On Christoffel and standard words and their derivatives

Alma D'Aniello[a], Aldo de Luca[a,1], Alessandro De Luca[b,2,*]

[a] *Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi di Napoli Federico II*
[b] *DIETI, Università degli Studi di Napoli Federico II*

Abstract

We introduce and study natural derivatives for Christoffel and finite standard words, as well as for characteristic Sturmian words. These derivatives, which are realized as inverse images under suitable morphisms, preserve the aforementioned classes of words. In the case of Christoffel words, the morphisms involved map a to $a^{k+1}b$ (resp., ab^k) and b to $a^k b$ (resp., ab^{k+1}) for a suitable $k > 0$. As long as derivatives are longer than one letter, higher-order derivatives are naturally obtained. We define the depth of a Christoffel or standard word as the smallest order for which the derivative is a single letter. We give several combinatorial and arithmetic descriptions of the depth, and (tight) lower and upper bounds for it.

Keywords: Christoffel word, Standard word, Central word, Characteristic word, Derivative of a word
2010 MSC: 68R15

1. Introduction

Since the first systematic study by M. Morse and G. A. Hedlund [28], Sturmian words have been among the most studied infinite words in combinatorics as they are the simplest aperiodic words in terms of *factor complexity*, and enjoy many beautiful characterizations and properties (see, for instance, [27, Chap. 2]).

Sturmian words are of interest in several fields of mathematics such as combinatorics, algebra, number theory, dynamical systems, and differential equations. They are also of great importance in theoretical physics as basic examples of 1-dimensional quasicrystals (cf. [12] and references therein) and in computer science where they are used in computer graphics as digital approximation of straight lines (cf. [25]).

A basic tool in the study of Sturmian words is the *palindromization map* ψ, first introduced by the second author [14]. It maps any finite binary word v (called *directive word* in this context) to a palindrome $\psi(v)$ called *central word*. The definition can be naturally extended to infinite directive words; when v spans among all binary words where both letters occur infinitely often, $\psi(v)$ gives exactly all *characteristic* Sturmian words (or infinite standard Sturmian words). An infinite word is Sturmian if it has the same set of factors as some characteristic Sturmian word.

Central words are thus all palindromic prefixes of characteristic Sturmian words; they can also be defined in a purely combinatorial way, as words having two coprime periods p, q and length $p + q - 2$. If w is a central word over the alphabet $\{a, b\}$, then awb is a (lower) *Christoffel word* and wab, wba are *standard words*. These classes of words, which also include the letters a and b, represent a finite counterpart to Sturmian words and are well studied in their own right as they satisfy remarkable and surprising combinatorial properties (see for instance [4, 17, 27]).

In a previous paper [16] the second and third author have studied an important connection between the combinatorics of these words and the famous Stern sequence. In this paper, which can be considered as a continuation of the previous one, we consider new combinatorial properties which are mainly related to the notion of derivative of a word.

*Corresponding author.
Email addresses: alma.daniello@unina.it (Alma D'Aniello), aldo.deluca@unina.it (Aldo de Luca), alessandro.deluca@unina.it (Alessandro De Luca)

[1] Partially supported by GNSAGA, INdAM.
[2] Partially supported by the Italian Ministry of Education (MIUR), under the PRIN project number 2010LYA9RH, titled "Automi e Linguaggi Formali: Aspetti Matematici e Applicativi".

Word derivation, meant as inverse image under some injective morphism (also called "desubstitution", or "inflation" in [11]), is a known topic in combinatorics on words. A well-known instance is the notion of *derivated word* of a recurrent word, introduced by F. Durand [22] along with the important concept of *return words*.

The main objective of this paper is to study some natural derivatives for noteworthy classes of finite Sturmian words, such as Christoffel and standard words. The paper is organized as follows. In Section 2.1 we consider the palindromization map. A well-known result by J. Justin [24], known as *Justin's formula*, links the palindromization map with *pure standard Sturmian morphisms*, i.e., morphisms of the monoid $\{\mu_a, \mu_b\}^*$ where for $x \in \{a, b\}$, and $y \neq x$, μ_x is defined as follows: $\mu_x : x \mapsto x,\ y \mapsto xy$. Setting for $v = v_1 \cdots v_n$, $\mu_v = \mu_{v_1} \circ \cdots \circ \mu_{v_n}$, one derives from Justin's formula, that every standard word $\psi(v)xy$ with $\{x, y\} = \{a, b\}$ is obtained as the image of xy under the morphism μ_v,

$$\psi(v)xy = \mu_v(xy)\,. \tag{1}$$

In Section 2.2, some basic relations existing between central, standard, and Christoffel words are recalled and new combinatorial properties are proved.

In Section 3, we discuss *Christoffel morphisms*, i.e., morphisms preserving Christoffel words. We provide a simple combinatorial proof for the known fact [4] that the monoid of Christoffel morphisms is generated by λ_a and λ_b, defined by

$$\lambda_a = \mu_a \ \text{ and } \ \lambda_b : a \mapsto ab,\ b \mapsto b.$$

Setting $\lambda_v = \lambda_{v_1} \circ \cdots \circ \lambda_{v_n}$ for $v = v_1 \cdots v_n$, this gives an analogue of formula (1) in the case of Christoffel words, namely

$$a\psi(v)b = \lambda_v(ab).$$

We also prove that the *inverse* image of a Christoffel word under a Christoffel morphism is a Christoffel word; again, this mirrors a well-known result for standard words and morphisms.

With such knowledge about Christoffel morphisms, in Section 4 we define a derivative for proper Christoffel words. In fact, for each such word w there exists some nonnegative integer k (the *index* of w) such that w can be uniquely factored over $X_k = \{a^k b, a^{k+1} b\}$ or $Y_k = \{ab^k, ab^{k+1}\}$; hence, w is the image, under the morphism $\varphi_k = \lambda_{a^k b}$ or $\hat{\varphi}_k = \lambda_{b^k a}$, of a word ∂w that we call the *derivative* of w. Since φ_k and $\hat{\varphi}_k$ are Christoffel morphisms, this derivative is still a Christoffel word.

Our choice of morphisms φ_k and $\hat{\varphi}_k$ for the definition is motivated by the following arguments. First, the factorization over X_k or Y_k is quite natural and has been used in well-known algorithms for recognizing factors of Sturmian words (or *digital straight segments*, in the computer graphics terminology; cf. [25]). Second, if $w = a\psi(v)b$ and v is not a power of a letter, then

$$\partial w = a\psi({}_+v)b,$$

where ${}_+v$ is the longest suffix of v immediately preceded by a letter different from the first letter of v. The operator $v \mapsto {}_+v$ was introduced by the last two authors in [16] and appears in some interesting results on Christoffel words; for instance, if v starts with the letter x and $\{x, y\} = \{a, b\}$, then the length $|a\psi({}_+v)b| = |\partial w|$ equals the number of occurrences of y in $a\psi(v)b$. Finally, a Christoffel word is determined by its derivative and the value of its index.

Further results on the derivatives of Christoffel words are proved. In particular, if a Christoffel word w is factored as $w = w_1 w_2$ with w_1 and w_2 proper Christoffel words, then $\partial w = \partial w_1 \partial w_2$. Moreover, the length of a Christoffel word $w = a\psi(v_1 v_2 \cdots v_n)b$ with $v_i \in \mathcal{A}$, $1 \leq i \leq n$, is equal to 2 plus the sum of the lengths of derivatives $\partial a\psi(v_i \cdots v_n)b$, $i = 1, \ldots, n$.

In Section 5, we naturally define higher order derivatives, by letting $\partial^{i+1} w = \partial(\partial^i w)$ whenever $\partial^i w$ is still a proper Christoffel word (i.e., not just a letter). The *depth* of a Christoffel word w is then the smallest $i \geq 0$ such that $\partial^i w$ is a letter. We give several descriptions of the depth of $a\psi(v)b$ as a function $\delta(v)$ of its directive word. We prove that $\delta(uv)$ equals either $\delta(u) + \delta(v)$ or $\delta(u) + \delta(v) - 1$. Tight lower and upper bounds of the depth are given; moreover, we characterize the directive words for which such bounds are attained. We give also a closed formula for the number $J_k(p)$ of the words v of length k such that $\delta(v) = p$.

In Section 6 we consider finite and infinite standard Sturmian words; using the standard morphisms $\mu_{a^k b}$ and $\mu_{b^k a}$ we define a natural derivative in these cases. This allows us to extend the previous results to standard words; in particular, the derivative of the standard word $\psi(v)xy$ with $\{x, y\} = \{a, b\}$ is either a letter or the proper standard word $\psi({}_+v)xy$, where ${}_+v$ is the same directive word found in the derivative of the Christoffel word $a\psi(v)b$. Hence, the depths of $\psi(v)ab$, $\psi(v)ba$, and $a\psi(v)b$ coincide. In the infinite case, the derivative Ds of a characteristic Sturmian word s is word isomorphic to a derivated word in the sense of Durand. We give a proof for the fact that a characteristic Sturmian word has only finitely many distinct higher order derivatives if and only if its directive word is ultimately periodic (see also [2]). Finally, we prove that there exists a simple relation between the derivative Ds of a characteristic word s and the derivative ∂s, namely $\partial s = bDs$.

2. Notation and Preliminaries

In the following, A will denote a finite non-empty set, or *alphabet* and A^* the *free monoid* generated by A. The elements of A are usually called *letters* and those of A^* *words*. The identity element of A^* is called *empty word* and denoted by ε. We set $A^+ = A^* \setminus \{\varepsilon\}$.

A word $w \in A^+$ can be written uniquely as a sequence of letters as $w = w_1 w_2 \cdots w_n$, with $w_i \in A$, $1 \le i \le n$, $n > 0$. The integer n is called the *length* of w and denoted $|w|$. The length of ε is 0. For any $w \in A^*$ and $x \in A$, $|w|_x$ denotes the number of occurrences of the letter x in w. For any word $v \in A^+$, we let $v^{(F)}$ (resp., $v^{(L)}$) denote the first (resp., last) letter of v.

Let $w \in A^*$. The word u is a *factor* of w if there exist words r and s such that $w = rus$. A factor u of w is called *proper* if $u \neq w$. If $w = us$, for some word s (resp., $w = ru$, for some word r), then u is called a *prefix* (resp., a *suffix*) of w. If u is a prefix of w, then $u^{-1}w$ denotes the word v such that $uv = w$.

Let p be a positive integer. A word $w = w_1 \cdots w_n$, $w_i \in A$, $1 \le i \le n$, has *period* p if the following condition is satisfied: for any integers i and j such that $1 \le i, j \le n$,

$$\text{if } i \equiv j \pmod{p}, \text{ then } w_i = w_j.$$

Let us observe that if a word w has a period p, then any non-empty factor of w has also the period p.

We let $\pi(w)$ denote the minimal period of w. Conventionally, we set $\pi(\varepsilon) = 1$. A word w is said to be *constant* if $\pi(w) = 1$, i.e., $w = z^k$ with $k \ge 0$ and $z \in A$. Two words v and w are *conjugate* if there exist words r and s such that $v = rs$ and $w = sr$.

Let $w = w_1 \cdots w_n$, $w_i \in A$, $1 \le i \le n$. The *reversal* of w is the word $w^\sim = w_n \cdots w_1$. One defines also $\varepsilon^\sim = \varepsilon$. A word is called *palindrome* if it is equal to its reversal. We let PAL denote the set of all palindromes on the alphabet A.

In the following, we let the alphabet A be totally ordered. We let $<_{lex}$ denote the lexicographic order induced on A^*. A word is called a *Lyndon word* if it is lexicographically less than any of its proper suffixes (cf. [26, Chap. 5]). As is well-known a Lyndon word $w \notin A$ can be factored (*standard factorization*) as $w = lm$ where l is a Lyndon word and m is the longest suffix of w which is a Lyndon word.

A right-infinite word x, or simply *infinite word*, over the alphabet A is just an infinite sequence of letters:

$$x = x_1 x_2 \cdots x_n \cdots \text{ where } x_i \in A, \text{ for all } i \ge 1\,.$$

For any integer $n \ge 0$, we let $x_{[n]}$ denote the prefix $x_1 x_2 \cdots x_n$ of x of length n. A factor of x is either the empty word or any sequence $x_i \cdots x_j$ with $i \le j$. The set of all infinite words over A is denoted by A^ω. An infinite word x is called *ultimately periodic* if there exist words $u \in A^*$ and $v \in A^+$ such that $x = uv^\omega$. The word x is called (purely) *periodic* if $u = \varepsilon$, i.e., $x = v \cdot v \cdot v \cdots$. A periodic word with $v \in A$ will be called *constant*. The word x is called *aperiodic* if it is not ultimately periodic.

We say that two finite or infinite words $x = x_1 x_2 \cdots$ and $y = y_1 y_2 \cdots$ on the alphabets A and A' respectively are *word isomorphic*, or simply *isomorphic*, if there exists a bijection $\phi : A \to A'$ such that $y = \phi(x_1)\phi(x_2)\cdots$.

We set $A^\infty = A^* \cup A^\omega$. For any $w \in A^\infty$ we let Fact(w) denote the set of all distinct factors of the word w.

In the following, we shall mainly concern with two-letter alphabets. We let $\mathcal{A}$ denote the alphabet whose elements are the letters a and b, totally ordered by setting $a < b$.

We let E denote the automorphism of $\mathcal{A}^*$ defined by $E(a) = b$ and $E(b) = a$. For each $w \in \mathcal{A}^\infty$, the word $E(w)$ is called the *complementary* word, or simply the *complement* of w. We shall often use for $E(w)$ the simpler notation $\bar{w}$.

We say that a word $v \in \mathcal{A}^k$, $k \ge 0$, is *alternating* if for $x, y \in \mathcal{A}$ and $x \neq y$, $v = (xy)^{\frac{k}{2}}$ if k is even and $v = (xy)^{\lfloor \frac{k}{2} \rfloor}x$ if k is odd, i.e., v is a single letter or if $|v| > 1$ any non-terminal letter in v is immediately followed by its complementary.

The *slope* $\eta(w)$ of a word $w \in \mathcal{A}^+$ is the fraction $\eta(w) = \frac{|w|_b}{|w|_a}$ if $|w|_a > 0$. We set $\eta(w) = \infty$ if $|w|_a = 0$.

If we identify the letters a and b of $\mathcal{A}$ respectively with the digits 0 and 1, for each $w \in \mathcal{A}^*$ we let $\langle w \rangle_2$, or simply $\langle w \rangle$, denote the *standard interpretation* of w as an integer at base 2. For instance, $\langle a \rangle = 0$, $\langle b \rangle = 1$, $\langle babba \rangle = 22$.

We represent a non-empty binary word $v \in \mathcal{A}^+$ as

$$v = x_0^{\alpha_0} \cdots x_n^{\alpha_n},$$

where $\alpha_i \geq 1$, $x_i \in \mathcal{A}$, $0 \leq i \leq n$, and $x_{i+1} = \bar{x_i}$ for $0 \leq i \leq n-1$. We call the list $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ the *integral representation* of the word v. Hence, the integral representation of a word v and its first letter $v^{(F)}$determine uniquely v. We set $\mathrm{ext}(v) = |x_0 \cdots x_n| = n + 1$ and call it *extension* of v. Moreover, we define $\mathrm{ext}(\varepsilon) = 0$.

For all definitions and notation concerning words not explicitly given in the paper, the reader is referred to the book of Lothaire [26]; for Sturmian words see [27, Chap. 2] and [1, Chap.s 9-10].

2.1. The palindromization map

We consider in A^* the operator $^{(+)} : A^* \to \mathrm{PAL}$ which maps any word $w \in A^*$ into the palindrome $w^{(+)}$ defined as the shortest palindrome having the prefix w (cf. [14]). The word $w^{(+)}$ is called the *right palindromic closure* of w. If Q is the longest palindromic suffix of $w = vQ$, then one has

$$w^{(+)} = vQv^{\sim} .$$

Let us now define the map

$$\psi : A^* \to \mathrm{PAL},$$

called *right iterated palindromic closure*, or simply *palindromization map*, over A^*, as follows: $\psi(\varepsilon) = \varepsilon$ and for all $u \in A^*$, $x \in A$,

$$\psi(ux) = (\psi(u)x)^{(+)} .$$

For instance, if $u = aaba$, one has $\psi(a) = a$, $\psi(aa) = (\psi(a)a)^{(+)} = aa$, $\psi(aab) = (aab)^{(+)} = aabaa$, and $\psi(u) = \psi(aaba) = aabaaabaa$.

The following proposition collects some basic properties of the palindromization map (cf., for instance, [14, 21]):

Proposition 2.1. *The palindromization map ψ satisfies the following properties:*

P1. *The palindromization map is injective.*
P2. *If u is a prefix of v, then $\psi(u)$ is a palindromic prefix (and suffix) of $\psi(v)$.*
P3. *If p is a prefix of $\psi(w)$, then $p^{(+)}$ is a prefix of $\psi(w)$.*
P4. *Every palindromic prefix of $\psi(v)$ is of the form $\psi(u)$ for some prefix u of v.*
P5. $|\psi(u^{\sim})| = |\psi(u)|$, *for any $u \in A^*$.*
P6. *The palindromization map ψ over $\{a, b\}^*$ commutes with the automorphism E, i.e., $\psi \circ E = E \circ \psi$.*

For any $w \in \psi(A^*)$ the unique word u such that $\psi(u) = w$ is called the *directive word* of w.

One can extend ψ to A^∞ defining ψ on A^ω as follows: let $x \in A^\omega$ be an infinite word

$$x = x_1 x_2 \cdots x_n \cdots , \quad x_i \in A,\ i \geq 1.$$

Since by property P2 of Proposition 2.1 for all n, $\psi(x_{[n]})$ is a prefix of $\psi(x_{[n+1]})$, we can define the infinite word $\psi(x)$ as:

$$\psi(x) = \lim_{n\to\infty} \psi(x_{[n]}).$$

The map $\psi : A^\omega \to A^\omega$ is injective. The word x is called the *directive word* of $\psi(x)$. It has been proved in [14] that if $x \in \{a, b\}^\omega$ the word $\psi(x)$ is a *characteristic Sturmian word* (or infinite standard Sturmian word) if and only if both the letters a and b occur infinitely often in the directive word x.

Example 2.2. Let $\mathcal{A} = \{a, b\}$. If $x = (ab)^\omega$, then the characteristic Sturmian word $f = \psi((ab)^\omega)$ having the directive word x is the famous *Fibonacci word*

$$f = abaababaabaab \cdots .$$

If $A = \{a, b, c\}$ the word $t = \psi((abc)^\omega)$ is the so-called *Tribonacci word*:

$$t = abacabaabacaba \cdots .$$

For any $x \in A$, we let μ_x denote the injective endomorphism of A^* defined by

$$\mu_x(x) = x,\ \mu_x(y) = xy,\ \text{for } y \in A \setminus \{x\}. \tag{2}$$

If $v = x_1x_2 \cdots x_n$, with $x_i \in A$, $i = 1, \ldots, n$, then we set:

$$\mu_v = \mu_{x_1} \circ \cdots \circ \mu_{x_n};$$

moreover, if $v = \varepsilon$, $\mu_\varepsilon = \mathrm{id}$. The following interesting theorem, due to Justin [24] and usually referred to as *Justin's formula*, relates the palindromization map to the morphisms μ_v.

Theorem 2.3. *For all $v, u \in A^*$*

$$\psi(vu) = \mu_v(\psi(u))\psi(v).$$

An important consequence of Justin's formula is the following lemma [8], which will be useful in the following.

Lemma 2.4. *For each $w \in A^*$ and $v \in A^\omega$, $\psi(wv) = \mu_w(\psi(v))$.*

For instance, if we take $w = a$, $x = (ab)^\omega$, then as one easily verifies

$$\psi(a(ab)^\omega) = \mu_a(f) = aabaaabaabaaab \cdots .$$

The case of a binary alphabet $\mathcal{A} = \{a, b\}$ deserves a special consideration. The following remarkable proposition holds (see, for instance [15, Prop. 4.10]).

Proposition 2.5. *For any $v \in \mathcal{A}^*$ and $x, y \in \mathcal{A}$, $x \neq y$,*

$$\mu_v(xy) = \psi(v)xy.$$

Corollary 2.6. *For any $w, v \in \mathcal{A}^*$ and $x, y \in \mathcal{A}$, $x \neq y$,*

$$\psi(wv)xy = \mu_w(\psi(v)xy).$$

Proof. By the preceding proposition one has:

$$\psi(wv)xy = \mu_{wv}(xy) = \mu_w(\mu_v(xy)) = \mu_w(\psi(v)xy).$$ □

Let v be a non-empty word. We let v^- (resp., ${}^-v$) denote the word obtained from v by deleting the last (resp., first) letter. If v is not constant, we let v_+ (resp., ${}_+v$) denote the longest prefix (resp., suffix) of v which is immediately followed (resp., preceded) by the complementary of the last (resp., first) letter of v. For instance, if $v = abbabab$, one has $v^- = abbaba$, $v_+ = abbab$, ${}^-v = bbabab$, and ${}_+v = babab$. From the definition one has

$${}_+(E(v)) = E({}_+v),\ (E(v))_+ = E(v_+),\ ({}_+v)^\sim = (v^\sim)_+\ . \tag{3}$$

As shown in [16], and as we shall see in some details in the next sections, the words v^-, v_+ and ${}_+v$, ${}^-v$ play an essential role in the combinatorics of Christoffel words.

Proposition 2.7. *Let $v \in \mathcal{A}^*$ be non-constant. Then*

$$\mu_v(a) = \mu_{v_+}(ba) = \psi(v_+)ba,\ \mu_v(b) = \mu_{v^-}(ab) = \psi(v^-)ab,\ \text{if } v^{(L)} = a$$

and

$$\mu_v(a) = \mu_{v^-}(ba) = \psi(v^-)ba,\ \mu_v(b) = \mu_{v_+}(ab) = \psi(v_+)ab,\ \text{if } v^{(L)} = b.$$

Proof. We shall prove the result only when $v^{(L)} = a$. The case $v^{(L)} = b$ is similarly dealt with. We can write $v = v_+ba^r$ for a suitable $r > 0$. Therefore, by Proposition 2.5 one has:

$$\mu_v(a) = \mu_{v_+ba^r}(a) = \mu_{v_+b}(a) = \mu_{v_+}(ba) = \psi(v_+)ba, \tag{4}$$

and

$$\mu_v(b) = \mu_{v^-a}(b) = \mu_{v^-}(ab) = \psi(v^-)ab, \tag{5}$$

which proves the assertion. □

Example 2.8. Let $v = abbab$. One has $v_+ = abb, v^- = abba$, and $v^{(L)} = b$. Hence, $\psi(v_+) = ababa$, $\psi(v^-) = ababaababa$, $\mu_{abbab}(a) = ababaabababa = \psi(v^-)ba$, and $\mu_{abbab}(b) = ababaab = \psi(v_+)ab$.

An immediate consequence of Proposition 2.7 is the following (see also [16]):

Corollary 2.9. *Let $v \in \mathcal{A}^*$ be non-constant, x the last letter of v, and $y = \bar{x}$. Then*

$$\psi(v) = \psi(v_+)yx\psi(v^-) = \psi(v^-)xy\psi(v_+).$$

Proof. Since v is not constant, $\psi(v)xy = \mu_v(xy) = \mu_v(x)\mu_v(y)$, and by Proposition 2.7, $\mu_v(x) = \psi(v_+)yx$ and $\mu_v(y) = \psi(v^-)xy$. The result follows. □

2.2. Central, standard, and Christoffel words

In the study of combinatorial properties of Sturmian words a crucial role is played by the set PER of all finite words w having two periods p and q such that $\gcd(p, q) = 1$ and $|w| = p + q - 2$.

The set PER was introduced in [17] where its main properties were studied. It has been proved that PER is equal to the set of the palindromic prefixes of all standard Sturmian words, i.e.,

$$\mathrm{PER} = \psi(\mathcal{A}^*).$$

The words of PER have been called *central* in [27, Chap.2].

The following structural characterization of central words was proved in [14] (see, also [9]).

Proposition 2.10. *A word w is central if and only if w is a constant or it satisfies the equation:*

$$w = w_1abw_2 = w_2baw_1$$

with $w_1, w_2 \in \mathcal{A}^*$. Moreover, in this latter case, w_1 and w_2 are central words, $p = |w_1| + 2$ and $q = |w_2| + 2$ are coprime periods of w, and $\min\{p, q\}$ is the minimal period of w.

The following lemma, which will be useful in the following, is in [14].

Lemma 2.11. *For any $w \in \mathrm{PER}$, one has $(wa)^{(+)}, (wb)^{(+)} \in \mathrm{PER}$. More precisely, if $w = w_1abw_2 = w_2baw_1$, then*

$$(wa)^{(+)} = w_2baw_1abw_2\,, \quad (wb)^{(+)} = w_1abw_2baw_1\,.$$

If $w = x^n$ with $\{x, y\} = \mathcal{A}$, then $(wx)^{(+)} = x^{n+1}$ and $(wy)^{(+)} = x^nyx^n$.

Characteristic Sturmian words can be equivalently defined in the following way. Let $c_0, c_1, \ldots, c_n, \ldots$ be any sequence of integers such that $c_0 \geq 0$ and $c_i > 0$ for $i > 0$. We define, inductively, the sequence of words $(s_n)_{n\geq0}$, where

$$s_0 = b,\ s_1 = a,\ \text{and } s_{n+1} = s_n^{c_{n-1}}s_{n-1} \text{ for } n \geq 1\,.$$

The sequence $(s_n)_{n\geq0}$ converges to a limit s which is a characteristic Sturmian word (cf. [27]). Every characteristic Sturmian word is obtained in this way. The Fibonacci word is obtained when $c_i = 1$ for $i \geq 0$.

We let Stand denote the set of all the words s_n, $n \geq 0$ of any sequence $(s_n)_{n\geq0}$. Any element of Stand is called *standard Sturmian word*, or simply *standard word*. A standard word different from a single letter is called *proper*.

The following remarkable relation existing between standard and central words has been proved in [17]:

$$\mathrm{Stand} = \mathcal{A} \cup \mathrm{PER}\{ab, ba\}.$$

More precisely, the following holds (see, for instance [15, Prop. 4.9]):

Proposition 2.12. *Any proper standard word can be uniquely expressed as $\mu_v(xy)$ with $\{x, y\} = \{a, b\}$ and $v \in \mathcal{A}^*$.*

Hence, by Proposition 2.5 one has

$$\mu_v(xy) = \psi(v)xy.$$

Let us set for any $v \in \mathcal{A}^*$ and $x \in \mathcal{A}$,

$$p_x(v) = |\mu_v(x)|. \tag{6}$$

From Justin's formula one derives (cf. [18, Prop. 3.6]) that $p_x(v)$ is the minimal period of $\psi(vx)$ and then a period of $\psi(v)$. Moreover, one has (cf. [19, Lemma 5.1])

$$p_x(v) = \pi(\psi(vx)) = \pi(\psi(v)x) \tag{7}$$

and $\gcd(p_x(v), p_y(v)) = 1$, so that

$$\pi(\psi(v)) = \min\{p_x(v), p_y(v)\}.$$

Moreover, if v is not constant, as v_+ is a proper prefix of v^-, by Proposition 2.7 one derives:

$$\pi(\psi(v)) = p_{v^{(L)}}(v) = |a\psi(v_+)b|. \tag{8}$$

Since $|\mu_v(xy)| = |\mu_v(x)| + |\mu_v(y)|$, from Proposition 2.12 and (6) one has

$$|\psi(v)| = p_x(v) + p_y(v) - 2. \tag{9}$$

Let us now introduce the important notion of *Christoffel word* [10] (see also [6]). Let p and q be non-negative coprime integers, and $n = p + q > 0$. The (lower) Christoffel word w of slope $\frac{p}{q}$ is defined as $w = x_1 \cdots x_n$ with

$$x_i = \begin{cases} a & \text{if } ip \bmod n > (i-1)p \bmod n \\ b & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, n$, where $k \bmod n$ denotes the remainder of the Euclidean division of k by n. Observe that the words a and b are the Christoffel words with slope $\frac{0}{1}$ and $\infty = \frac{1}{0}$, respectively.

The Christoffel words of slope $\frac{p}{q}$ with p and q coprime positive integers are called *proper Christoffel words*. The term slope given to the fraction $\frac{p}{q}$ is due to the circumstance that one easily derives from the definition that $p = |w|_b$ and $q = |w|_a$.

We observe that lower Christoffel words have also an interesting geometric interpretation in terms of suitable paths in the integer lattice $\mathbb{N} \times \mathbb{N}$ (cf. [4]). It is then natural to introduce the so-called upper Christoffel words, which can also be defined similarly to lower Christoffel words, by interchanging a and b, as well as p and q, in the previous definition. We shall not consider these latter words in the paper, since they are simply the reversal of lower Christoffel words.

Example 2.13. Let $p = 3$ and $q = 8$. The Christoffel construction is represented by the following diagram

$$0 \xrightarrow{a} 3 \xrightarrow{a} 6 \xrightarrow{a} 9 \xrightarrow{b} 1 \xrightarrow{a} 4 \xrightarrow{a} 7 \xrightarrow{a} 10 \xrightarrow{b} 2 \xrightarrow{a} 5 \xrightarrow{a} 8 \xrightarrow{b} 0$$

Let CH denote the class of Christoffel words. The following important result, proved in [3], shows a basic relation existing between central and Christoffel words:

$$\mathrm{CH} = a\mathrm{PER}b \cup \mathcal{A}.$$

Moreover, one has [3, 7]

$$\mathrm{CH} = \mathrm{St} \cap \mathrm{Lynd},$$

where Lynd denotes the set of Lyndon words and St the set of (finite) factors of all Sturmian words. Thus CH equals the set of all factors of Sturmian words which are Lyndon words. The following theorem summarizes some results on Christoffel words proved in [3, 6, 7].

Theorem 2.14. *Let w be a proper Christoffel word. Then the following hold:*

1. *There exist and are unique two Christoffel words w_1 and w_2 such that $w = w_1 w_2$. Moreover, $w_1 <_{lex} w_2$, and (w_1, w_2) is the standard factorization of w in Lyndon words.*
2. *If w has the slope $\frac{p}{q}$, then $|w_1| = p'$, $|w_2| = q'$, where p' and q' are the respective multiplicative inverse of p and q, modulo $|w|$.*
3. *Let $w = a\psi(v)b$ have the slope $\frac{p}{q}$. Then $p = p_a(v^\sim)$, $q = p_b(v^\sim)$ and $p' = p_a(v)$, $q' = p_b(v)$.*

Example 2.15. The Christoffel word w of the Example 2.13 having slope $\frac{3}{8}$ is

$$w = aaabaaabaab = aub,$$

where $u = aabaaabaa = \psi(a^2ba)$ is the central word of length 9 having the two coprime periods $p_a(v) = 4$ and $p_b(v) = 7$ with $v = a^2ba$. The word w can be uniquely factored as $w = w_1 w_2$, where w_1 and w_2 are the Lyndon words $w_1 = aaab$ and $w_2 = aaabaab$. One has $w_1 <_{lex} w_2$ with $|w_1| = 4 = p_a(v)$ and $|w_2| = 7 = p_b(v)$. Moreover, w_2 is the proper suffix of w of maximal length which is a Lyndon word. Finally, $\psi(v^\sim) = \psi(aba^2) = abaabaaba$, $p_a(v^\sim) = 3 = |w|_b$, $p_b(v^\sim) = 8 = |w|_a$, and $|w|_b p_a(v) = 3 \cdot 4 = 12 \equiv |w|_a p_b(v) = 8 \cdot 7 = 56 \equiv 1 \pmod{11}$.

The following proposition is an immediate consequence of item 1 of Theorem 2.14 and of Corollary 2.9 (see also [16]).

Proposition 2.16. *For any non-constant word $v \in \mathcal{A}^*$, the standard factorization of $a\psi(v)b$ in Lyndon words is*

$$(a\psi(v_+)b, a\psi(v^-)b) \ \text{ if } \ v^{(L)} = a \text{ and } (a\psi(v^-)b, a\psi(v_+)b) \ \text{ if } \ v^{(L)} = b.$$

By Proposition 2.16 we have that if v is not constant, then for any $x \in \mathcal{A}$

$$|a\psi(v)b|_x = |a\psi(v^-)b|_x + |a\psi(v_+)b|_x. \tag{10}$$

The following proposition is a direct consequence of (10). It gives a remarkable interpretation of the pair of words v_+ and v^- in the combinatorics of Christoffel words. Recall that the *mediant* of the two fractions a/b and c/d is the fraction $(a + c)/(b + d)$.

Proposition 2.17. *If $v \in \mathcal{A}^*$ is not constant, then the slope of the Christoffel word $a\psi(v)b$ is the mediant of the slopes of $a\psi(v_+)b$ and $a\psi(v^-)b$.*

Remark 2.18. Recall [16] that the slope of the Christoffel word $a\psi(v)b$ is equal to the reduced fraction $\mathrm{SB}(v)$ labeling the node (word) v in the Stern-Brocot tree. From the construction of this tree $\mathrm{SB}(v) = \mathrm{SB}(v_1) \oplus \mathrm{SB}(v_2)$, where $\oplus$ denotes the mediant operation, and v_1 and v_2 are the nearest ancestors of v above and to the right, and above and to the left respectively. It is readily verified that $\{v_1, v_2\} = \{v_+, v^-\}$ so that in any case $\mathrm{SB}(v) = \mathrm{SB}(v_+) \oplus \mathrm{SB}(v^-)$.

The following Propositions 2.19, 2.20, and 2.21 have been proved in [16].

Proposition 2.19. *For any $v \in \mathcal{A}^+$, $\pi(\psi(v^\sim)) = |a\psi(v)b|_{\overline{v^{(F)}}}$.*

Proposition 2.20. *If $v \in \mathcal{A}^*$ is not constant, then*

$$|a\psi(v)b| = |a\psi(v^-)b| + |a\psi(v_+)b| = |a\psi({}^-v)b| + |a\psi({}_+v)b|.$$

Moreover, $|a\psi({}_+v)b| = |a\psi(v)b|_{\overline{v^{(F)}}}$.

Proposition 2.21. *For any word $v = v_1 \cdots v_n$, with $n > 0$, $v_i \in \mathcal{A}$, $i = 1, \ldots, n$, one has*

$$|\psi(v)| = \sum_{i=1}^{n} \pi(\psi(v_1 \cdots v_i)) = \sum_{i=1}^{n} |a\psi(v_i \cdots v_n)b|_{\overline{v_i}}.$$

For any $v \in \mathcal{A}^*$ let $\mathrm{Ra}(v)$ denote the ratio $\mathrm{Ra}(v) = \frac{p_a(v)}{p_b(v)}$. We recall [16] that the reduced fraction $\mathrm{Ra}(v)$ labels the node (word) v in the Raney tree. The following remarkable proposition, which is readily derived from Propositions 2.7 and 2.20, holds:

Proposition 2.22. *Let v be a non-constant word over $\mathcal{A}$. If $v^{(L)} = a$ (resp., $v^{(L)} = b$), then*

$$\mathrm{Ra}(v) = \frac{|a\psi(v_+)b|}{|a\psi(v^-)b|}, \quad \left(\textit{resp.}, \ \mathrm{Ra}(v) = \frac{|a\psi(v^-)b|}{|a\psi(v_+)b|}\right).$$

If $v^{(F)} = a$ (resp., $v^{(F)} = b$), then

$$\mathrm{SB}(v) = \frac{|a\psi({}_+v)b|}{|a\psi({}^-v)b|}, \quad \left(\textit{resp.}, \ \mathrm{SB}(v) = \frac{|a\psi({}^-v)b|}{|a\psi({}_+v)b|}\right).$$

An interesting interpretation of the extension $\mathrm{ext}(v)$ of a directive word v of the central word $\psi(v)$ is given by the following:

Proposition 2.23. *Let $v = v_1v_2\cdots v_m$, $v_i \in \mathcal{A}$, $i = 1,\ldots,m$, be a word of $\mathcal{A}^+$. Let $w = \psi(v) = w_1\cdots w_k$ with $k = |\psi(v)|$ and $w_i \in \mathcal{A}$, $i = 1,\ldots,k$. Then one has*

$$\mathrm{ext}(v) = \mathrm{card}\{\pi(\psi(v_1\cdots v_i)) \mid 1 \le i \le m\} = \mathrm{card}\{\pi(w_1\cdots w_i) \mid 1 \le i \le k\}.$$

Proof. Let $i = 1,\ldots,m$ and set $u = v_1\cdots v_i$. For $x \in \mathcal{A}$ one has by (7) $\pi(\psi(ux)) = p_x(u)$. If $x = u^{(L)} = v_i$, then by (8), $p_{v_i}(u) = \pi(\psi(u))$ and the minimal period is unchanged. If $x = \overline{u}^{(L)} = \overline{v}_i$, then $p_{\overline{v}_i}(u) > \pi(\psi(u))$. Hence, if $v = x_0^{\alpha_0}\cdots x_n^{\alpha_n}$, the set of distinct minimal periods of $\psi(v_1\cdots v_i)$, $i = 1,\ldots,m$, is formed by the minimal periods of the words

$$\psi(x_0), \psi(x_0^{\alpha_0}x_1), \ldots, \psi(x_0^{\alpha_0}\cdots x_{n-1}^{\alpha_{n-1}}x_n)$$

whose number is $n + 1 = \mathrm{ext}(v)$.

Now let $w_1w_2\cdots w_r$ with $r \le k$ be a non-empty prefix of w. There exists $1 \le i < m$ such that

$$\psi(v_1\cdots v_i)v_{i+1} \le_p w_1w_2\cdots w_r \le_p \psi(v_1\cdots v_{i+1}),$$

where we let $\le_p$ denote the prefixal ordering. Hence, $\pi(\psi(v_1\cdots v_i)v_{i+1}) \le \pi(w_1w_2\cdots w_r) \le \pi(\psi(v_1\cdots v_{i+1}))$. By (7),

$$\pi(\psi(v_1\cdots v_i)v_{i+1}) = \pi(\psi(v_1\cdots v_{i+1})) = \pi(w_1w_2\cdots w_r).$$

Thus between $\pi(\psi(v_1\cdots v_i))$ and $\pi(\psi(v_1\cdots v_{i+1}))$ there are no new minimal periods. From this the result follows. □

Corollary 2.24. *For each $k > 0$ and $v \in \mathcal{A}^k$ the word $w = \psi(v)$ has the maximum number of distinct minimal periods of its prefixes if and only if v is alternating, i.e., w is a palindromic prefix of f or of $E(f)$.*

Proof. By the previous proposition the number of distinct minimal periods of $w = \psi(v)$ is given by $\mathrm{ext}(v)$. A word $v \in \mathcal{A}^k$ attains the maximum value k of $\mathrm{ext}(v)$ if and only if v is alternating. □

If $v = x_0^{\alpha_0}\cdots x_n^{\alpha_n}$ we set

$$\pi_i(v) = \pi(\psi(x_0^{\alpha_0}\cdots x_{i-1}^{\alpha_{i-1}}x_i)), \ 0 \le i \le n.$$

Moreover, we let $\overline{\pi}$ denote the arithmetic mean of the distinct minimal periods π_i, $0 \le i \le n$.

Corollary 2.25. *For $v \in \mathcal{A}^+$ one has:*

$$\frac{|\psi(v)|}{\mathrm{ext}(v)} \ge \overline{\pi},$$

where the equality holds if and only if v is alternating.

Proof. Let $n + 1 = \mathrm{ext}(v)$. By Proposition 2.21 one has

$$|\psi(v)| = \sum_{i=1}^{|v|} \pi(\psi(v_1\cdots v_i)) = \sum_{i=0}^{n} \alpha_i\pi_i \ge \sum_{i=0}^{n} \pi_i,$$

so that dividing for $n + 1$ we have

$$\frac{|\psi(v)|}{n+1} \ge \frac{\sum_{i=0}^{n}\pi_i}{n+1} = \overline{\pi}.$$

The equality holds if and only if $\alpha_i = 1$, $i = 0,\ldots,n$. From this the result follows. □

3. Christoffel morphisms

Let $x \in \mathcal{A}$ and $y = \bar{x}$, we consider the injective endomorphism $\mu_{\tilde{x}}$ of $\mathcal{A}^*$ defined by $\mu_{\tilde{x}}(x) = x$ and $\mu_{\tilde{x}}(y) = yx$. In the following, we shall set

$$\lambda_a = \mu_a \text{ and } \lambda_b = \mu_{\tilde{b}},$$

and for any $v = v_1 v_2 \cdots v_n$, $v_i \in \mathcal{A}$, $1 \le i \le n$, we define:

$$\lambda_v = \lambda_{v_1} \circ \lambda_{v_2} \circ \cdots \circ \lambda_{v_n}.$$

If $v = \varepsilon$, we set $\lambda_\varepsilon = \mathrm{id}$. Thus $\{\lambda_a, \lambda_b\}^* = \{\lambda_v \mid v \in \mathcal{A}^*\}$.

The following lemma shows that the morphism λ_b is *right conjugate* [27, Sect. 2.3.4] to μ_b.

Lemma 3.1. *For any $v \in \mathcal{A}^*$, $b\lambda_b(v) = \mu_b(v)b$.*

Proof. By induction on the length of v. The result is trivially verified if $|v| \le 1$. Let us then suppose $|v| > 1$ and write $v = ux$ with $x \in \mathcal{A}$. If $x = a$ then, by using the inductive hypothesis,

$$b\lambda_b(ua) = b\lambda_b(u)\lambda_b(a) = \mu_b(u)bab = \mu_b(ua)b.$$

If $x = b$, one has:

$$b\lambda_b(ub) = b\lambda_b(u)b = \mu_b(u)bb = \mu_b(ub)b.$$

□

Proposition 3.2. *For all $v \in \mathcal{A}^*$,*

$$\lambda_v(ab) = a\psi(v)b.$$

Proof. By induction on the length of v. If $|v| \le 1$, the result is trivially verified. Suppose $|v| > 1$ and write $v = xw$ with $x \in \mathcal{A}$ and $w \in \mathcal{A}^*$. By induction one has:

$$\lambda_{xw}(ab) = \lambda_x(\lambda_w(ab)) = \lambda_x(a\psi(w)b).$$

Let us first suppose that $x = a$. In such a case $\lambda_a = \mu_a$. By Justin's formula

$$\lambda_{aw}(ab) = \mu_a(a\psi(w)b) = a\mu_a(\psi(w))ab = a\psi(aw)b.$$

Let now $x = b$, so that $\lambda_b = \mu_{\tilde{b}}$. By Lemma 3.1 and Justin's formula one has:

$$\lambda_{bw}(ab) = ab\lambda_b(\psi(w))b = a\mu_b(\psi(w))bb = a\psi(bw)b.$$

□

Corollary 3.3. *For any $w, v \in \mathcal{A}^*$,*

$$a\psi(wv)b = \lambda_w(a\psi(v)b).$$

Proof. By the preceding proposition one has:

$$a\psi(wv)b = \lambda_{wv}(ab) = \lambda_w(\lambda_v(ab)) = \lambda_w(a\psi(v)b).$$

□

Proposition 3.4. *Let $v \in \mathcal{A}^*$ be non-constant. The following holds:*

$$\lambda_v(a) = \lambda_{v_+}(ab) = a\psi(v_+)b, \ \lambda_v(b) = \lambda_{v^-}(ab) = a\psi(v^-)b, \ \text{if } v^{(L)} = a$$

and

$$\lambda_v(a) = \lambda_{v^-}(ab) = a\psi(v^-)b, \ \lambda_v(b) = \lambda_{v_+}(ab) = a\psi(v_+)b, \ \text{if } v^{(L)} = b.$$

Proof. We shall prove the result only when $v^{(L)} = a$. The case $v^{(L)} = b$ is similarly dealt with. We can write $v = v_+ba^r$ for a suitable $r > 0$. Therefore, by Proposition 3.2 one has:

$$\lambda_v(a) = \lambda_{v_+ba^r}(a) = \lambda_{v_+b}(a) = \lambda_{v_+}(ab) = a\psi(v_+)b, \tag{11}$$

and

$$\lambda_v(b) = \lambda_{v^-a}(b) = \lambda_{v^-}(ab) = a\psi(v^-)b, \tag{12}$$

which proves the assertion. □

It is worth noting the similarity existing between Proposition 2.5, Corollary 2.6, and Proposition 2.7 concerning standard words and the morphisms μ_v, $v \in \mathcal{A}^*$, which preserve standard words [13], and Proposition 3.2, Corollary 3.3, and Proposition 3.4 concerning Christoffel words and the morphisms λ_v, $v \in \mathcal{A}^*$, which, as we shall see soon (cf. Proposition 3.5), preserve Christoffel words.

Let $\mathcal{M}_{CH}$ denote the monoid of all endomorphisms f of $\mathcal{A}^*$ which preserve Christoffel words, i.e., if $w \in \mathrm{CH}$, then $f(w) \in \mathrm{CH}$. Such a morphism f will be called *Christoffel morphism*. The following proposition was proved in [4] by a different (geometrical) technique[3].

Proposition 3.5. $\mathcal{M}_{CH} = \{\lambda_a, \lambda_b\}^*$.

Proof. Let $\lambda_v \in \{\lambda_a, \lambda_b\}^*$ and $w \in \mathrm{CH}$. We prove that $\lambda_v(w) \in \mathrm{CH}$. Let us first suppose that w is a proper Christoffel word. We can write $w = a\psi(u)b$ for a suitable $u \in \mathcal{A}^*$. Thus by Proposition 3.2

$$\lambda_v(w) = \lambda_v(a\psi(u)b) = \lambda_v(\lambda_u(ab)) = \lambda_{vu}(ab) = a\psi(vu)b \in \mathrm{CH}.$$

Let us now suppose that $w \in \mathcal{A}$. Let $w = a$. If v is not constant, then the result follows from Proposition 3.4. Let us suppose that v is constant. The result is trivial if $v = \varepsilon$. If $v = a^k$ with $k > 0$, we have $\lambda_{a^k}(a) = a \in \mathrm{CH}$. If $v = b^k$ one has $\lambda_{b^k}(a) = ab^k \in \mathrm{CH}$. In a similar way one proves the result if $w = b$.

Let now f be any Christoffel morphism. Since

$$f(a),\ f(b),\ f(ab) = f(a)f(b) \in \mathrm{CH},$$

one has that $(f(a), f(b))$ is the standard factorization of $f(ab)$ in Christoffel (Lyndon) words. As $f(ab)$ is a proper Christoffel word we can write $f(ab) = a\psi(v)b$. Let us suppose that v is not constant. If $v^{(L)} = a$, by Proposition 2.16 one has that the standard factorization of $a\psi(v)b$ in Lyndon words is $(a\psi(v_+)b, a\psi(v^-)b)$. This implies, in view of (11) and (12),

$$f(a) = a\psi(v_+)b = \lambda_{v_+}(ab) = \lambda_v(a),\ \ f(b) = a\psi(v^-)b = \lambda_{v^-}(ab) = \lambda_v(b).$$

Hence, in this case $f = \lambda_v$ and the result follows. The case $v^{(L)} = b$ is similarly dealt with.

Let us now suppose that v is constant. We suppose that $v = a^k$. One has $f(ab) = a\psi(a^k)b = \lambda_{a^k}(ab) = aa^kb$. In this case $f(a) = a = \lambda_{a^k}(a)$ and $f(b) = a^kb = \lambda_{a^k}(b)$. Hence, $f = \lambda_{a^k}$. In a similar way if $v = b^k$ one obtains $f = \lambda_{b^k}$. Thus the result is completely proved. ☐

Proposition 3.6. *Let* $v, w \in \mathcal{A}^*$. *If* $\lambda_v(w) \in \mathrm{CH}$, *then* $w \in \mathrm{CH}$.

Proof. Let us first prove that for $x \in \mathcal{A}$, if $\lambda_x(w) \in \mathrm{CH}$ then $w \in \mathrm{CH}$. If $\lambda_x(w) = y \in \mathcal{A}$, then the only possibility is $x = y$ and $w = x \in \mathrm{CH}$. Let us then suppose that $\lambda_x(w)$ is a proper Christoffel word $a\psi(v)b$ for a suitable $v \in \mathcal{A}^*$. We can write in view of Proposition 3.2

$$\lambda_x(w) = a\psi(v)b = \lambda_v(ab). \tag{13}$$

If $v = \varepsilon$, then one obtains $\lambda_x(w) = ab$ and $w \in \mathrm{CH}$. Let us then suppose $|v| > 0$. We wish to prove that $v^{(F)} = x$. To this end we show that $x = a$ if and only if $v^{(F)} = a$. Indeed, as $\lambda_a(w) \in \{a, ab\}^*$ and $\lambda_b(w) \in \{b, ab\}^*$ if $v^{(F)} = a$, as $\psi(v)$ begins with a, it follows that $x = a$. Conversely, suppose that $x = a$; one has that w has to terminate with b. Moreover, if $w = b^n$, with $n > 1$ one would have $\lambda_a(b^n) = (ab)^n \notin \mathrm{CH}$. If $n = 1$, then $\lambda_a(b) = ab$ and $v = \varepsilon$, a contradiction. Hence, in w there must be at least one occurrence of the letter a, so that we can write $w = w'ab^r$ with $r > 0$. Thus $ab \in \mathrm{Fact}(w)$. This implies that $aab \in \mathrm{Fact}(\lambda_a(w))$, so that $v^{(F)} = a$. We have then proved that $v^{(F)} = x$. Writing $v = xv'$ from (13) we have

$$\lambda_x(w) = a\psi(v)b = \lambda_{xv'}(ab) = \lambda_x(\lambda_{v'}(ab)).$$

As λ_x is injective, it follows $w = \lambda_{v'}(ab) \in \mathrm{CH}$.

The remaining part of the proof is obtained by induction on the length of v. If $|v| > 1$, set $v = xv'$ and suppose that $\lambda_v(w) \in \mathrm{CH}$. We can write $\lambda_v(w) = \lambda_x(\lambda_{v'}(w)) \in \mathrm{CH}$. It follows from we have previously proved that $\lambda_{v'}(w) \in \mathrm{CH}$ and by induction $w \in \mathrm{CH}$. ☐

[3] In [4] any Sturmian morphism, i.e., any endomorphism of $\mathcal{A}^*$ which preserves Sturmian words, is called Christoffel morphism.

The following lemma relates the morphisms $\lambda_{a^k b}$ and $\mu_{a^k b}$, $k \geq 0$.

Lemma 3.7. *For each $k \geq 0$ and $v \in \mathcal{A}^*$,*

$$\lambda_{a^k b}(bv) = \mu_{a^k b}(vb), \quad \lambda_{a^k b}(av) = a\mu_{a^k b}(vb).$$

Proof. By Lemma 3.1 one has $\lambda_{a^k b}(bv) = \lambda_{a^k}(b\lambda_b(v)) = \mu_{a^k}(\mu_b(v)b) = \mu_{a^k b}(vb)$ and the first equation is proved. By using again Lemma 3.1 one has $\lambda_{a^k b}(av) = \lambda_{a^k}(ab\lambda_b(v)) = \lambda_{a^k}(a\mu_b(v)b)) = \mu_{a^k}(a\mu_b(vb)) = a\mu_{a^k b}(vb)$. ☐

Lemma 3.8. *For all $v \in \mathcal{A}^*$ and $x \in \mathcal{A}$,*

$$|\lambda_v(x)| = |\mu_v(x)| = p_x(v).$$

Proof. Let us first suppose that v is constant. If $v = a^k$ with $k \geq 0$, then since $\lambda_{a^k} = \mu_{a^k}$, in view of (6) the result trivially follows. If $v = b^k$, then $|\lambda_{b^k}(a)| = |ab^k| = |b^k a| = |\mu_{b^k}(a)|$ and $|\lambda_{b^k}(b)| = |\mu_{b^k}(b)| = 1$ and the result is achieved. If v is not constant the result follows from Propositions 3.4 and 2.7. ☐

4. Derivative of a Christoffel word

Let $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ be an injective morphism. As is well-known (cf. [5]) the set $X = \varphi(\mathcal{A})$ is a code over the alphabet $\mathcal{A}$, i.e., any word of X^+ can be uniquely factored by the elements of X. Thus there exists an isomorphism, that we still denote by φ, of $\mathcal{A}^*$ and X^*. Let φ^{-1} be the inverse morphism of φ.

If $w \in X^+$, $\varphi^{-1}(w)$ is a uniquely determined word over the alphabet $\mathcal{A}$, that we call *derivative* of w with respect to φ. We shall denote $\varphi^{-1}(w)$ by $\partial_\varphi w$, or simply ∂w, when there is no ambiguity.

Example 4.1. Let $X = \{ab, ba\}$ and φ the Thue-Morse morphism defined by $\varphi(a) = ab$ and $\varphi(b) = ba$. One has that $\partial abbabaab = abba$.

Let w be the finite Sturmian word $w = aababaaba$. The word w can be decoded by the morphism $\mu_a : \{a, b\}^* \to \{a, ab\}^*$ or by the morphism $\mu_a^{\sim} : \{a, b\}^* \to \{a, ba\}^*$. In the first case one obtains the derivative $w_1 = abbaba$ which is still a finite Sturmian word, whereas in the second case one gets the derivative $w_2 = aabbab$ which is not a finite Sturmian word.

In the study of derivatives of finite words over $\mathcal{A}$ belonging to a given class C, we require that the set $\mathcal{M}$ of injective endomorphisms of $\mathcal{A}^*$ satisfies the two following basic conditions:

1. If $\varphi \in \mathcal{M}$, then for any $w \in C$, $\varphi(w) \in C$.
2. If $\varphi(v) = w$ and $w \in C$, then $v \in C$.

Moreover, one can restrict the class $\mathcal{M}$ of endomorphisms to some subclass $\hat{\mathcal{M}}$ assuring that the obtained derivatives satisfy suitable combinatorial properties.

In this section we shall consider the class C of Christoffel words. We define a derivative of a proper Christoffel word by referring to a suitable Christoffel morphism. A derivative in the case of finite (and infinite) standard Sturmian words and its relation with the previous one will be given in Section 6.

Let $u = \psi(v)$ be a central word. We define *index* of the central word u the integer 0 if $v = \varepsilon$ or, otherwise, the first element in the integral representation $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ of v, i.e., α_0. We let $\mathrm{ind}(u)$ denote the index of u. If $w = aub$ is a proper Christoffel word we define *index* (resp., *directive word*) of w the index (resp., directive word) of the central word u.

In the following, for $x \in \mathcal{A}$, we set $\mathrm{PER}_x = \mathrm{PER} \cap x\mathcal{A}^*$ and for any $k \geq 0$ we define the prefix code X_k and the suffix code Y_k:

$$X_k = \{a^k b, a^{k+1} b\} \text{ and } Y_k = \{ab^k, ab^{k+1}\}. \tag{14}$$

Lemma 4.2. *Let $w = aub$ be a proper Christoffel word with $u \neq \varepsilon$ and k be the index of u. If $u \in \mathrm{PER}_a$, then $w \in a^{k+1}bX_k^*$. If $u \in \mathrm{PER}_b$, then $w \in ab^kY_k^*$.*

Proof. We shall prove the lemma only in the case $u \in \mathrm{PER}_a$. The case $u \in \mathrm{PER}_b$ is dealt with in a similar way. We shall denote by Z_k the set $a^{k+1}bX_k^*$. The proof is by induction on the length of the directive word v of the central word $u = \psi(v)$ of index k. If $v = a^k$ then $u = a^k$ and $w = aub = a^{k+1}b$. If $v = a^k b$, then $u = a^k ba^k$ and $w = a^{k+1}ba^k b$ and we are done. Let us then suppose that the result is true for a directive word $v \in a^k b\mathcal{A}^*$ and prove it for the directive word vx with $x \in \mathcal{A}$. Since $\psi(v)$ begins with $a^k ba^k$, we can write from Proposition 2.10

$$u = \psi(v) = u_1 bau_2 = u_2 abu_1 = a^k ba^k \zeta, \tag{15}$$

where $\zeta \in \mathcal{A}^*$ and $u_1, u_2 \in \mathrm{PER}_a$. Moreover, from Lemma 2.11 one has:

$$\psi(va) = u_1 bau_2 abu_1 \quad \text{and} \quad \psi(vb) = u_2 abu_1 bau_2.$$

Let z_1 and z_2 be the Christoffel words $z_1 = a\psi(va)b$ and $z_2 = a\psi(vb)b$. One has that

$$z_1 = au_1 bau_2 abu_1 b = au_1 ba\psi(v)b = au_1 bw \text{ and } z_2 = wau_2 b. \tag{16}$$

From (15) one has that $|u_1| \geq k$. This implies that the index of $u_1 = \psi(v_1)$ is k. Since au_1b is a Christoffel word and $|v_1| < |v|$, one has by induction $au_1 b \in Z_k$. Also by induction $w \in Z_k$. Hence, by (16) one has $z_1 \in Z_k$.

As regards z_2 from (15) one has either $|u_2| \geq k$ or $|u_2| = k - 1$. In the first case since $\mathrm{ind}(u_2) = k$, in a way similar as above one derives by induction that the Christoffel word $au_2 b \in Z_k$, that implies by (16), as $w \in Z_k$, that $z_2 \in Z_k$. In the second case $au_2 b = a^k b \in X_k$, so that, as $w \in Z_k$, it follows $z_2 \in Z_k$ and this concludes the proof. □

If w is a proper Christoffel word, we can introduce a derivative of w as follows. If $w = aub$, where $u \in \mathrm{PER}_a$ is a central word of index k, we consider the prefix code X_k and the injective endomorphism φ_k of $\mathcal{A}^*$ defined by

$$\varphi_k(a) = a^{k+1}b, \quad \varphi_k(b) = a^k b. \tag{17}$$

By the previous lemma $w \in X_k^*$ and the derivative of w with respect to φ_k is $\partial_k w = \varphi_k^{-1}(w)$. Let us observe that from the definition for all $k \geq 0$ one has $\partial_k a^{k+1}b = a$ whereas $\partial_{k+1}a^{k+1}b = b$.

In the case $u \in \mathrm{PER}_b$, one can consider the injective endomorphism $\hat{\varphi}_k$ of $\mathcal{A}^*$ defined by

$$\hat{\varphi}_k(a) = ab^k, \quad \hat{\varphi}_k(b) = ab^{k+1}. \tag{18}$$

By the previous lemma $w \in Y_k^*$ and the derivative of w with respect to $\hat{\varphi}_k$ is $\hat{\partial}_k w = \hat{\varphi}_k^{-1}(w)$. Observe that for all $k \geq 0$ one has $\hat{\partial}_k ab^{k+1} = b$ whereas $\hat{\partial}_{k+1} ab^{k+1} = a$.

If $w = aub$ is a proper Christoffel word of index $k > 0$ the *derivative* of w is the word $\partial w = \partial_k w$ if $u \in \mathrm{PER}_a$ and $\partial w = \hat{\partial}_k w$ if $u \in \mathrm{PER}_b$. Finally, if $k = 0$, i.e., $w = ab$, we set $\partial ab = a$.

Let us observe that from the definition one has for each $k \geq 0$:

$$\varphi_k = \lambda_{a^k b} \quad \text{and} \quad \hat{\varphi}_k = \lambda_{b^k a},$$

so that by Proposition 3.5, φ_k and $\hat{\varphi}_k$ are Christoffel morphisms.

Example 4.3. Let $v = ab^2a^2b$ and w be the Christoffel word $a\psi(v)b$ where $\psi(v)$ is a central word of index 1. One has:

$$w = aababaababaababababaababaababab.$$

In this case one has $X_1 = \{a^2b, ab\}$ and $\partial w = \partial_1 w = abababbababb$.

If $w = a\psi(b^2a^2)b$, then $w = abbabbabbb$. The index of w is 2 and $Y_2 = \{ab^2, ab^3\}$ and $\partial w = \hat{\partial}_2 w = aab$.

Remark 4.4. Let us explicitly observe that two different Christoffel words can have the same derivative. For instance, the Christoffel words $w = a\psi(a^2b^2a)b$ and $w' = a\psi(b^2aba)b$ have both the derivative $ababb = a\psi(ba)b$. Moreover, from the definition it follows that all proper Christoffel words having directive words which are equal up to their index have the same derivative, i.e., for all $k > 0$, $x \in \mathcal{A}$ and $\xi \in \bar{x}\mathcal{A}^*$, one has $\partial(a\psi(x^k\xi)b) = \partial(a\psi(x\xi)b)$. A proper Christoffel word w is determined by its derivative ∂w and the value of its index.

Theorem 4.5. *The derivative of a proper Christoffel word is a Christoffel word.*

Proof. Let $w = aub$ be a proper Christoffel word of index k. If $k = 0$, i.e., $w = ab$, $\partial ab = a$ and in this case the result is trivially verified. Let us suppose $k > 0$. The derivative of w is $\partial w = \varphi_k^{-1}(w)$ if $u \in \mathrm{PER}_a$ and $\partial w = \hat{\varphi}_k^{-1}(w)$ if $u \in \mathrm{PER}_b$. In the first case $\varphi_k(\partial w) = w$ and in the second case $\hat{\varphi}_k(\partial w) = w$. Since φ_k and $\hat{\varphi}_k$ are Christoffel morphisms, by Proposition 3.6 it follows that in both cases $\partial w \in \mathrm{CH}$. □

Corollary 4.6. *Let $w = a\psi(v)b$ be a Christoffel word of index k, with v non-constant. If $w = w_1w_2$ with $w_1, w_2 \in \mathrm{CH}$ is the standard factorization of w in Lyndon words, then $w_1, w_2 \in X_k^*$ or $w_1, w_2 \in Y_k^*$ and $\partial w = \partial w_1 \partial w_2$ with $\partial w_1, \partial w_2 \in \mathrm{CH}$, is the standard factorization of ∂w in Lyndon words.*

Proof. By Theorem 4.5, ∂w is a Christoffel word. Since v is not constant, by Proposition 2.10 one has $\psi(v) = u_1bau_2 = u_2abu_1$ with $u_1, u_2 \in \mathrm{PER}$. Hence, $w = w_1w_2$ where $w_1 = au_1b$ and $w_2 = au_2b$ are two proper Christoffel words. In view of Theorem 2.14, w_1w_2 is the standard factorization of w in Lyndon words.

Let us suppose that $v \in a\mathcal{A}^*$. One has $\psi(v) = a^kba^k\xi$, with $\xi \in \mathcal{A}^*$ from which, as we have seen in the proof of Lemma 4.2, one derives that $w_1, w_2 \in X_k^*$. Thus $\partial w = \partial_k w = \partial_k w_1 \partial_k w_2$ with $\partial_k w_1, \partial_k w_2 \in \mathrm{CH}$. By Theorem 2.14 the result follows. The case $v \in b\mathcal{A}^*$ is similarly dealt with. □

Theorem 4.7. *Let $k \geq 1$ and $w \in X_k^* \cup Y_k^*$. If ∂w is a Christoffel word, then w is a proper Christoffel word.*

Proof. We shall suppose that $w \in X_k^*$. A similar proof can be done when $w \in Y_k^*$. One has that $w = \varphi_k(\partial w)$. Since φ_k is a Christoffel morphism, it follows that $w \in CH$. Moreover, it is readily verified that w is a proper Christoffel word. □

From Theorems 4.5 and 4.7 it follows:

Corollary 4.8. *Let $k > 0$ and $w \in X_k^* \cup Y_k^*$. Then w is a proper Christoffel word if and only if ∂w is a Christoffel word.*

Proposition 4.9. *If $w = a\psi(v)b$, then*

$$|\partial w| = \pi(\psi(v^{\sim})) = |a\psi(v)b|_{\overline{v}^{(F)}}.$$

Proof. From the definition of derivative of w one has $|\partial w| = |a\psi(v)b|_{\overline{v}^{(F)}}$, so that the result follows from Proposition 2.19. □

Corollary 4.10. *A proper Christoffel word $w = a\psi(v)b$ is uniquely determined by $v^{(F)}$, $|w|$, and $|\partial w|$.*

Proof. Let $w = a\psi(v)b$. By Proposition 4.9, $|\partial w| = |a\psi(v)b|_{\overline{v}^{(F)}}$, so that $|a\psi(v)b|_{v^{(F)}} = |w| - |\partial w|$. The Christoffel word w is uniquely determined by its slope $\eta(w) = |w|_b/|w|_a$. If $v^{(F)} = a$, then $\eta(w) = |\partial w|/(|w| - |\partial w|)$. If $v^{(F)} = b$, then $\eta(w) = (|w| - |\partial w|)/|\partial w|$. From this the result follows. □

From Propositions 2.21 and 4.9 one derives:

Corollary 4.11. *For any word $v = v_1 \cdots v_n$, with $n \geq 0$, $v_i \in \mathcal{A}$, $i = 1, \ldots, n$, one has*

$$|\psi(v)| = \sum_{i=1}^{n} |\partial a\psi(v_i \cdots v_n)b|.$$

Example 4.12. Let $w = a\psi(v)b$ with $v = a^2b^2a$. One has

$$w = a^3ba^2ba^3ba^2ba^2b.$$

Moreover, $\partial a\psi(ab^2a)b = \partial w = ababb$, $\partial a\psi(b^2a)b = \partial a\psi(ba)b = ab$, and $\partial a\psi(a)b = a$. Hence, $|\psi(v)| = 15 = 2 \cdot 5 + 2 \cdot 2 + 1$.

The following noteworthy theorem relates, through their directive words, the central word of a proper Christoffel word and the central word of its derivative.

Theorem 4.13. *If $w = a\psi(v)b$ and v is not constant, then*

$$\partial w = a\psi({}_+v)b.$$

Proof. Let $v = x_0^{\alpha_0}\cdots x_n^{\alpha_n}$ with $n \geq 1$ as v is not constant. We can write, setting $\alpha_0 = k$ and $\alpha_1 = h$, $v = x_0^k x_1^h \xi$. We shall first suppose that $\psi(v) \in \mathrm{PER}_a$, so that $v = a^k b^h \xi$. One has $a\psi(v)b \in X_k^*$. Since $\varphi_k = \lambda_{a^k b}$ and $(a^k b)^{-1}v = b^{h-1}\xi$, by Corollary 3.3 one obtains

$$a\psi(v)b = \varphi_k(\partial a\psi(v)b) = \varphi_k(a\psi(b^{h-1}\xi)b).$$

From the injectivity of φ_k it follows $\partial a\psi(v)b = a\psi(b^{h-1}\xi)b = a\psi({}_+v)b$.

Let us now suppose that $\psi(v) \in \mathrm{PER}_b$. We can write, $v = b^k a^h \xi$. One has $a\psi(v)b \in Y_k^*$. Since $\hat{\varphi}_k = \lambda_{b^k a}$ and $(b^k a)^{-1}v = a^{h-1}\xi$, by Corollary 3.3 one obtains

$$a\psi(v)b = \hat{\varphi}_k(\partial a\psi(v)b) = \hat{\varphi}_k(a\psi(a^{h-1}\xi)b).$$

From the injectivity of $\hat{\varphi}_k$ it follows $\partial a\psi(v)b = a\psi(a^{h-1}\xi)b = a\psi({}_+v)b$. □

A different proof of Theorem 4.13 based on continued fractions will be given at the end of the section.

Example 4.14. Let $w = aub$ with $u = \psi(a^2b^2a)$. One has

$$w = aaabaabaaabaabaab,$$

${}_+v = ba$, and $\partial w = ababb = a\psi(ba)b$. If $w = a\psi(ba^2b^2a)b$, one has ${}_+v = ab^2a$ and $\partial w = a\psi(ab^2a)b$. If $w = a\psi(abab)b$, then ${}_+v = ab$ and $\partial w = a\psi(ab)b$.

Corollary 4.15. *Let w be a Christoffel word $a\psi(v)b$ having the derivative $\partial w = a\psi({}_+v)b$. Then $\partial a\psi(E(v))b = a\psi(E({}_+v))b$ and $\partial a\psi(v^\sim)b = a\psi((v_+)^\sim)b$.*

Proof. The word v is not constant so that by the previous theorem and (3), one has $\partial a\psi(E(v))b = a\psi({}_+E(v))b = a\psi(E({}_+v))b$ and $\partial a\psi(v^\sim)b = a\psi({}_+(v^\sim))b = a\psi((v_+)^\sim)b$. □

Proposition 4.16. *Let $v = x_0^{\alpha_0}\cdots x_n^{\alpha_n}$. One has:*

$$|a\psi(v)b| = \sum_{i=0}^{n-1}\alpha_i|a\psi(x_{i+1}^{\alpha_{i+1}-1}x_{i+2}^{\alpha_{i+2}}\cdots x_n^{\alpha_n})b| + \alpha_n + 2.$$

Proof. Let $m = |v|$. By Corollary 4.11,

$$|\psi(v)| = \sum_{i=1}^{m}|\partial a\psi(v_i\cdots v_m)b|.$$

For any $0 \leq i \leq n-1$,

$$\partial a\psi(x_i^{\alpha_i}x_{i+1}^{\alpha_{i+1}}\cdots x_n^{\alpha_n})b = \partial a\psi(x_i^{\alpha_i-1}x_{i+1}^{\alpha_{i+1}}\cdots x_n^{\alpha_n})b = \\ \cdots = \partial a\psi(x_i x_{i+1}^{\alpha_{i+1}}\cdots x_n^{\alpha_n})b.$$

Moreover, $\partial a\psi(x_n^{\alpha_n})b = \partial a\psi(x_n^{\alpha_n-1})b = \cdots = \partial a\psi(x_n)b$ and $|\partial a\psi(x_n)b| = 1$. Hence, one has:

$$|\psi(v)| = \sum_{i=0}^{n-1}\alpha_i|\partial a\psi(x_i x_{i+1}^{\alpha_{i+1}}x_{i+2}^{\alpha_{i+2}}\cdots x_n^{\alpha_n})b| + \alpha_n.$$

By Theorem 4.13 for all $0 \leq i \leq n-1$,

$$\partial a\psi(x_i x_{i+1}^{\alpha_{i+1}}x_{i+2}^{\alpha_{i+2}}\cdots x_n^{\alpha_n})b = a\psi(x_{i+1}^{\alpha_{i+1}-1}x_{i+2}^{\alpha_{i+2}}\cdots x_n^{\alpha_n})b,$$

from which the result follows. □

Proposition 4.17. *Let $k \geq 0$ and f be the function which maps any $v \in \mathcal{A}^k$ into $\partial a\psi(v)b$. For $v, v' \in \mathcal{A}^k$ with $v \neq v'$ if $f(v) = f(v')$, then v and v' are not constant, ${}_+v = {}_+v'$, and $v = x^r y({}_+v), v' = y^r x({}_+v)$ with $r > 0$ and $\{x, y\} = \{a, b\}$. As a consequence the restrictions of f to $a\mathcal{A}^{k-1}$ and to $b\mathcal{A}^{k-1}$ are injective.*

Proof. Suppose $f(v) = f(v')$ with $v \neq v'$. If v is a constant, say $v = a^k$, then $f(v) = \partial a^{k+1}b = a$. As it is readily verified for no other word v' of $\mathcal{A}^k$ one can have $f(v') = f(v) = a$ which is a contradiction. Since both v and v' are not constant, one has:

$$\partial a\psi(v)b = a\psi({}_+v)b = \partial a\psi(v')b = a\psi({}_+v')b.$$

Hence, $\psi({}_+v) = \psi({}_+v')$. Since ψ is injective, it follows ${}_+v = {}_+v'$. Since v and v' have the same length and $v \neq v'$, $v = x^r y({}_+v), v' = y^r x({}_+v)$ with $r > 0$ and $\{x, y\} = \{a, b\}$. The remaining part of the proof trivially follows. □

The following important and well-known theorem concerning the slope of a proper Christoffel word holds (cf.[3]):

Theorem 4.18. *Let $w = aub$ be a proper Christoffel word with $u = \psi(v)$ and $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ be the integral representation of v. Then the slope of w is given by the continued fraction*

$$[\alpha_0; \alpha_1, \ldots, \alpha_{n-1}, \alpha_n + 1] \text{ if } v^{(F)} = b$$

and

$$[0; \alpha_0, \alpha_1, \ldots, \alpha_{n-1}, \alpha_n + 1] \text{ if } v^{(F)} = a.$$

Example 4.19. Let $v = a^2b^2a$. One has $w = a^3ba^2ba^3ba^2ba^2b$ and $\eta(w) = [0; 2, 2, 2] = \frac{5}{12}$. If $v = ba^2b$, then $w = abababbababb$ and $\eta(w) = [1; 2, 2] = \frac{7}{5}$.

As a consequence of Theorems 4.13 and 4.18 one obtains:

Corollary 4.20. *Let $w = aub$ be a proper Christoffel word with $u = \psi(v)$ and $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ the integral representation of v. The slope of ∂w is given by the continued fraction*

$$[\alpha_1 - 1; \alpha_2, \ldots, \alpha_n + 1] \text{ if } v^{(F)} = a$$

and

$$[0; \alpha_1 - 1, \alpha_2, \ldots, \alpha_n + 1] \text{ if } v^{(F)} = b.$$

We remark that the slope of a Christoffel word $w = a\psi(v)b$ determines uniquely the directive word v of $\psi(v)$ and then w. Now we can give a different proof of Theorem 4.13 by using continued fractions and Theorem 4.18.

Second proof of Theorem 4.13. We shall suppose that $\psi(v) \in \mathrm{PER}_a$ and $\alpha_0 = \mathrm{ind}(v)$. The case $\psi(v) \in \mathrm{PER}_b$ is similarly dealt with. From the construction of the derivative of w one has:

$$|\partial w|_a(\alpha_0 + 1) + |\partial w|_b\alpha_0 = |w|_a,$$

and

$$|\partial w|_b + |\partial w|_a = |w|_b.$$

From these relations one easily obtains:

$$\frac{1}{\eta(w)} = \alpha_0 + \frac{1}{1 + \eta(\partial w)}.$$

Let $\eta(w) = [0; \alpha_0, \ldots, \alpha_n + 1]$. One derives from the previous equation:

$$[0; \alpha_1, \ldots, \alpha_n + 1] = \frac{1}{1 + \eta(\partial w)},$$

from which one obtains:

$$\eta(\partial w) = [\alpha_1 - 1; \alpha_2, \ldots, \alpha_n + 1].$$

Therefore, one has $\partial w = a\psi(v')b$ where v' has the integral representation $(\alpha_1 - 1, \alpha_2, \ldots, \alpha_n)$ and therefore is equal to ${}_+v$, which proves the assertion. □

5. Depth of a Christoffel word

From Theorem 4.5 any proper Christoffel word w has a derivative $w' = \partial w$ which is still a Christoffel word. Therefore, if w' is proper one can consider $\partial w' \in \mathrm{CH}$ that we shall denote $\partial^2 w$. In general, for any $p \geq 1$, $\partial^p w$ will denote the derivative of order p of w. Since $|\partial^p w| > |\partial^{p+1} w|$, there exists an integer d such that $\partial^d w \in \mathcal{A}$; we call d the *depth* of w.

Example 5.1. Let w be the Christoffel word $w = a\psi(ab^2a^2b)b$ of Example 4.3. One has

$$\partial w = abababbababb,$$

which is the Christoffel word $a\psi({}_+v)b$ where ${}_+v = ba^2b$. The central word $\psi({}_+v)$ is of order 1 and $\partial^2 w = aabab = a\psi(ab)b$. Moreover, one has $\partial^3 w = ab$ and $\partial^4 w = a$, so that the depth of w is 4.

As we have previously seen, if $v \in \mathcal{A}^*$ is not constant, ${}_+v$ is the longest suffix of v which is immediately preceded by the complementary of the first letter of v. If ${}_+v$ is not constant one can consider ${}_+({}_+v)$ and so on. Thus for any $v \in \mathcal{A}^*$ we can define inductively $v_{(1)} = v$ and, if $v_{(n)}$ is not constant and $n \geq 1$,

$$v_{(n+1)} = {}_+(v_{(n)}).$$

Since $|v_{(n+1)}| < |v_{(n)}|$, there exists an integer $h = h(v)$ called *height* of v, such that $v_{(h)}$ is constant. For instance, if $v = a^2b^2a$, one has $v_{(1)} = a^2b^2a$, $v_{(2)} = ba$, and $v_{(3)} = \varepsilon$. Hence, $h(a^2b^2a) = 3$.

Proposition 5.2. *Let $w = a\psi(v)b$ be a proper Christoffel word. The depth of w is equal to the height of v.*

Proof. If v is constant, then $h = h(v) = 1$ and $\partial w \in \mathcal{A}$, so that the depth of w is 1. Let us then suppose that v is not a constant. This implies $h(v) > 1$. From Theorem 4.13 one derives that for $n \leq h - 1$

$$\partial^n w = a\psi(v_{(n+1)})b.$$

Since $v_{(h)}$ is constant, it follows that $\partial^h w \in \mathcal{A}$, so that the depth of w is h. $\square$

Let $v = x_0^{\alpha_0} \cdots x_n^{\alpha_n}$. For $i \in \{0, \ldots, n\}$ we define a map

$$\delta_i(v) : \{0, \ldots, n\} \to \{0, 1\}$$

as follows: $\delta_0(v) = \delta_n(v) = 1$. For $0 < i < n$, if $\alpha_i > 1$ we set $\delta_i(v) = 1$. Let $\alpha_i = 1$. If $\alpha_{i-1} > 1$ we set $\delta_i(v) = 0$. If $\alpha_{i-1} = 1$, then we set $\delta_i(v) = 1$ if and only if $\delta_{i-1}(v) = 0$. Let us define for any $v \in \mathcal{A}^+$

$$\delta(v) = \sum_{i=0}^{n} \delta_i(v).$$

Moreover, we set $\delta(\varepsilon) = 1$.

Example 5.3. Let $v = a^2bab^2aba$. In this case $n = 6$. Denoting $\delta_i(v)$ simply by δ_i, the sequence $\delta_0\delta_1 \cdots \delta_n$ is given by 1011011 and $\delta(v) = 5$.

Proposition 5.4. *Let $v \in \mathcal{A}^*$. Then $h(v) = \delta(v)$.*

Proof. If $v = \varepsilon$ the result is trivially true. Let $v \neq \varepsilon$. We can write $v = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\alpha_i \geq 1$, $0 \leq i \leq n$. We proceed by induction on n. If $n = 0$ then $h(v) = \delta(v) = 1$. If $n = 1$ then $h(v) = \delta(v) = 2$. Let $n = 2$, then $v = x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} = v_{(1)}$. There are two cases:

(1) $\alpha_1 = 1$. One has $v_{(2)} = x_2^{\alpha_2}$ and $h(v) = \delta(v) = 2$;

(2) $\alpha_1 > 1$. One has $v_{(2)} = x_1^{\alpha_1 - 1} x_2^{\alpha_2}$, $v_{(3)} = x_2^{\alpha_2 - 1}$ and $h(v) = \delta(v) = 3$.

Let $n > 2$, then ${}_+v = x_1^{\alpha_1 - 1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} = v_{(2)}$. Since by the definition of height, $h(v) = h({}_+v) + 1$ and, by induction, $h({}_+v) = \delta({}_+v)$, it suffices to prove that $\delta(v) = \delta({}_+v) + 1$. There are two possibilities:

(1) $\alpha_1 = 1$. In this case ${}_+v = x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $\delta_i(v) = \delta_{i-2}({}_+v)$ if $i \geq 2$. Indeed, if $i = 2$, as $\delta_1(v) = 0$, one has $\delta_2(v) = 1 = \delta_0({}_+v)$. From the definition of δ it follows that $\delta_i(v) = \delta_{i-2}({}_+v)$ if $i > 2$. Hence,

$$\delta(v) = 1 + \sum_{i=2}^{n} \delta_i(v) = 1 + \sum_{i=0}^{n-2} \delta_i({}_+v) = 1 + \delta({}_+v).$$

(2) $\alpha_1 > 1$. In this case ${}_+v = x_1^{\alpha_1-1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $\delta_i(v) = \delta_{i-1}({}_+v)$ for $i \geq 1$. Indeed, if $i = 1$ as $\alpha_1 > 1$ one has $\delta_1(v) = 1 = \delta_0({}_+v)$. For $i = 2$ if $\alpha_2 > 1$ then $\delta_2(v) = \delta_1({}_+v) = 1$. If $\alpha_2 = 1$, then $\delta_2(v) = 0 = \delta_1({}_+v)$ because $\delta_1(v) = \delta_0({}_+v) = 1$. From the definition of δ it follows that $\delta_i(v) = \delta_{i-1}({}_+v)$ for $i > 2$. Hence,

$$\delta(v) = \sum_{i=0}^{n} \delta_i(v) = 1 + \sum_{i=1}^{n} \delta_i(v) = 1 + \sum_{i=0}^{n-1} \delta_i({}_+v) = 1 + \delta({}_+v). \qquad \square$$

Example 5.5. Let $v = a^3bab^2aba$. One has $h(v) = 5$ and $\delta_0\delta_1\delta_2\delta_3\delta_4\delta_5\delta_6 = 1011011$, so that $\delta(v) = 5$.

Let $v = x_0^{\alpha_0} \cdots x_n^{\alpha_n}$. If $\alpha_i > 1$ for all $0 < i < n$, then from the definition of δ one has $\delta(v) = \mathrm{ext}(v) = n + 1$. Let us suppose on the contrary that $\alpha_i = 1$ for all $0 < i < n$. We can write $v = x_0^{\alpha_0} u x_n^{\alpha_n}$ where u is an alternating word $u = x_1x_2 \cdots x_{n-1}$. In this case it is easy derive that

$$\delta(v) = 2 + \left\lfloor \frac{n-1}{2} \right\rfloor = \mathrm{ext}(v) - \left\lceil \frac{|u|}{2} \right\rceil.$$

In general, by grouping together consecutive x_i, $0 < i < n$, having $\alpha_i = 1$ we can rewrite v uniquely as

$$v = v_0u_1v_2u_3 \cdots u_{k-1}v_k, \tag{19}$$

where all terms of the integral representation of u_i (resp., v_i), $1 \leq i \leq k-1$ are equal to 1 (resp., > 1) and all terms of the integral representation of v_0 (resp., v_k) are > 1, with the possible exception of the first (resp., last).

We call the u_i, $i = 1, 3, \ldots, k-1$, the *alternating components* of v. For example, if $v = a^3b^2abab^2aba^2ba$, then we can factore it as $v = (a^3b^2)(aba)(b^2)(ab)(a^2)(b)(a)$. In this case the alternating components of v are $u_1 = aba$, $u_3 = ab$, and $u_5 = b$.

Proposition 5.6. *Let $v \in \mathcal{A}^+$ and u_i, $i = 1, 3, \ldots, k-1$, be the alternating components of v. Then one has:*

$$\delta(v) = \mathrm{ext}(v) - \sum_{i=0}^{\frac{k-2}{2}} \left\lceil \frac{|u_{2i+1}|}{2} \right\rceil.$$

Proof. Let $v = v_0u_1v_2u_3 \cdots u_{k-1}v_k$. Since $\delta(v_{2i}) = \mathrm{ext}(v_{2i})$, $0 \leq i \leq k/2$ and $|u_{2i+1}| = \mathrm{ext}(u_{2i+1})$, $0 \leq i \leq k/2 - 1$, one has

$$\delta(v) = \sum_{i=0}^{\frac{k}{2}} \delta(v_{2i}) + \sum_{i=0}^{\frac{k-2}{2}} \left\lfloor \frac{|u_{2i+1}|}{2} \right\rfloor =$$

$$\sum_{i=0}^{\frac{k}{2}} \mathrm{ext}(v_{2i}) + \sum_{i=0}^{\frac{k-2}{2}} \left(\mathrm{ext}(u_{2i+1}) - \left\lceil \frac{|u_{2i+1}|}{2} \right\rceil \right) = \mathrm{ext}(v) - \sum_{i=0}^{\frac{k-2}{2}} \left\lceil \frac{|u_{2i+1}|}{2} \right\rceil. \qquad \square$$

Example 5.7. If $v = (a^3b^2)(aba)(b^2)(ab)(a^2)(b)(a)$, we have $\mathrm{ext}(v) = 11$, $\lceil |aba|/2 \rceil = 2$, $\lceil |ab|/2 \rceil = 1 = \lceil |b|/2 \rceil$, so that $\delta(v) = 11 - 4 = 7$.

In the following, for each word $v \in \mathcal{A}^*$ we let $[v]$ denote the set $[v] = \{v, v^{\sim}, \bar{v}, \bar{v}^{\sim}\}$. From Proposition 2.1 all Christoffel words $a\psi(z)b$ with a directive word $z \in [v]$ have the same length. The next proposition shows that they have the same depth.

Proposition 5.8. *Let $v \in \mathcal{A}^*$. All Christoffel words $a\psi(z)b$ with $z \in [v]$ have the same depth.*

Proof. The result is trivially true if v is constant. Let us then suppose that v is not constant. From Propositions 5.2 and 5.4 it is sufficient to prove that $\delta(z) = \delta(v)$ for all $z \in [v]$. It is readily verified that $\delta(\overline{v}) = \delta(v)$ as $\mathrm{ext}(\overline{v}) = \mathrm{ext}(v) = n$ and for each $0 \le i \le n$, $\delta_i(\overline{v}) = \delta_i(v)$. Let us now prove that $\delta(v^\sim) = \delta(v)$. Let us write v as in Eq. (19), so that

$$v^\sim = v_k^\sim u_{k-1}^\sim \cdots u_1^\sim v_0^\sim .$$

Since all $u_i^\sim$, $i = k-1, k-3, \ldots, 3, 1$ are the *alternating components* of $v^\sim$, by the fact that $\mathrm{ext}(v^\sim) = \mathrm{ext}(v) = n$ and $|u_i| = |u_i^\sim|$ in view of Proposition 5.6 it follows that $\delta(v^\sim) = \delta(v)$. From the previous results it follows immediately that $\delta(\overline{v}^\sim) = \delta(v)$. □

Let $u = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. We define u_1 as: $u_1 = \varepsilon$ if $\alpha_0 > 1$ or $n = 0$, and, otherwise, u_1 is the longest proper prefix of u such that $u = u_1 u'$, u_1 is alternating, and $u'^{(F)} \neq u_1^{(L)}$, i.e., the longest prefix $x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_i^{\alpha_i}$ of u with $i < n$ and $\alpha_0 = \alpha_1 = \ldots = \alpha_i = 1$.

Similarly, we define u_2 as $u_2 = \varepsilon$ if $\alpha_n > 1$ or $n = 0$ and, otherwise, u_2 is the longest proper suffix of u such that $u = u' u_2$, u_2 is alternating, and $u'^{(L)} \neq u_2^{(F)}$. For instance, if $u = ab^2aba$, then $u_1 = a$ and $u_2 = aba$; if $u = a^2b$, then $u_1 = \varepsilon$ and $v_2 = b$.

Proposition 5.9. *Let $u = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $v = y_0^{\beta_0} y_1^{\beta_1} \cdots y_m^{\beta_m}$. Then*

$$\delta(u) + \delta(v) - 1 \le \delta(uv) \le \delta(u) + \delta(v).$$

Moreover, $\delta(uv) = \delta(u) + \delta(v)$ if and only if $u^{(L)} \neq v^{(F)}$ and $|u_2|$, $|v_1|$ are both even.

Proof. If $u^{(L)} = v^{(F)}$, then $uv = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_{n-1}^{\alpha_{n-1}} x_n^{\alpha_n + \beta_0} y_1^{\beta_1} \ldots y_m^{\beta_m}$ and trivially $\delta(uv) = \delta(u) + \delta(v) - 1$. Let us then suppose $u^{(L)} \neq v^{(F)}$. We consider two cases: $|u_2|$ even and $|u_2|$ odd.

If $|u_2|$ is even, then $\delta_i(u) = \delta_i(uv)$, $i = 0, \ldots, n$. If $|v_1| = 0$, then $\delta_i(v) = \delta_{n+i+1}(uv)$, $i = 0, \ldots, m$ so that $\delta(uv) = \delta(u) + \delta(v)$. Let then $|v_1| = r \ge 1$. For each $i = 0, \ldots, r-1$, one has $(\delta_i(v), \delta_{n+i+1}(uv)) = (1, 0)$ if i is even and $(\delta_i(v), \delta_{n+i+1}(uv)) = (0, 1)$ if i is odd. Moreover, $\delta_i(v) = \delta_{n+i+1}(uv)$ for each $i = r, \ldots, m$.

It follows that if r is even, then the number of pairs $(1, 0)$ is equal to the number of pairs $(0, 1)$ that implies $\delta(uv) = \delta(u) + \delta(v)$. If r is odd, then the number of pairs $(1, 0)$ is equal to the number of pairs $(0, 1)$ plus 1, so that $\delta(uv) = \delta(u) + \delta(v) - 1$.

Let $|u_2|$ be odd. In this case $\delta_i(u) = \delta_i(uv)$ if $i = 0, \ldots, n-1$, $\delta_n(u) = 1$, and $\delta_n(uv) = 0$, $\delta_i(v) = \delta_{n+i+1}(uv)$, if $i = 0, \ldots, m$. It follows $\delta(uv) = \delta(u) + \delta(v) - 1$ and then the assertion. □

Example 5.10. Let $u = a^3b^2aba$, $w = a^3b^2a^2ba$, and $v = babab^2$. One has $u_2 = aba$, $w_2 = ba$, and $v_1 = baba$. One has $\delta(u) = \delta(w) = 4$, and $\delta(v) = 3$. One verifies that $\delta(uv) = 6$ and $\delta(wv) = 7$.

From Proposition 5.9 one readily derives:

Corollary 5.11. *Let $u = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $v = x \in \mathcal{A}$. Then*

$$\delta(u) \le \delta(ux) \le \delta(u) + 1.$$

Moreover, $\delta(ux) = \delta(u)$ if and only if $u^{(L)} = x$ or $\overline{u}^{(L)} = x$ and $|u_2|$ is odd.

Lemma 5.12. *If v is a non-constant word, then $h(v) = h(v_+) + 1$.*

Proof. By the definition of height, we have $h(v) = h({}_+v) + 1$. Moreover, Proposition 5.8 implies that $h(u) = h(u^\sim)$ for any $u \in \mathcal{A}^*$. Hence, to obtain the assertion it suffices to observe by (3) that $v_+ = ({}_+(v^\sim))^\sim$. □

We shall now give another equivalent definition for the function $h = \delta$. Let $H : \mathbb{N}_+ \to \mathbb{N}$ be the sequence defined by $H(1) = 0$ and, for all $n > 0$,

$$H(2n) = H(n) \quad \text{and} \quad H(4n \pm 1) = H(n) + 1\,.$$

The first few values of $H(n)$ are

$$0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 2, 1, 2, 1, 1, 0, 1, \ldots .$$

As an immediate consequence of the definition, for any $k \geq 1$ we have

$$H(2^{k+1} n \pm 1) = H(n) + 1 . \tag{20}$$

Note that the sequence $\hat{H}$ given by $\hat{H}(n) = H(n) + 1$ is the sequence A007302 in [29].

Proposition 5.13. *For all $v \in \mathcal{A}^*$, $h(v) = H(\langle bvb \rangle)$.*

Proof. We proceed by induction on $h(v)$. If $h(v) = 1$, then v is constant, so that $\langle bvb \rangle = 2^{k+1} \pm 1$ for some $k \geq 1$. From (20), $H(\langle bvb \rangle) = H(1) + 1 = 1$ follows.

Let now $h(v) > 1$, so that v contains both a and b as letters. By Lemma 5.12 and the induction hypothesis, we have $h(v) = H(\langle b(v_+)b \rangle) + 1$.

Now, if $v^{(L)} = a$, then there exists $k \geq 1$ such that $v = (v_+)ba^k$, so that

$$\langle bvb \rangle = \langle b(v_+)b \cdot a^k b \rangle = 2^{k+1} \langle b(v_+)b \rangle + 1 .$$

On the other hand, if $v^{(L)} = b$, then there exists $k \geq 1$ with $v = (v_+)ab^k$, so that

$$\langle bvb \rangle = \langle b(v_+)ab^{k+1} \rangle = 2^{k+1} \langle b(v_+)b \rangle - 1 .$$

In both cases, by (20) it follows $H(\langle bvb \rangle) = H(\langle b(v_+)b \rangle) + 1 = h(v)$, as desired. □

Example 5.14. Let $v = a^2bab$. One has $\langle ba^2bab^2 \rangle = 75$ and $H(75) = H(19) + 1 = H(5) + 2 = H(1) + 3 = 3$. Hence, $h(v) = 3$.

Proposition 5.15. *For all $v \in \mathcal{A}^*$,*

$$\left\lfloor \frac{\mathrm{ext}(v)}{2} \right\rfloor + 1 \leq h(v) \leq \left\lfloor \frac{|v|}{2} \right\rfloor + 1 .$$

The set of words v over $\mathcal{A} = \{a, b\} = \{x, y\}$ for which the lower bound is attained is $Y = x^+(yx^+)^*$ if $\mathrm{ext}(v)$ is odd and $Y = \{\varepsilon\} \cup x^+(yx^+)^*(y^+x)^*y^+$ if $\mathrm{ext}(v)$ is even. The set of words for which the upper bound is attained is $X = \{ab, ba\}^*\{\varepsilon, a, b\}\{ab, ba\}^*$.

Proof. Let us first prove the lower bound. Let $\mathrm{ext}(v) = n + 1$. One has that $h(v) = \delta(v) = \mathrm{ext}(v) - \mathrm{card}\{0 < i < n \mid \delta_i(v) = 0\}$. Since by the definition of δ in the sequence $\Delta_v = \delta_0(v) \cdots \delta_n(v)$ one cannot have two consecutive 0, it follows that the maximal value of $\mathrm{card}\{0 < i < n \mid \delta_i(v) = 0\}$ is attained if and only if $\Delta_v \in 1(01)^*$ if $\mathrm{ext}(v)$ is odd and $\Delta_v \in 1(01)^*(10)^*1$ if $\mathrm{ext}(v)$ is even. In both the cases the previous maximal value is equal to $\left\lceil \frac{n-1}{2} \right\rceil$. From this one has $\delta(v) \geq n + 1 - \left\lceil \frac{n-1}{2} \right\rceil = \left\lfloor \frac{\mathrm{ext}(v)}{2} \right\rfloor + 1$.

To complete the first part of the proof it is sufficient to observe that $\Delta_v \in 1(01)^*$ if and only if $v \in x^+(yx^+)^*$, and $\Delta_v \in 1(01)^*(10)^*1$ if and only if $v \in x^+(yx^+)^*(y^+x)^*y^+$.

Let us now prove the upper bound. If v is constant, then $h(v) = 1$ and the result is trivially true. Let us then suppose that v is not constant. Let $n = h(v) > 1$. By the definition of height, there exist $v_{(1)}, \ldots, v_{(n)}$ such that $v_{(1)} = v$, $v_{(n)}$ is constant, and $v_{(i+1)} = {}_+(v_{(i)})$ for $i = 1, \ldots, n-1$. Since for any non-constant word u one has $|u| \geq |{}_+u| + 2$, it follows $|v_{(i)}| \geq |v_{(i+1)}| + 2$, $i = 1, \ldots, n-1$, so that

$$|v| \geq |v_{(n)}| + 2n - 2 \geq 2n - 2; \tag{21}$$

hence $n = h(v) \leq \lfloor |v|/2 \rfloor + 1$.

Let us now prove that $h(v) = \lfloor |v|/2 \rfloor + 1$ if and only if $v \in X$. Suppose first that $|v|$ is even. The set of words of even length within X is $\{ab, ba\}^*$. Clearly, from (21), $n = 1 + |v|/2$ if and only $v_{(n)} = \varepsilon$ and $|v_{(i)}| - |v_{(i+1)}| = 2$ for $i = 1, \ldots, n-1$. Now, $|u| - |{}_+u| = 2$ if and only if $u = ab({}_+u)$ or $u = ba({}_+u)$. It follows that $h(v) = 1 + |v|/2$ if and only if $v \in \{ab, ba\}^*$.

Let now $|v|$ be odd. The subset of X made by words of odd length is

$$X' = X \setminus \{ab, ba\}^* = \{ab, ba\}^*\{a, b\}\{ab, ba\}^*.$$

It is not difficult to see that $X' = X_1' \cup X_2'$ where

$$X_1' = \{ab, ba\}^*\{a, b\} \text{ and } X_2' = \{ab, ba\}^*\{aab, bba\}\{ab, ba\}^*.$$

One has $|u| - |_+u| = 3$ if and only if $u = aab(_+u)$ or $u = bba(_+u)$. It follows that if $v \in X_1'$, then $v_{(n)} \in \{a, b\}$ and $|v_{(i)}| - |v_{(i+1)}| = 2$ for $i = 1, \ldots, n-1$. If $v \in X_2'$, then $v_{(n)} = \varepsilon$ and $|v_{(i)}| - |v_{(i+1)}| = 2$ for all i in $\{1, \ldots, n-1\}$ except exactly one j for which $|v_{(j)}| - |v_{(j+1)}| = 3$. From (21) one has that in both cases $n = (|v| + 1)/2 = \lfloor |v|/2 \rfloor + 1$.

Conversely, from (21) if v is such that $n = (|v| + 1)/2$, then we must have either $v_{(n)} \in \{a, b\}$ and $|v_{(i)}| - |v_{(i+1)}| = 2$ for $i = 1, \ldots, n-1$, or $v_{(n)} = \varepsilon$ and $|v_{(i)}| - |v_{(i+1)}| = 2$ for all i in $\{1, \ldots, n-1\}$ except exactly one j for which $|v_{(j)}| - |v_{(j+1)}| = 3$. In the former case, we obtain $v \in X_1'$, and in the latter $v \in X_2'$. □

We say that a word $v \in \mathcal{A}^+$ is *quasi-alternating* if each letter of v but exactly one, is immediately followed by its complementary. For instance, the words ab^2ab and aba^2bab are quasi-alternating.

Corollary 5.16. *Let $v \in \mathcal{A}^+$. Then*

$$h(v) = \left\lfloor \frac{\mathrm{ext}(v)}{2} \right\rfloor + 1 = \left\lfloor \frac{|v|}{2} \right\rfloor + 1 \tag{22}$$

if and only if v is alternating or v is quasi-alternating with $\mathrm{ext}(v)$ *equal to an even integer.*

Proof. (⇒) If (22) is satisfied, then $|v| = \mathrm{ext}(v)$ or $|v| = \mathrm{ext}(v) + 1$. In the first case v is alternating and in the second case quasi-alternating. Moreover, in the latter case $\mathrm{ext}(v)$ has to be even, otherwise $\frac{|v|}{2} = \lfloor \frac{\mathrm{ext}(v)}{2} \rfloor + 1$, a contradiction.
(⇐) If v is alternating or quasi-alternating, then by the preceding proposition $v \in X$ so that $h(v) = \lfloor \frac{|v|}{2} \rfloor + 1$. Moreover, if v is alternating, then $|v| = \mathrm{ext}(v)$ and we are done. If v is quasi-alternating, then $|v| = \mathrm{ext}(v) + 1$. If $\mathrm{ext}(v)$ is even, then $\lfloor |v|/2 \rfloor = \frac{\mathrm{ext}(v)}{2}$ and the result is obtained. □

Example 5.17. Let $v = a^2ba^3baba^3b$. One has $\Delta_v = 10101011$, so that $h(v) = \delta(v) = 5$. Since $\mathrm{ext}(v) = 8$, one has that $v \in Y$ and $h(v) = \mathrm{ext}(v)/2 + 1$. Let $v = abab^2a^2b \in X$; one has $h(v) = 5 = |v|/2 + 1$. Let v be the quasi-alternating word $v = abab^2ab$; one has $h(v) = \mathrm{ext}(v)/2 + 1 = 4 = \lfloor |v|/2 \rfloor + 1$.

For each pair k, p of positive integers we let $X_k(p)$ denote the set of all words of length k having a height equal to p, i.e.,

$$X_k(p) = \{v \in \mathcal{A}^k \mid h(v) = p\}.$$

Moreover, we set $J_k(p) = \mathrm{card}(X_k(p))$. From the definition one has $X_k(1) = \{a^k, b^k\}$. By Proposition 5.15 one has $X_k(p) = \emptyset$ if $p > \lfloor \frac{k}{2} \rfloor + 1$.

In order to give an exact formula for $J_k(p)$, we need some notation and preparatory results. We recall that for any $v = x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, the word v_2 is defined as $v_2 = \varepsilon$ if $\alpha_n > 1$ or $n = 0$ and, otherwise, v_2 is the longest proper suffix of v such that $v = v'v_2$, v_2 is alternating, and $v'^{(L)} \neq v_2^{(F)}$.

Let E be the set of words v such that v_2 is of even length, i.e., $E = \{v \in \mathcal{A}^* \mid |v_2| \equiv 0 \pmod 2\}$, and let

$$e_k(p) = \mathrm{card}(X_k(p) \cap E), \quad o_k(p) = \mathrm{card}(X_k(p) \setminus E),$$

so that $J_k(p) = e_k(p) + o_k(p)$.

The following proposition gives a recursive procedure allowing to computing $X_k(p)$ and then $J_k(p)$, for all k and p.

Proposition 5.18. *For all $k > 0$ and $p > 0$,*

$$X_{k+1}(p) \cap E = \{vv^{(L)} \mid v \in X_k(p)\} \cup \{v\overline{v}^{(L)} \mid v \in X_k(p) \setminus E\},$$
$$X_{k+1}(p) \setminus E = \{v\overline{v}^{(L)} \mid v \in X_k(p-1) \cap E\}.$$

Hence,

$$e_{k+1}(p) = J_k(p) + o_k(p) = e_k(p) + 2o_k(p), \tag{23}$$
$$o_{k+1}(p) = e_k(p-1). \tag{24}$$

Proof. By Corollary 5.11, for any $x \in \mathcal{A}$ one has $h(vx) = \delta(vx) = \delta(v) = h(v) = p$ if and only if $v^{(L)} = x$ or $\overline{v}^{(L)} = x$ and $|v_2|$ is odd. If $\overline{v}^{(L)} = x$ and $|v_2|$ is even, then $h(v) = p - 1$. Moreover, it is clear from the definition that $vv^{(L)} \in E$ for all v, whereas $v\overline{v}^{(L)}$ is in E if and only if v is not. From this the result follows. □

Example 5.19. Since if $v \in X_k(p)$ then $\overline{v} \in X_k(p)$, we set $X'_k(p) = \{v \in X_k(p) \mid v^{(F)} = a\}$. For $k = 2$ one has $X'_2(1) = \{a^2\}$ and $X'_2(2) = \{ab\}$. For $k = 3$, $X'_3(1) = \{a^3\}$ and $X'_3(2) = \{ab^2, a^2b, aba\}$. For $k = 4$, $X'_4(1) = \{a^4\}$, $X'_4(2) = \{ab^3, a^2b^2, aba^2\} \cup \{a^2ba\} \cup \{a^3b\}$, $X'_4(3) = \emptyset \cup \emptyset \cup \{ab^2a, abab\}$.

Lemma 5.20. *For all $k > 0$, one has $o_k(1) = 0$ and for $p > 1$,*

$$o_k(p) = 2^{p-1}\binom{k-p}{p-2}$$

with the usual convention that $\binom{n}{m} = 0$ whenever $n < m$.

Proof. Clearly $o_k(1) = 0$ since $X_k(1) = \{a^k, b^k\} \subseteq E$. Moreover, from (24) it follows $o_k(2) = e_{k-1}(1) = 2 = 2^{2-1}\binom{k-2}{2-2}$.

Let now $p > 2$. The assertion is trivially verified if $k < 2(p-1)$, since this implies $p > \lfloor k/2 \rfloor + 1$ and then $0 = J_k(p) \geq o_k(p)$. If $k = 2(p-1)$, we have $p = \lfloor k/2 \rfloor + 1$, so that $X_k(p) = \{ab, ba\}^{p-1}$ and $J_k(p) = 2^{p-1}$. By (23), $e_k(p) = J_{k-1}(p) + o_{k-1}(p) \leq 2J_{k-1}(p) = 2J_{2p-3}(p) = 0$; hence,

$$o_k(p) = J_k(p) = 2^{p-1} = 2^{p-1}\binom{k-p}{p-2}.$$

We can now assume, by (double) induction, that the assertion is verified for all smaller values of k and p. Substituting (23) in (24), we obtain

$$\begin{aligned} o_k(p) = e_{k-1}(p-1) &= e_{k-2}(p-1) + 2o_{k-2}(p-1) \\ &= e_{k-3}(p-1) + 2o_{k-3}(p-1) + 2o_{k-2}(p-1) \\ &= \cdots = 2\sum_{i=2(p-2)}^{k-2} o_i(p-1), \end{aligned}$$

where the last equality holds because $e_{2(p-2)}(p-1) = 0$. Therefore, by induction we have

$$o_k(p) = 2\sum_{i=2(p-2)}^{k-2} 2^{p-2}\binom{i-p+1}{p-3} = 2^{p-1}\sum_{i=0}^{k-2(p-1)}\binom{i+p-3}{p-3}.$$

The assertion now follows from the identity (see, for instance, [23])

$$\sum_{j=0}^{n-1}\binom{j+m}{m} = \binom{n+m}{m+1}.$$

□

Theorem 5.21. *For all $k, p > 0$,*

$$J_k(p) = 2^{p-1}\left(\binom{k-p+1}{p-1} + \binom{k-p}{p-1}\right).$$

Proof. If $k < 2(p-1)$, then $p > \lfloor k/2 \rfloor + 1$, so that $J_k(p) = 0$ as desired.

Let now $k \geq 2(p-1)$. The assertion is trivially verified for $p = 1$, so let us suppose $p > 1$. Using (24) and Lemma 5.20, we obtain

$$J_k(p) = e_k(p) + o_k(p) = o_{k+1}(p+1) + o_k(p) = 2^p\binom{k-p}{p-1} + 2^{p-1}\binom{k-p}{p-2} = 2^{p-1}\left(2\binom{k-p}{p-1} + \binom{k-p}{p-2}\right).$$

The proof is completed by Pascal's rule. □

From the preceding theorem one derives a simple formula for the number of words of length k for which the height reaches its maximal value $\lfloor \frac{k}{2} \rfloor + 1$.

Corollary 5.22. *Let $k > 0$. If k is even, $J_k(\frac{k}{2}+1) = 2^{\frac{k}{2}}$ and if k is odd, $J_k(\lfloor \frac{k}{2} \rfloor + 1) = 2^{\frac{k+1}{2}}(1 + \frac{1}{2}\lfloor \frac{k}{2} \rfloor)$.*

6. Derivative of a standard word

In this section we shall see that any finite or infinite standard Sturmian word w has, with respect to a suitable endomorphism of $\mathcal{A}^*$, a derivative which is still a standard Sturmian word.

For any $k \geq 0$ we define

$$X'_k = \{a^k b, a^k ba\} \text{ and } Y'_k = \{b^k a, b^k ab\}.$$

The sets X'_k and Y'_k are codes having a finite deciphering delay [5], so that any word $x \in X'^{\infty}_k$ (resp., $x \in Y'^{\infty}_k$) can be uniquely factored by the elements of X'_k (resp., Y'_k).

Let $w = uxy$, with $u \in$ PER and $\{x, y\} = \{a, b\}$, be a proper standard Sturmian word. We define index of w the index of the central word u.

Lemma 6.1. *Let w be a proper standard Sturmian word of index k. Then $w \in X'^*_k \cup \{a^{k+1}b\}$ if $w^{(F)} = a$ and $w \in Y'^*_k \cup \{b^{k+1}a\}$ if $w^{(F)} = b$.*

Proof. Let $w = \psi(v)xy$ be a proper standard Sturmian word of index k. We suppose that $w^{(F)} = v^{(F)} = a$. The case $w^{(F)} = v^{(F)} = b$ is symmetrically dealt with. If v is constant, i.e., $v = a^k$, then $vab = a^{k+1}b$ and $vba = a^k ba \in X'_k$ and the result is achieved. Let us then assume that v is not constant. By Lemma 4.2 one has that $a\psi(v)b \in a^{k+1}b\{a^k b, a^{k+1}b\}^*$, so that, as $\psi(v)$ is a palindrome,

$$\psi(v) \in a^k b\{a^k b, a^{k+1}b\}^* a^k.$$

As is readily verified $a^k b\{a^k b, a^{k+1}b\}^* \subseteq X'^*_k$, so that $\psi(v) \in X'^*_k a^k$. Hence, $\psi(v)ba \in X'^*_k a^k ba \subseteq X'^*_k$. As $\psi(v)$ terminates with $a^k ba^k$ it follows that $\psi(v)ab \in X'^*_k a^k ba^{k+1} b = X'^*_k (a^k ba) a^k b \subseteq X'^*_k$, which concludes the proof. □

If w is a proper standard Sturmian word, we can introduce a derivative of w as follows. For each $k \geq 0$ if $w^{(F)} = a$ we consider the code X'_k and the injective endomorphism $\mu_k = \mu_{a^k b}$ of $\mathcal{A}^*$ defined by

$$\mu_k(a) = a^k ba, \quad \mu_k(b) = a^k b. \tag{25}$$

By the previous lemma if $w \in X'^*_k$, we define the derivative Dw of w equal to the derivative $D_k w$ with respect to μ_k, i.e., $D_k w = \mu_k^{-1}(w)$. If $w = a^{k+1}b$, we define $Da^{k+1}b = D_{k+1}a^{k+1}b = b$. Let us observe that from the definition for all $k \geq 0$ one has $D_k a^k ba = a$.

If $w^{(F)} = b$, we consider the code Y'_k and the injective endomorphism $\hat{\mu}_k = \mu_{b^k a}$ of $\mathcal{A}^*$ defined by

$$\hat{\mu}_k(a) = b^k a, \quad \hat{\mu}_k(b) = b^k ab. \tag{26}$$

By the previous lemma if $w \in Y'^*_k$ we define the derivative Dw of w equal to the derivative $\hat{D}_k w$ with respect to $\hat{\mu}_k$, i.e., $\hat{D}_k w = \hat{\mu}_k^{-1}(w)$. If $w = b^{k+1}a$ we define $Db^{k+1}a = \hat{D}_{k+1}b^{k+1}a = a$. Observe that for all $k \geq 0$ one has $\hat{D}_k b^k ab = b$.

Finally, observe that if $k = 0$, i.e., $w = ba$ or $w = ab$, from the previous definition one has $Dba = a$ and $Dab = b$.

Example 6.2. Let $v = ab^2a^2b$ and w be the standard word $\psi(v)ba$ where $\psi(v)$ is a central word of index 1. One has:

$$w = ababaababaabababaababaababababa.$$

In this case one has $X'_1 = \{aba, ab\}$ and $Dw = D_1 w = (babab babab)ba = \psi(ba^2b)ba$. Similarly, one has $D\psi(v)ab = \psi(ba^2b)ab$.

If $w = \psi(b^2a^2)ba$, then $w = (bbabbabb)ba$. The index of w is 2 and $Y'_2 = \{b^2a, b^2ab\}$ and $Dw = \hat{D}_2 w = aba$. Similarly, one has $D\psi(b^2a^2)ab = aab$.

Let us recall (cf. [13, 27]) that an endomorphism f of $\mathcal{A}^*$ is called a *standard Sturmian morphism* if the image $f(s)$ of any finite or infinite standard Sturmian word s is a standard Sturmian word. This implies that if the image $f(s)$ of a binary word $s \in \mathcal{A}^{\infty}$ is a standard Sturmian word so is s. As is well-known standard Sturmian morphisms form a monoid generated by the morphisms μ_a, μ_b, and E. Hence, for each $k \geq 0$, $\mu_k, \hat{\mu}_k \in \{\mu_a, \mu_b\}^*$ are standard Sturmian morphisms called *pure*.

Theorem 6.3. *Let $k > 0$ and $w \in X'^*_k \cup Y'^*_k \cup \{a^{k+1}b\} \cup \{b^{k+1}a\}$. Then w is a proper standard Sturmian word if and only if Dw is a standard Sturmian word.*

Proof. (⇒) We shall suppose without loss of generality that $w^{(F)} = a$, so that, as w is a proper standard Sturmian word, $w \in X_k'^* \cup \{a^{k+1}b\}$. If $w \in X_k'^*$, then $Dw = \mu_k^{-1}(w)$. Since μ_k is a standard Sturmian morphism, it follows that $Dw \in$ Stand . Similarly, if $w = a^{k+1}b$, then $Dw = \mu_{k+1}^{-1}(a^{k+1}b) = b \in$ Stand.

(⇐) Let us now suppose that $Dw \in$ Stand. If $w \in X_k'^*$, then, as μ_k is a standard Sturmian morphism, $\mu_k(Dw) = w$ is a standard and proper Sturmian word. Similarly, if $w = a^{k+1}b$ one has $\mu_{k+1}(Dw) = \mu_{k+1}(b) = w$ which is a proper standard Sturmian word. □

The following theorem relates, through their directive words, the central word of a proper standard word and the central word of its derivative.

Theorem 6.4. *Let $w = \psi(v)xy$, with $\{x, y\} = \{a, b\}$ and v a non-constant word. Then*

$$D\psi(v)xy = \psi({}_+v)xy.$$

Proof. Let $v = x_0^{\alpha_0} \cdots x_n^{\alpha_n}$ with $n \geq 1$ as v is not constant. We can write, setting $\alpha_0 = k$ and $\alpha_1 = h$, $v = x_0^k x_1^h \xi$. We shall first suppose that $\psi(v) \in \mathrm{PER}_a$, so that $v = a^k b^h \xi$. Since ${}_+v = b^{h-1}\xi$, by using the Justin formula we can write:

$$\psi(v) = \psi(a^k b^h \xi) = \mu_{a^k b}(\psi(b^{h-1}\xi))\psi(a^k b) = \mu_{a^k b}(\psi({}_+v))a^k ba^k,$$

so that if $x = b$ and $y = a$

$$w_1 = \psi(v)ba = \mu_{a^k b}(\psi({}_+v))(a^k b)(a^k ba),$$

and if $x = a$ and $y = b$

$$w_2 = \psi(v)ab = \mu_{a^k b}(\psi(({}_+v))(a^k ba)(a^k b).$$

Hence,

$$Dw_1 = \psi({}_+v)ba,$$

and

$$Dw_2 = \psi({}_+v))ab.$$

which concludes the proof in the case $\psi(v) \in \mathrm{PER}_a$. The case $\psi(v) \in \mathrm{PER}_b$ can be proved in a similar way. □

Example 6.5. Let $w = uba$ with $u = \psi(a^2b^2a)$. One has

$$w = aabaabaaabaabaaba$$

and $Dw = babba = \psi(ba)ba$. If $w = \psi(ba^2b^2a)ab$, one easily obtains $Dw = \psi(ab^2a)ab$. If $w = \psi(abab)ba$, one derives $Dw = \psi(ab)ba$.

Corollary 6.6. *Let w be the standard word $w = \psi(v)ba$ where v is not constant and w' is the Christoffel word $w' = a\psi(v)b$. Then*

$$Dw = D\psi(v)ba = a^{-1}\partial(a\psi(v)b)a = a^{-1}(\partial w')a.$$

Proof. By the preceding theorem $D\psi(v)ba = \psi({}_+v)ba$. By Theorem 4.13 one has $\partial a\psi(v)b = a\psi({}_+v)b$. From this the result follows. □

Remark 6.7. The preceding corollary holds true also for the constant words a^k, $k \geq 0$. Indeed, $Da^kba = a = \partial a^{k+1}b$. However, it is not more true for b^k, $k > 0$. Indeed, $Db^{k+1}a = a$ whereas $\partial ab^{k+1} = b$.

From Theorem 6.3 any proper standard Sturmian word w has a derivative $w' = Dw \in$ Stand. Therefore, if w' is proper one can consider $Dw' \in$ Stand that we shall denote D^2w. In general, for any $p \geq 1$, D^pw will denote the derivative of order p of w. Since $|D^pw| > |D^{p+1}w|$, there exists an integer d such that $D^dw \in \mathcal{A}$; we call d the *depth* of the standard word w.

Example 6.8. Let w be the standard word $w = \psi(a^2b^2a)ba$ of Example 6.5. One has $Dw = babba = \psi(ba)ba$, $D^2w = ba$, and $D^3w = a$. Thus the depth of w is $d = 3$.

Proposition 6.9. *The depth of a standard word $w = \psi(v)xy$ with $\{x, y\} = \{a, b\}$ is equal to the depth of the Christoffel word $a\psi(v)b$.*

Proof. Let $n = h(v)$ be the height of v. The result is trivially true if v is constant or, equivalently, if $n = 1$. Let us then suppose that v is not constant. From Corollary 6.4 one derives that for all $p \leq n$

$$D^{p-1}\psi(v)xy = \psi(v_{(p)})xy.$$

Hence, $D^{n-1}\psi(v)xy = \psi(v_{(n)})xy$. Since $v_{(n)}$ is constant, one has $\psi(v_{(n)}) = v_{(n)}$ and $Dv_{(n)}xy \in \mathcal{A}$. Thus, the depth of w is equal to $n = h(v)$ and by Proposition 5.2 is equal to the depth of $a\psi(v)b$. □

Let s be now a characteristic, or infinite standard, Sturmian word. As we have seen in Sect. 2,

$$s = \psi(v) \text{ with } v \in \mathcal{A}^\omega \setminus \mathcal{A}^*(a^\omega \cup b^\omega).$$

Any word $v \in \mathcal{A}^\omega \setminus \mathcal{A}^*(a^\omega \cup b^\omega)$ can be uniquely represented as:

$$v = x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{n-1}^{\alpha_{n-1}} x_n^{\alpha_n} \cdots \qquad (27)$$

where for $i \geq 0$, $\alpha_i \geq 1$, $x_i \in \mathcal{A}$, and $x_{i+1} = \overline{x_i}$.

We define *index* of the characteristic word $s = \psi(v)$ the first exponent in the representation (27) of v, i.e., α_0. We let ind(s) denote the index of s.

Lemma 6.10. *Let s be a characteristic Sturmian word of index k. Then $s \in X_k'^\omega$ if $s^{(F)} = a$ and $s \in Y_k'^\omega$ if $s^{(F)} = b$.*

Proof. Let $s = \psi(v)$. We first suppose that $s^{(F)} = a$. Since s has index k, we can write $v = a^k b v'$ with $v' \in \mathcal{A}^\omega$. By Lemma 2.4 one has:

$$s = \psi(a^k b v') = \mu_{a^k b}(\psi(v')).$$

From (25), it follows that $s \in X_k'^\omega$. In a similar way one proves that $s \in Y_k'^\omega$ if $s^{(F)} = b$. □

We can now define the derivative Ds of a characteristic Sturmian word s of index k as follows:

$$Ds = \mu_k^{-1}(s) \text{ if } s^{(F)} = a, \ Ds = \hat{\mu}_k^{-1}(s) \text{ if } s^{(F)} = b.$$

Remark 6.11. As one easily verifies, Ds is word isomorphic to the derived word of s in the sense of Durand [22] constructed by factoring s in terms of the *first returns* to the prefix of length $k + 1$ of s. If $s^{(F)} = a$ (resp., $s^{(F)} = b$) the set of first returns to the prefix $a^k b$ (resp., $b^k a$) of s is $\{a^k b, a^k ba\}$ (resp., $\{b^k a, b^k ab\}$). We mention that a further notion of derivative for infinite words admitting a prefixal factorization, such as the characteristic Sturmian words, has been recently given in [20].

Theorem 6.12. *Let $s = Dt$ with $t \in X_k'^\omega \cup Y_k'^\omega$. Then s is a characteristic Sturmian word if and only if so is t.*

Proof. The result is an immediate consequence of the fact that the morphisms μ_k and $\hat{\mu}_k$ are standard Sturmian morphisms. □

Recall that an infinite word $v \in \mathcal{A}^\omega$ is *constant* if $v = x^\omega$ with $x \in \mathcal{A}$. If v is not constant one can consider the greatest suffix $_+v$ of v with respect to the suffixal ordering, which is immediately preceded by a letter different from $v^{(F)}$.

Theorem 6.13. *Let $w = \psi(v)$ be a characteristic Sturmian word. Then*

$$D\psi(v) = \psi(_+v).$$

Proof. Let us write v as $v = x_0^{\alpha_0} x_1^{\alpha_1}\xi$ and set $k = \alpha_0$ and $h = \alpha_1$. We first suppose that $w^{(F)} = a$. We can write $v = a^k b^h \xi$. Hence, $\psi(v) = \psi(a^k b^h \xi)$. By Lemma 2.4 we can write:

$$\psi(v) = \psi(a^k b b^{h-1}\xi) = \mu_{a^k b}(\psi(b^{h-1}\xi)) = \mu_k(\psi(b^{h-1}\xi)).$$

Hence, $D\psi(v) = \psi(b^{h-1}\xi) = \psi({}_+v)$. If $w^{(F)} = b$ one has $v = b^k a^h \xi$. We can write:

$$\psi(v) = \psi(b^k a^h \xi) = \psi(b^k a a^{h-1}\xi) = \mu_{b^k a}(\psi(a^{h-1}\xi)) = \hat{\mu}_k(\psi(a^{h-1}\xi)).$$

Hence, $D\psi(v) = \psi(a^{h-1}\xi) = \psi({}_+v)$. □

Let $s = \psi(v)$ be a characteristic Sturmian word. From Theorems 6.12 and 6.13, $Ds = \psi({}_+v)$ is a characteristic Sturmian word, so that ${}_+v$ is not constant. We can consider the infinite sequence $(D^p s)_{p\geq 0}$ of successive derivatives of s where

$$D^0 s = s \ \text{ and } D^p s = D(D^{p-1}s), \ \text{for } p > 0.$$

Similarly to the finite case, one can introduce a sequence of infinite words $(v_{(n)})_{n>0}$, where $v_{(1)} = v$, and for all $n \geq 1$, $v_{(n+1)} = {}_+(v_{(n)})$. If $s = \psi(v)$, then by the preceding theorem one has for each $p \geq 0$, $D^p s = \psi(v_{(p+1)})$ having for all $k > 0$, $v = u_k v_{(k)}$ with $u_k \in \mathcal{A}^*$ and $|u_k| < |u_{k+1}|$.

We say that a characteristic Sturmian word s is *stable* if there exist $m, n \geq 0$, $m \neq n$, such that $D^m s = D^n s$, i.e., $\mathrm{card}\{D^m s \mid m \geq 0\} < \infty$.

Theorem 6.14. *A characteristic Sturmian word s is stable if and only if its directive word is ultimately periodic.*

Proof. (⇒) Let m be the first integer > 0 such that there exists $n > m$ for which $D^{m-1}s = D^{n-1}s$. Hence, $\psi(v_{(m)}) = \psi(v_{(n)})$. Since ψ is injective, it follows $v_{(m)} = v_{(n)}$ and therefore $v = u_m v_{(m)} = u_n v_{(n)} = u_n v_{(m)}$. As $|u_m| < |u_n|$ one has $u_n = u_m \zeta$ with $\zeta \in \mathcal{A}^+$ and $v_{(m)} = \zeta v_{(m)}$, so that $v_{(m)}$ is the periodic word $v_{(m)} = \zeta^\omega$ and $v = u_m \zeta^\omega$.

(⇐) Suppose that $s = \psi(v)$ with $v = pq^\omega$, $p, q \in \mathcal{A}^*$, and $q \neq \varepsilon$. There exists an integer k such that for all $j > k$, $D^{j-1}s = \psi(v_{(j)})$ with $v_{(j)}$ suffix of q^ω. Hence, $v_{(j)} = q_j^\omega$, where q_j is a conjugate of q. By the pigeonhole principle it follows that there exist two distinct integers $m, n > k$ such that $q_n = q_m$ and therefore $v_{(m)} = v_{(n)}$. This implies $D^{n-1}s = D^{m-1}s$. □

Example 6.15. Let $f = \psi((ab)^\omega)$ be the Fibonacci word. One has that for all $p \geq 1$, $D^p f = f$, so that f is stable. Let $s = \psi(a^k(ab)^\omega)$ where k is a fixed integer ≥ 1. One has $Ds = \psi((ab)^\omega) = f$. Thus $Ds = D^p s = f$ for all $p \geq 1$ and s is stable. Let $s = \psi(aba^2ba^3b\cdots ba^nb\cdots)$. For any $p > 0$ one has $D^p s = \psi(a^{p+1}ba^{p+2}ba^{p+3}\cdots)$, so that s is not stable.

Let $v = x_0^{\alpha_0} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{n-1}^{\alpha_{n-1}} x_n^{\alpha_n} \cdots$ and $s = \psi(v)$ be the characteristic Sturmian word with the directive word v. The slope of s is the limit $\lim_{n\to\infty} \frac{|s_{[n]}|_b}{|s_{[n]}|_a}$. As is well-known (cf. [3, 27]), since s is Sturmian, this limit exists and is an irrational number equal to the continued fraction

$$[\alpha_0; \alpha_1, \ldots, \alpha_n, \ldots].$$

One can easily prove that the directive word v is periodic if and only if there exist integers $r > 0$ and $q \geq 0$ such that $\alpha_n = \alpha_{n+r}$ for all $n \geq q$, or, equivalently, the previous continued fraction is periodic. From Theorem 6.14 and [2, Theorem 20], one derives that *a characteristic Sturmian word s is stable if and only if the set of all derivated words in the sense of Durand (with respect to prefixes of s) is finite.*

For each $k \geq 0$, let X_k and Y_k be the sets defined by (14).

Lemma 6.16. *Let s be a characteristic Sturmian word of index k. Then* $s \in X_k^\omega$ *if* $s^{(F)} = a$ *and* $s \in Y_k^\omega$ *if* $s^{(F)} = b$.

Proof. Let us suppose $s^{(F)} = a$. As one readily verifies, for each $k \geq 0$ one has $X_k'^\omega = \{a^kba, a^kb\}^\omega = a^kb\{a^kb, a^{k+1}b\}^\omega = a^kbX_k^\omega$. By Lemma 6.10, one has $s \in X_k'^\omega$, so that $s \in X_k^\omega$. The case $s^{(F)} = b$ is dealt with in a similar way. □

We can define the derivative ∂s of a characteristic Sturmian word s of index k by

$$\partial s = \varphi_k^{-1}(s) \text{ if } \ s^{(F)} = a \ \text{ and } \partial s = \hat{\varphi}_k^{-1}(s) \text{ if } \ s^{(F)} = b.$$

Lemma 6.17. *Let $x = x_1x_2 \cdots x_n \cdots$ be an infinite word over $\mathcal{A}$. Then for each $k \geq 0$*

$$\varphi_k^{-1}(\mu_k(x)) = bx.$$

Proof. By Lemma 3.7, one has for all $n \geq 1$

$$\varphi_k^{-1}(\mu_k(x_{[n]}b)) = \varphi_k^{-1}(\mu_k(x_{[n]})a^k b) = \varphi_k^{-1}(\mu_k(x_{[n]})b = bx_{[n]}.$$

Thus,

$$\varphi_k^{-1}(\mu_k(x_{[n]}) = bx_{[n]}b^{-1}$$

and

$$\varphi_k^{-1}(\mu_k(x)) = \lim_{n\to\infty} \varphi_k^{-1}(\mu_k(x_{[n]}) = \lim_{n\to\infty} bx_{[n]}b^{-1} = bx.$$ □

Theorem 6.18. *Let s be a characteristic Sturmian word. Then*

$$\partial s = bDs.$$

Proof. If s is a characteristic Sturmian word of index k, then by the preceding lemma one has:

$$\partial(\mu_k(Ds)) = \partial s = bDs.$$ □

7. Concluding remarks

We have studied new combinatorial properties of Christoffel, central, and standard words, which are related to a suitable notion of derivative of a word. In this analysis, the palindromization map that allows to construct all central words, as well as all infinite standard words, plays an essential role. Indeed, it allows one to give a unified treatment for the previous classes of words. Moreover, one can make use of the important combinatorial tool represented by Justin's formula which links the palindromization map with pure standard morphisms. By this palindromization map, from one side one can obtain a very simple formula giving the derivative of a Christoffel word. From the other one can extend the previous results to the case of standard words. Finally, new interesting combinatorial problems arose from considering higher order derivatives and the depth of a Christoffel word and of a standard word. This gives a new insight on these noteworthy classes of words.

An interesting open problem is to try to extend some of the previous results to the case of alphabets with more than two letters, i.e., to the case of standard episturmian words. This extension seems to be quite hard since some basic combinatorial properties hold only for a binary alphabet.

Acknowledgments

We would like to thank the anonymous referees for many valuable comments and suggestions.

References

[1] J.-P. Allouche, J. Shallit, Automatic Sequences, Cambridge University Press, Cambridge UK, 2003.
[2] I. Araújo, V. Bruyere, Words derivated from Sturmian words, Theoret. Comput. Sci. 340 (2005) 204–219.
[3] J. Berstel, A. de Luca, Sturmian words, Lyndon words and trees, Theoret. Comput. Sci. 178 (1997) 171–203.
[4] J. Berstel, A. Lauve, C. Reutenauer, F. Saliola, Combinatorics on Words: Christoffel Words and Repetition in Words, volume 27 of *CRM monograph series*, American Mathematical Society, Providence, RI, 2008.
[5] J. Berstel, D. Perrin, Theory of Codes, Academic Press, New York, 1985.
[6] V. Berthé, A. de Luca, C. Reutenauer, On an involution of Christoffel words and Sturmian morphisms, European J. Combin. 29 (2008) 535–553.
[7] J.-P. Borel, F. Laubie, Quelques mots sur la droite projective réelle, J. Théor. Nombres Bordeaux 5 (1993) 23–51.
[8] M. Bucci, A. de Luca, A. De Luca, Characteristic morphisms of generalized episturmian words, Theoret. Comput. Sci. 410 (2009) 2840–2859.
[9] A. Carpi, A. de Luca, Codes of central Sturmian words, Theoret. Comput. Sci. 340 (2005) 220–239.
[10] E. B. Christoffel, Observatio arithmetica, Ann. Mat. Pur. Appl. 6 (1875) 148–152.

[11] N. G. de Bruijn, Sequences of zeros and ones generated by special production rules, Nederl. Akad. Wetensch. Indag. Math. 43 (1981) 27–37.
[12] N. G. de Bruijn, Updown generation of Beatty sequences, Nederl. Akad. Wetensch. Indag. Math. 51 (1989) 385–407.
[13] A. de Luca, Standard Sturmian morphisms, Theoret. Comput. Sci. 178 (1997) 205–224.
[14] A. de Luca, Sturmian words: structure, combinatorics, and their arithmetics, Theoret. Comput. Sci. 183 (1997) 45–82.
[15] A. de Luca, A standard correspondence on epicentral words, European J. Combin. 33 (2012) 1514–1536.
[16] A. de Luca, A. De Luca, Sturmian words and the Stern sequence, Theoret. Comput. Sci. 581 (2015) 26–44.
[17] A. de Luca, F. Mignosi, Some combinatorial properties of Sturmian words, Theoret. Comput. Sci. 136 (1994) 361–385.
[18] A. de Luca, L. Q. Zamboni, On graphs of central episturmian words, Theoret. Comput. Sci. 411 (2010) 70–90.
[19] A. de Luca, L. Q. Zamboni, Involutions of epicentral words, European J. Combin. 31 (2010) 867–886.
[20] A. de Luca, L. Q. Zamboni, On prefixal factorizations of words, European J. Combin., 52 (2016) 59–73.
[21] X. Droubay, J. Justin, G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (2001) 539–553.
[22] F. Durand, A characterization of substitutive sequences using return words, Discrete Math. 179 (1998) 89–101.
[23] R.L. Graham, D.E. Knuth, O. Patashnik, Concrete Mathematics, 2nd ed., Addison-Wesley, Reading Mass., 1994.
[24] J. Justin, Episturmian morphisms and a Galois theorem on continued fractions, Theor. Inform. Appl. 39 (2005) 207–215.
[25] R. Klette, A. Rosenfeld, Digital straightness — a review, Discrete Appl. Math. 139 (2004) 197–230.
[26] M. Lothaire, Combinatorics on Words, Addison-Wesley, Reading MA, 1983. Reprinted by Cambridge University Press, Cambridge UK, 1997.
[27] M. Lothaire, Algebraic Combinatorics on Words, volume 90 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge UK, 2002.
[28] M. Morse, G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, Amer. J. Math. 62 (1940) 1–42.
[29] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, `http://oeis.org/`, 2011.