RETURN WORDS AND DERIVATED SEQUENCES TO ROTE SEQUENCES

KATEŘINA MEDKOVÁ

1. Introduction

We study complementary symmetric Rote sequences, which are sequences over the binary alphabet $\{0,1\}$ with factor complexity $C(n)=2n$ and with language closed on exchange of letters $0 \leftrightarrow 1$. We refer about the work in progress [7].

Rote in [8] proved that a sequence $\mathbf{v}=v_0v_1\cdots$ is complementary symmetric Rote sequence if and only if its first difference sequence $\mathbf{u}=u_0u_1\cdots$, which is defined by $u_i=v_i-v_{i+1} \mod 2$, is Sturmian sequence.

Our aim is to describe return words and derivated sequences of complementary symmetric Rote sequences. In the sequel, we will call them simply Rote sequences. The study is based on the link between Rote and Sturmian sequences and recent work about derivated sequences of Sturmian sequences [1, 6].

Let $\mathbf{u}=u_0u_1u_2\cdots$ be an infinite sequence and 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 $i<j$.

Let w be a prefix of $\mathbf{u}$ which has k return words $r_0,r_1,\ldots,r_{k-1}$. Then the sequence $\mathbf{u}$ can be written as a concatenation of these return words: $\mathbf{u}=r_{d_0}r_{d_1}r_{d_2}\cdots$ and the derivated sequence of $\mathbf{u}$ to prefix w is the sequence $\mathbf{d}_{\mathbf{u}}(w)=d_0d_1d_2\cdots$. For simplicity, we consider two derivated sequences to be the same if they differ only by a permutation of letters. We work only with sequences 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.

Recall that the factor w of $\mathbf{u}$ is right special, if both words $w0$ and $w1$ are factors of $\mathbf{u}$. Analogously the left special factor is defined. The factor is bispecial, if it is both left and right special. To find all derivated sequence it suffices to study only right special prefixes. Indeed, if the prefix w is not right special, then there is a unique letter a such that wa is a factor of $\mathbf{u}$. Thus the occurrences of w and wa in $\mathbf{u}$ coincides and they have the same return words and derivated sequences. If $\mathbf{u}$ is aperiodic, then w is always a prefix of some right special prefix of $\mathbf{u}$.

We focus on standard Sturmian sequences, i.e. the sequences whose each prefix is left special. In that case, we can take into consideration only bispecial prefixes to find all derivated sequences.

2. Rote sequences and associated Sturmian sequences

We define the mapping $\mathcal{S}$ which maps factors of Rote sequence to factors of associated Sturmian sequence.

Definition 1. *The mapping $\mathcal{S}:\{0,1\}^+\to\{0,1\}^*$ is for every $v=v_0v_1\cdots v_n\in\{0,1\}^+$ of length at least 2 defined by $\mathcal{S}(v_0v_1\cdots v_n)=u_0u_1\cdots u_{n-1}$, where $u_i=v_i+v_{i+1} \mod 2$ for all $i\in\{0,1,\ldots,n-1\}$, $\mathcal{S}(v_0)=\varepsilon$.*

We can naturally extend the domain of $\mathcal{S}$ to $\{0,1\}^{\mathbb{N}}$ and write the associated Sturmian sequence $\mathbf{u}$ to the Rote sequence $\mathbf{v}$ as $\mathbf{u}=\mathcal{S}(\mathbf{v})$. To each Sturmian sequence, there are two associated Rote sequences $\mathbf{v}$ and $\mathbf{v}'$. Since we have $\mathbf{v}'=E(\mathbf{v})$, we work only with Rote sequences starting with the letter 0 without lose of generality. The factors of Rote sequence $\mathbf{v}$ and associated Sturmian sequence $\mathbf{u}$ are closely related.

Proposition 2. *Let $\mathbf{u}$ be a Sturmian sequence and $\mathbf{v}$ be a Rote sequence such that $\mathbf{u}=\mathcal{S}(\mathbf{v})$. The word u is a factor of $\mathbf{u}$ if and only if both words v,v' such that $u=\mathcal{S}(v)=\mathcal{S}(v')$ are the factors of $\mathbf{v}$. Moreover, for every $i\in\mathbb{N}$, i is an occurrence of u in $\mathbf{u}$ if and only if i is an occurrence of v in $\mathbf{v}$ or an occurrence of v' in $\mathbf{v}$.*

To study return words and derivated sequences of a given Rote sequence, we have to examine these objects in the case of associated Sturmian sequence. In [6] the authors describe derivated sequences of fixed points of Sturmian morphisms. Their basic idea is to suitably decompose a given Sturmian morphism onto some elementary morphisms.

A morphism is a mapping $\psi : \mathcal{A}^* \to \mathcal{B}^*$ such that $\psi(uv) = \psi(u)\psi(v)$ for all $u, v \in \mathcal{A}^*$. If $\mathcal{A} = \mathcal{B}$, ψ is a morphism over $\mathcal{A}^*$. The domain of the morphism ψ can be naturally extended to $\mathcal{A}^{\mathbb{N}}$. The matrix of a morphism ψ over $\mathcal{A}^*$ is a matrix M_ψ defined by $(M_\psi)_{ab} = |\psi(a)|_b$ for all $a, b \in \mathcal{A}$. The morphism is primitive, if there is a positive integer k such that all elements of $(M_\psi)^k$ are positive.

A fixed point of a morphism ψ is a sequence $\mathbf{u}$ such that $\psi(\mathbf{u}) = \mathbf{u}$. The sequence $\mathbf{u}$ is substitutive if $\mathbf{u} = \sigma(\mathbf{v})$ for a morphism σ and a sequence $\mathbf{v}$ which is a fixed point of morphism θ. Moreover, $\mathbf{u}$ is substitutive primitive if θ is primitive. Durand in [3] proved that sequence $\mathbf{u}$ is substitutive primitive if and only if $\mathbf{u}$ has finite number of distinct derivated sequences.

A morphism ψ is Sturmian if $\psi(\mathbf{u})$ is Sturmian sequence for any Sturmian sequence $\mathbf{u}$. Consider the following elementary Sturmian morphisms φ_b and φ_β defined by

$$\varphi_b : \begin{array}{l} 0 \to 0 \\ 1 \to 01 \end{array} \quad \text{with} \quad M_b = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \varphi_\beta : \begin{array}{l} 0 \to 10 \\ 1 \to 1 \end{array} \quad \text{with} \quad M_\beta = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

By [4] to a given standard Sturmian sequence $\mathbf{u}$ we can uniquely assign the pair: directive sequence $\mathbf{z} \in \{b, \beta\}^{\mathbb{N}}$ and the sequence $(\mathbf{u}^{(n)})_{n \geq 0}$, where $\mathbf{u}^{(n)} \in \{0,1\}^{\mathbb{N}}$ is a standard Sturmian sequence, such that for every $n \in \mathbb{N}$ we have

$$\mathbf{u} = \varphi_{z_0 z_1 \ldots z_{n-1}}(\mathbf{u}^{(n)}).$$

The directive sequence $\mathbf{z}$ contains infinitely many letters b and infinitely many letters β. In addition, $\mathbf{z}$ is purely periodic, i.e. $\mathbf{z} = z^\infty$, if and only if $\mathbf{u}$ is the fixed point of the morphism φ_z. This fixing morphism φ_z is always primitive.

3. Return words to prefixes of Rote sequences

Vuillon in [13] showed that an infinite sequence is Sturmian if and only if each non-empty factor has exactly two distinct return words. Using some results from [2] we show that all derivated sequences of Rote sequences to non-empty prefixes are over a ternary alphabet.

Theorem 3. *Let $\mathbf{v}$ be a Rote sequence. Then every non-empty prefix x of $\mathbf{v}$ has exactly three distinct return words.*

In addition, we can construct the return words in Rote sequence using the relevant return words in associated Sturmian sequence. We need an auxiliary definition of stability.

Definition 4. *The word $u = u_0 u_1 \cdots u_{n-1} \in \{0,1\}^*$ is called stable (S) if $\sum_{i=0}^{n-1} u_i = 0 \mod 2$. Otherwise, u is unstable (U).*

Definition 5. *Let w be a prefix of a Sturmian sequence $\mathbf{u}$ with return words r, s. Due to Vuillon's result [9] $\mathbf{u}$ is a concatenation of blocks $r^k s$ and $r^{k+1} s$ or blocks sr^k and sr^{k+1} for some positive integer k. With respect to the return words of w we distinguish three cases:*

i) *w is of type $SU(k)$, if r is stable and s is unstable;*
ii) *w is of type $US(k)$, if r is unstable and s is stable;*
iii) *w is of type $UU(k)$, if both r and s are unstable.*

The type of prefix w is denoted $\mathcal{T}_w$.

We do not define the type SS since it cannot appear in Strumian sequences. We use these prefix types to describe the return words to corresponding prefixes in Rote sequences.

Theorem 6. *Let x be a prefix of a Rote sequence $\mathbf{v}$. Denote by $\mathbf{u}$ the Sturmian sequence $S(\mathbf{v})$ and by w its prefix $S(x)$ with return words r and s. Then the prefix x of $\mathbf{v}$ has three return words $A, B, C \in \{0,1\}^+$ such that:*

i) *$r = \mathcal{S}(A0)$, $sr^{k+1}s = \mathcal{S}(B0)$ and $sr^k s = \mathcal{S}(C0)$ if w is of type $SU(k)$;*

ii) *$rr = \mathcal{S}(A0)$, $rsr = \mathcal{S}(B0)$ and $s = \mathcal{S}(C0)$ if w is of type $US(k)$;*
iii) *$rr = \mathcal{S}(A0)$, $rs = \mathcal{S}(B0)$ and $sr = \mathcal{S}(C0)$ if w is of type $UU(k)$.*

Remark 7. We can also determine the type of w using the matrix P_w composed of the Parikh vectors of return words to w. Let w be a prefix of a Sturmian sequence $\mathbf{u}$ with return words r, s, where r is the more frequent return word. Then the matrix P_w is defined by:

$$P_w = \begin{pmatrix} |r|_0 & |s|_0 \\ |r|_1 & |s|_1 \end{pmatrix} \mod 2\,,$$

where $|u|_a$ denotes the number of letters a in the word u. Then the type $\mathcal{T}_w$ of the prefix w is

i) SU if $P_w = \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix}$ ii) US if $P_w = \begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix}$ iii) UU if $P_w = \begin{pmatrix} p & q \\ 1 & 1 \end{pmatrix}$

for some numbers $p, q \in \{0, 1\}$.

4. Derivated sequences of Rote sequences

Theorem 8. *Let $\mathbf{v}$ be a Rote sequence with non-empty bispecial prefix x. Then the derivated sequence $\mathbf{d_v}(x)$ is uniquely determined by derivated sequence $\mathbf{d_u}(w)$ of $\mathbf{u} = \mathcal{S}(\mathbf{v})$ to the prefix $w = \mathcal{S}(x)$ and by type $\mathcal{T}_w$ of prefix w.*

Moreover, we are able to construct derivated sequence $\mathbf{d_v}(x)$ of a given Rote sequence $\mathbf{v}$ to a prefix x if we know the type of $S(x)$ and the derivated sequences of $S(\mathbf{v})$ to $S(x)$.

First, we explain how to determine the type of a given prefix w of a standard Sturmian sequence $\mathbf{u}$. Clearly it suffices to study only bispecial prefixes. Let us order the bispecial prefixes of $\mathbf{u}$ by their length starting from the shortest one. Their types can be determined using results from Section 3 of [6], where the authors explain how prefixes and their return words change under application of morphisms φ_b and φ_β.

Proposition 9. *Let $\mathbf{u}$ be a standard Sturmian sequence with directive sequence $\mathbf{z} \in \{b, \beta\}^{\mathbb{N}}$. Then the type of its n-th bispecial prefix w is given by the natural number k and the matrix P_w, where*

i) *$P_w = M_{z_0} M_{z_1} \cdots M_{z_{n-1}} \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \mod 2$ if the sequence $z_n z_{n+1} z_{n+2} \cdots$ has a prefix $b^k \beta$,*
ii) *$P_w = M_{z_0} M_{z_1} \cdots M_{z_{n-1}} \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \mod 2$ if the sequence $z_n z_{n+1} z_{n+2} \cdots$ has a prefix $\beta^k b$.*

The derivated sequences of a given Sturmian sequence $\mathbf{u}$ are described in detail in [6]. In particular, if $\mathbf{u}$ is a standard Sturmian sequence with directive sequence $\mathbf{z} = z_0 z_1 z_2 \cdots$, then the derivated sequence $\mathbf{d_u}(w)$ to the n-th bispecial prefix w is standard Sturmian sequence with directive sequence $z_n z_{n+1} z_{n+2} \cdots$.

It is well known that all Sturmian sequences are 2iet sequences. We show that derivated sequences of Rote sequences are 3iet sequences. A three interval exchange transformation $T : [0, 1) \to [0, 1)$ is given by partition of interval $[0, 1)$ into three subintervals $I_A = [0, \alpha)$, $I_B = [\alpha, \alpha + \beta)$ and $I_C = [\alpha + \beta, 1)$ and by permutation π on the set $\{1, 2, 3\}$ which expresses how these subintervals are rearranged, see [5] for more details. The 3iet sequence $\mathbf{u} = u_0 u_1 u_2 \cdots \in \{A, B, C\}^{\mathbb{N}}$ of transformation T with initial point $\rho \in [0, 1)$ is defined by $u_n = L$ if $T^n(\rho) \in I_L$ for all $n \in \mathbb{N}$.

Proposition 10. *Let $\mathbf{v}$ be a Rote sequence with non-empty prefix x, let $\mathbf{u} = \mathcal{S}(\mathbf{v})$ be a standard Sturmian sequence with prefix $w = \mathcal{S}(x)$. If $\mathbf{d_u}(w)$ is 2iet sequence of transformation with intervals $[0, \alpha)$ and $[\alpha, 1)$ and initial point ρ, then the derivated sequence $\mathbf{d_v}(x)$ is 3iet sequence of transformation T with initial point ρ, where T is given by:*

i) *intervals $[0, \alpha)$, $[\alpha, 2\alpha - k(1 - \alpha))$, $[2\alpha - k(1 - \alpha), 1)$ and permutation $(3, 2, 1)$ if w is of type $SU(k)$;*
ii) *intervals $[0, 2\alpha - 1)$, $[2\alpha - 1, \alpha)$, $[\alpha, 1)$ and permutation $(3, 2, 1)$ if w is of type $US(k)$;*
iii) *intervals $[0, 2\alpha - 1)$, $[2\alpha - 1, \alpha)$, $[\alpha, 1)$ and permutation $(2, 3, 1)$ if w is of type $UU(k)$.*

Now we suppose that directive sequence $\mathbf{z}$ is purely periodic, i.e. $\mathbf{u}$ is fixed point of morphism φ_z.

By [6], sequence $\mathbf{u}$ has at most $|z|$ distinct derivated sequences. We prove that the associated Rote sequence $\mathbf{v}$ has a finite number of distinct derivated sequences too and so by Durand's result from [3] it is substitutive primitive.

Theorem 11. *Let* $\mathbf{v}$ *be a Rote sequence and let* $\mathcal{S}(\mathbf{v}) = \mathbf{u}$ *be a standard Sturmian sequence which is a fixed point of a morphism* φ_z. *Then*

i) $\mathbf{v}$ *has at most* $3|z|$ *distinct derivated sequences, each of them is fixed by a morphism;*
ii) $\mathbf{v}$ *is substitutive primitive.*

Moreover, using our results and Durand's construction from [3] we can construct the fixing morphisms of these derivated sequences algorithmically.

Acknowledgements

Kateřina Medková acknowledges financial support by the Czech Technical University in Prague grant SGS17/193/OHK4/3T/14.

References

[1] I. M. Araújo, V. Bruyère, *Words derivated from Sturmian words*, Theoret. Comput. Sci., 340 (2005), pp. 204–219.
[2] L. Balková, E. Pelantová, and W. Steiner, *Sequences with constant number of return words*, Monatsh. Math., 155 (2008), pp. 251–263.
[3] F. Durand, *A characterization of substitutive sequences using return words*, Discrete Math., 179 (1998), pp. 89–101.
[4] J.Justin, G. Pirillo, *Episturmian words and episturmian morphisms*,Theoret. Comput. Sci., 276 (2002), pp. 281–313.
[5] M. Keane, *Interval exchange transformations*, Mathematische Zeitschrift, 141 (1975), pp. 25–31.
[6] K. Klouda, K. Medková, E. Pelantová and Š. Starosta, *Fixed points of Sturmian morphisms and their derivated words*, preprint: arXiv:1801.09203.
[7] K. Medková, E. Pelantová, L. Vuillon, *Return words and derivated sequences to Rote sequences*, in preparation.
[8] G. Rote, *Sequences with subword complexity* $2n$, J. of Number Theory, 46 (1993), pp. 196–213.
[9] L. Vuillon, *A characterization of Sturmian words by return words*, Eur. J. Combin., 22 (2001), pp. 263–275.