FIXED POINTS OF STURMIAN MORPHISMS AND THEIR DERIVATED WORDS

ABSTRACT

KAREL KLOUDA, KATEŘINA MEDKOVÁ, EDITA PELANTOVÁ, AND ŠTĚPÁN STAROSTA

1. Introduction

Sturmian words are probably the most studied object in combinatorics on words. They are aperiodic words over a binary alphabet having the least factor complexity possible, in other words, their factor complexity satisfies $\mathcal{C}_{\mathbf{u}}(n) = n+1$ for each $n \in \mathbb{N}$. Many properties, characterizations and generalizations are known, see for instance [4, 3, 2].

One of their characterizations is in terms of return words to their factors. Let $\mathbf{u} = u_0u_1u_2\cdots$ be a binary infinite word with $u_i \in \{0,1\}$. Let $w = u_iu_{i+1}\cdots u_{i+n-1}$ be its factor. The integer i is called an *occurrence* of the factor w. A return word to a factor w is a word $u_iu_{i+1}\cdots u_{j-1}$ with i and j being two consecutive occurrences of w such that $i < j$. In [13], Vuillon showed that an infinite word $\mathbf{u}$ is Sturmian if and only if each nonempty factor w has exactly two distinct return words. A straightforward consequence of this characterization is that if w is a prefix of $\mathbf{u}$, we may write

$$\mathbf{u} = r_{s_0}r_{s_1}r_{s_2}r_{s_3}\cdots$$

with $s_i \in \{0,1\}$ and r_0 and r_1 being the two return words to w. The coding of these return words, the word $d_{\mathbf{u}}(w) = s_0s_1s_2\cdots$ is called the *derivated word of* $\mathbf{u}$ *with respect to* w, introduced in [6]. A simple corollary of the characterization by return words and a result of [6] is that the derivated word $d_{\mathbf{u}}(w)$ is also a Sturmian word. This simple corollary follows also from other results. For instance, it follows from [1], where the authors investigate the derivated word of a standard Sturmian word and give its precise description. It also follows from the investigation of a more general setting in [5], which may in fact be used to describe derivated words of any episturmian word — generalized Sturmian words [7].

By the main result of [6], if $\mathbf{u}$ is a fixed point of a primitive morphism, the set of all derivated words of $\mathbf{u}$ is finite (the result also follows from [8]). In this case, again by [6], a derivated word itself is a fixed point of a primitive morphism.

In this article we study derivated words of fixed points of primitive Sturmian morphisms. By the results of [10], any primitive Sturmian morphism may be decomposed using elementary Sturmian morphisms — generators of the Sturmian monoid. We user the elementary Sturmian morphisms to describe the relation between the derivated words of a Sturmian sequence. The main result of our article is an exact description of the morphisms fixing the derivated words $d_{\mathbf{u}}(w)$ of $\mathbf{u}$, where $\mathbf{u}$ is fixed by a Sturmian morphism ψ and w is its prefix. For this purpose, we introduce an operation Δ acting on the set of Sturmian morphisms with unique fixed point, see Definition 4. Iterating this operation we create the desired list of the morphisms as stated in Theorem 5. The Sturmian morphisms with two fixed points are treated separately, see Proposition 6.

We continue our study by counting the number of derivated words, in particular by counting the distinct elements in the sequence $\left(\Delta^k(\psi)\right)_{k\geq 1}$. This number depends on the decomposition of ψ into the generators of the special Sturmian monoid, see below in Section 2. Using this decomposition, Propositions 8 and 9 provide the exact number of derivated words for two specific classes of Sturmian morphisms.

(K. Medková, E. Pelantová) Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19, Prague 1, Czech Republic

(K. Klouda, Š. Starosta) Department of Applied Mathematics, Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00, Prague 6, Czech Republic

E-mail: stepan.starosta@fit.cvut.cz (Štěpán Starosta).

For a general Sturmian morphism ψ, Proposition 7 gives a sharp upper bound on their number. The upper bound depends on the number of the elementary morphisms in the decomposition of ψ.

For our purposes, we do not fix the alphabet of a derivated word; two derivated words which differ only by a permutation of letters are identified one with another. Moreover, in the sequel, we work only with infinite words which are *uniformly recurrent*, i.e. each prefix w of $\mathbf{u}$ occurs in $\mathbf{u}$ infinitely many times and the set of all return words to w is finite (and, thus, the alphabet of the derivated word of $\mathbf{u}$ with respect to w is finite). Our aim is to describe the set

$$\mathrm{Der}(\mathbf{u}) = \{d_{\mathbf{u}}(w) \colon w \text{ is a prefix of } \mathbf{u}\}.$$

2. Sturmian morphisms

Let $\mathcal{A}$ be a finite alphabet. A *morphism* over $\mathcal{A}^*$ is a mapping $\psi : \mathcal{A}^* \mapsto \mathcal{A}^*$ such that $\psi(vw) = \psi(v)\psi(w)$ for all $v, w \in \mathcal{A}^*$. The domain of the morphism ψ can be naturally extended to $\mathcal{A}^{\mathbb{N}}$ by

$$\psi(u_0u_1u_2\cdots) = \psi(u_0)\psi(u_1)\psi(u_2)\cdots .$$

A morphism ψ is *primitive* if there exists a positive integer k such that the letter a occurs in the word $\psi^k(b)$ for each pair of letters $a, b \in \mathcal{A}$. A *fixed point* of a morphism ψ is an infinite word $\mathbf{u}$ such that $\psi(\mathbf{u}) = \mathbf{u}$.

A morphism ψ is a *Sturmian morphism* if $\psi(\mathbf{u})$ is a Sturmian word for any Sturmian word $\mathbf{u}$. The set of Sturmian morphisms together with composition forms the so-called *Sturmian monoid* usually denoted *St*. We work with these four elementary Sturmian morphisms:

$$\varphi_a : \begin{cases} 0 \to 0 \\ 1 \to 10 \end{cases} \qquad \varphi_b : \begin{cases} 0 \to 0 \\ 1 \to 01 \end{cases} \qquad \varphi_\alpha : \begin{cases} 0 \to 01 \\ 1 \to 1 \end{cases} \qquad \varphi_\beta : \begin{cases} 0 \to 10 \\ 1 \to 1 \end{cases}$$

and with the monoid $\mathcal{M}$ generated by them, i.e. $\mathcal{M} = \langle \varphi_a, \varphi_b, \varphi_\alpha, \varphi_\beta \rangle$. The monoid $\mathcal{M}$ is also called *special Sturmian monoid*. For a nonempty word $u = u_0 \cdots u_{n-1}$ over the alphabet $\{a, b, \alpha, \beta\}$ we put

$$\varphi_u = \varphi_{u_0} \circ \varphi_{u_1} \circ \cdots \circ \varphi_{u_{n-1}}.$$

The monoid $\mathcal{M}$ is not free. It is easy to show that for any $k \in \mathbb{N}$ we have

$$\varphi_{\alpha a^k \beta} = \varphi_{\beta b^k \alpha} \quad \text{and} \quad \varphi_{a \alpha^k b} = \varphi_{b \beta^k a}.$$

We can equivalently say that the following rewriting rules hold on the set of words from $\{a, b, \alpha, \beta\}^*$:

$$\alpha a^k \beta = \beta b^k \alpha \quad \text{and} \quad a \alpha^k b = b \beta^k a \qquad \text{for any } k \in \mathbb{N}\,. \tag{1}$$

In [12], the author reveals a presentation of the Sturmian monoid which includes the special Sturmian monoid $\mathcal{M} = \langle \varphi_a, \varphi_b, \varphi_\alpha, \varphi_\beta \rangle$. A presentation of the special Sturmian monoid follows from this result. It is also given explicitly in [9]:

Theorem 1. *Let $w, v \in \{a, b, \alpha, \beta\}^*$. The morphism φ_w is equal to φ_v if and only if the word v can be obtained from w by applying the rewriting rules (1).*

Note that the presentation of a generalization of the Sturmian monoid, the so-called *episturmian monoid*, is also known, see [11]. The next lemma summarizes several simple and well-known properties of Sturmian morphisms we exploit in the sequel.

Lemma 2. *Let $w \in \{a, b, \alpha, \beta\}^+$.*

(i) The morphism φ_w is primitive if and only if w contains at least one Greek letter α or β and at least one Latin letter a or b.
(ii) If φ_w is primitive, then each of its fixed points is aperiodic and uniformly recurrent.
(iii) If φ_w is primitive, then it has two fixed points if and only if w belongs to $\{a, \alpha\}^*$.

For $w \in \{a, b, \alpha, \beta\}^*$ the rules (1) preserve positions in w occupied by Latin letters $\{a, b\}$ and positions occupied by Greek letters $\{\alpha, \beta\}$. We define that $a < b$ and $\alpha < \beta$ which allows the following definition.

Definition 3. *Let $w \in \{a,b,\alpha,\beta\}^*$. The lexicographically greatest word in $\{a,b,\alpha,\beta\}^*$ which can be obtained from w by application of rewriting rules (1) is denoted $N(w)$. If $\psi = \varphi_w$, then the word $N(w)$ is the* normalized name *of the morphism ψ and it is also denoted by $N(\psi) = N(w)$.*

3. Derivated words of fixed points of Sturmian morphisms

Let $\psi \in \langle \varphi_a, \varphi_b, \varphi_\alpha, \varphi_\beta \rangle$ and $N(\psi) = w \in \{a,b,\alpha,\beta\}^* \setminus \{a,\alpha\}^*$ be the normalized name of the morphism ψ. The word w has a prefix $a^k\beta$ or $\alpha^k b$ for some $k \in \mathbb{N}$. This property enables us to define a transformation on the set of morphisms from $\mathcal{M} \setminus \langle \varphi_a, \varphi_\alpha \rangle$. This transformation is in fact the desired algorithm returning the morphisms $\psi_1, \psi_2, \ldots, \psi_\ell$ mentioned above.

Definition 4. *Let $w \in \{a,b,\alpha,\beta\}^* \setminus \{a,\alpha\}^*$ be the normalized name of a morphism ψ, i.e., $\psi = \varphi_w$. We put*

$$\Delta(w) = \begin{cases} N(w'a^k\beta) & \text{if } \ w = a^k\beta w', \\ N(w'\alpha^k b) & \text{if } \ w = \alpha^k b w' \end{cases}$$

and, moreover, $\Delta(\psi) = \varphi_{\Delta(w)}$.

Theorem 5. *Let $\psi \in \langle \varphi_a, \varphi_b, \varphi_\alpha, \varphi_\beta \rangle$ be a primitive morphism and $N(\psi) = w \in \{a,b,\alpha,\beta\}^* \setminus \{a,\alpha\}^*$ be its normalized name. Denote $\mathbf{u}$ the fixed point of ψ. The word $\mathbf{x}$ is (up to a permutation of letters) a derivated word of $\mathbf{u}$ with respect to one of its prefixes if and only if $\mathbf{x}$ is the fixed point of the morphism $\Delta^j(\psi)$ for some $j \geq 1$.*

Given a finite word u, we define the *cyclic shift* of $u = u_0u_1\cdots u_{n-1}$ to be the word

$$\mathrm{cyc}(u) = u_1u_2\cdots u_{n-1}u_0.$$

Proposition 6. *Let $w \in \{a,\alpha\}^*$ be the normalized name of a primitive morphism ψ and let a be its first letter.*

(i) Let $\mathbf{u}$ be the fixed point of ψ starting with 0. Denote $v = b^{-1}N(wb) \in \{a,\beta\}^*$. We have $\mathrm{Der}(\mathbf{u}) = \{\mathbf{v}\} \cup \mathrm{Der}(\mathbf{v})$, where $\mathbf{v}$ is the unique fixed point of the morphism φ_v.

(ii) Let $\mathbf{u}$ be the fixed point of ψ starting with 1. Put $v = \mathrm{cyc}(w)$. We have $\mathrm{Der}(\mathbf{u}) = \mathrm{Der}(\mathbf{v})$, where $\mathbf{v}$ is the fixed point of the morphism φ_v.

4. The number of derivated words

Proposition 7. *If $w \in \{a,b,\alpha,\beta\}^* \setminus \{a,\alpha\}^*$ it the normalized name of a primitive Sturmian morphism $\psi = \varphi_w$ and $\mathbf{u}$ is a fixed point of ψ, then*

$$1 \leq \#\mathrm{Der}(\mathbf{u}) \leq 3|w| - 4\,. \tag{2}$$

Moreover, for any length $n \geq 2$ there exist normalized names $w', w'' \in \{a,b,\alpha,\beta\}^* \setminus \{a,\alpha\}^*$ of length n such that

(i) $\varphi_{w'}$ and $\varphi_{w''}$ are not powers of other Sturmian morphisms,

(ii) for the fixed points $\mathbf{u}'$ and $\mathbf{u}''$ of the morphism $\varphi_{w'}$ and $\varphi_{w''}$, the lower resp. the upper bound in (2) is attained.

We also provide precise numbers of distinct derivated words for these three types of morphisms:

(1) ψ is a standard morphism from $\mathcal{M}$, i.e. $\psi \in \langle \varphi_b, \varphi_\beta \rangle$,
(2) ψ is a standard morphism from $\mathcal{M} \circ E$, i.e. $\psi \in \langle \varphi_b, \varphi_\beta \rangle \circ E$,
(3) ψ is a morphism from $\langle \varphi_a, \varphi_\alpha \rangle$.

To describe these numbers, we introduce the following morphism $F : \{a,b,\alpha,\beta\}^* \mapsto \{a,b,\alpha,\beta\}^*$ determined by

$$F(a) = \alpha, \quad F(\alpha) = a, \quad F(b) = \beta, \quad F(\beta) = b,$$

and we set

$$\mathrm{cyc_F}(w_1w_2w_3\cdots w_n) = w_2w_3\cdots w_nF(w_1)$$

for a finite word $w_1w_2w_3\cdots w_n$.

Proposition 8. *Let $\mathbf{u}$ be a fixed point of a standard Sturmian morphism ψ which is not a power of any other Sturmian morphism.*

(i) If $\psi = \varphi_w$, then $\mathbf{u}$ has $|w|$ distinct derivated words, each of them (up to a permutation of letters) is fixed by one of the morphisms

$$\varphi_{v_0}, \varphi_{v_1}, \varphi_{v_2}, \ldots, \varphi_{v_{|w|-1}}, \quad \textit{where } v_k = \operatorname{cyc}^k(w) \textit{ for } k = 0, 1, \ldots, |w| - 1.$$

(ii) If $\psi = \varphi_w \circ E$, then $\mathbf{u}$ has $|w|$ distinct derivated words, each of them (up to a permutation of letters) is fixed by one of the morphisms

$$\varphi_{v_0} \circ E, \varphi_{v_1} \circ E, \varphi_{v_2} \circ E, \ldots, \varphi_{v_{|w|-1}} \circ E, \quad \textit{where } v_k = \operatorname{cyc}_F^k(w) \textit{ for } k = 0, 1, \ldots, |w| - 1.$$

Proposition 9. *Let $w \in \{\alpha, a\}^*$ be the normalized name of a primitive morphism ψ such that the letter a is a prefix of w. Moreover, assume that ψ is not a power of any other Sturmian morphism.*

(i) The fixed point of ψ starting with 0 has exactly $1 + |w|_\alpha$ distinct derivated words.

(ii) The fixed point of ψ starting with 1 has exactly $1 + |w|_a$ distinct derivated words.

Acknowledgements

K.M., E.P. and Š.S. acknowledge financial support by the Czech Science Foundation grant GAČR 13-03538S. K. M. also acknowledges financial support by the Czech Technical University in Prague grant SGS17/193/OHK4/3T/14.

References

[1] I. M. Araújo and V. Bruyère, *Words derivated from Sturmian words*, Theoret. Comput. Sci., 340 (2005), pp. 204–219.

[2] L. Balková, E. Pelantová, and Š. Starosta, *Sturmian jungle (or garden?) on multiliteral alphabets*, RAIRO-Theor. Inf. Appl., 44 (2010), pp. 443–470.

[3] J. Berstel, *Sturmian and Episturmian words*, in Algebraic Informatics, S. Bozapalidis and G. Rahonis, eds., vol. 4728 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 23–47.

[4] J. Berstel and P. Séébold, *Sturmian words*, in Algebraic Combinatorics on Words, M. Lothaire, ed., Cambridge University Press, 2002, pp. 45–110.

[5] M. Bucci and A. De Luca, *On a family of morphic images of Arnoux-Rauzy words*, in LATA '09: Proceedings of the 3rd International Conference on Language and Automata Theory and Applications, Springer-Verlag, 2009, pp. 259–266.

[6] F. Durand, *A characterization of substitutive sequences using return words*, Discrete Math., 179 (1998), pp. 89–101.

[7] A. Glen and J. Justin, *Episturmian words: a survey*, Theor. Inf. Appl., 43 (2009), pp. 403–442.

[8] C. Holton and L. Q. Zamboni, *Descendants of primitive substitutions*, Theoret. Comput. Syst., 32 (1999), pp. 133–157.

[9] C. Kassel and C. Reutenauer, *Sturmian morphisms,the braid group B_4, Christoffel words and bases of F_2*, Annali di Matematica Pura ed Applicata, 186 (2007), pp. 317–339.

[10] F. Mignosi and P. Séébold, *Morphismes sturmiens et règles de Rauzy*, J. Théor. Nombres Bordeaux, 5 (1993), pp. 221–233.

[11] G. Richomme, *Conjugacy and episturmian morphisms*, Theoret. Comput. Sci., 302 (2003), pp. 1–34.

[12] P. Séébold, *Fibonacci morphisms and Sturmian words*, Theoret. Comput. Sci., 88 (1991), pp. 365–384.

[13] L. Vuillon, *A characterization of Sturmian words by return words*, Eur. J. Combin., 22 (2001), pp. 263–275.