Glen, A. (2006) Occurrences of palindromes in characteristic Sturmian words. Theoretical Computer Science, 352 (1-3). pp. 31-46.

http://researchrepository.murdoch.edu.au/3880

Occurrences of palindromes in characteristic Sturmian words

Amy Glen*

June 24, 2005

School of Mathematical Sciences, Discipline of Pure Mathematics, University of Adelaide,
South Australia, Australia, 5005

Abstract

This paper is concerned with palindromes occurring in characteristic Sturmian words c_α of slope α, where $\alpha \in (0,1)$ is an irrational. As c_α is a uniformly recurrent infinite word, any (palindromic) factor of c_α occurs infinitely many times in c_α with bounded gaps. Our aim is to completely describe where palindromes occur in c_α. In particular, given any palindromic factor u of c_α, we shall establish a decomposition of c_α with respect to the occurrences of u. Such a decomposition shows precisely where u occurs in c_α, and this is directly related to the continued fraction expansion of α.

Keywords: Combinatorics on words; Characteristic Sturmian word; Singular word; Palindrome; Morphism; Return word; Overlap.

2000 Mathematical Subject Classifications: primary 68R15; secondary 11B85.

---

1 Introduction

The fascinating family of Sturmian words consists of all aperiodic infinite words having exactly $n+1$ distinct factors of length n for each $n \in \mathbb{N}$. Such words have many applications in various fields of mathematics, such as symbolic dynamics, the study of continued fraction expansion, and also in some domains of physics (crystallography) and computer science (formal language theory, algorithms on words, pattern recognition). Sturmian words admit several equivalent definitions and have numerous characterizations; in particular, they can be characterized by their palindrome or return word structure [10, 16]. For a comprehensive introduction to Sturmian words, see for instance [1, 2, 23] and references therein.

Sturmian words have exactly two factors of length 1, and thus are infinite sequences over a two-letter alphabet $\mathcal{A} = \{a, b\}$, say. Here, an *infinite word* (or *sequence*) $\mathbf{x}$ over $\mathcal{A}$ is a map $\mathbf{x} : \mathbb{N} \to \mathcal{A}$. For any $i \geq 0$, we set $x_i = \mathbf{x}(i)$ and write $\mathbf{x} = x_0x_1x_2\cdots$, each $x_i \in \mathcal{A}$. Central to our study is the following characterization of Sturmian words, which was originally proved by Morse and Hedlund [21]. An infinite word $\mathbf{s}$ over $\mathcal{A} = \{a, b\}$ is Sturmian if and only if there exists an irrational $\alpha \in (0,1)$, and a real number ρ, such that $\mathbf{s}$ is equal to one of the following two infinite words:

$$s_{\alpha,\rho},\ s'_{\alpha,\rho} : \mathbb{N} \to \mathcal{A}$$

defined by

$$\begin{aligned} s_{\alpha,\rho}(n) &= \begin{cases} a & \text{if } \lfloor (n+1)\alpha + \rho \rfloor - \lfloor n\alpha + \rho \rfloor = 0, \\ b & \text{otherwise;} \end{cases} \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (n \geq 0) \\ s'_{\alpha,\rho}(n) &= \begin{cases} a & \text{if } \lceil (n+1)\alpha + \rho \rceil - \lceil n\alpha + \rho \rceil = 0, \\ b & \text{otherwise.} \end{cases} \end{aligned}$$

---

[*] E-mail: `amy.glen@adelaide.edu.au`

The irrational α is called the *slope* of $\mathbf{s}$ and ρ is the *intercept.* If $\rho = 0$, we have

$$s_{\alpha,0} = ac_\alpha \quad \text{and} \quad s'_{\alpha,0} = bc_\alpha,$$

where c_α is called the *characteristic Sturmian word* of slope α (see [2]).

Our focus will be on palindromic factors of c_α. In general terms, a *palindrome* is a finite word that reads the same backwards as forwards. Palindromes are important tools used in the study of factors of Sturmian words (e.g., [6, 8, 9, 10]), and they have also become objects of great interest in computer science. The aim of this current paper is to completely describe where palindromes occur in c_α (and hence $s_{\alpha,0}$, $s'_{\alpha,0}$). In order to do this, we shall make use of some previous results concerning factorizations of c_α into *singular words*, which are particular palindromes. Singular words were first defined for the Fibonacci word $\mathbf{f}$ (a special example of a Sturmian word) by Wen and Wen [25], who established a decomposition of $\mathbf{f}$ with respect to such words. This result was later extended by Melançon [19] to characteristic Sturmian words. More recently, Levé and Séébold [17] have generalized Wen and Wen's 'singular' decomposition of $\mathbf{f}$, by establishing a similar decomposition for each *conjugate* of $\mathbf{f}$ into what they called *generalized singular words.* This last result has now been further extended by the present author [14] to c_α (and $c_{1-\alpha}$), where α has continued fraction expansion $[0; 2, r, r, r, \ldots]$ for some $r \geq 1$.

It is well-known that any Sturmian word $\mathbf{s}$ is *uniformly recurrent*, i.e., any factor of $\mathbf{s}$ occurs infinitely often in $\mathbf{s}$ with bounded gaps [5]. Accordingly, any palindromic factor u of c_α has infinitely many occurrences in c_α and, as we shall see later (Corollary 5.2), the distance between any two adjacent occurrences of u is bounded above by an integer depending on u. Given any palindromic factor u of c_α, we shall establish a decomposition of c_α with respect to the occurrences of u. Such a decomposition shows precisely at which positions u occurs in c_α, and this is directly related to the continued fraction expansion of the irrational slope α.

This paper is organized as follows. In Section 2, after some preliminaries on words and morphisms, we will recall some facts about c_α and consider some of its singular decompositions (Section 2.2). Then, in Section 3, we consider the structure of palindromic factors of c_α with respect to its singular factors. We also recall the important notion of a return word and the concept of overlapping occurrences of a word in c_α. Section 4 contains the lemmas we need in order to establish the main result of this paper, which appears in Section 5. Lastly, using results of Section 4, we obtain decompositions of c_α that show precisely where a given factor of length q_n occurs in c_α (where q_n is the denominator of the n-th convergent to $\alpha = [0; 1 + d_1, d_2, d_3, \ldots]$, $d_i \geq 1$).

2 Preliminaries

Any of the following terminology that is not further clarified can be found in either [18] or [2], which give more detailed presentations.

2.1 Words and morphisms

In what follows, let $\mathcal{A}$ denote the two-letter alphabet $\{a, b\}$. A (finite) *word* is an element of the free monoid $\mathcal{A}^*$ generated by $\mathcal{A}$, in the sense of concatenation. The identity ε of $\mathcal{A}^*$ is called the *empty word*, and the *free semigroup* over $\mathcal{A}$ is defined by $\mathcal{A}^+ := \mathcal{A}^* \setminus \{\varepsilon\}$. We denote by $\mathcal{A}^\omega$ the set of all infinite words over $\mathcal{A}$, and define $\mathcal{A}^\infty := \mathcal{A}^* \cup \mathcal{A}^\omega$. The *length* $|w|$ of a finite word w is defined to be the number of letters it contains. (Note that $|\varepsilon| = 0$.)

A finite word z is a *factor* of a word $w \in \mathcal{A}^\infty$ if $w = uzv$ for some $u \in \mathcal{A}^*$ and $v \in \mathcal{A}^\infty$. Furthermore, z is called a *prefix* (resp. *suffix*) of w if $u = \varepsilon$ (resp. $v = \varepsilon$), and we write $z \subseteq_p w$ (resp. $z \subseteq_s w$). The word z is said to have an *occurrence* (or occur) at position $|u|$ of $w = uzv$, i.e., z begins at the $|u|$-th position of w. We denote by $|w|_z$ the number of occurrences of z in w, i.e., the number of distinct positions at which z occurs in w. For example, $|ababa|_{aba} = 2$ since aba has two occurrences at positions 0 and 2 in $ababa$.

For any word $w \in \mathcal{A}^\infty$, $\Omega(w)$ denotes the set of all factors of w. Moreover, we denote by $\Omega_n(w)$ the set of all factors of w of length $n \in \mathbb{N}$ (where $n \leq |w|$ for w finite), i.e., $\Omega_n(w) = \Omega(w) \cap \mathcal{A}^n$. If $u \in \Omega(w)$, then we shall simply write $u \prec w$.

The *reversal operation* $\sim$ in $\mathcal{A}^*$ is defined inductively by: $\widetilde{\varepsilon} = \varepsilon$ and, for any $u \in \mathcal{A}^*$ and $x \in \mathcal{A}$, $(\widetilde{ux}) = x\widetilde{u}$. Thus, if $w = x_0x_1x_2\ldots x_n$, with each $x_i \in \mathcal{A}$, then $\widetilde{w} = x_nx_{n-1}\ldots x_1x_0$. If $w = \widetilde{w}$, then w is called a *palindrome*, and we define PAL to be the set of all palindromes over $\mathcal{A}$. It is useful to note that if $|w|$ is even, then w is a palindrome if and only if $w = v\widetilde{v}$ for some word v. Otherwise, w is a palindrome if and only if $w = vx\widetilde{v}$ for some word v and some letter $x \in \mathcal{A}$.

The free monoid $\mathcal{A}^*$ can be naturally embedded within a *free group*. We shall denote by $\mathcal{F}$ the free group generated by $\mathcal{A}$, which contains the *inverse* u^{-1} of each word $u \in \mathcal{A}^*$. For any u, $v \in \mathcal{F}$, we have $uu^{-1} = u^{-1}u = \varepsilon$ and $(uv)^{-1} = v^{-1}u^{-1}$. If u, $w \in \mathcal{A}^*$, we shall write $u^{-1}w$ (resp. wu^{-1}) only if u is a prefix (resp. suffix) of w, so that $u^{-1}w$ (resp. wu^{-1}) is a word in $\mathcal{A}^*$. In particular, if $w = uv \in \mathcal{A}^*$, then $u^{-1}w = v$ and $wv^{-1} = u$, and we have $|u^{-1}w| = |w| - |u| = |v|$, $|wv^{-1}| = |w| - |v| = |u|$.

An *endomorphism* (or simply *morphism*) of $\mathcal{A}^*$ is a map $\psi : \mathcal{A}^* \to \mathcal{A}^*$ such that $\psi(uv) = \psi(u)\psi(v)$ for all $u, v \in \mathcal{A}^*$. It is uniquely determined by its image on the alphabet $\mathcal{A}$. Any morphism ψ of $\mathcal{A}^*$ can be uniquely extended to an endomorphism of $\mathcal{F}$ by defining $\psi(a^{-1}) = (\psi(a))^{-1}$ and $\psi(b^{-1}) = (\psi(b))^{-1}$, from which it follows that $\psi(w^{-1}) = (\psi(w))^{-1}$ for any $w \in \mathcal{F}$.

2.1.1 Standard morphisms

Define the following two morphisms of $\mathcal{A}^*$:

$$E : \begin{array}{ccc} a & \mapsto & b \\ b & \mapsto & a \end{array}, \qquad \varphi : \begin{array}{ccc} a & \mapsto & ab \\ b & \mapsto & a \end{array}.$$

A morphism ψ of $\mathcal{A}^*$ is *standard* if $\psi(\mathbf{x})$ is a characteristic Sturmian word for any characteristic Sturmian word $\mathbf{x}$ [2]. In fact, a morphism ψ is standard if and only if $\psi \in \{E, \varphi\}^*$, i.e., if and only if it is a composition of E and φ in any number and order [7, 2]. The standard morphisms E and φE will play an important role in the proof of our main result.

2.2 Characteristic Sturmian words c_α and singular words

Note that every irrational $\alpha \in (0, 1)$ has a unique continued fraction expansion

$$\alpha = [0; a_1, a_2, a_3, \ldots] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

where each a_i is a positive integer. If the sequence $(a_i)_{i\geq1}$ is eventually periodic, with $a_i = a_{i+m}$ for all $i \geq n$, we use the notation $\alpha = [0; a_1, a_2, \ldots, a_{n-1}, \overline{a_n, a_{n+1}, \ldots, a_{n+m-1}}]$. The *n-th convergent* to α is defined by

$$\frac{p_n}{q_n} = [0; a_1, a_2, \ldots, a_n], \quad \text{for all } n \geq 1,$$

where the sequences $(p_n)_{n\geq0}$ and $(q_n)_{n\geq0}$ are given by

$$\begin{array}{llll} p_0 = 0, & p_1 = 1, & p_n = a_np_{n-1} + p_{n-2}, & n \geq 2; \\ q_0 = 1, & q_1 = a_1, & q_n = a_nq_{n-1} + q_{n-2}, & n \geq 2. \end{array}$$

Suppose $\alpha = [0; 1 + d_1, d_2, d_3, \ldots]$ with $d_1 \geq 0$ and all other $d_n > 0$. To the *directive sequence* $(d_1, d_2, d_3, \ldots)$, we associate a sequence $(s_n)_{n\geq-1}$ of words defined by

$$s_{-1} = b, \; s_0 = a, \; s_n = s_{n-1}^{d_n}s_{n-2}, \quad n \geq 1.$$

Such a sequence of words is called a *standard sequence*, and we have

$$|s_n| = q_n \quad \text{for all } n \geq 0.$$

Note that ab is a suffix of s_{2n-1} and ba is a suffix of s_{2n}, for all $n \geq 1$.

Standard sequences are related to characteristic Sturmian words in the following way. Observe that, for any $n \geq 0$, s_n is a prefix of s_{n+1}, which gives obvious meaning to $\lim_{n\to\infty} s_n$ as an infinite word. In fact, each s_n is a prefix of c_α, and we have

$$c_\alpha = \lim_{n\to\infty} s_n \quad \text{(see [13, 3])}. \tag{2.1}$$

2.2.1 Some singular decompositions of c_α

Note that if $\alpha = [0; 1, d_1, d_2, d_3, \ldots]$, then

$$1 - \alpha = \frac{1}{1 + 1/(1/\alpha - 1)} = [0; 1 + d_1, d_2, d_3, \ldots]. \tag{2.2}$$

For any irrational $\alpha \in (0,1)$, $E(c_\alpha) = c_{1-\alpha}$, i.e., $c_{1-\alpha}$ is obtained from c_α by exchanging a's and b's [22]. Thus, in light of the above observation (2.2), we shall hereafter restrict our attention to the case when $\alpha = [0; 1 + d_1, d_2, d_3, \ldots]$ with $d_1 \geq 1$.

Melançon [19] (also see [4, 25]) has introduced the singular words $(w_n)_{n\geq 0}$ of c_α defined by

$$w_n = \begin{cases} as_nb^{-1} & \text{if } n \text{ is odd,} \\ bs_na^{-1} & \text{otherwise.} \end{cases}$$

Moreover, for each $n \geq -1$, Melançon [19] defined the words

$$v_n = \begin{cases} as_{n+1}^{d_{n+2}-1}s_nb^{-1} & \text{if } n \text{ is odd,} \\ bs_{n+1}^{d_{n+2}-1}s_na^{-1} & \text{otherwise.} \end{cases}$$

Clearly, the word v_n differs from w_{n+2} by a factor s_{n+1}, and it is easily proved that all v_n and w_n are palindromes. Here, we will call w_n (resp. v_n) the *n-th singular word* (resp. *n-th adjoining singular word*) of c_α, and use the convention $w_{-2} = v_{-2} = \varepsilon$, $w_{-1} = a$.

Singular words play an important role in the study of factors of Sturmian words. In particular, as we shall see in the next section, the words w_n and v_{n-1} can be used to determine the structure of all palindromic factors of a Sturmian word of slope α. We have the following decomposition of c_α in terms of singular and adjoining singular words.

Proposition 2.1. [25, 19] $c_\alpha = \prod_{j=-1}^{\infty}(v_{2j}w_{2j+1})^{d_{2j+3}} = \prod_{j=-1}^{\infty} v_j$. □

Notation. In order to simplify proceedings, we introduce some notation.

(i) Let $\gamma \in (0,1)$ be irrational with $\gamma = [0; a_1, a_2, a_3, \ldots]$. For any $n \in \mathbb{N}$ and integer k such that $k \geq 1 - a_{n+1}$, define

$$\gamma_{n,k} := [0; a_{n+1} + k, a_{n+2}, a_{n+3}, \ldots]$$

and write $\gamma_{n,0} = \gamma_n$. Note that $\gamma_{0,0} = \gamma = [0; a_1, a_2, \ldots, a_n + \gamma_n]$ for all $n \geq 1$.

(ii) As c_α is uniformly recurrent, given any factor w of c_α, the occurrences of w in c_α can be arranged as a sequence $(w^{(i)})_{i\geq 1}$, where $w^{(i)}$ denotes the i-th occurrence of w in c_α.

With the above notation, we may now state a corollary of Proposition 2.1.

Corollary 2.2. *Let $n \in \mathbb{N}$ be fixed. The characteristic Sturmian word c_α has the following two decompositions:*

(1)
$$c_\alpha = \left(\prod_{j=-1}^{n-1} (v_{2j} w_{2j+1})^{d_{2j+3}} \right) w_{2n}^{(1)} z_1 w_{2n}^{(2)} z_2 w_{2n}^{(3)} z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{2n+1}}$ over the alphabet $\{v_{2n-1}, w_{2n+1}\}$.

(2)
$$c_\alpha = \left(\prod_{j=-1}^{n-1} (v_{2j} w_{2j+1})^{d_{2j+3}} \right) z_1 W_1 z_2 W_2 z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{2n+1,1}}$ over the alphabet $\{w_{2n}, v_{2n-2}\}$ and, for all $i \geq 1$,

$$W_i = \begin{cases} v_{2n-1} & \text{if } z_i = w_{2n}, \\ w_{2n-1} & \text{if } z_i = v_{2n-2}. \end{cases}$$

Proof. See [19, Corollary 4.6]. ☐

Example 2.1. The best known example of a characteristic Sturmian word is the infinite *Fibonacci word* $\mathbf{f}$, which has been extensively studied by many authors (see [6, 9], for example). It is well-known that

$$\mathbf{f} = \lim_{n \to \infty} f_n = abaababaabaababaababaabaababaabaab \cdots ,$$

where $(f_n)_{n \geq -1}$ is the sequence of *finite Fibonacci words* defined by

$$f_{-1} = b, \ f_0 = a, \ f_n = f_{n-1} f_{n-2}, \quad n \geq 1.$$

Clearly, $|f_n| = F_n$, where F_n is the *n-th Fibonacci number* defined by

$$F_{-1} = 1, \ F_0 = 1, \ F_n = F_{n-1} + F_{n-2}, \quad n \geq 1.$$

Note that $(f_n)_{n \geq -1}$ is a standard sequence with associated directive sequence $(1, 1, 1, \ldots)$, and hence $w_n = v_n$ for all $n \geq -1$. Moreover, in view of (2.1), $\mathbf{f} = c_\alpha$ where $\alpha = (3 - \sqrt{5})/2 = [0; 2, \overline{1}]$, in which case $\alpha = \alpha_{2n+1,1}$ and $1 - \alpha = \alpha_{2n+1}$, for all $n \in \mathbb{N}$. Hence, $\mathbf{f} = c_{\alpha_{2n+1,1}}$ and $E(\mathbf{f}) = c_{\alpha_{2n+1}}$. Accordingly, one deduces from the above corollary that

$$\mathbf{f} = \left(\prod_{j=-1}^{n-1} w_j \right) w_n^{(1)} z_1 w_n^{(2)} z_2 w_n^{(3)} z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is the Fibonacci word over the alphabet $\{w_{n+1}, w_{n-1}\}$ (also see [25, Theorem 2]). For instance, when $n = 2$, $w_{n-1} = w_1 = aa$, $w_n = w_2 = bab$, $w_{n+1} = w_3 = aabaa$, and $\mathbf{z}$ is the Fibonacci word over the alphabet $\{aabaa, aa\}$. Indeed, one may write

$$\mathbf{f} = abaa(bab)aabaa(bab)aa(bab)aabaa(bab)aabaa(bab)aa(bab)aabaa(bab)aa(bab)aabaa \cdots .$$

3 Palindromes, return words and overlap

3.1 Structure of palindromes in c_α

In [4], Cao and Wen considered the structure of palindromic factors of c_α with respect to singular words. Specifically, they proved the following result concerning palindromic factors u of c_α with $q_n < |u| \leq q_{n+1}$. (For technical reasons, we set $q_{-1} = 1$, so that $|s_n| = q_n$ for all $n \geq -1$.)

Proposition 3.1. [4] *Let* $u \in \mathrm{PAL}$ *with* $q_n < |u| \leq q_{n+1}$ *for some* $n \in \mathbb{N}$. *Then* $u \prec c_\alpha$ *if and only if* u *takes one of the following forms:*

(1) $u = vw_n\widetilde{v}$ *with* $v \subseteq_s v_{n-1}$ *and* $|v| \leq \frac{1}{2}|v_{n-1}| = \frac{1}{2}(q_{n+1} - q_n)$;

(2) $u = vv_{n-1}\widetilde{v}$ *with* $v \subseteq_s w_n$ *and* $|v| \leq \frac{1}{2}q_n$;

(3) $u = v(w_{n-1}v_{n-2})^k w_{n-1}\widetilde{v}$ *with* $v \subseteq_s v_{n-2}$, $v \neq v_{n-2}$, *and* $0 \leq k \leq d_{n+1} - 2$;

(4) $u = v(v_{n-2}w_{n-1})^k v_{n-2}\widetilde{v}$ *with* $v \subseteq_s w_{n-1}$, $v \neq w_{n-1}$, *and* $0 \leq k \leq d_{n+1} - 1$;

(5) $u = w_{n+1}$.

Moreover, if $k = 0$ *in* (3) (*resp.* (4)), *then* $|v| > \frac{1}{2}|v_{n-2}| = \frac{1}{2}(q_n - q_{n-1})$ (*resp.* $|v| > \frac{1}{2}q_{n-1}$). ☐

Hereafter, we will make frequent use of the following properties of singular words. Some of these properties may be used without referring to the given lemma.

Lemma 3.2. [19, 4] *Let* $x, y \in \mathcal{A}$ $(x \neq y)$ *with* $y \subseteq_s s_n$. *Then, for any* $n \in \mathbb{N}$,

(1) $yx^{-1}w_n = ys_ny^{-1} = w_{n-1}v_{n-2}, \quad w_nx^{-1}y = v_{n-2}w_{n-1}$;

(2) $w_{n+1} = w_{n-1}v_{n-2}v_{n-1} = v_{n-1}v_{n-2}w_{n-1}$;

(3) $v_{n-1} = (w_{n-1}v_{n-2})^{d_{n+1}-1}w_{n-1}$;

(4) $w_{n+1} = (w_{n-1}v_{n-2})^{d_{n+1}}w_{n-1}$;

(5) $w_{n+1} = y\prod_{j=-1}^{n-1} v_j$;

(6) $w_n \not\prec w_{n+1}$;

(7) $v_{n-1} \not\prec w_n$.

☐

Now, for each $n \in \mathbb{N}$ and $0 \leq k \leq d_{n+1} - 1$, let us denote by $U_{n,k}$ and $\overline{U}_{n,k}$ the palindromes given by

$$U_{n,k} := (w_{n-1}v_{n-2})^k w_{n-1} \quad \text{and} \quad \overline{U}_{n,k} := (v_{n-2}w_{n-1})^k v_{n-2}.$$

Note that $U_{n,k} = w_{n-1}\overline{U}_{n,k}(v_{n-2})^{-1}$. Also observe that the singular words $(w_n)_{n\geq -1}$ and $(v_n)_{n\geq -1}$ are given by

$$w_{n-1} = U_{n,0} \quad \text{and} \quad v_{n-1} = U_{n,d_{n+1}-1} = \overline{U}_{n+1,0} \quad \text{for all } n \geq 0.$$

From the preceding proposition and Lemma 3.2, we easily deduce the following result, which gives the structure of all palindromic factors of c_α in terms of $U_{n,k}$ and $\overline{U}_{n,k}$. The proof is left to the reader.

Corollary 3.3. *Let* $u \in \mathrm{PAL}$ *with* $|u| \geq 2$. *Then* u *is a factor of* c_α *if and only if, for some* $n \in \mathbb{N}$, *we have*

$$u = vU_{n,k}\widetilde{v}, \text{ where } v \subseteq_s v_{n-2},\ v \neq v_{n-2} \text{ and } 0 \leq k \leq d_{n+1} - 2 \tag{3.1}$$

or

$$u = v\overline{U}_{n,k}\widetilde{v}, \text{ where } v \subseteq_s w_{n-1},\ v \neq w_{n-1} \text{ and } 0 \leq k \leq d_{n+1} - 1. \tag{3.2}$$

☐

Note. Let us point out that $\overline{U}_{0,0} = \varepsilon$ and $U_{0,k-1} = a^k = \overline{U}_{0,k}$ for $1 \leq k \leq d_1 - 1$. Therefore, if u takes the form (3.1) or (3.2) for $n = 0$, then $u = a^k$ for some $k \in [2, d_1 - 1]$. So, if $d_1 \leq 2$, a palindromic factor of c_α is given by (3.1) or (3.2) for some $n \geq 1$.

Remark 3.1. It is important to note that Corollary 3.3 (and also Proposition 3.1) gives the structure of all palindromic factors of any Sturmian word of slope α. Indeed, Mignosi [20] proved that any two Sturmian words $\mathbf{s}$, $\mathbf{t}$ of the same slope are *equivalent*, i.e., $\Omega(\mathbf{s}) = \Omega(\mathbf{t})$. Whence, for any real number ρ, we have

$$\Omega(s_{\alpha,\rho}) = \Omega(s'_{\alpha,\rho}) = \Omega(c_\alpha),$$

i.e., a palindrome is a factor of some Sturmian word of slope α if and only if it is a factor of c_α.

3.2 Return words and overlapping occurrences

Let us write $c_\alpha = x_0x_1x_2\cdots$, each $x_i \in \mathcal{A}$, and let $w \prec c_\alpha$. Suppose $n_1 < n_2 < n_3 < \cdots$ are all the natural numbers n_i such that $w = x_{n_i}x_{n_i+1}\cdots x_{n_i+|w|-1}$. Then the word $x_{n_i}\cdots x_{n_{i+1}-1}$ is a *return word* of w in c_α. That is, we define the set $\mathcal{R}_w(c_\alpha)$ of return words of w to be the set of all distinct words beginning with an occurrence of w and ending exactly before the next occurrence of w in c_α. This notion was introduced independently by Durand [11], and Holton and Zamboni [15]. Clearly, $\mathcal{R}_w(c_\alpha)$ is finite since the distance between two adjacent occurrences of w in c_α is bounded. In fact, Vuillon [24] has proved that an infinite word $\mathbf{s}$ over $\mathcal{A}$ is Sturmian if and only if, for any factor w of $\mathbf{s}$, there are exactly two return words of w in $\mathbf{s}$. Suppose $\mathcal{R}_w(c_\alpha) = \{u_1, u_2\}$. Then c_α can be uniquely factorized as $c_\alpha = vu_{i_1}u_{i_2}\cdots u_{i_k}\cdots$, where each $i_k \in \{1,2\}$ and the first occurrence of w in c_α is at position $|v|$. The infinite word $\mathcal{D}_w(c_\alpha) := u_{i_1}u_{i_2}\cdots u_{i_k}\cdots$, called the *derived word of* c_α *with respect to* w, can be viewed as an infinite word over the alphabet $\{u_1, u_2\}$. In particular, $\mathcal{D}_w(c_\alpha)$ is a Sturmian word over the alphabet $\mathcal{R}_w(c_\alpha)$ [12]. For example, the return words of w_n in $\mathbf{f}$ are w_nw_{n+1} and w_nw_{n-1}, and $\mathcal{D}_{w_n}(\mathbf{f})$ is the Fibonacci word over the alphabet $\{w_nw_{n+1}, w_nw_{n-1}\}$ (see Example 2.1).

Given $w \prec c_\alpha$, a return word of w in c_α is not necessarily longer than w, in which case w has overlapping occurrences in c_α. More precisely, if there exist non-empty words u, v and z such that $w = uz = zv$ and $uzv \prec c_\alpha$, then w is said to have *overlap* in c_α with *overlap factor* z. Further, one can write $uzv = wz^{-1}w$; whence w has overlap in c_α if $wz^{-1}w \prec c_\alpha$ for some $z \in \mathcal{A}^+$. In this case, wz^{-1} is a return word of c_α that has length less than that of w. Clearly, since any factor w of c_α has exactly two return words, w has at most two different overlap factors.

Return words, and the concept of overlap, are fundamentally important to our study of occurrences of palindromes in c_α. Indeed, we shall be establishing decompositions of c_α with respect to certain palindromic factors that have overlap, i.e., palindromic factors u that have a return word (or return words) of length(s) less than $|u|$. Specifically, given any palindromic factor u of c_α, we can write

$$c_\alpha = z_0u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots ,$$

where $z_0 \in \mathcal{A}^*$ and all other z_i are such that $z_i^{-1} \in \mathcal{A}^+$ or $z_i \in \mathcal{A}^*$, according to whether the occurrences $u^{(i)}$ and $u^{(i+1)}$ do or do not overlap each other, respectively. For instance, if $u = w_n$ is the n-th singular factor of the Fibonacci word, then, as shown in Example 2.1, each $z_i \in \{w_{n+1}, w_{n-1}\}$ $(i \geq 1)$; in which case u does not have overlap in $\mathbf{f}$.

The following result shows precisely which factors of c_α have no overlapping occurrences in c_α.

Proposition 3.4. [4, Theorem 10] *Let $u \prec c_\alpha$ with $q_n < |u| \leq q_{n+1}$ for some $n \in \mathbb{N}$. Then u has no overlap in c_α if and only if $u = w_{n+1}$, or $w_n \prec u$.* ☐

Accordingly, one easily deduces from Proposition 3.1 and Lemma 3.2 which palindromic factors of c_α do not have overlap.

Corollary 3.5. *Let $u \in$ PAL and $u \prec c_\alpha$ with $q_n < |u| \leq q_{n+1}$ for some $n \in \mathbb{N}$. Then u is a palindrome without overlap in c_α if and only if $u = w_{n+1}$, or $u = vw_n\widetilde{v}$ with $v \subseteq_s v_{n-1}$ and $|v| \leq \frac{1}{2}|v_{n-1}|$.* ☐

4 Decompositions of c_α into palindromes

In this section, we prove some lemmas which lead us to the main result of this paper (Theorem 5.1).

4.1 Useful results

In what follows, let us denote by G the standard morphism of $\mathcal{A}^*$ given by

$$G = \varphi E : \begin{array}{ccc} a & \mapsto & a \\ b & \mapsto & ab \end{array} .$$

Lemma 4.1. [22] *For any irrational* $\gamma \in (0,1)$, $E(c_\gamma) = c_{1-\gamma}$ *and* $G(c_\gamma) = c_{\gamma/(1+\gamma)}$. ☐

The following simple, yet useful, corollary (and the remark to follow) will be needed in our proofs.

Corollary 4.2. *For any irrational* $\gamma \in (0,1)$ *and* $k \in \mathbb{N}$, $G^k(c_\gamma) = c_{\gamma/(1+k\gamma)}$.

Proof. Induction on k. ☐

Remark 4.1. Recall that we are restricting our attention to c_α where α has continued fraction expansion $[0; 1+d_1, d_2, d_3, \ldots]$, $d_1 \geq 1$. Let us note that $\frac{\alpha}{1+k\alpha} = \frac{1}{k+1/\alpha} = [0; 1+d_1+k, d_2, d_3, \ldots] = \alpha_{0,k}$ and, more generally, $\frac{\alpha_n}{1+k\alpha_n} = [0; d_{n+1}+k, d_{n+2}, d_{n+3}, \ldots]$ for $n \geq 1$. Consequently,

$$G^k(c_{\alpha_n}) = c_{\alpha_{n,k}} \quad \text{for all } n \geq 0.$$

It is also easily checked that $1 - \alpha_{n+1,1} = [0; 1, d_{n+2}, d_{n+3}, \ldots] = \alpha_{n,1-d_{n+1}}$, for any $n \geq 1$; whence

$$E(c_{\alpha_{n+1,1}}) = c_{\alpha_{n,1-d_{n+1}}} \quad \text{for all } n \geq 1. \tag{4.1}$$

(Note that $E(c_{\alpha_{1,1}}) = c_{\alpha_{0,-d_1}}$.)

4.2 Some lemmas

Here, we simplify Melançon's decompositions of c_α, given in Corollary 2.2. In particular, we obtain two different decompositions of c_α with respect to occurrences of the palindromes

$$U_{n,k} = (w_{n-1}v_{n-2})^k w_{n-1} \quad \text{and} \quad \overline{U}_{n,k} = (v_{n-2}w_{n-1})^k v_{n-2} \quad (0 \leq k \leq d_{n+1}-1),$$

which form the basis of all palindromic factors of c_α (see Corollary 3.3). From the first of these decompositions, we easily deduce decompositions of c_α that show exactly where the singular words w_n and v_n occur in c_α, for any $n \in \mathbb{N}$. (Recall that Corollary 2.2 gives a decomposition of c_α which shows all of the occurrences of w_{2n}, but this result does not provide information as to the exact positions of w_{2n-1} in c_α.)

Notation. For any morphism ψ of $\mathcal{A}^*$ such that $\psi(a) = u$ and $\psi(b) = v$ for some u, $v \in \mathcal{A}^*$, we shall write $\psi = (u,v)$ to indicate the image of ψ on the alphabet $\mathcal{A}$. If $\mathbf{x} = x_0x_1x_2\cdots \in \mathcal{A}^\omega$, then $\psi(\mathbf{x}) = \psi(x_0)\psi(x_1)\psi(x_2)\cdots$ is the word obtained from $\mathbf{x}$ by replacing the letters a and b in $\mathbf{x}$ by the words u and v, respectively. We shall denote by $\mathbf{x}\{u,v\}$ the word $\psi(\mathbf{x})$. In particular, $c_\alpha\{u,v\}$ denotes the characteristic Sturmian word of slope α over the alphabet $\{u,v\}$.

Lemma 4.3. [4] *For any* $n \geq 1$, $c_\alpha = c_{\alpha_{n,1}}\{s_n, s_{n-1}\}$. ☐

Lemma 4.4. *For any* $n \in \mathbb{N}$,

$$\prod_{j=-1}^{n-1} (v_{2j}w_{2j+1})^{d_{2j+3}} = \prod_{j=-1}^{2n-1} v_j.$$

Proof. Using Lemma 3.2(3), observe that for any integer $j \geq 0$,

$$(v_{2j}w_{2j+1})^{d_{2j+3}} = v_{2j}w_{2j+1}(v_{2j}w_{2j+1})^{d_{2j+3}-1} = v_{2j}(w_{2j+1}v_{2j})^{d_{2j+3}-1}w_{2j+1} = v_{2j}v_{2j+1},$$

from which the result is readily deduced. ☐

Lemma 4.5. *For any* $n \geq 1$ *and* $0 \leq k \leq d_{n+1}-1$, *we have*

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) U_{n,k}^{(1)} z_1 U_{n,k}^{(2)} z_2 U_{n,k}^{(3)} z_3 \cdots,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ *is given by* $c_{\alpha_{n,-k}}$ *over the alphabet* $\{(U_{n,k-1})^{-1}, w_n\}$.

Note. We set $U_{n,-1} = (v_{n-2})^{-1}$ and $\overline{U}_{n,-1} = (w_{n-1})^{-1}$; whence if $k = 0$, then

$$(U_{n,k-1})^{-1} = (U_{n,-1})^{-1} = v_{n-2} \quad \text{and} \quad (\overline{U}_{n,k-1})^{-1} = (\overline{U}_{n,-1})^{-1} = w_{n-1}.$$

Proof of Lemma 4.5. We first prove the result for odd $n = 2m+1$, $m \geq 0$. By Corollary 2.2(1) and Lemma 4.4, we have

$$c_\alpha = \left(\prod_{j=-1}^{2m-1} v_j\right) \psi(c_{\alpha_{2m+1}}),$$

where $\psi = (w_{2m}v_{2m-1}, w_{2m}w_{2m+1})$. Further, by Remark 4.1, we have

$$\psi G^k(c_{\alpha_{2m+1,-k}}) = \psi(c_{\alpha_{2m+1}}) \quad \text{for } 0 \leq k \leq d_{2m+2} - 1.$$

Therefore,

$$c_\alpha = \left(\prod_{j=-1}^{2m-1} v_j\right) \psi G^k(c_{\alpha_{2m+1,-k}}),$$

where $G^k = (a, a^k b)$, and hence

$$\begin{aligned}\psi G^k &= (w_{2m}v_{2m-1}, (w_{2m}v_{2m-1})^k w_{2m}w_{2m+1}) \\ &= (w_{2m}v_{2m-1}, U_{2m+1,k}w_{2m+1}) \\ &= (U_{2m+1,k}(U_{2m+1,k-1})^{-1}, U_{2m+1,k}w_{2m+1}).\end{aligned}$$

Clearly, $U_{2m+1,k}w_{2m+1}$ and $w_{2m}v_{2m-1}$ $(= U_{2m+1,k}(U_{2m+1,k-1})^{-1})$ must be the two return words of $U_{2m+1,k}$. Also, using Lemma 3.2, we find that $U_{2m+1,k}$ is not a factor of the prefix $(\prod_{j=-1}^{2m-1} v_j)U_{2m+1,k-1}$ of c_α, since

$$\left(\prod_{j=-1}^{2m-1} v_j\right) U_{2m+1,k-1} = x^{-1}w_{2m}v_{2m-1}U_{2m+1,k-1} = x^{-1}U_{2m+1,k} \quad (x \in \mathcal{A}).$$

Thus, the derived word of c_α with respect to $U_{2m+1,k}$ is given by $\mathcal{D}_{U_{2m+1,k}}(c_\alpha) = c_{\alpha_{2m+1,-k}}\{w_{2m}v_{2m-1}, U_{2m+1,k}w_{2m+1}\}$, and we can write

$$c_\alpha = \left(\prod_{j=-1}^{2m-1} v_j\right) U^{(1)}_{2m+1,k} z_1 U^{(2)}_{2m+1,k} z_2 U^{(3)}_{2m+1,k} z_3 \cdots,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{2m+1,-k}}$ over the alphabet $\{(U_{2m+1,k-1})^{-1}, w_{2m+1}\}$. This completes the proof for odd n.

Let us now prove that the assertion holds for even $n = 2m$, $m \geq 1$. By considering occurrences of $U_{2m-1,d_{2m}-1}$ $(= v_{2m-2})$ in c_α, one deduces from the above that, for any integer $m \geq 1$,

$$c_\alpha = \left(\prod_{j=-1}^{2m-3} v_j\right) v_{2m-2}\phi(c_{\alpha_{2m-1,1-d_{2m}}}), \tag{4.2}$$

where $\phi = ((U_{2m-1,d_{2m}-2})^{-1}v_{2m-2}, w_{2m-1}v_{2m-2})$. In fact, using (2) and (3) of Lemma 3.2, we can write $\phi = (v_{2m-3}w_{2m-2}, w_{2m-1}v_{2m-2}) = ((v_{2m-2})^{-1}w_{2m}, w_{2m-1}v_{2m-2})$. Again, using Remark 4.1, we have $EG^{k+1}(c_{\alpha_{2m,-k}}) = E(c_{\alpha_{2m,1}}) = c_{\alpha_{2m-1,1-d_{2m}}}$, where $EG^{k+1} = (b, b^{k+1}a)$. Whence, it follows from (4.2) that

$$c_\alpha = \left(\prod_{j=-1}^{2m-2} v_j\right) \phi EG^{k+1}(c_{\alpha_{2m,-k}}),$$

where

$$\phi EG^{k+1} = (w_{2m-1}v_{2m-2}, (w_{2m-1}v_{2m-2})^{k+1}(v_{2m-2})^{-1}w_{2m}) = (U_{2m,k}(U_{2m,k-1})^{-1}, U_{2m,k}w_{2m}).$$

The result now follows (as for the odd case) since $U_{2m,k}w_{2m}$ and $w_{2m-1}v_{2m-2}$ $(= U_{2m,k}(U_{2m,k-1})^{-1})$ are the two return words of $U_{2m,k}$.

□

Remark 4.2. From Lemma 4.5, we readily deduce two 'singular' decompositions of c_α with respect to the occurrences of w_n and v_n, for any $n \in \mathbb{N}$. Indeed, we have $U_{n,0} = w_{n-1}$ and $U_{n,d_{n+1}-1} = v_{n-1}$. Therefore, taking $k = 0$ in the above lemma, we obtain a decomposition that shows exactly where the n-th singular word w_n occurs in c_α. That is, for any $n \geq 0$,

$$c_\alpha = \left(\prod_{j=-1}^{n-1} v_j\right) w_n^{(1)} z_1 w_n^{(2)} z_2 w_n^{(3)} z_3 \cdots , \tag{4.3}$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{n+1}}$ over the alphabet $\{v_{n-1}, w_{n+1}\}$.

Now, taking $k = d_{n+1} - 1$, we find that, for any $n \geq 0$,

$$c_\alpha = \left(\prod_{j=-1}^{n-1} v_j\right) v_n^{(1)} z_1 v_n^{(2)} z_2 v_n^{(3)} z_3 \cdots , \tag{4.4}$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{n+2,1}}$ over the alphabet $\{w_{n+1}, (U_{n+1,d_{n+2}-2})^{-1}\}$ (since $E(c_{\alpha_{n+1,1-d_{n+2}}}) = c_{\alpha_{n+2,1}}$). This also holds for $n = -1$ since, from Lemma 4.3, we have

$$c_\alpha = c_{\alpha_{1,1}}\{s_1, s_0\} = c_{\alpha_{1,1}}\{a^{d_1}b, a\} = c_{\alpha_{1,1}}\{v_{-1}w_0, w_{-1}v_{-2}\}.$$

The following simple decomposition of c_α (which has also been proved independently in [4]) is a direct consequence of (4.4).

Proposition 4.6. *For any $n \in \mathbb{N}$, we have*

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) c_{\alpha_{n+1,1}}\{v_{n-1}w_n, w_{n-1}v_{n-2}\}.$$

□

Lemma 4.7. *For any $n \geq 1$ and $0 \leq k \leq d_{n+1} - 1$, we have*

$$c_\alpha = \left(\prod_{j=-1}^{n-3} v_j\right) \overline{U}_{n,k}^{(1)} z_1 \overline{U}_{n,k}^{(2)} z_2 \overline{U}_{n,k}^{(3)} z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{n,1-k}}$ over the alphabet $\{(\overline{U}_{n,k-1})^{-1}, (U_{n-1,d_n-2})^{-1}\}$.

Proof. Follows almost immediately from Proposition 4.6. Indeed, $G^k(c_{\alpha_{n,1-k}}) = c_{\alpha_{n,1}}$ for $0 \leq k \leq d_{n+1} - 1$, and hence

$$c_\alpha = \left(\prod_{j=-1}^{n-3} v_j\right) c_{\alpha_{n,1}}\{v_{n-2}w_{n-1}, w_{n-2}v_{n-3}\} = \left(\prod_{j=-1}^{n-3} v_j\right) c_{\alpha_{n,1-k}}\{v_{n-2}w_{n-1}, (v_{n-2}w_{n-1})^k w_{n-2}v_{n-3}\},$$

where $v_{n-2}w_{n-1} = \overline{U}_{n,k}(\overline{U}_{n,k-1})^{-1}$ and

$$(v_{n-2}w_{n-1})^k w_{n-2}v_{n-3} = \overline{U}_{n,k}(v_{n-2})^{-1} w_{n-2}v_{n-3} = \overline{U}_{n,k}(U_{n-1,d_n-2})^{-1}.$$

□

5 Main result

We are now equipped with the necessary tools to prove the main result of this paper, which, in view of Corollary 3.3, completely describes occurrences of palindromes in c_α.

Theorem 5.1. *Let u be a palindromic factor of c_α with $|u| \geq 2$.*

(1) *Suppose $u = vU_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s v_{n-2}$, $v \neq v_{n-2}$, and $0 \leq k \leq d_{n+1} - 2$. Then*

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) v^{-1}u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots ,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{(vU_{n,k-1}\widetilde{v})^{-1}, \widetilde{v}^{-1}w_nv^{-1}\}$.

(2) *Suppose $u = v\overline{U}_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s w_{n-1}$, $v \neq w_{n-1}$, and $0 \leq k \leq d_{n+1} - 1$. Then*

$$c_\alpha = \left(\prod_{j=-1}^{n-3} v_j\right) v^{-1}u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots ,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{n,1-k}}$ over the alphabet $\{(v\overline{U}_{n,k-1}\widetilde{v})^{-1}, (vU_{n-1,d_n-2}\widetilde{v})^{-1}\}$.

Moreover, if $u = a^k$ for some $k \in [2, d_1 - 1]$, then $c_\alpha = u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots$, where $z_1z_2z_3\cdots$ is given by $c_{\alpha_{0,-k}}$ over the alphabet $\{(a^{k-1})^{-1}, b\}$.

Note. In regards to assertion (1), let us point out that v is a suffix (and $\widetilde{v}$ is a prefix) of w_n since $w_n = w_{n-2}v_{n-3}v_{n-2} = v_{n-2}v_{n-3}w_{n-2}$. Therefore, $\widetilde{v}^{-1}w_nv^{-1} \in \mathcal{A}^*$ since $|v| < |v_{n-2}| = q_n - q_{n-1} \leq \frac{1}{2}q_n = \frac{1}{2}|w_n|$.

Proof of Theorem 5.1. Assertions (1) and (2) are proved in a similar fashion, using Lemmas 4.5 and 4.7 respectively, so we just give the proof of (1). The last statement is trivial since $c_\alpha = G^k(c_{\alpha_{0,-k}}) = c_{\alpha_{0,-k}}\{a, a^kb\}$.

Suppose $u = vU_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s v_{n-2}$, $v \neq v_{n-2}$ and $0 \leq k \leq d_{n+1} - 2$. From Lemma 4.5, it follows that

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) v^{-1}(vU_{n,k}^{(1)}\widetilde{v})\widetilde{v}^{-1}z_1v^{-1}(vU_{n,k}^{(2)}\widetilde{v})\widetilde{v}^{-1}z_2v^{-1}(vU_{n,k}^{(3)}\widetilde{v})\widetilde{v}^{-1}z_3\cdots ,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{(U_{n,k-1})^{-1}, w_n\}$. Consequently, since each occurrence of u in c_α corresponds to an occurrence of $U_{n,k}$ in c_α, we have

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) v^{-1}u^{(1)}\hat{z}_1u^{(2)}\hat{z}_2u^{(3)}\hat{z}_3\cdots ,$$

where $\hat{z}_i = \widetilde{v}^{-1}z_iv^{-1}$, for all $i \geq 1$. (Note that $\left(\prod_{j=-1}^{n-2} v_j\right)v^{-1} \in \mathcal{A}^*$ since $v \subseteq_s v_{n-2}$.) Thus, $\hat{\mathbf{z}} := \hat{z}_1\hat{z}_2\hat{z}_3\cdots$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{(vU_{n,k-1}\widetilde{v})^{-1}, \widetilde{v}^{-1}w_nv^{-1}\}$. Indeed, $\hat{z}_i = \widetilde{v}^{-1}w_nv^{-1}$ if $z_i = w_n$, and $\hat{z}_i = \widetilde{v}^{-1}(U_{n,k-1})^{-1}v^{-1} = (vU_{n,k-1}\widetilde{v})^{-1}$ if $z_i = (U_{n,k-1})^{-1}$. This completes the proof of (1).

In part (2), note that $\left(\prod_{j=-1}^{n-3} v_j\right)v^{-1} \in \mathcal{A}^*$, since v is a proper suffix of w_{n-1}, and hence a suffix of $\prod_{j=-1}^{n-3} v_j = x^{-1}w_{n-1}$, where $x \in \mathcal{A}$ (by Lemma 3.2(5)).

□

Example 5.1. Let us now demonstrate Theorem 5.1 for c_α with $\alpha = [0; \overline{2, 1, 3, 1}] = (4\sqrt{5} - 5)/11$. In this case, we have

$$c_\alpha = abaabaabaababaabaabaabaababaabaabaabaab\cdots.$$

Also note that

$$w_{-1} = v_{-1} = a,\ w_0 = v_0 = b,\ w_1 = aa,\ w_2 = bab,\ w_3 = aabaabaabaa,\ v_1 = aabaabaa,\ v_2 = bab.$$

(i) Consider the palindromic factor $u = baaw_2aab = baababaab$, where $v = baa \subseteq_s v_1$. By Theorem 5.1(1),

$$\begin{aligned} c_\alpha &= v_{-1}v_0v_1(baa)^{-1}u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots \\ &= abaabaa(baababaab)z_1(baababaab)z_2(baababaab)z_3\cdots, \end{aligned}$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by c_{α_3} over the alphabet $\{aa, aabaa\}$. We have $c_{\alpha_3} = c_{\alpha_{3,0}} = [0; \overline{1, 2, 1, 3}] = \sqrt{5} - 57/38$, and hence $c_{\alpha_3} = bbabbbabbba\cdots$. Thus, we can write

$$c_\alpha = abaabaa(baababaab)aabaa(baababaab)aabaa(baababaab)aa(baababaab)aabaa(baababaab)\cdots.$$

(ii) Now consider the palindromic factor $u = U_{2,1} = (w_1v_0)^1w_1 = aabaa$. By Theorem 5.1(1),

$$c_\alpha = v_{-1}v_0u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots = ab(aabaa)z_1(aabaa)z_2(aabaa)z_3\cdots,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{2,-1}}$ over the alphabet $\{(aa)^{-1}, bab\}$. We have $\alpha_{2,-1} = [0; 2, \overline{1, 2, 1, 3}] = (4\sqrt{5} - 2)/19$, and therefore

$$c_{\alpha_{2,-1}} = abaabaababaabaabaababaabaabaa\cdots.$$

Hence, we can write

$$\begin{aligned} c_\alpha = {} & ab(aab\underline{aa)baa}bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)baa} \\ & bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)baa}\cdots. \end{aligned}$$

Notice that u has a unique overlap factor aa.

(iii) Let us now consider the palindromic factor $u = aU_{2,2}a = a(v_0w_1)^2v_0a = abaabaaba$, where $a \subseteq_s w_1 = aa$. Observe that

$$(vU_{2,1}\widetilde{v})^{-1} = (a(baa)^1ba)^{-1} = (abaaba)^{-1} \quad \text{and} \quad (vU_{1,d_2-2}\widetilde{v})^{-1} = (a(ba)^{-1}ba) = a^{-1}.$$

Thus, by Theorem 5.1(2), we have

$$c_\alpha = v_{-1}a^{-1}u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots = (abaabaaba)z_1(abaabaaba)z_2(abaabaaba)z_3\cdots,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{2,-1}}$ over the alphabet $\{(abaaba)^{-1}, a^{-1}\}$; whence

$$\begin{aligned} c_\alpha = {} & (aba\underline{abaaba)ab(a}ba\underline{aba[aba)aba}ab\underline{a]ba(aba[aba}aba)ab\underline{a]ba(abaaba}a \\ & b\underline{a)ba(aba[aba}aba)ab\underline{a]baabaaba}\cdots. \end{aligned}$$

In this case, u has two overlap factors: $abaaba$ and a.

$\square$

Let us now denote by $\mathrm{occ}_i(u)$ the position of the i-th occurrence of u in c_α, i.e., if $c_\alpha = zu\mathbf{x}$ for some $z \in \mathcal{A}^*$, $\mathbf{x} \in \mathcal{A}^\omega$ such that $|zu|_u = i$, then $\mathrm{occ}_i(u) = |z|$. With this notation, a given factor u of c_α occurs at precisely the positions $(\mathrm{occ}_i(u))_{i\geq 1}$ in c_α.

The following corollary of Theorem 5.1 gives the exact positions at which palindromes occur in c_α.

Corollary 5.2. *Let u be a palindromic factor of c_α with $|u| \geq 2$.*

(1) *Suppose $u = vU_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s v_{n-2}$, $v \neq v_{n-2}$, and $0 \leq k \leq d_{n+1} - 2$. Then $\mathrm{occ}_1(u) = \frac{1}{2}((k+2)q_n + q_{n-1} - |u| - 2)$ and, for all $i \geq 1$,*

$$\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + P_i,$$

where $(P_i)_{i\geq 1}$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{q_n, (k+1)q_n + q_{n-1}\}$.

(2) *Suppose $u = vU_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s w_{n-1}$, $v \neq w_{n-1}$, and $0 \leq k \leq d_{n+1} - 1$. Then $\mathrm{occ}_1(u) = \frac{1}{2}((k+1)q_n + q_{n-1} - |u| - 2)$ and, for all $i \geq 1$,*

$$\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + P_i,$$

where $(P_i)_{i\geq 1}$ is given by $c_{\alpha_{n,1-k}}$ over the alphabet $\{q_n, kq_n + q_{n-1}\}$.

Moreover, if $u = a^k$ for some $k \in [2, d_1 - 1]$, then $\mathrm{occ}_1(u) = 0$ and $\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + P_i$ for all $i \geq 1$, where $(P_i)_{i\geq 1}$ is given by $c_{\alpha_{0,-k}}$ over the alphabet $\{1, k+1\}$.

Proof. As with Theorem 5.1, the proofs of (1) and (2) are much the same, so we just give the proof of (1). The proof of the last statement is trivial.

Suppose $u = vU_{n,k}\widetilde{v}$ for some $n \geq 1$, where $v \subseteq_s v_{n-2}$, $v \neq v_{n-2}$ and $0 \leq k \leq d_{n+1} - 2$. Theorem 5.1(1) shows that

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j\right) v^{-1}u^{(1)}z_1u^{(2)}z_2u^{(3)}z_3\cdots, \tag{5.1}$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{(vU_{n,k-1}\widetilde{v})^{-1}, \widetilde{v}^{-1}w_nv^{-1}\}$. Observe that

$$|U_{n,k}| = (k+1)|w_{n-1}| + k|v_{n-2}| = (k+1)q_{n-1} + k(q_n - q_{n-1}) = kq_n + q_{n-1},$$

and hence

$$|v| = \frac{1}{2}(|u| - |U_{n,k}|) = \frac{1}{2}(|u| - kq_n - q_{n-1}).$$

Also recall that if x is the first letter of w_n, then $x^{-1}w_n = \prod_{j=-1}^{n-2} v_j$. Therefore, since v is a proper suffix of v_{n-2}, we have

$$\left|\left(\prod_{j=-1}^{n-2} v_j\right) v^{-1}\right| = |x^{-1}w_n| - |v| = q_n - 1 - \frac{1}{2}(|u| - kq_n - q_{n-1}).$$

Hence, the first occurrence of u in c_α is at position

$$\mathrm{occ}_1(u) = \frac{1}{2}((k+2)q_n + q_{n-1} - |u| - 2).$$

Furthermore,

$$|vU_{n,k-1}\widetilde{v}| = (k-1)q_n + q_{n-1} + 2|v| = (k-1)q_n + q_{n-1} + (|u| - kq_n - q_{n-1}) = |u| - q_n$$

and

$$|\widetilde{v}^{-1}w_nv^{-1}| = q_n - 2|v| = q_n - (|u| - kq_n - q_{n-1}) = (k+1)q_n + q_{n-1} - |u|.$$

Thus, it follows from (5.1) that $\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + P_i$ for all $i \geq 1$, where $(P_i)_{i\geq 1}$ is the characteristic Sturmian word of slope $\alpha_{n,-k}$ over the alphabet $\{q_n, (k+1)q_n + q_{n-1}\}$. ☐

Example 5.2. Let $\alpha = [0; \overline{2, 1, 3, 1}] = (4\sqrt{5} - 5)/11$ and consider the palindromic factor u of c_α given by $u = U_{2,1} = aabaa$. According to Corollary 5.2, one should find that u first occurs at position

$$\mathrm{occ}_1(u) = \frac{1}{2}(3q_2 - q_1 - 5 - 2) = \frac{1}{2}(9 + 2 - 7) = 2,$$

followed by the positions $\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + P_i$ for each $i \geq 1$, where $(P_i)_{i\geq 1}$ is the characteristic Sturmian word of slope $\alpha_{2,-1} = [0; 2, \overline{1, 2, 1, 3}]$ over the alphabet $\{q_2, 2q_2+q_1\} = \{3, 8\}$; that is, $(P_i)_{i\geq 1} = (3, 8, 3, 3, 8, 3, 3, 8, 3, 8, 3, 3, 8, 3, 3, 8, \ldots)$. Indeed, from Example 5.1(2), we have

$$\begin{aligned} c_\alpha = {} & ab(aab\underline{aa)baa}bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)baa} \\ & bab(aab\underline{aa)b(aa}baa)bab(aab\underline{aa)baa}\cdots , \end{aligned}$$

from which it is evident that $u = aabaa$ occurs at positions 2, 5, 13, 16, 19, 27, 30, 33, 41, 44, 52, 55, 58, 66, $\ldots$.

Remark 5.1. In general, if u_1 and u_2 are the two return words of a factor u of c_α, it is clear that

$$\mathrm{occ}_{i+1}(u) = \mathrm{occ}_i(u) + |u_{j_i}| \text{ where } j_i = 1 \text{ or } 2.$$

In particular, the sequence $(j_i)_{i\geq 1}$ is a Sturmian word over the alphabet $\{1, 2\}$ (see [12] or Section 3.2). In the case when u is a palindromic factor of c_α, Corollary 5.2 shows that the sequence $(j_i)_{i\geq 1}$ is given by $c_{\alpha_{n,-k}}$ over the alphabet $\{1, 2\}$, for some $n \in \mathbb{N}$ and $0 \leq k \leq d_{n+1} - 1$. For example, the two return words of w_n $(= U_{n+1,0})$ are $u_1 = w_n v_{n-1}$ and $u_2 = w_n w_{n+1}$, where

$$|u_1| = q_{n+1} \quad \text{and} \quad |u_2| = q_{n+1} + q_n.$$

From Corollary 5.2, $\mathrm{occ}_1(w_n) = q_{n+1} - 1$ and $\mathrm{occ}_{i+1}(w_n) = \mathrm{occ}_i(w_n) + |u_{j_i}|$ for each $i \geq 1$, where $(j_i)_{i\geq 1}$ is given by $c_{\alpha_{n+1}}$ over the alphabet $\{1, 2\}$.

6 Occurrences of factors of length q_n in c_α

In this last section, we determine the structure of all factors of length q_n of c_α with respect to the singular words w_n, w_{n-1}, and v_{n-2}. Subsequently, using some results from Section 4, we completely describe where factors of length q_n occur in c_α.

Let $w = x_1 x_2 \cdots x_m \in \mathcal{A}^*$, each $x_i \in \mathcal{A}$, and let $k \in \mathbb{N}$ with $0 \leq k \leq m - 1$. The *k-th conjugate* of w is the word $C_k(w) := x_{k+1} x_{k+2} \cdots x_m x_1 x_2 \cdots x_k$. Further, we conventionally set $C_{-k}(w) = C_{|w|-k}(w)$ and define $C(w) := \{C_k(w) : \ 0 \leq k \leq |w| - 1\}$.

One can easily prove that any conjugate of s_n is a factor of c_α. Certainly, $C(s_{-1}) = \{b\}$ and, for $n \geq 0$,

$$s_{n+3} = s_{n+2}^{d_{n+3}} s_{n+1} = (s_{n+1}^{d_{n+2}} s_n)^{d_{n+3}} s_n^{d_{n+1}} s_{n-1} = (s_{n+1}^{d_{n+2}} s_n)^{d_{n+3}-1} s_{n+1}^{d_{n+2}} s_n^{d_{n+1}+1} s_{n-1},$$

where $d_{n+1} + 1 \geq 2$. Thus, s_n^2 is a factor of c_α, and hence the claim is proved since any conjugate of s_n is a factor of s_n^2.

Now, each s_n is a primitive word [8], i.e., s_n cannot be written as a non-trivial integer power of a shorter word. Consequently, s_n has q_n distinct conjugates, i.e., $|C(s_n)| = q_n$. Furthermore, from the above observation, $C(s_n)$ is a set of factors of c_α. It is therefore deduced that the set of all factors of length q_n of c_α consists of $C(s_n)$ and w_n. That is,

$$\Omega_{q_n}(c_\alpha) = C(s_n) \cup \{w_n\}.$$

Indeed, since c_α is a Sturmian word, it must have exactly $q_n + 1$ distinct factors of length q_n.

Lemma 6.1. *For any $n \geq 1$, $C_{q_n-1}(s_n) = w_{n-1} v_{n-2}$ and $C_{q_{n-1}-1}(s_n) = v_{n-2} w_{n-1}$. Moreover,*

(1) *for $0 \leq k \leq q_{n-1} - 2$, $C_k(s_n) = u v_{n-2} v$, where $vu = w_{n-1}$ and $|v| = k + 1$;*

(2) *for $q_{n-1} - 1 \leq k \leq q_n - 1$, $C_k(s_n) = u w_{n-1} v$, where $vu = v_{n-2}$ and $|v| = k + 1 - q_{n-1}$.*

Proof. By Lemma 3.2(1), $C_{-1}(s_n) = C_{q_n-1}(s_n) = w_{n-1} v_{n-2}$. Therefore, $C_{q_{n-1}-1}(s_n) = v_{n-2} w_{n-1}$ since $|w_{n-1}| = q_{n-1}$. Assertions (1) and (2) follow immediately. ☐

Accordingly, a factor of length q_n of c_α is either w_n, or has at least one of the words v_{n-2} and w_{n-1} as a factor. We shall now establish two different decompositions of c_α, which show exactly where conjugates of s_n occur in c_α.

Theorem 6.2. *Let $n \geq 1$.*

(1) *Suppose $w = C_k(s_n)$ for some $k \in [0, q_{n-1} - 2]$, so that $w = uv_{n-2}v$, where $vu = w_{n-1}$ and $|v| = k + 1$. Then*

$$c_\alpha = \left(\prod_{j=-1}^{n-3} v_j \right) u^{-1} w^{(1)} z_1 w^{(2)} z_2 w^{(3)} z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{n,1}}$ over the alphabet $\{\varepsilon, (uU_{n-1,d_n-2}v)^{-1}\}$.

(2) *Suppose $w = C_k(s_n)$ for some $k \in [q_{n-1} - 1, q_n - 1]$, so that $w = uw_{n-1}v$, where $vu = v_{n-2}$ and $|v| = k + 1 - q_{n-1}$. Then*

$$c_\alpha = \left(\prod_{j=-1}^{n-2} v_j \right) u^{-1} w^{(1)} z_1 w^{(2)} z_2 w^{(3)} z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by c_{α_n} over the alphabet $\{\varepsilon, v^{-1} w_n u^{-1}\}$.

Proof. Using decompositions (4.4) and (4.3) (consequences of Lemma 4.5), the proof follows along exactly the same lines as the proof of Theorem 5.1. ☐

In light of Theorem 6.2 and the w_n-decomposition of c_α given by (4.3) (together with the fact that $\Omega_{q_n}(c_\alpha) = C(s_n) \cup \{w_n\}$), we have now shown precisely where each factor of length q_n occurs in c_α. It is important to note that it follows from Proposition 3.4 that a factor w of length q_n does not have overlap in c_α if and only if $w = w_n$, or $w = C_k(s_n)$ for some $k \in [q_{n-1} - 1, q_n - 1]$. Certainly, if w takes the latter form, then $w = uw_{n-1}v$ with $vu = v_{n-2}$ and $|v| = k + 1 - q_{n-1}$. In this case, Theorem 6.2(2) shows that w does not have overlapping occurrences since $w_n = (vu)v_{n-3}w_{n-2} = w_{n-2}v_{n-3}(vu)$, where $vu = v_{n-2}$, and hence $v^{-1}w_n u^{-1} \in \mathcal{A}^*$.

Example 6.1. Suppose $\alpha = [0; \overline{2, 1}] = (\sqrt{3} - 1)/2$. Then

$$c_\alpha = abaabaababaabaabaababaabaabaababaabaababaabaabaababaabaabaababa \cdots .$$

Let us demonstrate the above theorem by considering the first two conjugates of $s_3 = abaabaab$; namely, $C_1(s_3)$ $(= C_{q_2-2}(s_3))$ and $C_2(s_3)$ $(= C_{q_2-1}(s_3))$. First observe that

$$w_{-1} = v_{-1} = a,\ w_0 = v_0 = b,\ w_1 = aa,\ w_2 = bab,\ w_3 = aabaabaa,\ v_1 = aabaa,\ v_2 = bab.$$

(1) Let $w = C_1(s_3)$; the first conjugate of s_3. We have $w = baabaaba = uv_1v$, where $u = b$, $v = ba$ and $vu = bab = w_2$. Hence, by Theorem 6.2(1),

$$c_\alpha = v_{-1}v_0 b^{-1} w^{(1)} z_1 w^{(2)} z_2 w^{(3)} z_3 \cdots = a(baabaaba)z_1(baabaaba)z_2(baabaaba)z_3 \cdots ,$$

where $\mathbf{z} := z_1 z_2 z_3 \cdots$ is given by $c_{\alpha_{3,1}}$ over the alphabet $\{\varepsilon, (baaba)^{-1}\}$. (Note that $(uU_{2,d_3-2}v)^{-1} = (b(aab)^0 aaba)^{-1} = (baaba)^{-1}$.) Since $\alpha_{3,1} = [0; 2, \overline{2, 1}] = 1 - \sqrt{3}/3$, we have $c_{\alpha_{3,1}} = ababaabababaab \cdots$, and thus we can write

$$\begin{aligned} c_\alpha = a(baabaaba)(baa\underline{baaba)aba}(baa\underline{baaba)aba}(baabaaba)(baa\underline{baaba)aba} \\ (baa\underline{baaba)aba}(baa\underline{baaba)aba}(baabaaba)(baa\underline{baaba)aba} \cdots . \end{aligned}$$

(2) Now let $w = C_2(s_3)$; the second conjugate of s_3. Then $w = aabaabab = v_1w_2$, and it follows from Theorem 6.2(2) that

$$c_\alpha = v_{-1}v_0v_1{v_1}^{-1} w^{(1)} z_1 w^{(2)} z_2 w^{(3)} z_3 \cdots = ab(aabaabab)z_1(aabaabab)z_2(aabaabab)z_3 \cdots ,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by c_{α_3} over the alphabet $\{\varepsilon, w_3(aabaa)^{-1}\} = \{\varepsilon, aab\}$. Note that $\alpha_3 = [0;\overline{1,2}] = \sqrt{3}-1$, and hence $c_{\alpha_3} = bbabbbabbba\cdots$. Therefore,

$$\begin{aligned} c_\alpha &= ab(aabaabab)aab(aabaabab)aab(aabaabab)(aabaabab)aab(aabaabab) \\ &\quad aab(aabaabab)aab(aabaabab)(aabaabab)aab(aabaabab)aab\cdots. \end{aligned}$$

Remark 6.1. From Theorem 6.2, one can easily deduce Lemma 4.3 (i.e., $c_\alpha = c_{\alpha_{n,1}}\{s_n, s_{n-1}\}$ for all $n \geq 1$), as follows. Observe that $s_n = C_0(s_n) = uv_{n-2}v$, where $u = y^{-1}w_{n-1}$ and $v = y$. Similarly, $s_{n-1} = x^{-1}w_{n-2}v_{n-3}x$, where $x \in \mathcal{A}$ $(x \neq y)$. Hence,

$$c_\alpha = \left(\prod_{j=-1}^{n-3} v_j\right)(y^{-1}w_{n-1})^{-1}s_n^{(1)}z_1s_n^{(2)}z_2s_n^{(3)}z_3\cdots = s_n^{(1)}z_1s_n^{(2)}z_2s_n^{(3)}z_3\cdots,$$

where $\mathbf{z} := z_1z_2z_3\cdots$ is given by $c_{\alpha_{n,1}}$ over the alphabet $\{\varepsilon, (y^{-1}w_{n-1}U_{n-1,d_n-2}y)^{-1}\}$. That is,

$$c_\alpha = c_{\alpha_{n,1}}\{s_n, s_n(y^{-1}w_{n-1}(w_{n-2}v_{n-3})^{-1}v_{n-2}y)^{-1}\}.$$

Using Lemma 3.2, we have

$$\begin{aligned} s_n(y^{-1}w_{n-1}(w_{n-2}v_{n-3})^{-1}v_{n-2}y)^{-1} &= s_n(s_{n-1}x^{-1}(w_{n-2}v_{n-3})^{-1}v_{n-2}y)^{-1} \\ &= s_ny^{-1}(v_{n-2})^{-1}(w_{n-2}v_{n-3})x(s_{n-1})^{-1} \\ &= x^{-1}w_n(v_{n-2})^{-1}xs_{n-1}(s_{n-1})^{-1} \\ &= x^{-1}w_{n-2}v_{n-3}x \\ &= s_{n-1}, \end{aligned}$$

and therefore $c_\alpha = c_{\alpha_{n,1}}\{s_n, s_{n-1}\}$, as required.

We finish with a corollary of Theorem 6.2 (*cf.* Corollary 5.2).

Corollary 6.3. *Let $n \geq 1$ and suppose $w = C_k(s_n)$ for some $k \in [0, q_n - 1]$. Then $\mathrm{occ}_1(w) = k$ and, for all $i \geq 1$, $\mathrm{occ}_{i+1}(w) = \mathrm{occ}_i(w) + P_i$, where $(P_i)_{i\geq 1}$ is given by:*

- *$c_{\alpha_{n,1}}$ over the alphabet $\{q_n, q_{n-1}\}$ if $k \in [0, q_{n-1} - 2]$, or*
- *c_{α_n} over the alphabet $\{q_n, q_n + q_{n-1}\}$ if $k \in [q_{n-1} - 1, q_n - 1]$.*

□

7 Acknowledgements

Special thanks to Bob Clarke and Alison Wolff for their support and encouragement. Thanks also to the two anonymous referees for their helpful suggestions and comments. This research was supported by the George Fraser Scholarship of The University of Adelaide.

References

[1] J.-P. Allouche, J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, *Cambridge University Press*, 2003.

[2] J. Berstel, P. Séébold, Sturmian words, in: *M. Lothaire, Algebraic Combinatorics On Words, Encyclopedia of Mathematics and its Applications*, vol. 90, *Cambridge University Press*, 2002, pp. 45–110.

[3] T.C. Brown, Descriptions of the characteristic sequence of an irrational, *Canad. Math. Bull.* **36** (1) (1993), 15–21.

[4] W.-T. Cao, Z.-Y. Wen, Some properties of the factors of Sturmian sequences, *Theoret. Comput. Sci.* **304** (2003), 365–385, doi:10.1016/S0304–3975(03)00213–5.

[5] E.M. Coven, G.A. Hedlund, Sequences with minimal block growth, *Math. Systems Theory* **7**, (2) (1973), 138–153.

[6] A. de Luca, A combinatorial property of the Fibonacci words, *Inform. Process. Lett.* **12**, (4) (1981), 193–195, doi:10.1016/0020–0190(81)90099–5.

[7] A. de Luca, Standard Sturmian morphisms, *Theoret. Comput. Sci.* **178** (1997), 205–224, doi:10.1016/S0304–3975(96)00054–0.

[8] A. de Luca, F. Mignosi, Some combinatorial properties of Sturmian words, *Theoret. Comput. Sci.* **136** (1994), 361–385, doi:10.1016/0304-3975(94)00035-H.

[9] X. Droubay, Palindromes in the Fibonacci word, *Inform. Process. Lett.* **55** (1995), 217–221, doi:10.1016/0020–0190(95)00080–V.

[10] X. Droubay, G. Pirillo, Palindromes and Sturmian words, *Theoret. Comput. Sci.* **223** (1999), 73–85, doi:10.1016/S0304-3975(97)00188-6.

[11] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math.* **179** (1998), 89–101, doi:10.1016/S0012-365X(97)00029-0.

[12] I. Fagnot, L. Vuillon, Generalized balances in Sturmian words, *Discrete Appl. Math.* **121** (2002), 83–101, doi:10.1016/S0166-218X(01)00247-5.

[13] A.S. Fraenkel, M. Mushkin, U. Tassa, Determination of $[n\theta]$ by its sequence of differences, *Canad. Math. Bull.* **21** (4) (1978), 441–446.

[14] A. Glen, Conjugates of characteristic Sturmian words generated by morphisms, *European J. Combin.* **25** (7) (2004), 1025–1037, doi:10.1016/j.ejc.2003.12.012.

[15] C. Holton, L.Q. Zamboni, Descendants of primitive substitutions, *Theory Comput. Systems* **32** (1999), 133–157.

[16] J. Justin, L. Vuillon, Return words in Sturmian and episturmian words, *Theor. Inform. Appl.* **34** (5) (2000), 343–356.

[17] F. Levé, P. Séébold, Conjugation of standard morphisms and a generalization of singular words, *Bull. Belg. Math. Soc. Simon Stevin* **10** (5) (2003), 737–747.

[18] M. Lothaire, *Combinatorics On Words, Encyclopedia of Mathematics and its Applications*, vol. 17, *Addison-Wesley*, Reading, Massachusetts, 1983.

[19] G. Melançon, Lyndon words and singular factors of Sturmian words, *Theoret. Comput. Sci.* **218** (1999), 41–59, doi:10.1016/S0304–3975(98)00249–7.

[20] F. Mignosi, Infinite words with linear subword complexity, *Theoret. Comput. Sci.* **65** (1989), 221–242, doi:10.1016/0304-3975(89)90046-7.

[21] M. Morse, G.A. Hedlund, Symbolic Dynamics II: Sturmian Trajectories, *Amer. J. Math.* **62** (1940), 1–42.

[22] B. Parvaix, Propriétés d'invariance des mots sturmiens, *J. Théorie Nombres Bordeaux* **9** (1997), 351–369.

[23] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics, Lecture Notes in Mathematics*, vol. 1794, *Springer-Verlag*, Berlin, 2002.

[24] L. Vuillon, A characterization of Sturmian words by return words, *European J. Combin.* **22** (2001), 262–275, doi:10.1006/eujc.2000.0444.

[25] Z.-X. Wen, Z.-Y. Wen, Some properties of the singular words of the Fibonacci word, *European J. Combin.* **15** (6) (1994), 587–598, doi:10.1006/eujc.1994.1060.