Section 3 and Theorem 9] (2002). Moreover, for a general alphabet, in [22] (2008) the function ω page 13 is a symmetric discrete interval exchange, as shows the formula there, similar to (1). It turns out that, for perfectly clustering words w, the intervals I_h and J_h defining the associated symmetric discrete interval exchange may be read directly on the first and last columns of the Burrows-Wheeler matrix of w.

Note that the Burrows-Wheeler transform is a particular case, discovered independently, of a bijection due to Gessel and the second author [6] (1993).

3 A characterization of perfectly clustering words

Let w be a word on a totally ordered alphabet A. We call *special factorization of* w a factorization of w of the form

$$a_1\pi_1 a_2\pi_2 \cdots \pi_{k-1}a_k, \tag{2}$$

where $Alph(w) = \{a_1 < a_2 < \cdots < a_k\}$ and $\pi_1, \ldots, \pi_{k-1} \in A^*$. If w is equal to (2), we set

$$W = a_k\pi_{k-1}\cdots\pi_2 a_2\pi_1 a_1.$$

We call this special factorization *palindromic* if each word π_i is a palindrome; in this case $W = \tilde{w}$.

Theorem 7. *The following conditions are equivalent, for a primitive word w on the totally ordered finite alphabet A:*

(i) w is a perfectly clustering Lyndon word;

(ii) w is a product of two palindromes and w has a palindromic special factorization;

(iii) w has a special factorization (2) such that w is conjugate to W.

This theorem generalizes known results on perfectly clustering Lyndon words on two letters, which are the Christoffel words, according to [15]. The equivalence of (i) and (ii) is a generalization of a result of Aldo de Luca and Filippo Mignosi, in [14, Proposition 8] (see also [12, Theorem 2.2.4] or [21, Theorem 12.2.10]). The equivalence of (i) and (iii) is a generalization of a result of Giuseppe Pirillo [20] (see also [21, Theorem 15.2.5]).

The theorem will be proved in Section 5. The proof uses crucially the automorphisms of the free group introduced by the first author [10] (see Subsection 2.2). It would be interesting to find a direct proof, using only the definition of pefectly clustering words, and/or discrete interval exchanges.

4 Several lemmas

Lemma 8. *If $g \in F(A)$ is a palindrome, then so are $\lambda_\ell(g)\ell$ and $\ell^{-1}\lambda_\ell(g)$ (resp. $\ell\rho_\ell(g)$ and $\rho_\ell(g)\ell^{-1}$).*

Proof. We prove it for ρ_ℓ, the proof for λ_ℓ being similar. If $g \in \{1\} \cup A \cup A^{-1}$, this is easily verified. If g is a palindrome in $F(A)$, not of the previous form, then $g = xhx$ with $x \in A \cup A^{-1}$, and h a palindrome, shorter than g. By induction, $\ell\rho_\ell(h)$ is a palindrome, hence $\rho_\ell(h)\ell^{-1}$ too. We have $\ell\rho_\ell(g) = \ell\rho_\ell(x)\rho_\ell(h)\rho_\ell(x)$.

Suppose first that $x = a \in A$. If $a < \ell$, then $\ell\rho_\ell(g) = \ell a\ell\rho_\ell(h)a\ell$; if $a = \ell$, then $\ell\rho_\ell(g) = \ell\ell\rho_\ell(h)\ell$; if $a > \ell$, then $\ell\rho_\ell(g) = \ell\ell^{-1}a\rho_\ell(h)\ell^{-1}a = a\rho(h)\ell^{-1}a$.

Suppose now that $x = a^{-1}$ and $a \in A$. If $a < \ell$, then $\ell\rho_\ell(g) = \ell\ell^{-1}a^{-1}\rho_\ell(h)\ell^{-1}a^{-1} = a^{-1}\rho_\ell(h)\ell^{-1}a^{-1}$. If $a = \ell$, then $\ell\rho_\ell(g) = \ell\ell^{-1}\rho_\ell(h)\ell^{-1} = \rho_\ell(h)\ell^{-1}$; if $a > \ell$, then $\ell\rho_\ell(g) = \ell a^{-1}\ell\rho_\ell(h)a^{-1}\ell$.

In each case $\ell\rho_\ell(g)$ is a palindrome. □

Lemma 9. *Let $u \in A^*$, $b, \ell \in A$, with $b > \ell$. Suppose that $\rho_\ell(u)$ is positive, and that u has a factor πb, where π is a nonempty palindrome. Then $\rho_\ell(\pi)\ell^{-1}$ is a positive palindrome.*

Proof. We know that each letter $> \ell$ in u is preceded by a letter $\leqslant \ell$. Hence the last letter of π is $a \leqslant \ell$, and equal to its first letter: $\pi = ava$, or $\pi = a$. If $\pi = a$, then $\rho_\ell(\pi)\ell^{-1} = a$ if $a < \ell$, and $\rho_\ell(\pi)\ell^{-1} = 1$ if $a = \ell$, which settles this case.

Suppose now that $\pi = ava$. Then each letter $> \ell$ in av is preceded by a letter $\leqslant \ell$; therefore $\rho_\ell(av)$ is positive. If $a < \ell$, then $\rho_\ell(\pi)\ell^{-1} = \rho_\ell(av)a\ell\ell^{-1} = \rho_\ell(av)a$ is positive; if $a = \ell$, $\rho_\ell(\pi)\ell^{-1} = \rho_\ell(av)\ell\ell^{-1} = \rho_\ell(av)$ is positive. Moreover, $\rho_\ell(\pi)\ell^{-1}$ is a palindrome by Lemma 8. □

The easy proofs of the two following lemmas are left to the reader.

Lemma 10. *Let E be a totally ordered finite set, with some element M. Suppose that there exists an injective function $\nu : E\backslash\{M\} \to E$ such that $\forall x \in E\backslash\{M\}, x < \nu(x)$. Then $M = \max(E)$ and ν is the next element function of E (that is $\nu(x) = \min\{y \in E, y > x\}$).*

Lemma 11. *Let w be a word with a special factorization as in (2). Consider $i = 1, \ldots, k-1$ and some factorization $\pi_i = uv$. Then*

$$va_i \cdots \pi_1 a_1 a_k \pi_{k-1} \cdots a_{i+1} u < va_{i+1} \cdots \pi_{k-1} a_k a_1 \pi_1 \cdots a_i u,$$

the left-hand side word is conjugate to W, and the right-hand side word is conjugate to w.

Lemma 12. *Let w be a primitive word with a special factorization (2), such that w is conjugate to W. Consider two consecutive rows of its Burrows-Wheeler matrix, viewed as two words w', w''. Then*

$$w' = va_i \cdots \pi_1 a_1 a_k \pi_{k-1} \cdots a_{i+1} u$$

and

$$w'' = va_{i+1} \cdots \pi_{k-1} a_k a_1 \pi_1 \cdots a_i u,$$

for some $i = 1, \ldots, k-1$, and some factorization $\pi_i = uv$. Moreover, w is the smallest, and W the largest, element in their conjugation class.

Proof. Let $\mathcal{C}$ denote the set of conjugates of w and W. Note that, by Lemma 11, the elements of $\mathcal{C}$ are w, together with all words w'', and they are all distinct since w is primitive. Similarly, the elements of $\mathcal{C}$ are W, together with all the words w', and these words are distinct.

Define a mapping $\nu : \mathcal{C} \setminus W \to \mathcal{C}$ by

$$\nu(w') = w''. \tag{3}$$

This mapping is well-defined, by the previous paragraph, and by Lemma 11. Moreover, ν is injective since the words at the right of (3) are all distinct. By Lemma 11, ν maps each element of $\mathcal{C} \setminus W$ onto a larger element in $\mathcal{C}$.

Thus by Lemma 10, $W = \max(\mathcal{C})$ and ν is the next function in $\mathcal{C}$. Moreover, $w = \min(\mathcal{C})$, since w is not in the image of ν. □

Lemma 13. *Let w be a perfectly clustering Lyndon word w with a special palindromic factorization (2). Then for any i, $a_i\pi_i \cdots \pi_{k-1}a_k a_1 \pi_1 \cdots a_{i-1}\pi_{i-1}$ is the smallest conjugate of w beginning by a_i.*

Proof. The conclusion is clear for $i = 1$, since w is a Lyndon word. Suppose that $i \geqslant 2$. It is enough to prove that, in the conjugation class of w, the word preceding $a_i\pi_i \cdots \pi_{k-1}a_k a_1 \pi_1 \cdots a_{i-1}\pi_{i-1}$ for the lexicographical order begins by a_{i-1}. This follows from the next paragraph.

Since w is a perfectly clustering word, it is a product of two palindromes ([22] Corollary 4.4), hence the word $\tilde{w}$ is conjugate to w; moreover, $\tilde{w}$ is the largest conjugate of w for lexicographical order ([22] Theorem 4.3), and clearly $W = \tilde{w}$. By Lemma 12, for two consecutive rows of the Burrows-Wheeler matrix of w, viewed as two words w', w'', one has

$$w' = va_{i-1} \cdots \pi_1 a_1 a_k \pi_{k-1} \cdots a_i u$$

and

$$w'' = va_i \cdots \pi_{k-1} a_k a_1 \pi_1 \cdots a_{i-1} u,$$

for some $i = 2, \dots, k-1$, with $\pi_{i-1} = uv$. Hence w', w'' begin by the same letter, except if $v = 1$, in which case w' begins by a_{i-1} and $w'' = a_i \cdots \pi_{k-1}a_k a_1 \pi_1 \cdots a_{i-1}\pi_{i-1}$ begins by a_i. □

Lemma 14. *Let w be a perfectly clustering Lyndon word, which is an encoding of some circular symmetric discrete interval exchange, corresponding to the composition $(c_1, \dots, c_k)$. Suppose that w has a special palindromic factorization (2). Let $s = 1, \dots, k-1$; if π_s is the empty word, then $c_1 + \cdots + c_s = c_{s+1} + \cdots + c_k$.*

Proof. By the remarks following Lemma 6, taking into account Lemma 13, the hypothesis implies that $\sigma(m_s) = m_{s+1}$, hence we conclude using Lemma 5. □

Lemma 15. *Let u be a perfectly clustering Lyndon word, with a special palindromic factorization (2). Suppose that $\rho_\ell(u)$ is a positive word, longer than u. Let i be maximum such that $a_i \leqslant \ell$. Then π_j, $j = i, \dots, k-1$ are nonempty.*

Proof. Since $\rho_\ell(u) \in A^*$, we know by Lemma 3 that: (*) in u, each letter $> \ell$ is preceded by some letter $\leqslant \ell$.

Suppose first that $j = i+1, \dots, k-1$. Then $a_j, a_{j+1} > \ell$. In u, a_{j+1} is preceded by $a_j\pi_j$. Hence, by (*), π_j cannot be empty.

Suppose now that $j = i$. We have two cases: $a_i = \ell$, $a_i < \ell$. By Corollary 2, we have $\sum_{a<\ell} |u|_a > \sum_{a>\ell} |u|_a$. Suppose by contradiction that π_i is the empty word; then by Lemma 14, we have $|u|_{a_1} + \cdots + |u|_{a_i} = |u|_{a_{i+1}} + \cdots + |u|_{a_k}$.

Consider the first case: $a_i = \ell$. Then the former inequality becomes $|u|_{a_1} + \cdots + |u|_{a_{i-1}} > |u|_{a_{i+1}} + \cdots + |u|_{a_k}$, contradicting the equality since $|u|_{a_i} > 0$. Consider the second case: $a_i < \ell$, hence $a_{i+1} > \ell$. The inequality becomes $|u|_{a_1} + \cdots + |u|_{a_i} > |u|_{a_{i+1}} + \cdots + |u|_{a_k}$, contradicting the equality too. □

5 Proof of Theorem 7

(i) ⇒ (ii). It is known that w is conjugate to its reversal $\tilde{w}$ (see [22, Corollary 4.4]). If w is of length at most 2, then (ii) holds clearly. Suppose now that w is of length at least 3. Then by Theorem 1, $w = f(u)$ for some shorter perfectly clustering word u and some automorphism $f = \lambda_\ell$ or ρ_ℓ, $\ell \in A$, with the further property: either a) $A = Alph(w) = Alph(u)$, or b) $A = Alph(w) = Alph(u) \cup \ell$ and $\ell \notin Alph(u)$.

By induction, we may assume that $u = a_1\pi_1a_2\pi_2 \cdots \pi_{k-1}a_k$, where $Alph(u) = \{a_1 < a_2 < \cdots < a_k\}$ and where the word π_i are palindromes. We may assume that $f = \rho_\ell$, the other case being similar.

Write ρ for ρ_ℓ. Since $\rho(u) \in A^*$, we know by Lemma 3 that: (*) in u, each letter $> \ell$ is preceded by some letter $\leqslant \ell$.

In case a), let $\ell = a_i$. We know that $\pi_i, \dots, \pi_{k-1}$ are nonempty, by Lemma 15.

We have $u = \left(\prod_{1\leqslant j\leqslant i-1} a_j\pi_j\right) a_i \left(\prod_{i\leqslant j\leqslant k-1} \pi_j a_{j+1}\right)$. Hence

$$w = \left(\prod_{1\leqslant j\leqslant i-1} a_j a_i \rho(\pi_j)\right) a_i \left(\prod_{i\leqslant j\leqslant k-1} \rho(\pi_j) a_i^{-1} a_{j+1}\right).$$

Since π_j is nonempty for $j = i, \dots, k-1$, and is followed by a_{j+1} in u, it follows from Lemma 9 that $\rho(\pi_j)a_i^{-1}$ is a positive palindrome. For $j = 1, \dots, i-1$, $\rho(\ell\pi_j) = \rho(a_i\pi_j) = a_i\rho(\pi_j)$ is a positive palindrome, by (*) (which implies that in $\ell\pi_j$, each letter $> \ell$ is preceded by a letter $\leqslant \ell$), and Lemma 8. This concludes case a).

In case b), we have for some i, $a_1 < \cdots < a_i < \ell < a_{i+1} < \ldots < a_k$.

We know that $\pi_{i+1}, \dots, \pi_k$ are nonempty, by Lemma 15.

Note that $i \geqslant 1$, since otherwise $\ell < a_1$, and since a_1 is not preceded in u by any letter, this contradicts (*).

We have therefore

$$u = a_1 \prod_{1\leqslant j\leqslant k-1} \pi_j a_{j+1} = a_1 \left(\prod_{1\leqslant j\leqslant i-1} \pi_j a_{j+1}\right) \left(\prod_{i\leqslant j\leqslant k-1} \pi_j a_{j+1}\right),$$

hence

$$\begin{aligned} w &= a_1\ell\left(\prod_{1\leqslant j\leqslant i-1}\rho(\pi_j)a_{j+1}\ell\right)\left(\prod_{i\leqslant j\leqslant k-1}\rho(\pi_j)\ell^{-1}a_{j+1}\right) \\ &= a_1\left(\prod_{1\leqslant j\leqslant i-1}\ell\rho(\pi_j)a_{j+1}\right)\ell\left(\prod_{i\leqslant j\leqslant k-1}\rho(\pi_j)\ell^{-1}a_{j+1}\right) \\ &= a_1q_1\cdots a_{i-1}q_{i-1}a_iq_i\ell q_{i+1}\cdots a_{k-1}q_ka_k, \end{aligned}$$

with $q_1 = \ell\rho(\pi_1), \ldots, q_{i-1} = \ell\rho(\pi_{i-1}), q_i = 1, q_{i+1} = \rho(\pi_i)\ell^{-1}, \ldots, q_k = \rho(\pi_{k-1})\ell^{-1}$. The elements q_j are all positive palindromes: this is proved as in case a).

(ii) $\Rightarrow$ (iii). By hypothesis w has a palindromic special factorization (2) and w is conjugate to $\tilde{w}$, since w is a product of two palindromes. Since the π_i are palindromes, $\tilde{w} = W$. Thus (iii) holds.

(iii) $\Rightarrow$ (i) Let $\mathcal{C}$ denote the set of conjugates of w. By hypothesis, $W \in \mathcal{C}$. Define a mapping $\nu : \mathcal{C} \setminus W \to \mathcal{C}$ by

$$\nu(va_i\cdots\pi_1a_1a_k\pi_{k-1}\cdots a_{i+1}u) = va_{i+1}\cdots\pi_{k-1}a_ka_1\pi_1\cdots a_iu, \tag{4}$$

where $\pi_i = uv$. This mapping is the next element function, by Lemma 12, W is the largest element of $\mathcal{C}$, and w its smallest, and in particular a Lyndon word. Thus $\mathcal{C} = \{w < \nu(w) < \nu^2(w) < \cdots < \nu^{n-1}(w)\}$ with n the length of w.

Now, by Equation (4), we see that the last letter of $x = va_i\cdots\pi_1a_1a_k\pi_{k-1}\cdots a_{i+1}u$ and that of $\nu(x)$ are equal, except if u is the empty word. This can happen exactly for $k-1$ elements x of $C \setminus W$, namely those of the form $x_i = \pi_ia_i\cdots\pi_1a_1a_k\pi_{k-1}\cdots a_{i+1}$ ($u = 1$, $v = \pi_i$). The last letter of x_i is a_{i+1} and that of $\nu(x_i) = \pi_ia_{i+1}\cdots\pi_{k-1}a_ka_1\pi_1\cdots a_i$ is a_i. Since the last letter of w is a_k, this implies that the last letters of the words $w, \nu(w), \nu^2(w), \cdots, \nu^{n-1}(w)$ form a weakly decreasing sequence. Thus the last column of the Burrows-Wheeler matrix of w is weakly decreasing, and w is therefore perfectly clustering.

6 Consequences

From Lemma 13, we may deduce the uniqueness of the special palindromic factorization of a perfectly clustering Lyndon word, whose existence is asserted in Theorem 7.

Corollary 16. *The factorization (2) of a perfectly clustering Lyndon word w is unique. Precisely: for any i, $a_i\pi_i\cdots\pi_{k-1}a_k$ is the smallest suffix of w (in lexicographic order) beginning by a_i, and $a_i\pi_i\cdots\pi_{k-1}a_ka_1\pi_1\cdots a_{i-1}\pi_{i-1}$ is the smallest conjugate of w (in lexicographic order) beginning by a_i.*

Proof. The fact that the two last assertions are equivalent follows for example from [7, Proposition 2.1]. □

a	a	a	b	a	a	b
a	a	b	a	a	a	b
a	a	b	a	a	b	a
a	b	a	a	a	b	a
a	b	a	a	b	a	a
b	a	a	a	b	a	a
b	a	a	b	a	a	a

Figure 3: Burrows-Wheeler matrix of the Christoffel word $aaabaab$

We study now the transition from a row to the next row, in the Burrows-Wheeler matrix of a perfectly clustering word. We recall first what happens in the case of two letters.

Consider an alphabet $A = \{a < b\}$ with two-letters. A perfectly clustering Lyndon word w on A^* is a lower Christoffel word, by a theorem of Mantaci, Restivo and Sciortino [15, Theorem 9]. Thus the last column of the Burrows-Wheeler matrix of w is of the form $b, \dots, b, a, \dots, a$. A further striking property is that two consecutive rows of this matrix, viewed as two words u, v, are always of the form $u = yabx, v = ybax$, with moreover $xy = p$, where p is the palindrome such that $w = apb$; see [1, Corollary 5.1] (see also [19, Theorem 2], or [21, Theorem 15.2.4]). For an example, see Figure 3, where $w = aaabaab$, $p = aabaa$ and for example the rows 3 and 4: $u = a\mathbf{ab}aaba, v = a\mathbf{ba}aaba$, $x = aaba, y = a$.

We may generalize this result as follows.

Theorem 17. *Let w be a perfectly clustering word. Then for any two consecutive rows of its Burrows-Wheeler matrix, viewed as two words w', w'', one has $w' = ymx, w'' = y\tilde{m}x$, for some words x, y, m such that xy is one of the palindromes of the special palindromic factorization (2) of w.*

See Figure 2 for an example: the alphabet is $a < b < c$, $w = \mathbf{a}cacac\mathbf{b}bb\mathbf{c}$ (special palindromic factorization, with palindromes $cacac$ and bb); rows 1 and 2 are $acac\mathbf{acbbb}c$, $acac\mathbf{bbbca}c$ with the reversed factor in bold, and corresponding palindrome $cacac$; rows 4 and 5 are $bb\mathbf{bcacacac}, bb\mathbf{cacacacb}$ and palindrome bb.

Proof. By Lemma 12, $w' = va_i \cdots \pi_1 a_1 a_k \pi_{k-1} \cdots a_{i+1} u$ and

$$w'' = va_{i+1} \cdots \pi_{k-1} a_k a_1 \pi_1 \cdots a_i u,$$

for some $i = 1, \dots, k-1$, with $\pi_i = uv$. Then the theorem follows by taking $y = v, x = u, m = a_i \cdots \pi_1 a_1 a_k \pi_{k-1} \cdots a_{i+1}$, so that $\tilde{m} = a_{i+1} \cdots \pi_{k-1} a_k a_1 \pi_1 \cdots a_i$, since the π_j are palindromes. □

Since the arguments in the proof of Lemma 14 may be reversed, we obtain the following corollary.

Corollary 18. *Let w be a perfectly clustering word, with its special palindromic factorization (2). Then π_s is empty if and only if $|w|_{a_1} + \cdots + |w|_{a_s} = |w|_{a_{s+1}} + \cdots + |w|_{a_k}$.*

This corollary is a first result on the palindromes appearing in factorization (2). We intend to study these palindromes elsewhere. Note that for a two-letter alphabet, these palindromes, called *central words*, where first studied by Aldo de Luca. He gave, among others, the following beautiful characterization: the central words are the image of the mapping called *iterated palindromization* (see [13], [21, Chapter 12]).

Acknowledgements

The authors thank Antonio Restivo for useful mail exchanges. This work was partially supported by NSERC, Canada. We are thankful to the anonymous referees for their valuable comments.

References

[1] J.-P. Borel, C. Reutenauer, On Christoffel classes, *RAIRO Informatique théorique et appliquée*, **40**, 2006, 15–27.

[2] S. Brlek, S. Hamel, M. Nivat and C. Reutenauer, On the palindromic complexity of infinite words, *International Journal of Foundations of Computer Science*, **15**, 2004, 2, 293–306.

[3] M. Burrows, D.J. Wheeler, A Block-sorting lossless data compression algorithm, Systems Research Center research report, 1994, 24 pages.

[4] M.G. Castelli, F. Mignosi, A. Restivo, Fine and Wilf's theorem for three periods and a generalization of Sturmian words, *Theoretical Computer Science*, **218**, 1999, 1, 83–94.

[5] S. Ferenczi, L. Q. Zamboni, Clustering words and interval exchanges, *Journal of Integer Sequences* **16**, 2013, 2, Article 13.2.1, 9 pages.

[6] I. Gessel, C. Reutenauer, Counting permutations with given cycle structure and descent set, *Journal of Combinatorial Theory. Serie A* **64**, 1993 189–215.

[7] C. Hohlweg, C. Reutenauer, Lyndon words, permutations and trees, *Theoretical Computer Science* **307**, 2003, 173–178.

[8] J. Justin, Episturmian morphisms and a Galois theorem on continued fractions, *RAIRO Informatique théorique et applications* **39**, 2005, 207–215.

[9] S. Labbé, C. Reutenauer, A d-dimensional extension of Christoffel words, *Discrete & Computational Geometry. An International Journal of Mathematics and Computer Science* **54**, 2015, 1, 152–181.

[10] M. Lapointe, Combinatoire des mots : mots parfaitement amassants, triplets de Markoff et graphes chenilles, Thèse de doctorat, Université du Québec àMontréal, 2020.

[11] M. Lapointe, Perfectly clustering words are primitive positive elements of the free group, *Lecture Notes in Computer Science* **12847**, 2021, 117–128.

[12] M. Lothaire, Combinatorics on words, Cambridge University Press, 2002.

[13] A. de Luca: Sturmian words: structure, combinatorics, and their arithmetics, Theoretical Computer Science 183, 1997, 45–82.

[14] A. de Luca, F. Mignosi, Some combinatorial properties of Sturmian words, *Theoretical Computer Science* **136**, 1994, 361–385.

[15] S. Mantaci, A. Restivo, M. Sciortino, Burrows-Wheeler transform and Sturmian words, *Information Processing Letters* **86**, 2001, 241–246.

[16] G. Melançon, C. Reutenauer, On a class of Lyndon words extending Christoffel words and related to a multidimensional continued fraction algorithm, *Journal of Integer Sequences* **16**, 2023, 9, Article 13.9.7, 30 pages.

[17] V. Oseledets, On the spectrum of ergodic automorphisms, Doklady Akademii Nauk. SSSR **168**, 10091011, 1966. In Russian. English translation in Soviet Math. Doklady **7**, 776-779, 1966.

[18] G. Paquin, On a generalization of Christoffel words: epichristoffel words, *Theoretical Computer Science* **410**, 2009, 38-40, 3782–3791.

[19] D. Perrin, A. Restivo, A note on Sturmian words, *Theoretical Computer Science* **429**, 2012, 265-272.

[20] G. Pirillo: A new characteristic property of the palindrome prefixes of a standard Sturmian word, *Séminaire Lotharingien de Combinatoire* **43**, 1999, B43f.

[21] C. Reutenauer, From Christoffel words to Markoff numbers, Oxford University Press, 2019.

[22] J. Simpson, S.J. Puglisi, Words with simple Burrows-Wheeler Transforms, *The Electronic Journal of Combinatorics* **15**, 2008, #R83, 17 pp.