CALCULATION OF THE COMPLEXITIES OF SUBSTITUTIVE SEQUENCES OVER A BINARY ALPHABET

BO TAN, ZHI-XIONG WEN, AND YIPING ZHANG

ABSTRACT. We consider the complexities of substitutive sequences over a binary alphabet. By studying various types of special words, we show that, knowing some initial values, its complexity can be completely formulated via a recurrence formula determined by the characteristic polynomial.

Keywords: Substitution, Special Word, Complexity
2000 Mathematics Subject Classification: Primary 20M05; Secondary 68R15.

1. Introduction

The study of substitutions over a finite alphabet plays important roles in many fields such as finite automata, symbolic dynamics, formal languages, number theory, fractal geometry etc. It has various applications to quasi-crystals, computational complexity, information theory... (see [1, 2, 7, 9, 10] and the references therein). In addition, substitutions are also fundamental objects in combinatorial group theory [11, 12].

Given an infinite sequence $\xi = \xi_1\xi_2\xi_3\cdots$ ($\xi_i \in A$) over some finite alphabet $A$, we denote by $L_n(\xi)$ the set $\{\xi_i\cdots\xi_{i+n-1} \mid i \geq 1\}$ of factors of $\xi$ of length $n$ ($n \geq 1$), and by convention $L_0(\xi)$ is the singleton consisting of the empty word $\varepsilon$. The set $L(\xi) = \cup_{n\geq 0}L_n(\xi)$ is then called the language of $\xi$, and the function $p_\xi(n) := \#L_n(\xi)$ the complexity of $\xi$, here and hereafter $\#$ denotes the cardinality of a finite set.

Let $A^*$ be the free monoid generated by $A$ (with $\varepsilon$ as the neutral element). A morphism $\sigma : A^* \rightarrow A^*$ is called a substitution. We deal with only the non-erasing substitutions (the image of any letter in $A$ is not the empty word), whence the substitution can be extended naturally to $A^N$, the set of infinite sequences over $A$. Denote by $\xi_\sigma$ any one of the fixed points of $\sigma$ (that is $\sigma(\xi_\sigma) = \xi_\sigma$), if it exists.

The study of the complexity of $\xi_\sigma$ (also called the complexity of $\sigma$) has a long history. In general, it is very difficult to find out the explicit formula for $p_\xi(n)$ for a given $\sigma$; only some calculations for specific classes of substitutions can be found in the literature. Here are some known results :

- $p_\xi(n) \leq n$ for some $n$ if and only if $\xi$ is ultimately periodic, and in this case the complexity is bounded [13];
• A sequence $\xi$ of complexity $p_\xi(n) = n + 1$ is called Sturmian. There are many equivalent characterizations and interesting properties of Sturmian sequences (see, e.g. [9, 18, 22]);
• Rote [17] constructed a class of sequences with complexity $2^n$ by using graphs;
• Mossé [14] studied the case of $q$-automata (which correspond to substitutions of constant length). A method to compute $p(n)$ with linear recurrence formula was given under some technical conditions;
• Over a ternary alphabet, a class of Tribonacci type substitutions with complexity $2^n + 1$ was introduced by Arnoux and Rauzy [3]. An example of substitution (Triplex Substitution) with complexity $3^n$ is presented by the authors [21].
• For a fixed point of some substitution, the complexity can only be of the following five different asymptotic forms: $\Theta(1)$, $\Theta(n)$, $\Theta(n \log \log n)$, $\Theta(n \log n)$ or $\Theta(n^2)$, where $\Theta(g(n))$ means a function $f(n)$ satisfying $0 < \lim \inf \frac{f(n)}{g(n)} \leq \lim \sup \frac{f(n)}{g(n)} < \infty$ [15].
• For a survey and more general computation of factor complexity of word (on an alphabet of cardinality more than 2), we suggest to see [6, 8].

In this paper, we consider general substitutions $\sigma$ over a binary alphabet. Using Mossé’s theory of identifiability ([14]) and by studying various types of special words ([5, 6]), we show that the complexity $p(n)$ can be completely formulated knowing some initial values, and a recurrence formula is given.

2. Notations and Preliminary

We fix the binary alphabet $\mathbb{A} = \{a, b\}$ consisting of two letters $a$ and $b$. Let $\mathbb{A}^*$ be the free monoid generated by $\mathbb{A}$ (with the empty word $\varepsilon$ as the neutral element), and $\mathbb{A}^\mathbb{N}$ be the set of all infinite sequences (also called infinite words) over $\mathbb{A}$.

If $w \in \mathbb{A}^*$, we denote by $|w|$ its length and by $|w|_a$ (resp. $|w|_b$) the number of occurrences of the letter $a$ (resp. $b$) in $w$. The abelian Parikh vector of $w$ is then defined to be the column vector $L(w) = (|w|_a, |w|_b)^t \in \mathbb{N}^2$.

A word $v$ is a factor of a word $w$ (written as $v \in w$) if there exist $u, u' \in \mathbb{A}^*$, such that $w = uvu'$. It is sometimes convenient to use the notation “$\oslash$” to stand for some word which we don’t care so much. Thus $v$ is a factor of a word $w$ if and only if $w = \oslash v \oslash$ (remark that even within a formula, $\oslash$’s may represent different words). We say that $v$ is a prefix (resp. a suffix) of $w$ if $w = v \oslash$ (resp. $w = \oslash v$), and then we write $v \triangleleft w$ (resp. $v \triangleright w$). Two words $v$ and $w$ are said to be comparable, written $v \asymp w$, if either $v \triangleright w$ or $w \triangleleft v$. The notions of factor and prefix extend to infinite words in a natural way.

It is also convenient to put, e.g. $\mathbb{A}^* v := \{xv; x \in \mathbb{A}^*\}$, $\mathbb{A}^* v \mathbb{A}^* := \{xvy; x, y \in \mathbb{A}^*\}$, etc. Thus $w \in \mathbb{A}^* v \mathbb{A}^* \iff v \in w$; $w \in v \mathbb{A}^* \iff v \triangleleft w$, and so on.
When \( \xi = w_1 \cdots w_m \cdots \in A^* \cup A^N \) (\( w_i \in A \)), we also write \( \xi|_1 = w_1, \cdots, \xi|_m = w_m, \cdots \), and \( \xi[i, j] = w_i w_{i+1} \cdots w_j (i \leq j) \).

As already defined, a substitution \( \sigma \) over \( A \) is a morphism \( \sigma \) of \( A^* \). The matrix \( M = (L(\sigma(a)), L(\sigma(b))) \) is called the incidence matrix of \( \sigma \). The characteristic polynomial \( \lambda^2 - \text{tr}(M)\lambda + \det(M) \) of \( M \) is also called the characteristic polynomial of \( \sigma \).

If \( \sigma(a) \) and \( \sigma(b) \) have distinct first letters, we say that the substitution \( \sigma \) is marked, and if moreover \( \sigma(a) = a \odot \) and \( \sigma(b) = b \odot \), we say that \( \sigma \) is well-marked. It is easy to see that \( \sigma^2 \) is well-marked if \( \sigma \) is marked.

In this paper, all substitutions are assumed to be non-erasing, that is, the image of each letter is not empty. Whence, the substitution can be extended naturally to \( A^N \). An infinite word \( \xi = \xi_1 \xi_2 \cdots \) is a fixed point of \( \sigma \) if \( \sigma(\xi) = \xi \).

Hereafter, we suppose that the substitution \( \sigma \) is primitive (i.e. its incidence matrix \( M \) is primitive: \( M^n \) possesses positive coordinates for some positive integer \( n \)). The following easy facts for a primitive substitution \( \sigma \) are well known:

1. the fixed point of \( \sigma \) is recurrent, that is, every factor will occur for infinitely many times; and all the fixed points of \( \sigma \) have the same language;
2. a substitution \( \sigma \) and its powers \( \sigma^n \) (\( n \geq 1 \)) have the same fixed points, and thus have the same language;
3. if one substitution is a composition of an inner automorphism (of the free group) with another substitution, then the two substitutions have the same language.

We suppose also that the fixed point \( \xi \) of \( \sigma \) is not (ultimately) periodic; the periodic case are characterized completely by Séébold [19]. In particular, whence \( \{\sigma(a), \sigma(b)\} \) is a code, and thus \( \sigma \) is marked up to an inner automorphism (see [9]). For the sake of calculation of the complexity of a non-periodic primitive substitution, we may further suppose, without loss of generality, that the substitution is well-marked.

The notion of “special words” is a powerful tool for calculating the complexity. See [5, 6] and [4, 9, 10] for more information.

Let \( W \) be a factor of \( \xi \). If \( \delta \in A \) such that \( W\delta \) is a factor of \( \xi \), then we say that \( W\delta \) is a right extension of \( W \). A word is called a right special word (special word for short) of \( \xi \) if it has more than one extensions, that is, \( Wa \in \xi \) and \( Wb \in \xi \). Similarly we define “left extension” and “left special word”. It is easy to see that a suffix (resp. prefix) of a special (resp. left special) word is also special (resp. left special).

Let \( S_n \) (resp. \( LS_n \)) be the set of special words (resp. left special words) of length \( n \) of \( \xi \). Put \( S = \bigcup_{n \geq 0} S_n \) (resp. \( LS = \bigcup LS_n \)). It is easy to see that

\[
s(n) := \#S_n = \#LS_n = \Delta p(n + 1)(:= p(n + 1) - p(n)).
\]

Hence the study of \( p(n) \) is almost equivalent to the study of \( s(n) \).
2.1. The word $W_0$ and the letters $\delta_a, \delta_b$.

Write $A = \sigma(a), B = \sigma(b)$, and denote $\{A, B\}^*$ the set of words obtained by a finite concatenation of the words $A$ and $B$. Put, as before, e.g., $\{A, B\}^* A := \{VA; V \in \{A, B\}^*\}$. Remark that since $\sigma$ is non-periodic, $\{A, B\}$ is a code and $\{A, B\}^*$ is a disjoint union of $\{A, B\}^* A$ and $\{A, B\}^* B$.

Since $\sigma$ is non-periodic, the left-infinite words $A^\infty(= \cdots AA \cdots A)$ and $B^\infty$ are different. Let $W_0$ be the longest common suffix of $A^\infty$ and $B^\infty$ (see also [20]). Remark that $W_0$ is possibly empty.

The following lemma is a direct consequence of Fine-Wilf theorem [16].

**Lemma 2.1.** $|W_0| \leq |A| + |B| - 2$.

By the definition of $W_0$, for some $\delta_a, \delta_b \in \{a, b\}$ with $\{\delta_a, \delta_b\} = \{a, b\}$,

$$A^\infty = \oplus \delta_a W_0 \text{ and } B^\infty = \oplus \delta_b W_0. \tag{2.1}$$

Formula (2.1) shows that there exist $m \geq 0$ and $A' A (|A'| < |A|)$ such that

$$W_0 = A' A^m, \quad \text{and } \delta_a A' \triangleright A,$$

and similarly

$$W_0 = B' B^k, \quad \text{and } \delta_b B' \triangleright B.$$

The following lemma is essentially due to [20].

**Lemma 2.2.** (1) For $W \in \{A, B\}^*$, we have $W_0 \triangleright W$. Furthermore,

(2) If $W \in \{A, B\}^* A$ (resp. $\{A, B\}^* B$) and $|W| > |W_0|$, then $\delta_a W_0 \triangleright W$ (resp. $\delta_b W_0 \triangleright W$), where $\delta_a$ and $\delta_b$ are defined in (2.1).

(3) Let $W \in \{A, B\}^*$. If $\delta_a W_0 \triangleright W$ (resp. $\delta_b W_0 \triangleright W$), then $W \in \{A, B\}^* A$ (resp. $\{A, B\}^* B$).

In brief, any word in $\{A, B\}^*$ is comparable with $W_0$. Amongst them, the word in $\{A, B\}^* A$ is comparable with $\delta_a W_0$ and $\{A, B\}^* B$ is comparable with $\delta_b W_0$.

**Proof.** If $W = A$ or $W = B$, the lemma is obvious. Suppose $W \in \{A, B\}^*$ such that $W_0 \triangleright W$, we claim that $\delta_a W_0 \triangleright WA$ and $\delta_b W_0 \triangleright WB$. The two statements can be proven in the same way, and we only show the first one by considering the following two cases:

Case 1: $W_0 \triangleright W$. Then $W_0 A \triangleright WA$, and on the other hand, $\delta_a W_0 \triangleright W_0 A$ because both of them are suffixes of $A^\infty$. Hence $\delta_a W_0 \triangleright WA$.

Case 2: $W \triangleright W_0$. Then $WA \triangleright W_0 A$, while $W_0 A$ is a suffix of $A^\infty$, and thus $WA$ is a suffix of $A^\infty$. This yields that $WA \triangleright \delta_a W_0$ because both of them are suffixes of $A^\infty$. \hfill $\Box$

**Corollary 2.1.** Let $W \in \{A, B\}^*$. Then $W_0 \triangleright W_0 W$, $\delta_a W_0 \triangleright W_0 WA$, $\delta_b W_0 \triangleright W_0 WB$. In particular, $\delta_a W_0 \triangleright W_0 A$, $\delta_b W_0 \triangleright W_0 B$. 

2.2. Natural decomposition and identifiability.

Let \( \xi \) be a fixed sequence of \( \sigma \). Write \( \xi = \xi_1 \xi_2 \cdots \). Since \( \sigma(\xi) = \xi \), we have the following so called “natural decomposition” of \( \xi \)

\[
(2.3) \quad \xi = [\xi_1 \xi_2 \cdots \xi_{n-1}] [\xi_n \cdots \xi_{n+1}] \cdots \]

where \( \xi_k \in A = \{a, b\} \), \( \sigma(\xi_k) = \xi_{n_k} \cdots \xi_{n_{k+1} - 1} \in \{A, B\} \) \((k \geq 1)\), and \( n_1(:= 1), \ldots, n_k(:= |\sigma(\xi[1,k-1])| + 1), \ldots \) are called the “cutting positions” of \( \xi \). We denote

\[
(2.4) \quad E_1 = \{n_k; k \geq 1\}.
\]

Now consider the factors of \( \xi \). Let \( W = \xi_i \xi_{i+1} \cdots \xi_j \in \xi \), then (comparing to (2.3) for some integers \( k, l \) \((n_k - 1 < i \leq n_k \leq n_l \leq n_{l+1} - 1 \leq j < n_{l+2})\), we have

\[
W = \xi_i \cdots \xi_{n_k-1} [\xi_{n_k} \cdots \xi_{n_{k+1}-1}] \cdots [\xi_{n_l} \cdots \xi_{n_{l+1}-1}] [\xi_{n_{l+1}} \cdots \xi_j],
\]

that is, observing the cutting positions of \( W \) in \( \xi \) we can write out the following natural decomposition of \( W \)

\[
(2.5) \quad W = U \sigma(\xi_k) \cdots \sigma(\xi_l) V = U \sigma(W') V;
\]

where

\[
U = \xi_i \cdots \xi_{n_k-1} \triangleright \sigma(\xi_{k-1}), \quad |U| < |\sigma(\xi_{k-1})|,
\]

\[
\sigma(\xi_k) = \xi_{n_k} \cdots \xi_{n_{k+1}-1},
\]

\[
\vdots
\]

\[
\sigma(\xi_l) = \xi_{n_l} \cdots \xi_{n_{l+1}-1},
\]

\[
V = \xi_{n_{l+1}} \cdots \xi_j \triangleleft \sigma(\xi_{l+1}), \quad |V| < |\sigma(\xi_{l+1})|,
\]

\[
W' = \xi_k \cdots \xi_l \in \xi.
\]

We say that \( W' \) (resp. \( \xi_m, k \leq m \leq l \)) is the ancestor of \( \sigma(W') \) (resp. \( \sigma(\xi_m) \)). Sometimes, we also call \( \xi_{k-1} \xi_k \cdots \xi_l \xi_{l+1} \) the ancestor of \( W \).

We extend a little more the significance of “natural decomposition”: if \( W = U \sigma(W_1) W'' W_2 \) as in (2.5), we shall also say that \( W = U' \sigma(W'') V' \) is a “natural decomposition” (where \( U' = U \sigma(W_1), V' = \sigma(W_2) V \), and we write \( W = U'' \sigma(W'') \sigma(W''') \) \( V' \). Equivalently, the notation \( U'' \sigma(W'') \sigma(W''') \) \( V' \) means that there exist \( U'', V'' \in A^* \) such that

\[
(2.6) \quad U'' W'' V'' \in \xi, \quad U' \triangleright \sigma(U''), \quad \text{and} \quad V' \triangleleft \sigma(V'').
\]

Intuitively, \( U'' \sigma(W'') \sigma(W''') \) \( V' \) appears in \( \xi \) with “[ ]” and “[ ]” showing the interested natural cutting positions.

We call the decomposition as in (2.5) a strict natural decomposition of \( W \).

Remark that any natural decomposition can be extended to a strict one, and, in general, the natural decompositions of a factor are not unique; and that the fact \( U \sigma(W)V \in \xi \) does not always mean \( U[\sigma(W)]V \)!
From the theory of identifiability we have (recall that $\xi[i, j] = \xi_i \cdots \xi_j$):

**Lemma 2.3.** [14] There exists an integer $C$ (depending on $\sigma$) such that, if $W \in \xi$ can be written as $W = \xi[i - C, i + C] = \xi[j - C, j + C]$ with $i \in E_1$, then we have $j \in E_1$.

We shall say that $\xi[i - C, i + C]$ and $\xi[j - C, j + C]$ have a relative common cutting position (at the positions $i$ and $j$ respectively). As a consequence, if $W$ is long enough, say $|W| \geq L$ with

\[
L = \max\{2C + \max\{|A|, |B|\}, |A| + |B| - 1\} (> |W_0|)
\]

and it appears at different positions in $\xi$: $W = \xi[i_1, i_2] = \xi[j_1, j_2]$, then roughly speaking, at the middle position of $\xi[i_1, i_2]$ and $\xi[j_1, j_2]$, they have a relative common cutting position: for some integer $N \in ([W]/2 - \max\{|A|, |B|\}, [W]/2 + \max\{|A|, |B|\})$, $i_1 + N \in E_1$ and $j_1 + N \in E_1$.

3. The Operator $T$ and Structure of $\mathcal{LS}$

Define $T : A^* \to A^*$:

\[
T(W) = W_0\sigma(W).
\]

Notice that $T$ is not a morphism on $A^*$. It is readily checked that $T$ is injective and

\[
T^n(W) = W_0\sigma(W_0) \cdots \sigma^{n-1}(W_0)\sigma^n(W).
\]

**Lemma 3.1.** If $W \in \xi$, then $T(W) \in \xi$. Moreover, $T(W) = W_0[\sigma(W)]$.

**Proof.** Due to the primitivity of $\sigma$, the fixed sequence $\xi$ is recurrent. Thus for any $n \in \mathbb{N}$, $UW \in \xi$ for some $U \in A^*$ with $|U| = n$. Now by the $\sigma$-invariance of $\xi$, we have that $\sigma(U)\sigma(W) \in \xi$. When the length $n$ of $U$ is large, $W_0 \triangleright \sigma(U)$ by Lemma 2.2, therefore $T(W) = W_0[\sigma(W)] \in \xi$. \qed

**Lemma 3.2.** Let $W_1, W_2 \in A^*$. Then $T(W_1) = T(W_2)$ if and only if $W_1 = W_2$; $T(W_1) \prec T(W_2)$ if and only if $W_1 \prec W_2$; $T(W_1) \triangleright T(W_2)$ if and only if $W_1 \triangleright W_2$.

**Proof.** The first two easy statements hold since $\sigma$ is well marked, and the last one follows from Corollary 2.3. \qed

The following lemma tells us that if a factor $W$ appears at two positions with different natural decompositions, then, up to a prefix $W_0' \triangleright W_0$, they have the same relative cutting positions.

**Lemma 3.3.** Suppose that $W \in \xi$, $|W| \geq L$ with $L$ defined in (2.7), and that $W$ appears at two different positions in $\xi$, with $W = P_1[\sigma(U_1)]Q_1$ and $W = P_2[\sigma(U_2)]Q_2$ the corresponding strict natural decompositions. Then, denoting by $U$ the longest common suffix of $U_1$ and $U_2$ and thus writing $U_1 = U'_1U$, $U_2 = U'_2U$ (where $U'_1$ or $U'_2$ is possibly empty), we have that $U$ is nonempty and

\[
P_1\sigma(U_1)Q_1 = W_0'[\sigma(U)]Q = P_2\sigma(U_2)Q_2,
\]
where \( Q = Q_1 = Q_2 \), \( W'_0 = P_1 \sigma (U'_1) = P_2 \sigma (U'_2) \gg W_0 \). More precisely, either \( W'_0 \gg W_0 \), or \( U'_1 = U'_2 = \varepsilon \) and \( W'_0 \gg \sigma (\delta) \) for some \( \delta \in \mathbb{A} \).

**Proof.** By Lemma 2.3, the two strict natural decompositions share a relative cutting position, and thus all the cutting positions after this one. This implies that \( U_1 \) and \( U_2 \) have nonempty common suffix, i.e., \( U \) is not empty. Also this implies that \( Q_1 = Q_2 \), and consequently that \( P_1 \sigma (U'_1) = P_2 \sigma (U'_2) \gg W_0 \), where the last formula is due to Lemma 2.2. \( \square \)

**Lemma 3.4.** (1) If \( W \in \mathcal{LS} \) with \( |W| \geq L \). Then there exist unique \( U \in \mathbb{A}^* \), \( \delta \in \mathbb{A} \) and \( Q \ll \sigma (\delta) \) with \( U \delta \in \xi \) and \(|Q| < |\sigma (\delta)|\), such that

\[
aW = aW_0[\sigma (U)]Q \quad \text{and} \quad bW = bW_0[\sigma (U)]Q.
\]

(2) If \( W \in \mathcal{S} \) with \( |W| \geq L \). Then there exist \( U \in \mathbb{A}^* \), \( W'_0 \in \mathbb{A}^* \) with either \( W'_0 \gg W_0 \), or \( W'_0 \gg \sigma (\delta) \) and \(|W'_0| < |\sigma (\delta)|\) for some \( \delta \in \mathbb{A} \), such that

\[
Wa = W'_0[\sigma (U)]a \quad \text{and} \quad Wb = W'_0[\sigma (U)]b.
\]

(3) If \( W \in \mathcal{LS} \cap \mathcal{S} \) with \( |W| \geq L \). Then there exists a unique \( U \in \mathbb{A}^* \) such that \( W = T(U) \).

**Remark:** The word \( w \) in \( \mathcal{LS} \cap \mathcal{S} \) is called a bispecial word, which is developed in [5], see also [4].

**Proof.** (1) Consider the strict natural decompositions of \( aW \) and \( bW \):

\[
aW = aP_a[\sigma (U_a)]Q_a \quad \text{and} \quad bW = bP_b[\sigma (U_b)]Q_b,
\]

with \( U \) the longest common suffix of \( U_a \) and \( U_b \), \( U_a = U'_a U \), \( U_b = U'_b U \). Then, as in the previous proof, \( U \) is nonempty, \( Q_a = Q_b \), \( P_a \sigma (U'_a) = P_b \sigma (U'_b) \). Moreover, putting \( W'_0 = P_a \sigma (U'_a) \), we have that \( aW'_0 \gg \sigma (W_a) \) and \( bW'_0 \gg \sigma (W_b) \) with \( W_a, W_b \in \mathbb{A}^* \) and the last letters of \( W_a \) and \( W_b \) are distinct. Together with Lemma 2.2, these facts imply that \( W'_0 = W_0 \).

(2) The proof for this part is similar to the first part.

(3) This is a corollary of the first two parts. \( \square \)

**Lemma 3.5.** (1) \( W_0 \in \mathcal{LS} \);

(2) Any prefix of a left special word is left special;

(3) If \( W \in \mathcal{LS} \), then \( T(W) \in \mathcal{LS} \);

(4) Let \( W \in \mathcal{LS} \) with \( |W| \geq L \), then there exist unique \( U \in \xi \), \( \delta \in \{a, b\} \) such that \( W = W_0[\sigma (U)]Q = T(U)Q \ll T(W') \) (see Lemma 3.4), where \( W' = U \delta \).

Further more, \( U, W' \in \mathcal{LS} \).

**Proof.** (1) and (2) are obvious.

(3) If \( aW \in \xi \), then \( T(aW) \in \xi \) by Lemma 3.1. By Lemma 2.2, \( \delta_a T(W) = \delta_a W_0 \sigma (W) \) is a suffix of \( T(aW) = W_0 \sigma (W) \), and thus \( \delta_a T(W') \in \xi \). From this, we see that \( W \in \mathcal{LS} \) implies \( T(W) \in \mathcal{LS} \).

(4) It follows from the proof of the preceding lemma. \( \square \)
Now let\[ \mathcal{LS} = \bigcup_{i=1}^{L} \mathcal{L}_i, \quad \mathcal{LS}_n = \{ W; W \triangleleft T^n(W'), W' \in \mathcal{LS} \}. \]

Remark that $\mathcal{LS}_n$ is monotone with respect to $n$. The following theorem follows directly from the above lemma:

**Theorem 3.1.** $\mathcal{LS} = \bigcup \mathcal{LS}_n = \lim_{n \to \infty} \mathcal{LS}_n$.

**Remark:** The above theorem tells us that all left special words (which determine the complexity) can be obtained from a finite set $\mathcal{LS}$ of left special words and by the operation $T$.

4. Structure of $S$ and Calculation of $\Delta^2 p(n)$

Knowing the initial values, calculating $p(n)$ boils down into calculating $\Delta s(n+1) = \#S_{n+1} - \#S_n$. Notice that any suffix of a special word is also special, hence if $W \in S_{n+1}$ then $W = \delta W'$ for some $W' \in S_n$ and $\delta \in \{a, b\}$. Thus the set of special words can be visualized as a tree showing clearly how $S_{n+1}$ derives from $S_n$ (see the example and the figure therein in the last section).

As usual, for studying the special words’ tree, we shall use the following notations for special words, see also [6]:

**Definition 4.1.** Let $W \in S$. If neither $aW$ nor $bW$ is in $S$, we say that $W$ is a weak special word; If both $aW$ and $bW$ are in $S$, we say that $W$ is a strong special word. We denote by $S^0$ and $S^2$ the set of weak special words and the strong weak special words respectively. The collection of other special words is denoted by $S^1$.

For $i \in \{0, 1, 2\}$, we write $S^n_i = S^i \cap L_n$. It is clear that
\[ S_n = S^0_n \cup S^1_n \cup S^2_n \quad \text{and} \quad S = S^0 \cup S^1 \cup S^2. \]

**Lemma 4.1.** (1) $\Delta s(n+1) = s(n+1) - s(n) = \#S^2_n - \#S^0_n$.

(2) $S^0_n \cup S^2_n \subset S_n \cap \mathcal{LS}_n$.

**Proof.** (see Theorem 4.5.4 [6]) (1) and the fact that $S^2_n \subset \mathcal{LS}_n$ are obvious. If a special word has only one left extension, then this left extension is also special. □

**Lemma 4.2.** Let $c, d \in A, W \in \xi$. If $cWd \in \xi$, then $\delta c T(W)d \in \xi$. Conversely, if $\delta c T(W)d \in \xi$ and $|T(W)| \geq L$, then $cWd \in \xi$.

**Proof.** If $cWd \in \xi$, then by Lemma 3.1, $T(cWd) \in \xi$, i.e., $W_0 \sigma(c) \sigma(W) \sigma(d) \in \xi$. This together with Corollary 2.1 and the fact that $\sigma$ is well marked implies that $\delta c W_0 \sigma(W)d = \delta c T(W)d \in \xi$.

Conversely, if $\delta c T(W)d \in \xi$ and $|T(W)| \geq L$, then by Lemma 3.3 we know that $\delta c T(W)d = \delta c W_0 \sigma(W)d$ is a natural decomposition. Considering the ancestor of $\delta c T(W)d$, we know, again by Corollary 2.1 and the fact that $\sigma$ is well marked, that $cWd \in \xi$. □
Lemma 4.3. If \( W \in S \), then \( T(W) \in S \) (thus \( \sigma(W) \in S \)); furthermore \( T(W)a = W_0[\sigma(W)]a \), and \( T(W)b = W_0[\sigma(W)]b \).

Conversely if \( W \in S \) and \( |W| \geq L \), then there exists \( U \in S \) such that \( W \triangleright T(U) \).

Proof. Let \( W \in S \), then \( Wa, Wb \in \xi \), and by Lemma 3.1

\[
W_0[\sigma(W)]A, W_0[\sigma(W)]B \in \xi. 
\]

Recalling \( A = a\oplus \) and \( B = b\oplus \), The first part of our lemma is thus proved.

The rest part is a restatement of Lemma 3.4(2). \( \square \)

We can say more on the structure of \( S^2 \) and \( S^0 \).

Lemma 4.4. If \( W \in S^2 \) then \( T(W) \in S^2 \). Conversely if \( W \in S^2 \) and \( |W| \geq L \), then there exists a unique \( U \in S^2 \) such that \( W \triangleright T(U) \).

Proof. Let \( W \in S^2 \). Then we have, by definition, that

\[
aWa, aWb, bWa, bWb \in \xi, 
\]

and, by Lemma 4.2 that

\[
\delta_aT(W)a, \delta_aT(W)b, \delta_bT(W)a, \delta_bT(W)b \in \xi, 
\]

i.e., \( T(W) \in S^2 \). The first part of the lemma is proved.

Now suppose \( W \in S^2 \) and \( |W| \geq L \). Then by Lemmas 4.1(2) and 3.4(3), \( W = T(U) \). By Lemma 4.2 \( U \in S^2 \). \( \square \)

Lemma 4.5. If \( W \in S^0 \) and \( |T(W)| \geq L \), then \( T(W) \in S^0 \). Conversely if \( W \in S^0 \) and \( |W| \geq L \), then there exists a unique \( U \in S^0 \) such that \( W = T(U) \).

Proof. By Lemma 4.2 when \( |T(W)| \geq L \) we know that \( cWd \in \xi \) if and only if \( \delta_cT(W)d \in \xi \). Whence \( W \in S^0 \) if and only if \( T(W) \in S^0 \). The remaining proof is almost same with the corresponding part for the preceding Lemma. \( \square \)

Now denote \( \tilde{S}^2 = \bigcup_{i=1}^{L} S^2_i \) the set of strong special words of length less than \( L \); \( \tilde{S}^2 \) the set of the words \( W \in \tilde{S}^2 \) such that \( |T(W)| > L \). The sets \( \tilde{S}^0 \) and \( \tilde{S}^0 \) are defined in a similar way. Let

\[
(4.11) \quad \tilde{S} = \tilde{S}^0 \cup \tilde{S}^2
\]

which will be considered as “initial special words”.

Lemma 4.6. For any \( n > L \), we have

\[
\#S^2_n = \sum_{W \in \tilde{S}^2} \sum_{k \geq 1} \delta(|T^k(W)|, n), \quad \#S^0_n = \sum_{W \in \tilde{S}^0} \sum_{k \geq 1} \delta(|T^k(W)|, n),
\]

where \( \delta(i, j) \) is the Kronecker symbol: \( \delta(i, j) = 1 \) if \( i = j \) and \( = 0 \) otherwise.
Proof. Let $U \in \mathcal{S}_n^2$. By Lemma 4.4, there exist $k \geq 1$ and $W \in \tilde{\mathcal{S}}^2$, which are unique, such that $U = T^k(W)$. Conversely if $|T^k(W)| = n$ for some $k \geq 1, W \in \tilde{\mathcal{S}}^2$, then $T^k(W) \in \mathcal{S}_n^2$. Thus we have
\[
\mathcal{S}_n^2 = \{U; U = T^k(W), |T^k(W)| = n, k \geq 1, W \in \tilde{\mathcal{S}}^2\}
\]
where $k$ and $W$ in the representation $U = T^k(W)$ are uniquely determined by $U$. The first equality is thus proved. The second is proved similarly. $\square$

The following formula then follows from the above lemma and Lemma 4.1:

Lemma 4.7. For any $n > L$, we have
\[
\Delta s(n + 1) = s(n + 1) - s(n) = \sum_{W \in \tilde{\mathcal{S}}^2} \sum_{k \geq 1} \delta(|T^k(W)|, n) - \sum_{W \in \tilde{\mathcal{S}}^0} \sum_{k \geq 1} \delta(|T^k(W)|, n).
\]

It can be written as
\[
\Delta s(n + 1) = \sum_{W \in \mathcal{S}} \sum_{k \geq 1} \text{sgn}(W) \delta(|T^k(W)|, n),
\]
where $\mathcal{S} = \bigcup_{i=1}^{L} \mathcal{S}_i$ (the special words of length less than $L$), and
\[
\text{(4.12)} \quad \text{sgn}(W) = \begin{cases} 
-1 & \text{if } W \in \tilde{\mathcal{S}}^0 \\
1 & \text{if } W \in \tilde{\mathcal{S}}^2 \\
0 & \text{otherwise.} 
\end{cases}
\]

Remark: 1. The function $\text{sgn}(\cdot)$ is equal to the bilateral multiplicity of a factor [6]). See Theorem 4.5.4 [6] for more general cases.

2. The above lemma tells us that the complexity $p(n)$ can be computed knowing a finite set $\mathcal{S}$ of special words. In the next section, we will find out a (non-linear) recurrence formula for the computation.

5. Recurrence Formula for the Complexity

Recall that $M$ denotes the incidence matrix of $\sigma$. Then $M^2$ is the incidence matrix of $\sigma^2$ which possess non-negative eigenvalues. Since $\sigma$ and $\sigma^2$ share the fixed sequence $\xi$, we may suppose without loss of generality that the eigenvalues of $M$ is non-negative.

Let $\lambda_1 \geq \lambda_2 \geq 0$ be the two eigenvalues, $V_1, V_2$ be the corresponding eigenvectors. Since $M$ is primitive, $\lambda_1 > \lambda_2$ and $V_1$ is positive.

Recall that: for $W \in \{a, b\}^*$, $L(W) = (|W|_a, |W|_b)^t$, \[
\text{(5.13)} \quad |\sigma^n(W)| = (1, 1)M^nL(W).
\]
Lemma 5.1. Let \( X, Y \in \mathbb{R}^2 \). Then there exists \( N = N(X, Y) \geq 1 \) such that 
\[(1, 1)M^{N+n}(X - Y) \ (n \in \mathbb{N}) \text{ is of constant sign. That is,}
(1, 1)M^{N+n}X > (\text{resp.} \ <) \ (1, 1)M^{N+n}Y \text{ for all } n \in \mathbb{N}.
\]

Proof. Let \( X - Y = \mu_1 V_1 + \mu_2 V_2 \) where \( \mu_1, \mu_2 \in \mathbb{R} \), then for \( k \geq 1 \),
\[(1, 1)M^k(X - Y) = \lambda_1^k \mu_1(1, 1) V_1 + \lambda_2^k \mu_2(1, 1) V_2.\]

Case 1. \( \mu_1 = 0 \). Then \((1, 1)M^k(X - Y) = \lambda_2^k \mu_2(1, 1) V_2\), which is obviously of the
sign of \( \lambda_2 \mu_2(1, 1) V_2 \) independent of \( k \geq 1 \).

Case 2. \( \mu_1 > 0 \). Since \( \lambda_1 > 0 \), \((1, 1)V_1 > 0 \) and \( \lambda_1 > \lambda_2 \geq 0 \), there exists \( N \geq 1 \)
such that for \( k \geq N \) we have \( \lambda_1^k \mu_1(1, 1) V_1 + \lambda_2^k \mu_2(1, 1) V_2 > 0 \).

Case 3. \( \mu_1 < 0 \). The similar proof as Case 2. \( \square \)

Corollary 5.1. Let \( W_1, W_2 \in A^* \). There exists \( N = N(W_1, W_2) \) such that 
\[|T^{N+n}(W_1)| - |T^{N+n}(W_2)| \ (n \in \mathbb{N}) \text{ is of constant sign. This sign (called the final sign) will be denoted by SGN}\{W_1, W_2\}.
\]

Proof. The lemma follows directly from the above lemma and \((5.13)\). \( \square \)

In fact, we can say more:

Corollary 5.2. Let \( W_1, W_2 \in A^* \). Then there exist \( m_1 = m_1(W_1, W_2), m_2 = m_2(W_1, W_2) \in \mathbb{N} \) such that one of the following alternatives holds:

1. \( |T^{m_1}(W_1)| = |T^{m_2}(W_2)| < |T^{m_1+1}(W_1)| = |T^{m_2+1}(W_2)| < |T^{m_1+2}(W_1)| = |T^{m_2+2}(W_2)| < \ldots \)

2. \( |T^{m_1}(W_1)| < |T^{m_2}(W_2)| < |T^{m_1+1}(W_1)| < |T^{m_2+1}(W_2)| < |T^{m_1+2}(W_1)| < |T^{m_2+2}(W_2)| < \ldots .\)

Proof. If SGN\((T^m(W_1), T^n(W_2)) = 0 \) for some \( m, n \in \mathbb{N} \), the alternative (1)
holds.

Otherwise, SGN\((T^m(W_1), T^n(W_2)) \neq 0 \) for any \( m, n \in \mathbb{N} \). We assume, without
loss of generality, that SGN\((W_1, W_2) = -1 \). Due to the primitivity, \( W_2 \) is a
factor of \( T^l(W_1) \) for \( l \) large enough, and it turns out that SGN\((T^l(W_1), W_2) = 1 \).
Now clearly \( m \mapsto \text{SGN}(T^m(W_1), W_2) \) is an increasing mapping from \( \mathbb{N} \) onto
\( \{-1, 1\} \), therefore there exists \( m \in \mathbb{N} \) such that SGN\((T^m(W_1), W_2) = -1 \), while
SGN\((T^{m+1}(W_1), W_2) = 1 \). Whence the alternative (2) holds for
\[m_2 = \max\{N(T^m(W_1), W_2), N(T^{m+1}(W_1), W_2)\}, \text{ and } m_1 = m + m_2. \square \]

Now we can deduce from the above lemma the recurrence properties of the
complexity. First let \( \tilde{S} = \{S_1, S_2, \ldots, S_K\} \) and denote
\[n_1 - 1 = \max \{\max\{m_1(W_1, W_2), m_2(W_1, W_2)\}; \ W_1, W_2 \in \tilde{S}\}, \]
where \( m_1(W_1, W_2), m_2(W_1, W_2) \) are defined in Lemma 5.2. \( \square \)
We start from $T^{n_1}(S_1)$. By Lemma 5.2, for each $j = 2, 3, \cdots, K$, there exists unique $n_j \in \mathbb{N}$ such that $|T^{n_1}(S_1)| \leq |T^{n_j}(S_j)| < |T^{n_1+1}(S_1)|$. Without loss of generality we may suppose that

$$|T^{n_1}(S_1)| \leq |T^{n_2}(S_2)| \leq \cdots \leq |T^{n_K}(S_K)| \leq |T^{n_1+1}(S_1)|.$$ 

Then for simplifying the notations let $N_j^j = |T^{j}(T^{n_h}(S_h))|$ $(1 \leq k \leq K, j \in \mathbb{N})$. We have by Lemma 5.2 the following unison property for the "jumps of $|T^j(W_k)|$":

$$\begin{align*}
N_0^0 & \leq N_0^1 \leq \cdots \leq N_K^0 \\
& \leq N_1^1 \leq N_1^2 \leq \cdots \leq N_K^1 \\
& \cdots \cdots \\
& \leq N_j^j \leq N_j^{j+1} \leq \cdots \leq N_K^j \\
& \leq N_{j+1}^{j+1} \leq \cdots 
\end{align*}$$

(5.14)

Now we can formulate the recurrence formula of the complexity. Let $\chi_{[m,n)}$ denote the indicator function of the integers’ interval $[m, n)$. Let $I^j = [N_j^j, N_{j+1}^j), j \in \mathbb{N}$. We see that $I^j$ is the disjoint union of the subintervals $I_k^j = [N_k^j, N_{k+1}^j)$ $(j \in \mathbb{N}, k \in \{1, 2, \cdots, K\})$, where $N_{K+1}^j = N_1(j + 1)$. That is

$$[N_1^0, \infty) = \bigcup_{j=0}^{\infty} I^j, \quad I^j = \bigcup_{k=1}^{K} I_k^j.$$ 

5.1. Initial values of the complexity.

Finally let $c_k = \sum_{i=1}^{K} \sgn(S_i) \delta(|T^{n_i}(S_i)|, |T^{n_h}(S_h)|) (k = 1, \cdots, K)$, where $\sgn(\cdot)$ is defined in (4.12). Then by Lemma 4.7 we have, $\Delta s(n+1) = c_k$ if $n = |T^{n_h}(S_h)|(k = 1, \cdots, K)$ and $= 0$ otherwise. In other words, $n \mapsto s(n+1) (n \in I^0)$ is a step function with jumps $c_k$ at $n = N_k(0) (k \in \{1, 2, \cdots, K\})$:

$$s(n+1) = s(N_1(0)) + \sum_{k=1}^{K} (c_1 + \cdots + c_k) \chi_{I_k^0}(n) (n \in I^0),$$

(5.15)  

5.2. Recurrence formula of $s(\cdot + 1)$ on $I^j$.

Notice that $I^j = \bigcup_{k=1}^{K} I_k^j (j \in \mathbb{N})$ can be calculated directly or by some easy recurrence formula as described in the following:

**Proposition 5.1.** We have for any $W \in \mathbb{A}^*$, $n \in \mathbb{N}$,

1. $|\sigma^{n+2}(W)| = \text{tr}(M) |\sigma^{n+1}(W)| - \text{det}(M) |\sigma^n(W)|$;

2. $|\sigma^n(W)| = \lambda_1^n \mu_1(1, 1)V_1 + \lambda_2^n \mu_2(1, 1)V_2$ if $L(W) = \mu_1 V_1 + \mu_2 V_2$;
\[ |T^n(W)| = \lambda_1^n \mu_1(1, 1)V_1 + \lambda_2^n \mu_2(1, 1)V_2 + b_n, \]
where \( b_n(n \in \mathbb{N}) \) is a fixed sequence given explicitly by \( L(W_0) \) and \( M \).

**Proof.** All the results can be deduced easily from (3.8), (5.13) and Cayley-Hamilton formula (with \( I \) denotes the identity matrix): \( M^2 = \text{tr}(M)M - \det(M)I \). □

We have just seen the recurrence properties of the intervals \( I_j(j \in \mathbb{N}) \). Still using Lemma 4.7 and the formula (5.14) and we see that what happens for \( s(n+1) \) \((n \in I_j, j \in \mathbb{N})\) is recurrently the same as \( s(n+1)(n \in I^0) \), i.e., similar to (5.15) we have proved the following

**Theorem 5.1.** Let \( \sigma \) be a well marked, primitive, non-periodic substitution having non-negative eigenvalues. Then for \( n \in [N^0_1, \infty) = \bigcup_{j=0}^{\infty} I^j \), the following recurrence formula holds:

\[ s(n+1) = s(N^j_1) + \sum_{k=1}^{K} (c_1 + \cdots + c_k)\chi_{I_k}(n) \quad (n \in I^j, j \geq 0). \]

**Remark:** 1. The conditions “primitive, well marked, non-periodic, having non-negative eigenvalues” are non-essential as have already mentioned.
2. \( s(N^j_{1+1}) - s(N^j_1) \equiv c_1 + \cdots + c_K \quad (j \in \mathbb{N}) \), which implies roughly \( s(\lambda_1^n) \approx n(c_1 + \cdots + c_K) \) for large \( n \).
3. Although the above mentioned \( N^0_1 \) can be more or less controlled in the proof of the theorem, but how to give efficiently this big integer \( N \) remains as an open problem.

Finally let us give briefly an example: consider the substitution \( \sigma = (aab, ba) \) i.e., \( a \mapsto aab, b \mapsto ba \).

For this substitution, we have \( W_0 = \varepsilon \) and thus \( T = \sigma \). The incidence matrix \( M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \) and the characteristic polynomial is \( \lambda^2 - 3\lambda + 1 \). The fixed point reads

\[ \xi = aabaabbaaabaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaabbaab baab \cdots \]

The tree of the special words is depicted in Figure 1.

The weak and strong special words (here \( \sigma^0 \) is the identity map):
- \( S^0 = \{aba, aabbaa, \cdots \} = \{\sigma^n(aba); n = 0, 1, 2, \cdots \} \),
- \( S^2 = \{\varepsilon, a, aab, aabbaabbbaa, \cdots \} = \{\varepsilon \} \cup \{\sigma^n(a); n = 0, 1, 2, \cdots \} \).

From the structure of special words, the numbers of special words \( s(n) \) and the complexity \( p(n) \) read
We can formulate $s(n)$ as

$$s(n) = \begin{cases} 
1 & \text{if } n = 0, \\
2 & \text{if } n = 1, \\
3 & \text{if } n \in \{2, 3\} \cup \bigcup_{k \geq 0} [d(k) + 1, g(k + 1)], \\
4 & \text{if } n \in \bigcup_{k \geq 0} [g(k) + 1, d(k)], 
\end{cases}$$

where the number sequences $g(k)$ and $d(k)$ are defined as

$$g(k) = (1, 1)M^k(2, 1)^t, \quad d(k) = (1, 1)M^k(3, 1)^t,$$

satisfying both the same recurrence:

$$\begin{cases} 
g(k + 2) = 3g(k + 1) - g(k), \\
d(k + 2) = 3d(k + 1) - d(k), 
\end{cases}$$

with $g(0) = 3, g(1) = 8$ and $d(0) = 4, d(1) = 11$. 

---

![Figure 1. Tree of Special Words](image)
ACKNOWLEDGEMENT  The authors would like to thank Prof. Z.Y. Wen (Tsinghua), J.P. Allouche (Jussieu) and others for helpful discussions, references and corrections.

REFERENCES

[1] J.P. Allouche, Automates finis en théorie des nombres. Exposition. Math. 5 (1987) 239–266.
[2] P. Arnoux and S. Ito, Pisot substitutions and Rauzy fractals. Journées Montoises d’Informatique Théorique (Marne-la-Vallée, 2000), Bull. Belg. Math. Soc. Simon Stevin 8 (2001) 181–207.
[3] P. Arnoux and G. Rauzy, Représentation géométrique de suites de complexité $2n+1$. Bull. Soc. Math. France, 119 (1991) 199–215.
[4] J. Cassaigne, Special factors of sequences with linear subword complexity. in Developments in Language Theory II (DLT’95), Magdeburg (Allemagne), 25–34, World Sci. 1996.
[5] J. Cassaigne. Complexité et facteurs spéciaux. Bull. Belg. Math. Soc. Simon Stevin, 4 (1997) 67–88, Journées Montoises (Mons, 1994).
[6] J. Cassaigne and F. Nicolas, Factor complexity. Combinatorics, automata and number theory, 163–247, Encyclopedia Math. Appl., 135, Cambridge Univ. Press, Cambridge, 2010.
[7] A. Cobham, Uniform tag sequences. Math. System Theory 6 (1972) 164–192.
[8] A.E. Frid, On the subword complexity of iteratively generated infinite words. Discrete Appl. Math. 6 114(1-3):115-120, 2001.
[9] M. Lothaire, Combinatorics on words. Encyclopedia of Mathematics and its applications 17, Cambridge, 1997.
[10] M. Lothaire, Algebraic combinatorics on words. Encyclopedia of Mathematics and its Applications 90, Cambridge University Press, 2002.
[11] R.C. Lyndon and P.E. Schupp, Combinatorial group theory, Spring-Verlag, 1977.
[12] W. Magnus, A. Karrass and D. Solitar, Combinatorial group theory: Presentations of groups in terms of generators and relations, Second Revised Edition by Dover Publications Inc. 1976.
[13] M. Morse and G.A. Hedlund, Symbolic dynamics II: Sturmian trajectories. Amer. J. Math. 62 (1940) 1–42.
[14] B. Mossé, Reconnaissabilité des substitutions et complexité des suites automatiques. Bull. Soc. Math. France, 124 (1996) 329–346.
[15] J.-J. Pansiot, Complexité des facteurs des mots infinis engendrés par morphismes itérés. Automata, languages and programming (Antwerp, 1984), 380–389, Lecture Notes in Comput. Sci., 172, Springer, Berlin, 1984.
[16] N. Pytheas Fogg, Substitutions in dynamics, arithmetics and combinatorics, V. Berthé, S. Ferenczi, C. Mauduit, A. Siegel, (eds.), Lecture Notes in Mathematics 1794, Springer, 2002.
[17] G. Rote, Sequences with subword complexity $2^n$. J. Number Theory, 46 (1994) 196–213.
[18] P. Séébold, Fibonacci morphisms and Sturmian words. Theoret. Comput. Sci. 88 (1991) 365–384.
[19] P. Séébold, An effective solution to the D0L periodicity problem in the binary case. EATCS Bull. 36 (1988) 137–151.
[20] B. Tan and Z-Y Wen, Invertible substitutions and Sturmian sequences. European Journal of Combinatorics 24 (2003) 983-1002.
[21] B. Tan, Z.-X. Wen and Y. Zhang, On the triplex substitution - combinatorial properties. C.R.Acad. Sci. Paris, 346 (2008) 813–818.
[22] Z.-X. Wen and Z.-Y. Wen, *Local isomorphisms of invertible substitutions*. C. R. Acad. Sci. Paris, t. 318 Série I, (1994) 299–304.

School of Mathematics and Statistics, Huazhong University of Science and Technology, 430074 Wuhan, P. R. China

E-mail address: tanbo@hust.edu.cn

School of Mathematics and Statistics, Huazhong University of Science and Technology, 430074 Wuhan, P. R. China

E-mail address: zhi-xiong.wen@hust.edu.cn

School of Mathematics and statistics, Wuhan University, 430072 Wuhan, P. R. China

E-mail address: ypzhang@whu.edu.cn