The subword complexity of smooth words on 2-letter alphabets

Yun Bao Huang
Department of Mathematics
Hangzhou Normal University
Xiasha Economic Development Area
Hangzhou, Zhejiang 310036, China
huangyunbao@sina.com
huangyunbao@gmail.com

2010.12.13

Abstract. Let $\gamma_{a,b}(n)$ be the number of smooth words of length n over the alphabet $\{a,b\}$ with $a<b$. Say that a smooth word w is *left fully extendable* (LFE) if both aw and bw are smooth. In this paper, we prove that for any positive number ξ and positive integer n_0 such that the proportion of b's is larger than ξ for each LFE word of length exceeding n_0, there are two constants c_1 and c_2 such that for each positive integer n, one has

$$c_1 \cdot n^{\frac{\log(2b-1)}{\log(1+(a+b-2)(1-\xi))}} < \gamma_{a,b}(n) < c_2 \cdot n^{\frac{\log(2b-1)}{\log(1+(a+b-2)\xi)}}.$$

In particular, taking $a=1$ and $b=2$ in the above inequalities arrives at Huang and Weakley's result. Moreover, for 2-letter even alphabet $\{a,b\}$, there are two suitable constants c_1, c_2 such that

$$c_1 \cdot n^{\frac{\log(2b-1)}{\log((a+b)/2)}} < \gamma_{a,b}(n) < c_2 \cdot n^{\frac{\log(2b-1)}{\log((a+b)/2)}} \textit{for each positive integer } n.$$

Keywords: Derivative; height; smooth word; LFE word.

1. Introduction

The curious Kolakoski sequence K which Kolakoski introduced in [19], is the infinite sequence over the alphabet $\Sigma = \{1, 2\}$, which starts with 2 and equals the sequence defined by its run lengths:

$$K = \underbrace{22}_{2}\underbrace{11}_{2}\underbrace{2}_{1}\underbrace{1}_{1}\underbrace{22}_{2}\underbrace{1}_{1}\underbrace{22}_{2}\underbrace{11}_{2}\underbrace{2}_{1}\underbrace{11}_{2}\underbrace{22}_{2}\underbrace{\cdots}_{\cdots}.$$

Here, a run is a maximal subsequence of consecutive identical symbols. The Kolakoski sequence K has received a remarkable attention in $[1, 2, 3, \cdots, 26]$. For research situations of the Kolakoski sequence K and related problems before 1996, readers can refer to Dekking [12].

Keane [17] asked whether the density of $1'$s in K is 0.5. Chvátal [9] proved that the upper density of $1'$s as well as the upper density of $2'$s in K is less than 0.500838. Steacy [24] studied the structure in the kolakoski sequence K and obtained some conditions which are equivalent to Keane$'$s problem.

In order to study wether the Kolakoski sequence K is recurrent and/or is closed under complement, Dekking [11] introduced the notion of C^{∞}-words over the alphabet $\{1, 2\}$ for the first time and noted that the finite factors of K must be C^{∞}-words. Moreover, he proved that there exists a suitable positive constant c such that $c \cdot n^{2.15} \leq \gamma(n) \leq n^{7.2}$ and conjectured that there are suitable constants c_1 and c_2 such that $c_1 n^q \leq P_K(n) \leq c_2 n^q$, where $\gamma(n)$ denote the number of C^{∞}-words of length n, $P_K(n)$ denote the number of subwords (factors) of length n which occur in the Kolakoski sequence K, $q = (\log 3)/\log(3/2)$.

Weakley [26] showed that there are positive constants C_1 and C_2 such that for each n satisfying $B(k-1) + 1 \leq n \leq A(k) + 1$ for some k, $C_1 n^q \leq \gamma(n) \leq C_2 n^q$, where $A(k)$, $B(k)$ denote respectively the minimum and the maximal length of FE words of height k ([26] Corollary 9).

Huang and Weakley [15] proved that for any positive number ϕ and positive integer n_0 satisfying $|u|_2/|u| > \frac{1}{2} - \phi$ for each LDE word u of length exceeding n_0, there are two suitable constants c_1 and c_2 such that

$$c_1 n^{\frac{\log 3}{\log((3/2)+\phi+(2/N))}} < \gamma(n) < c_2 n^{\frac{\log 3}{\log((3/2)-\phi)}} \text{ for each } n \in N.$$

With the best value known for ϕ, and large N, this gives

$$c_1 n^{2.7087} < \gamma_{1,2}(n) < c_2 n^{2.7102}.$$

A naturally arising question is whether or not we can establish the estimates of subword complexity function of smooth words for the other 2-letter alphabets. This paper is a study of subword complexity function of smooth words for any 2-letter alphabets (Theorem 10). We establish the bounds of minimal and maximal heights of smooth words of length n (Lemma 9), the best bounds of minimal and maximal heights of smooth words of length n for 2-letter even alphabets (Lemma 13) and the good lower and upper bounds of the subword complexity function $\gamma_{a,b}(n)$ for 2-letter even alphabet $\{a, b\}$ (Theorem 14), which would give $\gamma_{a,b}(n) \approx cn^{\log(2b-1)/\log\frac{a+b}{2}}$, where c is a suitable constant.

The paper is structured as follows. In Section 2, we shall first fix some notations and introduce some notions. Second in Section 3, we give some lemmas which are needed to establish the estimates of the complexity function for arbitrary 2-letter alphabets. Third, in Section 4, we obtain the lower and upper bounds of the subword complexity function of smooth words. Moreover, in Section 5, we establish the good lower and upper bounds of the subword complexity function $\gamma_{a,b}(n)$ for 2-letter even alphabets. Finally, in Section 6, we end this paper with some concluding remarks.

2. Definitions and notation

Let $\Sigma = \{a, b\}$ with $a < b$ and a, b being positive integers, Σ^* denotes the free monoid over Σ with ε as the empty word. A finite word over Σ is an element of Σ^*. If $w = w_1w_2\cdots w_n$, $w_i \in \Sigma$ for $i = 1, 2, \cdots, n$, then n is called the length of the word w and is denoted by $|w|$. Let $|w|_\alpha$ be the number of α which occur in w for $\alpha \in \Sigma$, then $|w| = |w|_a + |w|_b$.

Given a word $w \in \Sigma^*$, a *factor* (or *subword*) u of w is a word $u \in \Sigma^*$ such that there exist $x, y \in \Sigma^*$ such that $w = xuy$. If $x = \varepsilon$ then u is called *prefix*. A *run* (or *block*) is a maximal factor of the form $u = \alpha^k, \alpha \in \Sigma$. Finally, N is the set of positive integers and the cardinal number of A is denoted by $|A|$ for a set A.

The *reversal* (or *mirror image*) of $u = u_1u_2\cdots u_n \in \Sigma^*$ is the word $\tilde{u} = u_nu_{n-1}\cdots u_2$ u_1. The *complement* (or *permutation*) of $u = u_1u_2\cdots u_n \in \Sigma^*$ is the word $\bar{u} = \bar{u}_1\bar{u}_2\cdots\bar{u}_n$, where $\bar{a} = b, \bar{b} = a$.

Now we generalize the definition of differentiable words, which Dekking first introduced in [11], to over arbitrary 2-letter alphabet $\{a, b\}$ from the alphabet $\{1, 2\}$.

To do so, for $w \in \Sigma^*$, $r(w)$ denotes the number of runs of w, $fr(w)$ and $lr(w)$

denote the first and last runs of w respectively, and $lfr(w)$ and $llr(w)$ denote the lengths of the first and last run of w respectively. For example, if $w = a^2b^{2b}a^ab^3$, then $r(w) = 4$, $fr(w) = a^2$, $lr(w) = b^3$, $lfr(w) = 2$ and $llr(w) = 3$.

Then we first need to introduce the concept of the closure of a word w over Σ in order to establish the notion of differentiable word for arbitrary 2-letter alphabets.

Definition 1. *Let* $w \in \Sigma^*$ *and*

$$w = \alpha^{t_1}\bar{\alpha}^{t_2}\dots\beta^{t_k}, \tag{2.1}$$

where $\alpha \in \Sigma$, $\beta = \alpha$ *if* $2 \nmid k$, *or else* $\beta = \bar{\alpha}$, $1 \leq t_i \leq b$ *for* $1 \leq i \leq k$.

$$\hat{w} = \begin{cases} w, & lfr(w) \leq a \text{ and } llr(w) \leq a \\ \alpha^{b-t_1}w, & lfr(w) > a \text{ and } llr(w) \leq a \\ w\beta^{b-t_k}, & lfr(w) \leq a \text{ and } llr(w) > a \\ \alpha^{b-t_1}w\beta^{b-t_k}, & lfr(w) > a \text{ and } llr(w) > a \end{cases}$$

Then $\hat{w}$ *is said to be the closure of a word* w.

For example, let $w = 3311133313133311133$, $u = 3313133311$, then u is a factor of w, and $\hat{w} = 333111333131333111333$, $\hat{u} = 333131333111$. Thus $\hat{u}$ is a factor of $\hat{w}$, which also holds in general (see Lemma 3 (1)).

Definition 2. *Let* $w \in \Sigma^*$ *be of the form* (2.1). *If the length of every run of* w *only takes* a *or* b *except for the lengths of the first and last runs, then we call that* w *is differentiable, and its derivative, denoted by* $D(w)$, *is the word whose* j*th symbol equals the length of the* j*th run of* w, *discarding the first and/or the last run if its length is less than* b.

If $\hat{w}$ is differentiable, then we call that w is closurely differentiable. If a finite word w is arbitrarily often closurely differentiable, then we call w a $C^\infty_{a,b}$-word or a smooth word over the alphabet $\{a, b\}$, and the set of all smooth words over the alphabet $\{a, b\}$ is denoted by $C^\infty_{a,b}$ or C^∞.

Let $\rho(w) = D(\hat{w})$, then it is clear that w is a smooth word if and only if there is a positive integer k such that $\rho^k(w) = \varepsilon$.

Note that if $b = a + 1$ then $\hat{w} = w$. Thus, w is differentiable if and only if w is closurely differentiable, which suggests that w is a smooth word if and only if there is a positive integer k such that $D^k(w) = \varepsilon$.

By the definition 2, it is clear that if $b-a\geq 2$ and $a\neq 2$, then $a^{b-1}b^a a^a b^{b-1}$ is differentiable but not closurely differentiable. Moreover, D is an operator from Σ^* to Σ^*, $r(w)\leq |D(w)|+2$ and

$$D(\hat{w})=\begin{cases} bD(w), & b>lfr(w)>a \text{ and } llr(w)\leq a \\ D(w)b, & b>llr(w)>a \text{ and } lfr(w)\leq a \\ bD(w)b, & b>lfr(w)>a \text{ and } b>llr(w)>a \\ D(w), & \text{otherwise} \end{cases}. \tag{2.2}$$

From (2.2), it follows that if w is closurely differentiable, then it must be differentiable.

A word v such that $D(v)=w$ is said to be a primitive of w. The two primitives of w having minimal length are the shortest primitives of w. For example, b have $2b^2$ primitives of the form $\alpha^i\bar{\alpha}^b\alpha^j$, where $\alpha=a,b$, $i,j=0,1,\cdots b-1$, and a^b, b^b are the shortest primitives. It is easy to see that for any word $w\in C^\infty$, there are at most $2b^2$ primitives, and the difference of lengths of two primitives of w is at most $2(b-1)$.

The height of a smooth word w is the smallest integer k such that $D^{k+1}(w)=\varepsilon$. We write $ht(w)$ for the height of w. For example, if $w=32^3 3^3 2^3 3^2 2^2 3^2 2^3 3^3 2^3 3$, then $ht(w)=3$.

It immediately follows from the definition 2 that

(1) $D(\tilde{u})=\widetilde{D(u)}$, $D(\bar{u})=D(u)$ for each $u\in\Sigma^*$

(2) $w\in C^\infty \Longleftrightarrow \bar{w},\tilde{w}\in C^\infty$.

3. Some lemmas

The following Lemmas 3 to 5 reveal the relations among the operators mirror image, complement, closure and derivative.

Lemma 3 ([16], Lemma 5). *Let w be a differentiable word and u is a factor of w. Then*

(1) *both $\hat{u}$ and w are factors of $\hat{w}$;*

(2) $\hat{\tilde{w}}=\tilde{\hat{w}}, \hat{\bar{w}}=\bar{\hat{w}}$;

(3) *$D(u)$ is a factor of $D(w)$;*

(4) *If w is closurely differentiable, then both $\rho(u)$ and $D(w)$ are factors of $\rho(w)$, and $\rho(\bar{w})=\rho(w)$, $\rho(\tilde{w})=\widetilde{\rho(w)}$.*

Proof. (1) From the definition 1 of the closure of a word, it follows the assertion (1).

(2) It immediately follows from the definitions of the closure, complement and mirror image of a word w and the definition of the operators ρ.

(3) Since u is a factor of w, by the definition 2 of the derivative of a word w, we see that $D(u)$ is a factor of $D(w)$.

(4) Since w is closurely differentiable and $\rho(w) = D(\hat{w})$, by the assertion (1), $\hat{u}$ and w are both factors of $\hat{w}$. Moreover by the assertion (3), we see that $D(\hat{u})$ and $D(w)$ are factors of $D(\hat{w})$, that is, both $\rho(u)$ and $D(w)$ are factors of $\rho(w)$. Finally, by the assertion (2), we have $\rho(\bar{w}) = D(\hat{\bar{w}}) = D(\bar{\hat{w}}) = D(\hat{w}) = \rho(w)$. Similarly, $\rho(\tilde{w}) = D(\hat{\tilde{w}}) = D(\tilde{\hat{w}}) = \widetilde{D(\hat{w})} = \widetilde{\rho(w)}$. $\square$

From the definitions 1-2, it immediately follows that

Lemma 4 ([16], Lemma 6)**.** *Let $w = w_1 w_2 \cdots w_n$ be a differentiable word with $n \geq a+1$.*

(1) *If $lfr(w) = b$ then $w_1 w$ is not a differentiable word and $D(\bar{w}_1^i w) = D(w)$ for $i \leq b-1$;*

(2) *If $lfr(w) < b$ then $D(w_1^{b-lfr(w)} w) = bD(w)$;*

(3) *If $lfr(w) \leq a$ and $r(w) > 1$ then $D(\bar{w}_1 w_1^{a-lfr(w)} w) = aD(w)$.* $\square$

Lemma 5 ([16], Lemma 7)**.** (1) *Let $w = w_1 w_2 \cdots w_n$ be a smooth word. Then any factor of w is also a smooth word;*

(2) *Any smooth word $w = w_1 w_2 \cdots w_n$ has both a left and a right smooth extensions.*

Proof. (1) If w is a smooth word and u is a factor of w, then note that $w \in C^\infty \Longleftrightarrow \rho^k(w) = \varepsilon$ for some positive integer k, by Lemma 3 (4), we obtain that $\rho^i(u)$ is a factor of $\rho^i(w)$ for any positive integer $i \leq k$. And hence $\rho^k(w) = \varepsilon$ suggests $\rho^k(u) = \varepsilon$, so that u is a smooth word.

(2) We verify the assertion (2) by induction on $|w|$. Since $D(\tilde{w}) = \widetilde{D(w)}$, we only need to verify that w has a left smooth extension. It is clear that if $r(w) \leq 1$, where $r(w)$ is the number of runs of w, then the assertion (2) holds. We proceed to the induction step. Assume now that $r(w) \geq 2$ and the assertion (2) holds for smooth words shorter than w.

If $lfr(w) \leq a$ then by Lemma 4 (2-3), we have $D(\bar{w}_1 w_1^{a-lfr(w)} w) = aD(w)$ and $D(w_1^{b-lfr(w)} w) = bD(w)$. Thus by $|D(w)| < |w|$, we see that at least one of $aD(w)$ and $bD(w)$ is a smooth word, which means that w has a left smooth extension.

If $b > lfr(w) > a$, then by $w \in C^\infty$, we obtain that $\hat{w}$ is a left smooth extension of w.

If $b = lfr(w)$, then by Lemma 4 (1), we see that $\bar{w}_1 w$ is a left smooth extension of w. □

Now we are in a position to generalize the notion of LDE words to over arbitrary 2-letter alphabets from the alphabet $\{1, 2\}$, which Weakley first introduced in [26].

If aw and bw are both smooth, then the word w is said to be *left fully extendable* (LFE). Clearly, LFE words are closed under complement. For every nonnegative integer k, let LF_k denote the set of LFE words of length k.

Let $\gamma_{a,b}(k)$ denote the number of smooth words of length k over the alphabet $\{a, b\}$. Being similar to Weakley [26], define the differences of $\gamma_{a,b}$ by $\gamma'_{a,b}(k) = \gamma_{a,b}(k+1) - \gamma_{a,b}(k)$ for each $k \geq 0$. From the definition of LFE words, it immediately follows that $\gamma'_{a,b}(k) = |LF_k|$ for each nonnegative integer k. Since $\gamma(0) = \gamma'(0) = 1$, so

$$\gamma_{a,b}(k) = \gamma_{a,b}(0) + \sum_{i=0}^{k-1} \gamma'_{a,b}(i) = 1 + \sum_{i=0}^{k-1} |LF_i| = 2 + \sum_{i=1}^{k-1} |LF_i| \text{ for } k \geq 1. \quad (3.1)$$

Lemma 6. *Let $w = w_1 w_2 \cdots w_k$ be a smooth word, where $k \in N$. If w is a LFE word then $D(w)$ is also a LFE word, and if $k \geq b$ or $r(w) > 1$ then $w = w_1^a w_{a+1} \dots w_k$, where $w_1 \neq w_{a+1}$.*

Proof. Assume that w is a LFE word of length exceeding 0. If $k = |w| < b$ then it follows from both $w_1 w$ and $\bar{w}_1 w$ being smooth words that $w = \overbrace{\alpha^a \dots \beta^a}^{t} \bar{\beta}^j$, where $\alpha \in \Sigma$, $\beta = \alpha$ if $2 \nmid t$, otherwise $\beta = \bar{\alpha}$, $0 \leq j, t \leq b-1$, $j + t \geq 1$, $k = t \cdot a + j$. So $D(w) = a^{t-1}$ if $j, t \geq 1$, or else $D(w) = \varepsilon$. So, in view of $t < b$ we see that $aD(w)$ and $bD(w)$ are both smooth words, that is, $D(w)$ is a LFE word.

If $k \geq b$, since $w_1 w$ is a smooth word, we get $lfr(w) < b$, which suggests that $w = w_1^a w_{a+1} \dots w_k$ and $w_{a+1} \neq w_1$ by $\bar{w}_1 w \in C^\infty$. Moreover, note that each smooth word has a left smooth extension (Lemma 5 (2)), from $w_1 w, \bar{w}_1 w \in C^\infty$ it follows that $aD(w)(= D(\bar{w}_1 w))$ and $bD(w)(= D(w_1^{b-a} w))$ are both smooth words, that is, $D(w)$ is a LFE word. □

Let LF denote the set $\bigcup_{i=0}^{\infty} LF_i$ and $P(A) = \{u \in LF : |u| > 0 \text{ and } D(u) \in A\}$ for $A \subseteq \Sigma^*$. We now give the number of the elements contained in $P^j(\varepsilon)$ for $j \in N$.

Lemma 7. $|P^j(\varepsilon)| = 4(b-1)(2b-1)^{j-1}$ *for* $j \in N$.

Proof. By Lemma 6 and the definition of $P(A)$, we see that $P^{j+1}(\varepsilon)$ is exactly composed of all LFE primitives of $P^j(\varepsilon)$.

Since for each LFE words of the form $\alpha \dots b$ there are exactly $2b$ LFE primitives:

$$\bar{\beta}^a \Delta_\beta^{-1}(\alpha \dots b)\gamma^j,$$

where $\alpha,\ \beta \in \Sigma,\ j = 0, 1, \dots, b-1;\ \gamma = \beta$ if $2 \mid |\alpha \dots b|,$ or else $\gamma = \bar{\beta}$.

for each LFE words of the form $\alpha \dots a$ there are exactly $2(b-1)$ LFE primitives:

$$\bar{\beta}^a \Delta_\beta^{-1}(\alpha \dots a)\gamma^j,$$

where $\alpha,\ \beta \in \Sigma,\ j = 1, \dots, b-1;\ \gamma = \beta$ if $2 \mid |\alpha \dots a|,$ or else $\gamma = \bar{\beta}$.

In addition, because of $\overline{\alpha \dots b} = \bar{\alpha} \dots a$, we see that the numbers of LFE words of the form both $\alpha \dots b$ and $\alpha \dots a$ are equal in all LFE words of the same heights. It follows that

$$\begin{aligned} |P^j(\varepsilon)| &= 2b \cdot \frac{1}{2}|P^{j-1}(\varepsilon)| + 2(b-1) \cdot \frac{1}{2}|P^{j-1}(\varepsilon)| \\ &= (2b-1)|P^{j-1}(\varepsilon)| \text{ for } j \in N, \end{aligned}$$

which suggests that

$$|P^j(\varepsilon)| = (2b-1)^{j-1}|P(\varepsilon)|. \tag{3.2}$$

Since the primitives of ε are of the form $\alpha^i \bar{\alpha}^j$, where $0 \le i, j \le b-1$ and $i + j \ge 1$, so by $\bar{\alpha}(\alpha^i \bar{\alpha}^j) \in C^\infty$, we get that if $i \ge 1$ and $j \ge 1$ then $i = a,\ j = 1, 2, \dots, b-1$, if $i \ge 1$ and $j = 0$ then $\alpha^{i+1},\ \bar{\alpha}\alpha^i \in C^\infty$, which suggests $1 \le i \le b-1$. Thus ε have exactly $4(b-1)$ LFE primitives. Thus (3.2) gives the desired result. □

Lemma 8. *Let ξ be a positive real number and n_0 a positive integer such that*

$$|u|_b/|u| > \xi \text{ for every LFE word } u \text{ of length exceeding } n_0. \tag{3.3}$$

Then

(1) $|D(w)| \le \alpha|w|$ *for each LFE word w with $|w| > N_0$, where $\alpha = 1/(1 + (a + b - 2)\xi)$, N_0 is a suitable positive integer.*

(2) $|w| \le \beta|D(w)| + q$ *for each LFE word w, where $\beta = 1 + (a + b - 2)(1 - \xi)$, q is a suitable positive constant.*

Proof. (1) Since the complement of any smooth word is still a smooth word of the same length and $|u|_a = |\bar{u}|_b$, the hypothesis (3.3) of Lemma 8 means that

$$|u|_a/|u| > \xi \text{ for every LFE word } u \text{ with } |u| \ge n_0. \tag{3.4}$$

It is easy to see

$$|w| = |D(w)| + (a-1)|D(w)|_a + (b-1)|D(w)|_b + c, \text{ where } 0 \le c \le 2(b-1). \quad (3.5)$$

From (3.3) to (3.5), one has $|w| \ge (1+(a+b-2)\xi)|D(w)|$ for $|D(w)|_b/|D(w)| > \xi$, which implies $|D(w)| < \alpha|w|$ for every LFE word w with $|w| \ge N_0$, where N_0 is a suitable positive integer such that $|D(w)| \ge n_0$ as soon as $|w| \ge N_0$.

(2) As $|D(w)|_a/|D(w)| + |D(w)|_b/|D(w)| = 1$, from (3.3) and (3.4) ones get

$$|D(w)|_a/|D(w)| < 1-\xi \text{ for each LFE word } w \text{ with } |w| \ge N_0, \quad (3.6)$$

$$|D(w)|_b/|D(w)| < 1-\xi \text{ for each LFE word } w \text{ with } |w| \ge N_0. \quad (3.7)$$

So, from (3.5) to (3.7) it follows that $|w| \le \beta|D(w)| + 2(b-1)$ for $|w| \ge N_0$, which means that (2) also holds. □

The next lemma establishes the bounds of the heights of C^∞-words of length n, which is of independent interest.

Lemma 9. *Let $ht_{max}(n)$ and $ht_{min}(n)$ denote respectively the maximal and the minimal heights of LFE words of length n, then for any positive number ξ and positive integer n_0 satisfying $|u|_b/|u| > \xi$ for each LFE word u with $|u| > n_0$, there are two suitable constants t_1 and t_2 such that for every positive integer n, one has*

$$ht_{min}(n) > \frac{\log n}{\log(1+(a+b-2)(1-\xi))} + t_1, \quad (3.8)$$

$$ht_{max}(n) < \frac{\log n}{\log(1+(a+b-2)\xi)} + t_2, \quad (3.9)$$

where $t_1 = -\frac{\log(2(b-1)+\frac{q}{\beta-1})}{\log\beta}$, q and β are determined by Lemma 8 (2).

Proof. First, one checks (3.9). Since $|D(w)| < |w|$ for each $|w| > 0$, and $|D(w)| \le \alpha|w|$ for each LFE word w satisfying $|w| \ge N_0$ by Lemma 8 (1).

Let $k_0 - 1$ be the greatest height of all LFE words of length$< N_0$ and m_0 is the least positive integer such that if $|w| = m_0$, then the height of each LFE word w is no less than k_0. Thus for every LFE word w, if $|w| \ge m_0$, then one can get

$$|D^k(w)| < \alpha^{k-k_0}|w| \text{ for } k \ge k_0.$$

Hence

$$|D^k(w)| < 1 \quad \text{as soon as} \quad \alpha^{k-k_0}|w| \le 1$$
$$\iff \quad k \ge \log(|w|)/\log(1/\alpha) + k_0,$$

which means that the height $k-1$ of w is smaller than $\log(|w|)/\log(1+(a+b-2)\xi)+k_0$. Since there are only finite many LFE words satisfying $|w| < m_0$, so there is a suitable constant t_2 such that (3.9) holds for each LFE word.

Second, by Lemma 8 (2), one has $|w| \leq \beta|D(w)| + q$ for each LFE word w, where $\beta = 1 + (a+b-2)(1-\xi)$, q is a suitable constant, which means that

$$
\begin{aligned}
|w| &< \beta^k|D^k(w)| + q\frac{\beta^k - 1}{\beta - 1} \\
&< 2(b-1)\beta^k + \frac{q\beta^k}{\beta - 1} \\
&= (2(b-1) + \frac{q}{\beta - 1})\beta^k \\
&= m\beta^k,
\end{aligned}
$$

where $m = 2(b-1) + q/(\beta - 1)$, k is the height of w. Thus the length $|w|$ of a LFE word w of height k is less than $m\beta^k$, and it follows that

$$k > (\log|w| - \log m)/\log\beta,$$

which gives the desired lower bound of $ht_{min}(n)$, where $t_1 = -\log m/\log\beta$. □

Remark 1. (1) From (3.3) and (3.4) it immediately follows that the positive real number ξ satisfying the condition (3.3) must be smaller than 1/2.

(2) From the proof of Lemma 8 we easily see that if we substitute LFE words in Lemma 8 with some infinite subclass of smooth words, which is closed under complement, then the corresponding result also holds.

(3) From the proof of Lemma 9 we see that if we replace LFE words in Lemma 9 with some infinite subclass of smooth words, which is closed under both complement and the operator D, then the corresponding result still holds.

4. The subword complexity of smooth words

Now, we can establish our main result on subword complexity function $\gamma_{a,b}(n)$ of smooth words over 2-letter alphabets.

Theorem 10. *For any positive real number ξ and positive integer n_0 satisfying $|u|_b/|u| > \xi$ for every LFE word u with $|u| > n_0$, there exist two suitable constants c_1 and c_2 such that*

$$c_1 n^{\frac{\log(2b-1)}{\log(1+(a+b-2)(1-\xi))}} \leq \gamma_{a,b}(n) \leq c_2 n^{\frac{\log(2b-1)}{\log(1+(a+b-2)\xi)}}$$

for every positive integer n.

Proof. First, from the definition of $ht_{max}(n)$, one sees that the length of LFE words of the height larger than $ht_{max}(n)$ must be larger than n. Thus $\cup_{i=1}^{n-1} LF_i \subseteq \cup_{j=1}^{ht_{max}(n)} P^j(\varepsilon)$. So from (3.1) and Lemma 7, for any $n \in N$, one has

$$\begin{aligned}
\gamma_{a,b}(n) &= 2 + \sum_{i=1}^{n-1} |LF_i| \\
&\leq 2 + \sum_{j=1}^{ht_{max}(n)} |P^j(\varepsilon)| \\
&= 2 + \sum_{j=1}^{ht_{max}(n)} 4(b-1) \cdot (2b-1)^{j-1} \\
&= 2 \cdot (2b-1)^{ht_{max}(n)}. \qquad (4.1)
\end{aligned}$$

So combining (3.9) and (4.1) yields the desired upper bound of $\gamma_{a,b}(n)$, where $c_2 = 2(2b-1)^{t_2}$.

Second, from the definition of $ht_{min}(n)$, it follows that the length of all LFE words with the height no more than $ht_{min}(n) - 1$ must be less than n. Thus, again from (3.1) and Lemma 7, for any $n \in N$ one can get

$$\begin{aligned}
\gamma_{a,b}(n) &= 2 + \sum_{i=1}^{n-1} |LF_i| \\
&\geq 2 + \sum_{j=1}^{k} |P^j(\varepsilon)| \\
&= 2 + \sum_{j=1}^{k} 4(b-1) \cdot (2b-1)^{j-1} \\
&= 2 \cdot (2b-1)^k, \qquad (4.2)
\end{aligned}$$

where $k = ht_{min}(n) - 1$. Thus, the desired lower bound of $\gamma_{a,b}(n)$ is obtained from (3.8) and (4.2), where $c_1 = 2(2b-1)^{t_1-1}$, t_1 is decided by Lemma 9. □

Remark 2. Theorem 10 indicates that only if we could get lower and upper bounds of letters frequency of LFE words, then correspondingly we could obtain an estimate of subword complexity function $\gamma_{a,b}(n)$ of smooth words.

Taking $\Sigma = \{1, 2\}$ in Theorem 10, we obtain

Corollary 11. *For any positive number ξ and positive integer n_0 satisfying $|u|_2/|u| > \xi$ for each LDE word u with $|u| > n_0$, there exist two suitable constants c_1 and c_2 such that*

$$c_1 \cdot n^{\frac{\log 3}{\log(2-\xi)}} \leq \gamma_{1,2}(n) \leq c_2 \cdot n^{\frac{\log 3}{\log(1+\xi)}} \text{ for each } n \in N.$$

It is obvious that Corollary 11 suggests the main Theorem 1 in [15].

5. The subword complexity of smooth words on 2-letter even alphabets

Lemma 12. *If w is a 2-times differentiable finite word over 2-letter even alphabet $\{a, b\}$, then*

(1) $||w|_a - |w|_b| \leq b$;

(2) $\frac{1}{2} - \frac{b}{2|w|} \leq \frac{|w|_b}{|w|} \leq \frac{1}{2} + \frac{b}{2|w|}$;

(3) $\lim_{|w|\to\infty} \frac{|w|_a}{|w|} = \lim_{|w|\to\infty} \frac{|w|_b}{|w|} = \frac{1}{2}$;

(4) $\rho|D(w)| - q_2 \leq |w| \leq \rho|D(w)| + q_1$,

where $q_1 = (\rho - 1)b + 2(b - 1)$, $q_2 = (\rho - 1)b$, $\rho = \frac{a+b}{2}$.

Proof. It is obvious that (1) $\Rightarrow$ (2) $\Rightarrow$ (3). So we only need to check (1) and (4).

(1) Since $w \in C^2_{a,b}$, we have $D^2(w) \in \Sigma^*$. Thus $D^2(w) = \alpha^{t_1}\bar{\alpha}^{t_2}\cdots\beta^{t_k}$, where $\alpha \in \Sigma$, $t_i \in N$ for $i = 1, 2, \cdots, k$, and if $2 \mid k$ then $\beta = \bar{\alpha}$, otherwise $\beta = \alpha$. It follows that

$$\Delta^{-1}_{\gamma_1}(D^2(w)) = \overbrace{\gamma_1^{\alpha}\bar{\gamma_1}^{\alpha}\cdots\bar{\gamma_2}^{\alpha}}^{t_1}\overbrace{\gamma_2^{\bar{\alpha}}\bar{\gamma_2}^{\bar{\alpha}}\cdots\bar{\gamma_3}^{\bar{\alpha}}}^{t_2}\cdots\overbrace{\gamma_k^{\beta}\bar{\gamma_k}^{\beta}\cdots\bar{\gamma}_{k+1}^{\beta}}^{t_k}; \tag{5.1}$$

$$D(w) = \bar{\gamma}_1^{i}\Delta^{-1}_{\gamma_1}(D^2(w))\gamma_{k+1}^{j} \tag{5.2}$$

where $0 \leq i, j \leq b - 1$, $\gamma_i \in \Sigma$ and if $2 \mid t_m$ then $\gamma_{m+1} = \gamma_m$ or else $\gamma_{m+1} = \bar{\gamma}_m$.

Note that a and b are both even numbers, from (5.1) it immediately follows

$$\Delta^{-1}_{\gamma}(\Delta^{-1}_{\gamma_1}(D^2(w))) = \overbrace{\gamma^{\gamma_1}\bar{\gamma}^{\gamma_1}\cdots\bar{\gamma}^{\gamma_1}}^{\alpha}\overbrace{\gamma^{\bar{\gamma}_1}\bar{\gamma}^{\bar{\gamma}_1}\cdots\bar{\gamma}^{\bar{\gamma}_1}}^{\alpha}\cdots\overbrace{\gamma^{\bar{\gamma}_{k+1}}\bar{\gamma}^{\bar{\gamma}_{k+1}}\cdots\bar{\gamma}^{\bar{\gamma}_{k+1}}}^{\beta}, \tag{5.3}$$

where $\alpha, \beta, \gamma, \gamma_1, \cdots, \gamma_{k+1} \in \Sigma$.

Then (5.3) gives

$$|\Delta^{-1}_{\gamma}(\Delta^{-1}_{\gamma_1}(D^2(w)))|_a = |\Delta^{-1}_{\gamma}(\Delta^{-1}_{\gamma_1}(D^2(w)))|_b = \frac{1}{2}. \tag{5.4}$$

Now from (5.2) ones get

$$\begin{aligned} w &= \xi^{c_1}\Delta_{\bar{\xi}}^{-1}(D(w))\eta^{c_2} \\ &= \xi^{c_1}\Delta_{\bar{\xi}}^{-1}(\bar{\gamma}_1^i)\Delta_{\mu}^{-1}(\Delta_{\gamma_1}^{-1}(D^2(w)))\Delta_{\mu}^{-1}(\gamma_{k+1}^j)\eta^{c_2} \end{aligned} \tag{5.5}$$

where $0 \le i, j, c_1, c_2 \le b-1$, $\mu = \bar{\xi}$ if $2 \mid i$ or else $\mu = \xi$, $\eta = \bar{\xi}$ if $2 \mid (i+j)$ or else $\eta = \xi$. Note that

$$||\xi^{c_1}\Delta_{\bar{\xi}}^{-1}(\bar{\gamma}_1^i)|_a - |\xi^{c_1}\Delta_{\bar{\xi}}^{-1}(\bar{\gamma}_1^i)|_b| \le b,$$

$$||\Delta_{\mu}^{-1}(\gamma_{k+1}^j)\eta^{c_2}|_a - |\Delta_{\mu}^{-1}(\gamma_{k+1}^j)\eta^{c_2}|_b| \le b.$$

And if

$$|\xi^{c_1}\Delta_{\bar{\xi}}^{-1}(\bar{\gamma}_1^i)|_\alpha \ge |\xi^{c_1}\Delta_{\bar{\xi}}^{-1}(\bar{\gamma}_1^i)|_{\bar{\alpha}}$$

then

$$|\Delta_{\mu}^{-1}(\gamma_{k+1}^j)\eta^{c_2}|_\alpha \le |\Delta_{\mu}^{-1}(\gamma_{k+1}^j)\eta^{c_2}|_{\bar{\alpha}}.$$

Thus combining (5.4) and (5.5) produces the desired result (1).

(4) From (1) it immediately follows that

$$|w|_\alpha - b \le |w|_{\bar{\alpha}} \le |w|_\alpha + b \text{ for } \alpha \in \Sigma. \tag{5.6}$$

Since $|w| = |w|_\alpha + |w|_{\bar{\alpha}}$, from (5.6) we get

$$\frac{|w|}{2} - \frac{b}{2} \le |w|_\alpha \le \frac{|w|}{2} + \frac{b}{2} \text{ for } \alpha \in \Sigma.. \tag{5.7}$$

So, combining (3.5) and (5.7) gives the desired result (4). □

From Remark 1 and Lemma 12 (4), we can establish the following useful bounds of the heights of smooth words of length n for 2-letter even alphabets.

Lemma 13. *Let a, b be both even numbers. Then there are two constants t_1, t_2 such that for each positive integer n, ones have*

$$ht_{min}(n) > \frac{\log n}{\log \rho} + t_1, \tag{5.8}$$

$$ht_{max}(n) < \frac{\log n}{\log \rho} + t_2, \tag{5.9}$$

where

$$t_1 = -\frac{\log(3b - 2 + \frac{2(b-1)}{\rho-1})}{\log \rho},$$

$$t_2 = 2 - \frac{\log(\frac{\rho-2}{\rho-1}b)}{\log\rho}$$

$$\rho = \frac{a+b}{2},$$

$-2.3347 < -\frac{\log 13}{\log 3} \le t_1 < -1,\ 0.7944 < 2 - \frac{\log 20}{\log 12} \le t_2 \le 2 - \frac{\log 2}{\log 3} \approx 1.36907.$

Proof. First, from the proof of (3.8) and the right half part of Lemma 12 (4) it immediately follows the desired lower bound of $ht_{min}(n)$, where

$$t_1 = -\frac{\log(3b-2+\frac{4(b-1)}{a+b-2})}{\log\frac{b+a}{2}}.$$

Thus

$$t_1 < -\frac{\log b}{\log\frac{b+b}{2}} = -1,$$

and if $a = b-2$ then

$$t_1 = -\frac{\log(3b-2+\frac{2(b-1)}{b-2})}{\log(b-1)} \to -1\ (b\to\infty).$$

If $b=4$ then $a=2$, which means $t_1 = -\frac{\log 13}{\log 3}$. For $b \ge 4$, we have

$$\begin{aligned} t_1 &\ge -\frac{\ln(3b-2+\frac{4(b-1)}{b})}{\ln\frac{b+2}{2}} \\ &= -\frac{\ln\frac{3b^2+2b-4}{b}}{\ln\frac{b+2}{2}}. \end{aligned} \tag{5.10}$$

Let

$$g(b) = \ln 3\ln\frac{3b^2+2b-4}{b} - \ln 13\ln\frac{b+2}{2},$$

then

$$g'(b) = (\ln 3)\frac{3b^2+4}{3b^3+2b^2-4b} - \frac{\ln 13}{b+2}.$$

By Maple, we easily see that the roots of the equation $3(\ln 3-\ln 13)b^3 + (6\ln 3 - 2\ln 13)b^2 + 4(\ln 3+\ln 13)b + 8\ln 3 = 0$ are approximately equal to -1.003, -0.894, 2.229. Hence, since $g'(4) < 0$ and $g'(b)$ is continuous in $[4, +\infty)$, we obtain $g'(b) < 0$ for all $b \ge 4$. Therefore $g(b) \le g(4) = 0$, which suggests

$$\frac{\ln\frac{3b^2+2b-4}{b}}{\ln\frac{b+2}{2}} \le \frac{\log 13}{\log 3}.$$

Then (5.10) gives $t_1 \geq -\frac{\log 13}{\log 3}$.

Second, we use an argument similar to the proof of (3.8) to obtain the upper bound of $ht_{max}(n)$. Note that if $ht(w) \geq 2$, then $|w| > 2b$. Then from the left half part of Lemma 12 (4), we get

$$|D(w)| < \frac{1}{\rho}|w| + b. \tag{5.11}$$

Now assume w is a smooth word of length n with height k larger than or equal to 2. Since $ht(w) \geq 2$, from (5.11), we arrive at

$$\begin{aligned}
2b &< |D^{k-2}(w)| \\
&< \frac{1}{\rho}|D^{k-3}(w)| + b \\
&< \frac{1}{\rho^2}|D^{k-4}(w)| + \frac{1}{\rho}b + b \\
&\cdots \\
&< \frac{1}{\rho^{k-2}}|w| + \frac{1}{\rho^{k-3}}b + \cdots + \frac{1}{\rho^2}b + \frac{1}{\rho}b + b \\
&< \frac{1}{\rho^{k-2}}|w| + \frac{1}{1-\rho^{-1}}b.
\end{aligned}$$

Thus

$$\rho^{k-2} < \frac{|w|}{\tau}, \ \ \textit{where } \tau = \frac{\rho-2}{\rho-1}b,$$

which means

$$k < \frac{\log n}{\log \rho} + 2 - \frac{\log \tau}{\log \rho}. \tag{5.12}$$

Note that the length n of a smooth word of height 1 is greater than or equal to $a+2 \geq 4$, so

$$\frac{\log n}{\log \rho} + 2 - \frac{\log \tau}{\log \rho} \geq 2 + \frac{\log \frac{4(\rho-1)}{b(\rho-2)}}{\log \rho} > 1,$$

which means (5.12) holds for every smooth word. Now from (5.12) it immediately follows the desired upper bound (5.9) of $ht_{max}(n)$.

From

$$(b-3)^2 - (a-1)^2 - 8 \geq 0 \text{ for } b \geq a+4,$$

we get

$$(b-a)(a+b-4) \geq 2(a+b),$$

$$2b-(a+b)\geq \frac{2(a+b)}{a+b-4},$$

$$b\geq \rho+\frac{\rho}{\rho-2},$$

$$b\frac{\rho-2}{\rho-1}\geq \rho,$$

which means

$$\frac{\log(\frac{\rho-2}{\rho-1}b)}{\log\rho}>1>\frac{\log 2}{\log 3} \text{ for } b\geq a+4.$$

Thus if $b\geq a+4$ then $t_2\leq 2-\frac{\log 2}{\log 3}\approx 1.36907$.

If $b=a+2$ then $\rho=a+1$, so

$$t_2=2-\frac{\ln\frac{(a-1)(a+2)}{a}}{\ln(a+1)}.$$

Let

$$f(a)=\ln(3)\ln\frac{(a-1)(a+2)}{a}-\ln(2)\ln(a+1)$$

then

$$\begin{aligned}
f'(a) &= \ln(3)\frac{a^2+2}{(a-1)(a+2)a}-\frac{\ln 2}{a+1}\\
&> \ln(3)(\frac{a^2+2}{(a-1)(a+2)a}-\frac{1}{a+1})\\
&= \ln(3)\frac{4a+2}{(a^2-1)(a+2)a}\\
&> 0 \text{ for every } a>1.
\end{aligned}$$

Hence, $f(a)\geq f(2)=0$ for each $a\geq 2$, that is,

$$\frac{\ln\frac{(a-1)(a+2)}{a}}{\ln(a+1)}\geq\frac{\ln 2}{\ln 3} \text{ for each } a\geq 2,$$

which also gives the desired result $t_2\leq 2-\frac{\log 2}{\log 3}$.

Finally, machine computation shows

$$t_2\geq 2-\frac{\log 20}{\log 12} \text{ for } b\leq 58. \tag{5.13}$$

Moreover, in view of $a<b$, we obtain

$$t_2\geq 2-\frac{\log b}{\log\frac{b}{2}}. \tag{5.14}$$

And let

$$h(b) = 2 - \frac{\ln b}{\ln \frac{b}{2}}, \tag{5.15}$$

then

$$h'(b) = \frac{\frac{1}{b}(\ln b - \ln \frac{b}{2})}{(\ln \frac{b}{2})^2} > 0 \text{ for } b \geq 4.$$

which means

$$h(b) \geq h(60) \approx 0.7962 > 2 - \frac{\log 20}{\log 12} \text{ for } b \geq 60.$$

Thus (5.13), (5.14) and (5.15) give the desired lower bound of the constant t_2. $\square$

Theorem 14. *Let a, b be both even numbers. Then there exist two suitable constants c_1, c_2 such that*

$$c_1 n^{\frac{\log(2b-1)}{\log(a+b)-\log 2}} \leq \gamma_{a,b}(n) \leq c_2 n^{\frac{\log(2b-1)}{\log(a+b)-\log 2}},$$

where $c_1 = 2(2b-1)^{t_1-1}$, $c_2 = 2(2b-1)^{t_2}$, t_1, t_2 are determined by Lemma 13.

Proof. From the proof of Theorem 10 we easily see that (4.1) and (4.2) always hold. Thus combining (5.8) and (4.2) gives

$$\gamma_{a,b}(n) \geq \ c_1 n^{\frac{\log(2b-1)}{\log(a+b)-\log 2}}.$$

Similarly, from (5.9) and (4.1) it follows

$$\gamma_{a,b}(n) \leq c_2 n^{\frac{\log(2b-1)}{\log(a+b)-\log 2}}. \ \square$$

6. Concluding remarks

To establish the estimates of subword complexity function of smooth words to follow our thoughts and methods is an interesting problem for large alphabets Σ_n containing n letters, where $n \geq 3$.

For the 3-letter alphabet $\Sigma_3 = \{2, 4, 6\}$, let

$$\begin{aligned} w_1 &= 64^2 2^6 6^6 4^6 6^6 2^6 4^6, \\ w_2 &= 42^6 6^6 4^6 6^6 2^6 4^6, \\ w_3 &= 42^6 6^6 4^6 6^6 2^6, \end{aligned}$$

$$
\begin{aligned}
v_1 &= 4^6 2^2 6^2,\\
v_2 &= 2^6 6^6 2^6 6^6 2^6 6^6 4^4 6^2,\\
v_3 &= 2^6 4^6 2^2 6^2,\\
u_1 &= 2^2 6^2 4^6,\\
u_2 &= 4^4 2^2 6^2 2^2 6^2 2^2 6^2 4^6,\\
u_3 &= 2^2 4^2 6^2,
\end{aligned}
$$

then $D(w_1) = 26^6$, $D(w_2) = 6^6$, $D(w_3) = 6^5$, $D(v_1) = 62$, $D(v_2) = 6^6 4$, $D(v_3) = 6^2 2$, $D(u_1) = 26$, $D(u_2) = 2^6 6$, $D(u_3) = 2$, we easily see that each of w_1, w_2 and w_3 has only one left smooth extension and $D(w_i)$ has exactly i left smooth extensions for $i = 1, 2, 3$; each of v_1, v_2 and v_3 has exactly two left smooth extensions and $D(v_i)$ has exactly i left smooth extensions for $i = 1, 2, 3$; each of u_1, u_2 and u_3 has exactly three left smooth extensions and $D(u_i)$ has exactly i left smooth extensions for $i = 1, 2, 3$. Thus for large alphabets containing at least three letters, the estimates of factor complexity function of smooth words become more complicated than the case for 2-letter alphabets.

Acknowledgements

I would like to thank Professor William D. Weakley very much for providing me the opportunity to work with him in [15].

References

[1] **M. Baake, B. Sing,** *Kolakoski-(3,1) is a (deformed) model set,* Canad. Math. Bull., 47 (2) (2004), 168-190.

[2] **V. Berthé, S. Brlek, P. Choquette,** *Smooth words over arbitrary alphabets,* Theoret. Comput. Sci., 341 (2005), 293-310.

[3] **S. Brlek, S. Dulucq, A. Ladouceur, L. Vuillon,** *Combinatorial properties of smooth infinite words,* Theoret. Comput. Sci., 352 (2006), 306-317.

[4] **S. Brlek, D. Jamet, G. Paquin,** *Smooth words on 2-letter alphabets having same parity,* Theoret. Comput. Sci., 393 (2008), 166-181.

[5] **S. Brlek, A. Ladouceur,** *A note on differentiable Palindromes,* Theoret. Comput. Sci., 302 (2003), 167-178.

[6] **S. Brlek, G. Melançon , G. Paquin,** *Properties of the extremal infinite smooth words,* Discrete Math. Theoret. Comput. Sci., 9 (2) (2007), 33-50.

[7] **A. Carpi,** *Repetitions in the Kolakovski sequence,* Bull. of the EATCS, 50 (1993), 194-196.

[8] **A. Carpi,** *On repeated factors in C^∞-words,* Inf. Process. Lett., 52 (1994) 289-294.

[9] **V. Chvátal,** *Notes On the Kolakoski sequence,* DIMACS Tech. Rep., 93-84 (1994).

[10] **F. M. Dekking,** *Regularity and irragularity of sequences generated by automata,* Séminaire de Théorie des Nombres de Bordeaux, 1979-80, exposé n° 9, 901-910.

[11] **F. M. Dekking,** *On the structure of selfgenerating sequences,* Séminaire de Théorie des Nombres de Bordeaux, 1980-81, exposé n° 31, 3101-3106.

[12] **F. M. Dekking,** *What is the long range order in the kolakoski sequence?* in R.V.Moody (ed.), The mathematics of Long-Range Aperiodic order, Kluwer Academic Publishers (1997) 115-125.

[13] **Y. B. Huang,** *About the number of C^∞-words of form $\tilde{w}xw$,* Theoret. Comput. Sci., 393 (2008), 280-286.

[14] **Y. B. Huang,** *The complexity of $C^{b\omega}$-words of the form $\tilde{w}xw$,* Theoret. Comput. Sci., **410** (2009), 4892-4904.

[15] **Y. B. Huang, W. D. Weakley,** *A note on the complexity of C^∞-words,* Theoret. Comput. Sci., **411** (2010), 3731-3735.

[16] **Y. B. Huang,** *The powers of smooth words over arbitrary 2-letter alphabets,* arXiv:0904.0562v1 (2009).

[17] **M. S. Keane,** *Ergodic theory and subshifts of finite type,* in: Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces, T. Bedford, M. Keane, C. Series (Eds.), Oxford University Press (Oxford 1991), 350-370.

[18] **C. Kimberling,** *Problem 6287,* Amer. Math. Monthly, 86 (1979), 35-70.

[19] **W. Kolakoski,** *Self-generating runs, Problem 5304,* Amer. Math. Monthly, 72 (1965), 674. Solution: Amer. Math. Monthly, 73 (1966), 681-682.

[20] **A. Lepistö,** *Repetitions in the Kolakoski sequence,* Development in Language Theory (1993), 130-143.

[21] **G. Păun,** *How much Thue is Kolakovski?* Bull. of the EATCS, 49 (1993), 183-185.

[22] **B. Sing,** *Kolakoski sequences - an example of aperiodic order,* J. Non-Cryst. Solids, 334-335 (2004), 100-104.

[23] **B. Sing,** *Kolakoski-(2m,2n) are limit-periodic model sets,* J. Math. Phys. 44 (2) (2003), 899-912.

[24] **R. Steacy,** *Structure in the Kolakoski sequence,* Bull. of the EATCS, 59 (1996), 173-182.

[25] **B. Steinsky,** *A Recursive Formula for the Kolakoski Sequence,* Journal of Integer Sequences, 9(3) (2006), 06.3.7.

[26] **W. D. Weakley,** *On the number of C^∞-words of each length,* Jour. of comb. Theory, Ser.A, 51 (1989), 55-62.