Glen, A. (2004) Conjugates of characteristic Sturmian words generated by morphisms. European Journal of Combinatorics, 25 (7). pp. 1025-1037.

http://researchrepository.murdoch.edu.au/3881

Conjugates of characteristic Sturmian words generated by morphisms

Amy Glen*

December 15, 2003

School of Pure Mathematics, University of Adelaide, South Australia, 5005

Abstract

This article is concerned with characteristic Sturmian words of slope α and $1-\alpha$ (denoted by c_α and $c_{1-\alpha}$ respectively), where $\alpha \in (0,1)$ is an irrational number such that $\alpha = [0; 1+d_1, \overline{d_2, \ldots, d_n}]$ with $d_n \geq d_1 \geq 1$. It is known that both c_α and $c_{1-\alpha}$ are fixed points of non-trivial (standard) morphisms σ and $\hat{\sigma}$, respectively, if and only if α has a continued fraction expansion as above. Accordingly, such words c_α and $c_{1-\alpha}$ are *generated* by the respective morphisms σ and $\hat{\sigma}$. For the particular case when $\alpha = [0; 2, \overline{r}]$ ($r \geq 1$), we give a decomposition of each *conjugate* of c_α (and hence $c_{1-\alpha}$) into *generalized adjoining singular words*, by considering conjugates of powers of the standard morphism σ by which it is generated. This extends a recent result of Levé and Séébold on conjugates of the infinite Fibonacci word.

Keywords: combinatorics on words; characteristic Sturmian words; conjugation; Sturmian morphisms; standard morphisms; singular words.

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

1 Introduction

In recent years, combinatorial properties of finite and infinite words have become significantly important in fields of physics, biology, mathematics, and computer science. In particular, the fascinating family of Sturmian words has been the subject of many papers (see [2, 13, 11], for example). These words, which represent the simplest family of quasi-crystals, have numerous 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 are (aperiodic) infinite words with exactly $n+1$ distinct factors of length n, for each $n \in \mathbb{N}$. Since this implies that a Sturmian word has exactly two factors of length 1, then any such word is over a two-letter alphabet, say $\mathcal{A} = \{a, b\}$. There are many characterizations and numerous properties of Sturmian words. For a comprehensive study of the basic properties of Sturmian words, and of their transformations by morphisms, see Berstel and Séébold [2].

Here, an *infinite word* x over an alphabet $\mathcal{A}$ is a map $x : \mathbb{N} \longrightarrow \mathcal{A}$. For any $i \geq 0$, we set $x_i = x(i)$ and write $x = x_0x_1x_2\cdots$, where each $x_i \in \mathcal{A}$. In this paper, we will utilize the following characterization of Sturmian words, which was originally proved by Morse and Hedlund [13]. An infinite word 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 s is one of the following two infinite words:

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

*E-mail: amy.glen@adelaide.edu.au; URL: http://www.maths.adelaide.edu.au/∼aglen

defined by

$$s_{\alpha,\rho}(n) = \begin{cases} a & \text{if } \lfloor (n+1)\alpha + \rho \rfloor - \lfloor n\alpha + \rho \rfloor = 0, \\ b & \text{otherwise;} \end{cases}$$

$$(n \geq 0)$$

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

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

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

where c_α is called the *characteristic Sturmian word* of slope α.

The infinite Fibonacci word f is a special example of a characteristic Sturmian word of slope $\alpha = [0; 2, \overline{1}] = (3 - \sqrt{5})/2$, which is generated by a (standard) morphism. Wen and Wen [15] have established a factorization of the Fibonacci word into *singular words* and, in a similar fashion, Melançon [11] has proposed a generalization of singular words over a two-letter alphabet that allows for a decomposition of all the characteristic Sturmian words. More recently, Levé and Séébold [9] have obtained a generalization of Wen and Wen's 'singular' decomposition of the Fibonacci word, by establishing a similar decomposition for each *conjugate* of this infinite word into what they call *generalized singular words.* The aim of this current paper is to extend this latter result to any characteristic Sturmian word of slope $\alpha = [0; 2, \overline{r}]$ (resp. $1 - \alpha = [0; 1, 1, \overline{r}]$), $r \geq 1$, which we will show is generated by a particular *standard* morphism σ (resp. $\hat{\sigma}$).

This paper is organized in the following manner. In Section 2, after recalling some combinatorial notions used in the study of words and morphisms (§2.1), we will consider right conjugation of standard morphisms (§2.2). We shall then discuss characteristic Sturmian words c_α and a 'singular' decomposition of such words, which we will later generalize to each conjugate of c_α for particular α. In the section to follow (§3), we describe all irrationals $\alpha \in (0, 1)$ such that c_α is generated by a morphism, and subsequently obtain generalizations of Levé and Séébold's [9] results (on conjugates of the Fibonacci word) for c_α and $c_{1-\alpha}$ with $\alpha = [0; 2, \overline{r}]$.

2 Preliminaries

2.1 Words and Morphisms

Any of the following terminology that is not further clarified can be found in either [10] or [2], which give more detailed presentations.

Let $\mathcal{A}$ be a finite set of symbols that we shall call an *alphabet*, the elements of which are called *letters.* 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 semi-group*, denoted by $\mathcal{A}^+$, 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 $\mathcal{A}^\infty := \mathcal{A}^* \cup \mathcal{A}^\omega$.

A finite word w is a *factor* of $x \in \mathcal{A}^\infty$ if $x = uwv$ for some $u \in \mathcal{A}^*$ and $v \in \mathcal{A}^\infty$. Furthermore, w is called a *prefix* (resp. *suffix*) of x if $u = \varepsilon$ (resp. $v = \varepsilon$). An infinite word $z \in \mathcal{A}^\omega$ is called a *suffix* of $x \in \mathcal{A}^\omega$ if there is a word $w \in \mathcal{A}^*$ such that $x = wz$. The *length* $|w|$ of a finite word w is defined to be the number of letters it contains. (Note that $|\varepsilon| = 0$.)

The *inverse* of $w \in \mathcal{A}^*$, written w^{-1}, is defined by $ww^{-1} = w^{-1}w = \varepsilon$. It must be emphasized that this is merely notation, i.e. for $u, v, w \in \mathcal{A}^*$, the words $u^{-1}w$ and wv^{-1} are defined only if u (resp. v) is a prefix (resp. suffix) of w. Also note that if $w = uv$ then $wv^{-1} = u$ and $u^{-1}w = v$, and if $x = wx'$, where $w \in \mathcal{A}^*$ and $x' \in \mathcal{A}^\omega$, then $w^{-1}x = x'$.

Two words w, $z \in \mathcal{A}^*$ are said to be *conjugate* if there exist words u, v such that $w = uv$ and $z = vu$. If $|u| = k$, then z is called the k-th conjugate of w. This notion extends to infinite words as follows. For $k \in \mathbb{N}$, the k-th *conjugate* of an infinite word x over $\mathcal{A}$ is the infinite word x' such that $x = ux'$, where $u \in \mathcal{A}^*$ and $|u| = k$.

A *morphism on* $\mathcal{A}$ is a map $\psi : \mathcal{A}^* \longrightarrow \mathcal{A}^*$ 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 $\mathcal{A}$. If $\psi(c) = cw$, for some letter $c \in \mathcal{A}$ and some $w \in \mathcal{A}^+$, then ψ is said to be *prolongable* on c. In this case, the word $\psi^n(c)$ is a proper prefix of the word $\psi^{n+1}(c)$ for each $n \in \mathbb{N}$, and the sequence $(\psi^n(c))_{n\geq 0}$ converges to a unique infinite word

$$x = \lim_{n\to\infty} \psi^n(c) = \psi^\omega(c).$$

An infinite word x is *generated by a morphism* if $x = \psi^\omega(c)$ for some letter c and some morphism ψ.

In what follows, it is assumed that all words are over the two-letter alphabet $\mathcal{A} = \{a, b\}$.

2.2 Conjugation of Standard Morphisms

Define on $\mathcal{A}$ the following three morphisms

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

A morphism ψ is *Sturmian* if and only if $\psi \in \{E, \varphi, \widetilde{\varphi}\}^*$, i.e. if and only if it is a composition of E, φ, and $\widetilde{\varphi}$ in any number and order (see [12]). Furthermore, a morphism ψ is *standard* if and only if $\psi \in \{E, \varphi\}^*$ (see [6]). Note that a morphism is *non-trivial* if it is neither E nor $Id_{\mathcal{A}}$ (the identity morphism).

Suppose ψ and ξ are morphisms on $\mathcal{A}$. If there exists a word u such that

$$\psi(w)u = u\xi(w) \quad \text{for all words } w \in \mathcal{A}^*,$$

then ξ is called the $|u|$-th *right conjugate* of ψ, denoted $\psi_{|u|}$.

It has been shown by Séébold [14] that the number of distinct right conjugates of a standard morphism ψ is $|\psi(ab)| - 1$; namely, the morphisms ψ_0 to $\psi_{|\psi(ab)|-2}$. The following useful lemma is proved in [9].

Lemma 2.1. *Suppose the infinite word x is generated by the standard morphism ψ. Let $k \in \mathbb{N}$ with $0 \leq k \leq |\psi(ab)|-2$, and let v denote the prefix of x of length k. Then ψ_k is such that $\psi_k(x) = v^{-1}x$.* □

Thus, if ψ is a standard morphism that generates an infinite word x, one deduces from Lemma 2.1 that the result of applying ψ_k to x simply consists of deleting the first k letters of x, i.e. $\psi_k(x)$ is the k-th conjugate of x.

2.3 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 $(a_i)_{i\geq 1}$ 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* of 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 $(p_n)_{n\geq 0}$ and $(q_n)_{n\geq 0}$ are given by

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

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, one can prove [7, 3] that each s_n is a prefix of c_α, and we have

$$c_\alpha = \lim_{n\to\infty} s_n.$$

2.3.1 Singular Decomposition of c_α

Melançon [11] (also see [4]) has proposed a generalization of Wen and Wen's [15] singular factors of the Fibonacci word to the case of any characteristic Sturmian word c_α and, in doing so, has established a decomposition of c_α into *adjoining singular words*, as shown below.

For c_α such that $\alpha = [0; 1 + d_1, d_2, d_3, \ldots]$ with $d_1 \geq 1$, Melançon [11] introduced the singular words w_n of c_α defined by

$$w_n = \begin{cases} a s_n b^{-1} & \text{if } n \text{ is odd,} \\ b s_n a^{-1} & \text{otherwise,} \end{cases}$$

for $n \geq 1$, with $w_{-2} = \varepsilon, w_{-1} = a, w_0 = b$. Furthermore, the following words v_n are also defined in [11]. For all $n \geq -1$,

$$v_n = \begin{cases} a s_{n+1}^{d_{n+2}-1} s_n b^{-1} & \text{if } n \text{ is odd,} \\ b s_{n+1}^{d_{n+2}-1} s_n a^{-1} & \text{otherwise.} \end{cases}$$

Clearly, the word v_n differs from w_{n+2} by a factor s_{n+1}, and it can be proved that all v_n and w_n are palindromes (i.e. words that read the same backwards as forwards). We shall call v_n the n-th *adjoining singular word* of c_α, and set $v_{-2} = \varepsilon$. In terms of the singular and adjoining singular words of c_α, the following generalization of Wen and Wen's [15] singular decomposition of the Fibonacci word has been established.

Theorem 2.2. [11] $c_\alpha = \prod_{j=-1}^{\infty} (v_{2j} w_{2j+1})^{d_{2j+3}} = \prod_{j=-1}^{\infty} v_j$. □

In the next section, for the case $\alpha = [0; 2, \overline{r}]$, we will generalize this factorization of c_α (and hence $c_{1-\alpha}$), by showing that, for each prefix v of c_α, $v^{-1}c_\alpha$ can be decomposed into *generalized adjoining singular words*. Such a result has already been established (by Levé and Séébold [9]) for the case of the Fibonacci word $f = c_{(3-\sqrt{5})/2}$, where $\frac{3-\sqrt{5}}{2} = [0; 2, \overline{1}]$.

3 Decompositions of Conjugates of c_α

3.1 Characteristic Sturmian Words Generated by Morphisms

Here, we describe all irrationals $\alpha \in (0, 1)$ such that the characteristic Sturmian word c_α is generated by a morphism. In order to do this, we must first define a special set of irrational numbers. A *Sturm number* (see [2]) is an irrational number $\alpha \in (0, 1)$ that has a continued fraction expansion of one of the following types:

(i) $\alpha = [0; 1 + d_1, \overline{d_2, \ldots, d_n}] < \frac{1}{2}$ with $d_n \geq d_1 \geq 1$;

(ii) $\alpha = [0; 1, d_1, \overline{d_2, \ldots, d_n}] > \frac{1}{2}$ with $d_n \geq d_1$.

Observe that if $\alpha = [0; 1 + d_1, \overline{d_2, \ldots, d_n}]$ with $d_n \geq d_1 \geq 1$, then

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

Hence, α has an expansion of type (i) if and only if $1 - \alpha$ has an expansion of type (ii). Accordingly, α is a Sturm number if and only if $1 - \alpha$ is a Sturm number.

In what follows, we will always assume (unless otherwise stated) that α is a Sturm number of type (i). Also, we shall denote the standard sequence of c_α (resp. $c_{1-\alpha}$) by $(s_n)_{n \geq -1}$ (resp. $(\hat{s}_n)_{n \geq -1}$). Clearly, we have $\hat{s}_1 = \hat{s}_{-1} = b$ since $1 - \alpha = [0; 1, d_1, \overline{d_2, \ldots, d_n}]$. Consequently, $c_{1-\alpha}$ is obtained from c_α by exchanging all letters a and b in c_α, i.e. $c_{1-\alpha} = E(c_\alpha)$. Indeed, it is easily checked that

$$\hat{s}_n = E(s_{n-1}) \quad \text{for all } n \geq 0.$$

Hence,

$$E(c_\alpha) = E\left(\lim_{n \to \infty} s_n\right) = \lim_{n \to \infty} E(s_n) = \lim_{n \to \infty} \hat{s}_{n+1} = c_{1-\alpha}.$$

Therefore, we can restrict our attention to characteristic Sturmian words c_α such that α is a Sturm number of type (i). Later, an analogue of the main result of this paper (Theorem 3.7) will be deduced for $c_{1-\alpha}$.

We say that a morphism ψ *fixes* an infinite word x if $\psi(x) = x$, in which case x is called a *fixed point* of ψ. The following result describes all irrationals $\alpha \in (0, 1)$ such that c_α is a fixed point of a non-trivial morphism.

Theorem 3.1. [5, 8, 1] *Let $\alpha \in (0,1)$ be irrational. Then c_α is a fixed point of a non-trivial morphism σ if and only if α is a Sturm number. In particular, if $\alpha = [0; 1 + d_1, \overline{d_2, \ldots, d_n}]$ with $d_n \geq d_1 \geq 1$, then c_α is the fixed point of any power of the morphism*

$$\sigma : \begin{array}{ccc} a & \mapsto & s_{n-1} \\ b & \mapsto & s_{n-1}^{d_n - d_1} s_{n-2} \end{array}.$$

Further, $c_{1-\alpha}$ is a fixed point of any power of the morphism

$$\hat{\sigma} : \begin{array}{ccc} a & \mapsto & \hat{s}_n^{d_n - d_1} \hat{s}_{n-1} \\ b & \mapsto & \hat{s}_n \end{array}.$$

□

Note that, for any $m \in \mathbb{N}$, both σ^m and $\hat{\sigma}^m$ are standard morphisms. In fact, it was shown by Crisp et al. [5] that

$$\sigma = (\varphi E)^{d_1} E (\varphi E)^{d_2} E \cdots (\varphi E)^{d_{n-1}} E (\varphi E)^{d_n - d_1}; \quad \text{and}$$
$$\hat{\sigma} = E(\varphi E)^{d_1} E (\varphi E)^{d_2} E \cdots (\varphi E)^{d_{n-1}} E (\varphi E)^{d_n - d_1} E.$$

Observe that $\hat{\sigma} = E\sigma E$.

Now, Séébold [14] proved that a standard morphism ψ generates an infinite (characteristic Sturmian) word if and only if

$$\psi \in \{\varphi, E\varphi, \varphi E, E\varphi E\}^+ \backslash (\{E\varphi\}^+ \cup \{\varphi E\}^+).$$

Here, we prove that a characteristic Sturmian word c_γ is generated by a (standard) morphism if and only if γ is a Sturm number. Specifically, we prove that $c_\alpha = \sigma^\omega(a)$ and $c_{1-\alpha} = \hat{\sigma}^\omega(b) = (E\sigma E)^\omega(b)$. These are direct results of the following lemma and corollary.

Lemma 3.2. *For any $k \in \mathbb{N}$, $\sigma(s_k) = s_{k+(n-1)}$. Consequently, if $k \in \mathbb{N}$ is fixed, then*

$$\sigma^m(s_k) = s_{k+m(n-1)} \quad \textit{for all } m \geq 0.$$

Proof. Mathematical induction. □

Corollary 3.3. *For any integer* $m \geq 1$, $\sigma^m(a) = s_{m(n-1)}$ *and* $\sigma^m(b) = s_{m(n-1)}^{d_n - d_1} s_{m(n-1)-1}$. □

As an immediate consequence of the above corollary, we have the following result.

Corollary 3.4. *Let* $\alpha = [0; 1 + d_1, \overline{d_2, \cdots, d_n}]$ *with* $d_n \geq d_1 \geq 1$. *Then*

(i) $c_\alpha = \lim_{m\to\infty} \sigma^m(a) = \sigma^\omega(a)$;

(ii) $c_{1-\alpha} = \lim_{m\to\infty} \hat{\sigma}^m(b) = \hat{\sigma}^\omega(b)$, *where* $\hat{\sigma} = E\sigma E$.

Proof. The fact that $c_\alpha = \lim_{m\to\infty} \sigma^m(a)$ follows from Corollary 3.3 since $\sigma^m(a) = s_{m(n-1)}$, for any integer $m \geq 1$. Moreover, we know that $c_{1-\alpha} = E(c_\alpha)$, so that (ii) is obtained by realizing

$$c_{1-\alpha} = E\left(\lim_{m\to\infty} \sigma^m(a)\right) = \lim_{m\to\infty} E\sigma^m E(b) = \lim_{m\to\infty} (E\sigma E)^m(b).$$

□

In light of Corollary 3.4, one has that if γ is a Sturm number, then c_γ is generated by a (standard) morphism. The converse is trivially true in view of Theorem 3.1.

By considering Melançon's factorization of c_α into adjoining singular words (Theorem 2.2), we will now extend Levé and Séébold's result (Theorem 4.6 in [9]) to the case $\alpha = [0; 2, \overline{r}]$.

3.2 The case $\alpha = [0; 2, \overline{r}]$

Now, if $\alpha = [0; 2, \overline{r}]$, then for each $m \in \mathbb{N}$,

$$v_m = \begin{cases} a s_{m+1}^{r-1} s_m b^{-1} & \text{if } m \text{ is odd,} \\ b s_{m+1}^{r-1} s_m a^{-1} & \text{otherwise,} \end{cases}$$

and $v_{-1} = a s_0^0 s_{-1} b^{-1} = a$. Observe that, for any integer $m \geq 0$, $|\sigma^m(ab)| = |\sigma(s_1)| = |s_{m+1}| = q_{m+1}$. Furthermore, using Corollary 3.3, it is easily checked that, for any $m \geq -1$,

$$v_m = \begin{cases} a\sigma^{m+1}(b) b^{-1} & \text{if } m \text{ is odd,} \\ b\sigma^{m+1}(b) a^{-1} & \text{otherwise,} \end{cases}$$

since $\sigma^{m+1}(b) = s_{m+1}^{r-1} s_m$, for all $m \in \mathbb{N}$. Also note that $|v_m| = q_{m+2} - q_{m+1}$, for all $m \geq -1$.

Whence, we have the following special case of Lemma 2.1.

Lemma 3.5. *Suppose* $\alpha = [0; 2, \overline{r}]$ *and let* $k, m \in \mathbb{N}$ *be such that* $0 \leq k \leq q_{m+1} - 2$. *If* v *denotes the prefix of length* k *of* $c_\alpha (= (\sigma^m)^\omega(a))$, *then* $(\sigma^m)_k(c_\alpha) = v^{-1} c_\alpha$. □

The next lemma shows how to remove a prefix from the 'singular' decomposition of c_α (*cf.* Proposition 4.5 in [9]).

Lemma 3.6. *Suppose* $\alpha = [0; 2, \overline{r}]$ *and let* $k, m \in \mathbb{N}$ *be such that* $k = q_{m+1} - p$ *with* $2 \leq p \leq q_{m+1} - q_m + 1$. *Then*

$$(\sigma^m)_k(c_\alpha) = u^{-1} v_{m-1} \prod_{j=m}^{\infty} v_j,$$

where u *is the prefix of* v_{m-1} *of length* $|u| = q_{m+1} - q_m + 1 - p$.

Proof. We have $q_m - 1 \leq k \leq q_{m+1} - 2$. Thus, if $k = 0$, then we must have $m = 0$. Now, $k = q_1 - p = 0$ implies $p = q_1 = 2$, and therefore, $|u| = q_1 - q_0 + 1 - 2 = 0$. Further, from Theorem 2.2, we have

$$(\sigma^0)_0(c_\alpha) = c_\alpha = \prod_{j=-1}^{\infty} v_j,$$

so the result holds for $k = 0$.

Suppose $k \geq 1$, then $m \geq 1$ and, in this case, observe that

$$q_m - 1 = \sum_{j=-1}^{m-2} (q_{j+2} - q_{j+1}).$$

From Lemma 3.5, we know that $(\sigma^m)_k(c_\alpha)$ is the word obtained from c_α by removing its prefix of length k, i.e. of length at least $\sum_{j=-1}^{m-2}(q_{j+2} - q_{j+1})$. Whence, since $|v_j| = q_{j+2} - q_{j+1}$ for any $j \geq -1$, then $(\sigma^m)_k(c_\alpha)$ is obtained from $c_\alpha = \prod_{j=-1}^{\infty} v_j$ by first removing the prefix $v_{-1}v_0v_1 \cdots v_{m-2}$. Then, from the remaining infinite word $v_{m-1}\prod_{j=m}^{\infty} v_j$, we remove the prefix u of v_{m-1} of length

$$|u| = k - (q_m - 1) = q_{m+1} - p - q_m + 1 = q_{m+1} - q_m + 1 - p.$$

□

Example 3.1. Take $\alpha = [0; 2, \overline{3}] = (\sqrt{13} - 1)/6$, so that

$$c_\alpha = \underline{a}\ \underline{babab}\ \underline{aababababababaa}\ bababababababaababababaabababab\cdots$$

Note that $v_{-1} = a, v_0 = babab, v_1 = aabababaababaa$ since $s_1 = ab$ and $s_2 = abababa$. Hence, by the preceding lemma,

$$\begin{aligned}
c_\alpha &= (\sigma^0)_0(c_\alpha) = v_{-1}v_0v_1v_2v_3\cdots \\
a^{-1}c_\alpha &= (\sigma^1)_1(c_\alpha) = v_0v_1v_2v_3\cdots \\
(ab)^{-1}c_\alpha &= (\sigma^1)_2(c_\alpha) = b^{-1}v_0v_1v_2v_3\cdots \\
(aba)^{-1}c_\alpha &= (\sigma^1)_3(c_\alpha) = (ba)^{-1}v_0v_1v_2v_3\cdots \\
(abab)^{-1}c_\alpha &= (\sigma^1)_4(c_\alpha) = (bab)^{-1}v_0v_1v_2v_3\cdots \\
(ababa)^{-1}c_\alpha &= (\sigma^1)_5(c_\alpha) = (baba)^{-1}v_0v_1v_2v_3\cdots \\
(ababab)^{-1}c_\alpha &= (\sigma^2)_6(c_\alpha) = v_1v_2v_3\cdots\ , \text{ etc.}
\end{aligned}$$

For any $n \geq -1$, set $(\sigma^{n+1})_{-1}(b) = v_n$.

Theorem 3.7. *Let $k, m \in \mathbb{N}$ be such that $k = q_{m+1} - p$ with $2 \leq p \leq q_{m+1} - q_m + 1$. Then*

$$(\sigma^m)_k(c_\alpha) = \prod_{j=m-1}^{\infty} (\sigma^{j+1})_{q_{m+1}-q_m-p}(b).$$

Proof. It follows immediately from Lemma 3.6 that

$$(\sigma^m)_k(c_\alpha) = u^{-1}v_{m-1}\prod_{j=m}^{\infty} v_j = u^{-1}\prod_{j=m-1}^{\infty} v_j,$$

where u is the prefix of v_{m-1} such that $|u| = q_{m+1} - q_m + 1 - p$. Thus, if $u = \varepsilon$, then

$$(\sigma^m)_k(c_\alpha) = \prod_{j=m-1}^{\infty} v_j = \prod_{j=m-1}^{\infty} (\sigma^{j+1})_{-1}(b),$$

and so the result holds since $|u| = 0 = q_{m+1} - q_m + 1 - p$ implies $q_{m+1} - q_m - p = -1$.

If $k = 0$, then $m = 0$, and hence, $p = 2$ so that $|u| = q_1 - q_0 + 1 - 2 = 2 - 1 - 1 = 0$ (i.e. $u = \varepsilon$). So the result holds for $m = 0$, and we therefore take $k \geq 1$, so that $m \geq 1$.

Observe that, by definition of the adjoining singular word v_m, there exist letters $x, y \in \mathcal{A}$ $(x \neq y)$ such that, for any integer $p \geq -1$,

$$v_{m+p} = \begin{cases} x\sigma^{m+p+1}(b)y^{-1} & \text{if } p \text{ is odd,} \\ y\sigma^{m+p+1}(b)x^{-1} & \text{otherwise.} \end{cases}$$

Hence,

$$\begin{aligned}\prod_{j=m-1}^{\infty} v_j &= (x\sigma^m(b)y^{-1})(y\sigma^{m+1}(b)x^{-1})(x\sigma^{m+2}(b)y^{-1})\cdots \\ &= x\sigma^m(b)\sigma^{m+1}(b)\sigma^{m+2}(b)\cdots \\ &= x\prod_{j=m-1}^{\infty}\sigma^{j+1}(b),\end{aligned}$$

and therefore, $(\sigma^m)_k(c_\alpha) = u^{-1}x\prod_{j=m-1}^{\infty}\sigma^{j+1}(b)$. If $u \neq \varepsilon$, then there exists a word $\hat{u}$ such that $u^{-1}x = \hat{u}^{-1}$, which implies that $\hat{u} = x^{-1}u$ with $|\hat{u}| = q_{m+1} - q_m - p$.

For any $n \in \mathbb{N}$, $\sigma^{n+1}(b) = s_{n+1}^{r-1}s_n$ is a prefix of $\sigma^{n+2}(b) = s_{n+2}^{r-1}s_{n+1} = (s_{n+1}^{r-1}s_n)^{r-1}s_{n+1}$, and $|\sigma^{n+1}(b)| = |v_n| = q_{n+2} - q_{n+1}$. Whence, for any integer $r \geq m-1$, we have $\sigma^{r+1}(b) = \hat{u}u_{r+1}$, for some $u_{r+1} \in \mathcal{A}^*$ with $|u_{r+1}| \geq p$. Indeed, for $r \geq m-1$, we have

$$\begin{aligned}|u_{r+1}| = |\sigma^{r+1}(b)| - |\hat{u}| &= q_{r+2} - q_{r+1} - q_{m+1} + q_m + p \\ &= (q_{r+2} - q_{m+1}) - (q_{r+1} - q_m) + p \geq p.\end{aligned}$$

Consequently, by definition of right conjugation of morphisms,

$$(\sigma^{r+1})_{q_{m+1}-q_m-p}(b) = u_{r+1}\hat{u}.$$

From the above observations, we therefore find

$$\begin{aligned}(\sigma^m)_k(c_\alpha) &= u^{-1}x\prod_{j=m-1}^{\infty}\sigma^{j+1}(b) \\ &= \hat{u}^{-1}\prod_{j=m-1}^{\infty}\sigma^{j+1}(b) \\ &= \hat{u}^{-1}\sigma^m(b)\prod_{j=m}^{\infty}\sigma^{j+1}(b) \\ &= \hat{u}^{-1}\hat{u}u_m\prod_{j=m}^{\infty}\sigma^{j+1}(b) \\ &= u_m\prod_{j=m}^{\infty}\sigma^{j+1}(b) \\ &= u_m\hat{u}\hat{u}^{-1}\prod_{j=m}^{\infty}\sigma^{j+1}(b) \\ &= (\sigma^m)_{q_{m+1}-q_m-p}(b)\hat{u}^{-1}\prod_{j=m}^{\infty}\sigma^{j+1}(b) \\ &= \cdots \\ &= \prod_{j=m-1}^{\infty}(\sigma^{j+1})_{q_{m+1}-q_m-p}(b).\end{aligned}$$

□

So if v is a prefix of c_α, then $v^{-1}c_\alpha$ can be obtained by concatenating all the words $[(\sigma^{j+1})_i(b)]_{j\geq l}$, where i and l are integers depending only on $|v|$. The characteristic Sturmian word c_α is the special case when $i = l = -1$, so that

$$c_\alpha = \prod_{j=-1}^{\infty}(\sigma^{j+1})_{-1}(b),$$

where Melançon's adjoining singular words, v_j, are all the words $(\sigma^{j+1})_{-1}(b), j \geq -1$.

Recall that the Fibonacci word f is the characteristic Sturmian word c_α such that $\alpha = [0;2,\overline{1}] = (3-\sqrt{5})/2$. In this case, one has $\sigma = \varphi$, i.e.

$$f = \varphi^\omega(a) = abaababaabaababaababaabaababaabaab\cdots$$

If we set $\varphi^{-1}(a) = b$, then for any integer $n \geq -1$, $\varphi^n(a) = \varphi^{n+1}(b)$ with $|\varphi^n(a)| = F_n = |\varphi^{n+1}(b)|$, where F_n is the n-th *Fibonacci number* defined by

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

Note that $q_m = F_m$ for every $m \in \mathbb{N}$, and

$$F_{m+1} - F_m = F_{m+1} - F_m = F_{m-1} \quad \text{for all } m \in \mathbb{N}.$$

Hence, it is deduced from Theorem 3.7 that if $k, m \in \mathbb{N}$ are such that $k = F_{m+1} - p$ with $2 \leq p \leq F_{m-1} + 1$, then

$$(\varphi^m)_k(f) = \prod_{j=m-1}^{\infty} (\varphi^j)_{F_{m-1}-p}(a),$$

which is Levé and Séébold's result (Theorem 4.6 in [9]).

As an example, we list some decompositions of conjugates of the characteristic Sturmian word c_α for $\alpha = [0;\overline{2}] = \sqrt{2} - 1$.

		v_{-1}	v_0	v_1	v_2	v_3	$\cdots$
$m = 0$	$c_\alpha = (\sigma^0)_{-1}(b)$		$(\sigma^1)_{-1}(b)$	$(\sigma^2)_{-1}(b)$	$(\sigma^3)_{-1}(b)$	$(\sigma^4)_{-1}(b)$	$\cdots$
$m = 1$	$a^{-1}c_\alpha =$		$(\sigma^1)_{-1}(b)$	$(\sigma^2)_{-1}(b)$	$(\sigma^3)_{-1}(b)$	$(\sigma^4)_{-1}(b)$	$\cdots$
	$(ab)^{-1}c_\alpha =$		$(\sigma^1)_0(b)$	$(\sigma^2)_0(b)$	$(\sigma^3)_0(b)$	$(\sigma^4)_0(b)$	$\cdots$
	$(aba)^{-1}c_\alpha =$		$(\sigma^1)_1(b)$	$(\sigma^2)_1(b)$	$(\sigma^3)_1(b)$	$(\sigma^4)_1(b)$	$\cdots$
$m = 2$	$(abab)^{-1}c_\alpha =$			$(\sigma^2)_{-1}(b)$	$(\sigma^3)_{-1}(b)$	$(\sigma^4)_{-1}(b)$	$\cdots$
	$(ababa)^{-1}c_\alpha =$			$(\sigma^2)_0(b)$	$(\sigma^3)_0(b)$	$(\sigma^4)_0(b)$	$\cdots$
	$(ababaa)^{-1}c_\alpha =$			$(\sigma^2)_1(b)$	$(\sigma^3)_1(b)$	$(\sigma^4)_1(b)$	$\cdots$
	$(ababaab)^{-1}c_\alpha =$			$(\sigma^2)_2(b)$	$(\sigma^3)_2(b)$	$(\sigma^4)_2(b)$	$\cdots$
	$(ababaaba)^{-1}c_\alpha =$			$(\sigma^2)_3(b)$	$(\sigma^3)_3(b)$	$(\sigma^4)_3(b)$	$\cdots$
	$(ababaabab)^{-1}c_\alpha =$			$(\sigma^2)_4(b)$	$(\sigma^3)_4(b)$	$(\sigma^4)_4(b)$	$\cdots$
	$(ababaababa)^{-1}c_\alpha =$			$(\sigma^2)_5(b)$	$(\sigma^3)_5(b)$	$(\sigma^4)_5(b)$	$\cdots$
$m = 3$	$(ababaababaa)^{-1}c_\alpha =$				$(\sigma^3)_{-1}(b)$	$(\sigma^4)_{-1}(b)$	$\cdots$

3.3 The case $1 - \alpha = [0;1,1,\overline{r}]$

Now, if $\alpha = [0;2,\overline{r}]$, then $1-\alpha = [0;1,1,\overline{r}]$. By observing that $c_{1-\alpha} = \hat{\sigma}^\omega(b)$, where $\hat{\sigma} = E\sigma E$, it is clear that the following theorem is an immediate consequence of Theorem 3.7 and the lemma below.

Lemma 3.8. *For any standard morphism* ψ,

$$(E\psi E)_k = E\psi_k E; \quad 0 \leq k \leq |\psi(ab)| - 2.$$

Proof. Let $w \in \mathcal{A}^*$, with $|w| = k$, be such that $\psi(z)w = w\psi_k(z)$ for all $z \in \mathcal{A}$. Then, for some $z \in \mathcal{A}$,

$$E\psi E(z)E(w) = E(\psi E(z)w) = E(\psi(y)w), \quad \text{for some } y \in \mathcal{A},\ y \neq z.$$

Therefore, for $z \in \mathcal{A}$ and $0 \leq k \leq |\psi(ab)| - 2$, we have

$$E\psi E(z)E(w) = E(\psi(y)w) = E(w\psi_k(y)) = E(w)E\psi_k E(z).$$

Thus, there exists a word of length k, namely $w' = E(w)$, such that for any $u \in \mathcal{A}^*$,

$$E\psi E(u)w' = w'E\psi_k E(u).$$

□

Theorem 3.9. *Let $k, m \in \mathbb{N}$ be such that $k = q_{m+1} - p$ with $2 \leq p \leq q_{m+1} - q_m + 1$. Then*

$$(\hat{\sigma}^m)_k(c_{1-\alpha}) = \prod_{j=m-1}^{\infty} (\hat{\sigma}^{j+1})_{q_{m+1}-q_m-p}(a).$$ □

4 Concluding Remarks

Note that, by Corollary 3.3, for any integer $m \geq 1$,

$$\sigma^m(b) = s_{m(n-1)}^{d_n - d_1} s_{m(n-1)-1},$$

where σ is the standard morphism that generates c_α, for $\alpha = [0; 1 + d_1, \overline{d_2, \cdots, d_n}]$ with $d_n \geq d_1 \geq 1$. Now, by the periodicity of the continued fraction expansion of α, $d_i = d_{i+n-1}$, for all $i \geq 2$. Hence, it is easily deduced that

$$d_n = d_{m(n-1)+1}, \quad \text{for any integer } m \geq 1,$$

and one may write

$$\sigma^m(b) = s_{m(n-1)}^{d_{m(n-1)+1} - d_1} s_{m(n-1)-1}.$$

Whence, if $d_1 = 1$, then for each $m \geq 1$,

$$v_{m(n-1)-1} = \begin{cases} a\sigma^m(b)b^{-1} & \text{if } m \text{ is even,} \\ b\sigma^m(b)a^{-1} & \text{otherwise.} \end{cases}$$

Accordingly, we do not have a 'nice' expression for each v_k ($k \in \mathbb{N}$) in terms of a power of σ unless $n = 2$, i.e. unless $\alpha = [0; 2, \bar{r}]$ for some $r \in \mathbb{N}^+$. Therefore, we cannot establish an extension of Theorem 3.7 to the case of c_α (nor $c_{1-\alpha}$) with α (as above) having $d_1 \geq 2$ and $n \geq 3$.

5 Acknowledgements

The author would like to thank Bob Clarke and Alison Wolff for their careful reading of the paper and useful suggestions. The research was supported by the George Fraser Scholarship of the University of Adelaide.

References

[1] J. Berstel and P. Séébold, A characterization of Sturmian morphisms, in: A.M. Borzyszkowski and S. Sokolowski (Eds.), *Mathematical Foundations of Computer Science* 1993, *Lecture Notes in Computer Science*, vol. 711, *Springer-Verlag*, Berlin, 1993, pp. 281–290.

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

[3] T.C. Brown, Descriptions of the characteristic sequence of an irrational, *Canad. Math. Bull.* **36**, No.1 (1993), 15–21.

[4] W.-T. Cao and 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] D. Crisp, W. Moran, A. Pollington and P. Shiue, Substitution invariant cutting sequences, *J. Théorie Nombres Bordeaux* **5** (1993), 123–137.

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

[7] A.S. Fraenkel, M. Mushkin and U. Tassa, Determination of $[n\theta]$ by its sequence of differences, *Canad. Math. Bull.* **21**, No.4 (1978), 441–446.

[8] T. Komatsu and A.J. van der Poorten, Substitution invariant Beatty sequences, *Japan. J. Math.* **22**, No.2 (1996), 349–354.

[9] F. Levé and P. Séébold, Conjugation of standard morphisms and a generalization of singular words. Preprint (to appear in *Bull. Soc. Math. Belg.*).

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

[11] 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.

[12] F. Mignosi and P. Séébold, Morphismes Sturmiens et règles de Rauzy, *J. Théorie Nombres Bordeaux* **5** (1993), 221–233.

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

[14] P. Séébold, On the conjugation of standard morphisms, *Theoret. Comput. Sci.* **195** (1998), 91–109, doi:10.1016/S0304–3975(97)00159–X.

[15] Z.-X. Wen and Z.-Y. Wen, Some properties of the singular words of the Fibonacci word, *European J. Combin.* **15**, No.6 (1994), 587–598, doi:10.1006/eujc.1994.1060.