Powers in a class of $A$-strict standard
episturmian words

Amy Glen

LaCIM, Université du Québéco à Montréal, C.P. 8888, Succ. Centre-ville,
Montréal, Québec, CANADA, H3C 3P8

Abstract

This paper concerns a specific class of strict standard episturmian words whose
directive words resemble those of characteristic Sturmian words. In particular, we
explicitly determine all integer powers occurring in such infinite words, extending
recent results of Damanik and Lenz (2003), who studied powers in Sturmian words.
The key tools in our analysis are canonical decompositions and a generalization of
singular words, which were originally defined for the ubiquitous Fibonacci word. Our
main results are demonstrated via some examples, including the $k$-bonacci word, a
generalization of the Fibonacci word to a $k$-letter alphabet ($k \geq 2$).

Key words: episturmian word; Sturmian word; Arnoux-Rauzy sequence; $k$-bonacci
word; singular word; index; powers

2000 MSC: 68R15

1 Introduction

Introduced by Droubay, Justin and Pirillo [8], episturmian words are an interesting nat-
ural generalization of the well-known family of Sturmian words (aperiodic infinite words
of minimal complexity) to an arbitrary finite alphabet. Episturmian words share many
properties with Sturmian words and include the well-known Arnoux-Rauzy sequences, the
study of which began in [1] (also see [16,21] for example).

In this paper, the study of episturmian words is continued in more detail. In particular,
for a specific class of episturmian words (a typical element of which we shall denote by
$s$), we will explicitly determine all of the integer powers of words occurring in it. This has
recently been done in [6] for Sturmian words, which are exactly the aperiodic episturmian
words over a 2-letter alphabet.

Email address: amy.glen@gmail.com (Amy Glen).
A finite word \( w \) is said to have a (non-trivial) integer power in an infinite word \( x \) if \( w^p = uu\cdots w \ (p \text{ times}) \) is a factor of \( x \) for some integer \( p \geq 2 \). Here, our analysis of powers occurring in episturmian words \( s \) hinges on canonical decompositions in terms of their ‘building blocks’. Another key tool is a generalization of singular words, which were first defined in [23] for the ubiquitous Fibonacci word, and later extended to Sturmian words in [19] and the Tribonacci sequence in [22]. Our generalized singular words will prove to be useful in the study of factors of episturmian words, just as they have been for Sturmian words.

This paper is organized as follows. After some preliminaries (Section 2), we define, in Section 3, a restricted class of episturmian words upon which we will focus for the rest of the paper. A typical element of this class will be denoted by \( s \). In Section 4, we prove some simple results, which lead us to a generalization of singular words for episturmian words of the form \( s \). The index (i.e., maximal fractional power) of the building blocks of \( s \) is then studied in Section 5. Finally, in Section 6, we determine all squares (and subsequently higher powers) occurring in episturmian words, just as they have been for Sturmian words.

2 Definitions and notations

2.1 Words

Let \( \mathcal{A} \) denote a finite alphabet. A (finite) word is an element of the free monoid \( \mathcal{A}^* \) generated by \( \mathcal{A} \), in the sense of concatenation. The identity \( \varepsilon \) of \( \mathcal{A}^* \) is called the empty word, and the free semigroup, denoted by \( \mathcal{A}^+ \), is defined by \( \mathcal{A}^+ := \mathcal{A}^* \setminus \{\varepsilon\} \). An infinite word (or simply sequence) \( x \) is a sequence indexed by \( \mathbb{N} \) with values in \( \mathcal{A} \), i.e., \( x = x_0x_1x_2\cdots \), where each \( x_i \in \mathcal{A} \). The set of all infinite words over \( \mathcal{A} \) is denoted by \( \mathcal{A}^\omega \), and we define \( \mathcal{A}^\infty := \mathcal{A}^* \cup \mathcal{A}^\omega \). If \( u \) is a non-empty finite word, then \( w^u \) denotes the purely periodic infinite word \( uwu\cdots \).

If \( w = x_1x_2\cdots x_m \in \mathcal{A}^+ \), each \( x_i \in \mathcal{A} \), the length of \( w \) is \( |w| = m \) and we denote by \( |w|_a \) the number of occurrences of a letter \( a \) in \( w \). (Note that \( |\varepsilon| = 0 \).) The reversal of \( w \) is \( \overline{w} = x_mx_{m-1}\cdots x_1 \), and if \( w = \overline{w} \), then \( w \) is called a palindrome.

A finite word \( w \) is a factor of \( z \in \mathcal{A}^\infty \) if \( z = uvv \) for some \( u \in \mathcal{A}^*, \ v \in \mathcal{A}^\infty \), and we write \( w \prec z \). Further, \( w \) is called a prefix (resp. suffix) of \( z \) if \( u = \varepsilon \) (resp. \( v = \varepsilon \)), and we write \( w \prec_p z \) (resp. \( w \prec_s z \)).

An infinite word \( x \in \mathcal{A}^\omega \) is called a suffix of \( z \in \mathcal{A}^\omega \) if there exists a word \( w \in \mathcal{A}^+ \) such that \( z = wx \). A factor \( w \) of a word \( z \in \mathcal{A}^\infty \) is right (resp. left) special if \( wa, wb \) (resp. \( aw, bw \)) are factors of \( z \) for some letters \( a, b \in \mathcal{A}, a \neq b \).

For any word \( w \in \mathcal{A}^\infty \), \( \Omega(w) \) denotes the set of all its factors, and \( \Omega_n(w) \) denotes the set of all factors of \( w \) of length \( n \in \mathbb{N} \), i.e., \( \Omega_n(w) := \Omega(w) \cap \mathcal{A}^n \) (where \( |w| \geq n \) for \( w \) finite). Moreover, the alphabet of \( w \) is \( \text{Alph}(w) := \Omega(w) \cap \mathcal{A} \) and, if \( w \) is infinite, we denote by
Ult(w) the set of all letters occurring infinitely often in w. Two infinite words \( x, y \in \mathcal{A}^\omega \) are said to be equivalent if \( \Omega(x) = \Omega(y) \), i.e., if \( x \) and \( y \) have the same set of factors.

Recall that a finite word \( w \) is said to have a (non-trivial) integer power in an infinite word \( x = x_0x_1x_2x_3 \cdots \) if \( w^p = ww \cdots w \) (\( p \) times) is a factor of \( x \) (i.e., there exists an integer \( i \geq 0 \) such that \( w^p = x_{i}x_{i+1} \cdots x_{i+|w|-1} \) for some integer \( p \geq 2 \).

Let \( w = x_1x_2 \cdots x_m \in \mathcal{A}^* \), each \( x_i \in \mathcal{A} \), and let \( j \in \mathbb{N} \) with \( 0 \leq j \leq m - 1 \). The \( j \)-th conjugate of \( w \) is the word \( C_j(w) := x_{j+1}x_{j+2} \cdots x_m x_1x_2 \cdots x_j \), and we denote by \( C(w) \) the conjugacy class of \( w \), i.e., \( C(w) := \{ C_j(w) : 0 \leq j \leq |w| - 1 \} \). Observe that if \( w \) is primitive (i.e., not a power of a shorter word), then \( w \) has exactly \( |w| \) distinct conjugates.

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

A morphism on \( \mathcal{A} \) is a map \( \psi : \mathcal{A}^* \to \mathcal{A}^* \) such that \( \psi(uv) = \psi(u)\psi(v) \) for all \( u, v \in \mathcal{A}^* \). It is uniquely determined by its image on the alphabet \( \mathcal{A} \).

### 2.2 Episturmian words

An infinite word \( t \in \mathcal{A}^\omega \) is episturmian if \( \Omega(t) \) is closed under reversal and \( t \) has at most one right (or equivalently left) special factor of each length. Moreover, an episturmian word is standard if all of its left special factors are prefixes of it.

Standard episturmian words are characterized in [8] using the concept of the palindromic right-closure \( w^{(+)} \) of a finite word \( w \), which is the (unique) shortest palindrome having \( w \) as a prefix (see [7]). Specifically, an infinite word \( t \in \mathcal{A}^\omega \) is standard episturmian if and only if there exists an infinite word \( \Delta(t) = x_1x_2x_3 \cdots \), each \( x_i \in \mathcal{A} \), called the directive word of \( t \), such that the infinite sequence of palindromic prefixes \( u_1 = \varepsilon, u_2, u_3, \ldots \) of \( t \) (which exists by results in [8]) is given by

\[
 u_{n+1} = (u_nx_n)^{(+)} , \quad n \in \mathbb{N}^+ .
\]

**Note.** For any \( w \in \mathcal{A}^+ \), \( w^{(+)} = ww^{-1}\overline{w} \) where \( v \) is the longest palindromic suffix of \( w \).

An important point is that a standard episturmian word \( t \) can be constructed as a limit of an infinite sequence of its palindromic prefixes, i.e., \( t = \lim_{n \to \infty} u_n \).

Let \( a \in \mathcal{A} \) and denote by \( \Psi_a \) the morphism on \( \mathcal{A} \) defined by

\[
 \Psi_a : \begin{cases} 
 a & \mapsto a \\
 x & \mapsto ax \quad \text{for all } x \in \mathcal{A} \setminus \{a\}. 
\end{cases}
\]

Another useful characterization of standard episturmian words is the following (see [15]). An infinite word \( t \in \mathcal{A}^\omega \) is standard episturmian with directive word \( \Delta(t) = x_1x_2x_3 \cdots \) if and only if there exists an infinite sequence of infinite words \( t^{(0)} = t, t^{(1)}, t^{(2)}, \ldots \) such that \( t^{(i-1)} = \Psi_{x_i}(t^{(i)}) \) for all \( i \in \mathbb{N}^+ \). Moreover, each \( t^{(i)} \) is a standard episturmian word with directive word \( \Delta(t^{(i)}) = x_{i+1}x_{i+2}x_{i+3} \cdots \), the \( i \)-th shift of \( \Delta(t) \).
To the prefixes of the directive word $\Delta(t) = x_1x_2 \cdots$, we associate the morphisms
\[ \mu_0 := \text{Id}, \quad \mu_n := \Psi x_1 \Psi x_2 \cdots \Psi x_n, \quad n \in \mathbb{N}^+, \]
and define the words
\[ h_n := \mu_n(x_{n+1}), \quad n \in \mathbb{N}, \]
which are clearly prefixes of $t$. We have the following useful formula [15]
\[ u_{n+1} = h_{n-1}u_n; \]
and whence, for $n > 1$ and $0 < p < n$,
\[ u_n = h_{n-2}h_{n-3} \cdots h_1h_0 = h_{n-2}h_{n-3} \cdots h_{p-1}u_p. \quad (2) \]

Some useful properties of the words $h_n$ and $u_n$ are given by the following lemma.

**Lemma 2.1** [15] For all $n \in \mathbb{N}$,

(i) $h_n$ is a primitive word;

(ii) $h_n = h_{n-1}$ if and only if $x_{n+1} = x_n$;

(iii) if $x_{n+1} \neq x_n$, then $u_n$ is a proper prefix of $h_n$. □

Two functions can be defined with regard to positions of letters in a given directive word. For $n \in \mathbb{N}^+$, let $P(n) = \sup\{p < n : x_p = x_n\}$ if this integer exists, $P(n)$ undefined otherwise. Also, let $S(n) = \inf\{p > n : x_p = x_n\}$ if this integer exists, $S(n)$ undefined otherwise. By the definitions of palindromic closure and the words $u_n$, it follows that $u_{n+1} = u_nx_nu_n$ (whence $h_{n-1} = u_nx_n$) if $x_n$ does not occur in $u_n$, and $u_{n+1} = u_nu_{P(n)}^{-1}u_n$ (whence $h_{n-1}u_{P(n)} = u_n$) if $x_n$ occurs in $u_n$. Thus, if $P(n)$ exists, then
\[ h_{n-1} = h_{n-2}h_{n-3} \cdots h_{P(n)-1}, \quad n \geq 1. \quad (3) \]

**2.2.1 Strict episturmian words**

A standard episturmian word $t \in A^\omega$, or any equivalent (episturmian) word, is said to be $B$-strict (or $k$-strict if $|B| = k$, or strict if $B$ is understood) if $\text{Alph}(\Delta(t)) = \text{Ult}(\Delta(t)) = B \subseteq A$. In particular, a standard episturmian word over $A$ is $A$-strict if every letter in $A$ occurs infinitely many times in its directive word. The $k$-strict episturmian words have complexity $(k-1)n + 1$ for each $n \in \mathbb{N}$ (i.e., $(k-1)n + 1$ distinct factors of length $n$ for each $n \in \mathbb{N}$). Such words are exactly the $k$-letter Arnoux-Rauzy sequences.

**2.2.2 Return words**

Let $x \in A^\omega$ be recurrent, i.e., any factor $w$ of $x$ occurs infinitely often in $x$. A return word of $w \in \Omega(x)$ is a factor of $x$ that begins at an occurrence of $w$ in $x$ and ends exactly before the next occurrence of $w$ in $x$. Thus, a return word of $w$ is a non-empty factor $u$ of $x$ such that $w$ is a prefix of $uw$ and $uw$ contains two distinct occurrences of $w$. This notion was introduced independently by Durand [9], and Holton and Zamboni [14].
Episturmian words are recurrent and, according to [17, Corollary 4.5], each factor of an \(A\)-strict episturmian word has exactly \(|A|\) return words.

3 A class of strict standard episturmian words

Given any infinite sequence \(\Delta = x_1x_2x_3\cdots\) over a finite alphabet \(A\), we can define a standard episturmian word having \(\Delta\) as its directive word (using (1)). In this paper, however, we shall only consider a specific family of \(A\)-strict episturmian words.

Let \(A_k\) denote a \(k\)-letter alphabet, say \(A_k = \{a_1, a_2, \ldots, a_k\}\), and suppose \(t\) is a standard episturmian word over \(A_k\). Then the directive word of \(t\) can be expressed as:

\[
\Delta(t) = a_1^{d_1}a_2^{d_2}\cdots a_k^{d_k}a_{k+1}^{d_{k+2}}\cdots a_k^{d_{2k}}a_1^{d_{2k+1}}\cdots,
\]

where the \(d_i\) are non-negative integers. In what follows, we will restrict our attention to the case when all \(d_i > 0\); that is, we shall only study the class of \(k\)-strict standard episturmian words \(s \in A_k^w\) with directive words of the form:

\[
\Delta = a_1^{d_1}a_2^{d_2}\cdots a_k^{d_k}a_{k+1}^{d_{k+2}}\cdots a_k^{d_{2k}}a_1^{d_{2k+1}}\cdots, \quad \text{where all } d_i > 0.
\] (4)

This definition of \(s\) will be kept throughout the rest of this paper.

Let us define an infinite sequence \((s_n)_{n \geq 1 - k}\) of finite words associated with \(s\) as follows:

\[
s_{1-k} = a_2, \quad s_{2-k} = a_3, \quad \ldots, \quad s_{-1} = a_k, \quad s_0 = a_1,
\]

\[
s_n = s_{n-1}^{d_n}s_{n-2}^{d_{n-1}}\cdots s_0^{d_1}a_{n+1}, \quad 1 \leq n \leq k - 1,
\]

\[
s_n = s_{n-1}^{d_n}s_{n-2}^{d_{n-1}}\cdots s_{n-k+1}^{d_{n-k+1}}s_{n-k}, \quad n \geq k.
\] (5)

Clearly, \(s_n\) is a prefix of \(s_{n+1}\) for all \(n \geq 0\) (and hence \(|s_n|_{n \geq 0}\) is a strictly increasing sequence of positive integers).

**Example 3.1** It is well-known that the standard (or characteristic) Sturmian word \(c_\alpha\) of irrational slope \(\alpha = [0; 1 + d_1, d_2, d_3, \ldots]\), \(d_1 \geq 1\), (see [3] for a definition) is the standard episturmian word over \(A = \{a, b\}\) with directive word \(\Delta(c_\alpha) = a^{d_1}b^{d_2}a^{d_3}b^{d_4}a^{d_5}\cdots\). We have \(c_\alpha = \lim_{n \to \infty} s_n\), where \((s_n)_{n \geq 1}\) is the standard sequence associated with \(c_\alpha\), defined by

\[
s_{-1} = b, \quad s_0 = a, \quad s_n = s_{n-1}^{d_n}s_{n-2}, \quad n \geq 1.
\]

This coincides with our definition (5) above when \(k = 2\). Observe that, for all \(n \geq 0\), \(|s_n| = q_n\), where \(q_n\) is the denominator of the \(n\)-th convergent to \([0; 1 + d_1, d_2, d_3, \ldots]\).

For all \(m \geq 1\), let \(L_m := d_1 + d_2 + \cdots + d_m\). Then, writing \(\Delta(c_\alpha) = x_1x_2x_3\cdots\) with each \(x_i \in A\), we have \(x_{n+1} \neq x_n\) if and only if \(n\) is equal to some \(L_m\). One easily deduces that \(S(L_m) = L_{m+1} + 1\) and \(P(L_{m+1} + 1) = L_m\), and it can also be shown that the \(h_{L_m}\) satisfy the same recurrence relation as the \(q_m\). Hence, \(|h_{L_m}| = q_m\), and clearly we have \(h_{L_m} = s_m\) (see Proposition 3.2 below).
Notice that \( s \) has directive word resembling \( \Delta(c_n) \). In fact, as in the 2-letter case, \( s \) can be constructed as the limit, as \( n \) tends to infinity, of the sequence \((s_n)_{n \geq 1}\) given by (5), as shown below.

**Notation:** Hereafter, let \( L_n := d_1 + d_2 + \cdots + d_n \) for each \( n \geq 1 \).

**Proposition 3.2** For any \( n \geq 1 \), \( s_n = h_{L_n} \). Moreover, \( s = \lim_{n \to \infty} s_n \).

**PROOF.** The directive word of \( s \) is given by

\[
\Delta = a_1^{d_1} a_2^{d_2} \cdots a_k^{d_k} d_1 d_2^{d_{k+1}} \cdots d_k = x_1 x_2 x_3 x_4 \cdots, \quad x_i \in A_k.
\]

For \( n \geq 1 \), we have \( x_{n+1} \neq x_n \) (and hence \( h_n \neq h_{n-1} \)) if and only if \( n \) is equal to some \( L_m \). In particular, for any \( m \geq 1 \),

\[
h_{L_m} = h_{L_{m+1}-r}, \quad 1 \leq r \leq d_{m+1}.
\]

Furthermore, it is clear that, for all \( n \geq k \),

\[
P(L_n + 1) = L_{n-k+1},
\]

and \( P(L_n + 1) \) is undefined for \( 1 \leq n \leq k - 1 \).

First we show that \( s_n = h_{L_n} \) for \( 1 \leq n \leq k \). Observe that, for \( 1 \leq n \leq k - 1 \),

\[
h_{L_n} = \Psi_{a_1}^{d_1} \Psi_{a_2}^{d_2} \cdots \Psi_{a_n}^{d_n} (a_{n+1}),
\]

\[
= \Psi_{a_1}^{d_1} \Psi_{a_2}^{d_2} \cdots \Psi_{a_{n-1}}^{d_{n-1}} (a_n a_{n+1})
\]

\[
= h_{L_{n-1}}^{(d_n)} \Psi_{a_1 a_2}^{d_1} \cdots \Psi_{a_{n-1}}^{d_{n-1}} (a_{n+1})
\]

\[
= h_{L_{n-1}}^{(d_n)} h_{L_{n-2}}^{(d_{n-1})} \cdots h_{L_1}^{(d_1)} a_1 a_{n+1}
\]

\[
= h_{L_{n-1}}^{(d_n)} h_{L_{n-2}}^{(d_{n-1})} \cdots h_{L_1}^{(d_1)} a_0 a_{n+1}.
\]

Similarly, since \( h_{L_k} = \Psi_{a_1}^{d_1} \Psi_{a_2}^{d_2} \cdots \Psi_{a_k}^{d_k} (a_1) \), one finds that

\[
h_{L_k} = h_{L_{k-1}}^{d_k} h_{L_{k-2}}^{d_{k-1}} \cdots h_{L_1}^{d_1} a_1.
\]

Thus, we see that the \( s_n \) satisfy the same recurrence relation as the \( h_{L_n} \) for \( 1 \leq n \leq k \). Therefore, since \( h_0 = \mu_0(a_1) = a_1 = s_0 \), we have

\[
s_n = h_{L_n} \quad \text{for all} \ n, \ 1 \leq n \leq k.
\]

Now take \( n \geq k + 1 \). Then, by (3) and (7), we have

\[
h_{L_n} = h_{L_{n-1}} h_{L_{n-2}} \cdots h_{L_{n-k+1}} h_{L_{n-k+1}-1},
\]

and therefore it follows from (6) that

\[
h_{L_n} = h_{L_{n-1}}^{d_n} h_{L_{n-2}}^{d_{n-2}} \cdots h_{L_{n-k+1}}^{d_{n-k+2}} h_{L_{n-k}}.
\]
Whence, since \( s_n = h_{L_n} \) for \( 1 \leq n \leq k \), (9) shows that the \( s_n \) satisfy the same recurrence relation as the \( h_{L_n} \) for \( n \geq k + 1 \). Thus, by virtue of this fact and (8), we have

\[ s_n = h_{L_n} \quad \text{for all } n \geq 1, \]

as required.

The second assertion follows immediately from the first since \( s = \lim_{m \to \infty} h_m \) [15]. □

Accordingly, the words \( (s_n)_{n \geq 1} \) can be viewed as ‘building blocks’ of \( s \).

**Example 3.3** The Tribonacci sequence (or Rauzy word [20]) is the standard episturmian word over \( \{a, b, c\} \) directed by \( (abc)^\omega \). Since all \( d_i = 1 \), we have \( L_n = n \), and hence \( h_n = s_n = s_{n-1}s_{n-2}s_{n-3} \) for all \( n \geq 1 \).

### 3.1 Two special integer sequences

Set \( Q_n := |s_n| \) for all \( n \geq 0 \). Then the integer sequence \( (Q_n)_{n \geq 0} \) is given by:

\[
\begin{align*}
Q_0 &= 1, \\
Q_n &= d_nQ_{n-1} + d_{n-1}Q_{n-2} + \cdots + d_1Q_0 + 1, \quad 1 \leq n \leq k - 1, \\
Q_n &= d_nQ_{n-1} + d_{n-1}Q_{n-2} + \cdots + d_{n+2-k}Q_{n+1-k} + Q_{n-k}, \quad n \geq k.
\end{align*}
\]

Now, define the integer sequence \( (P_n)_{n \geq 0} \) by:

\[
\begin{align*}
P_0 &= 0, \\
P_n &= d_nP_{n-1} + d_{n-1}P_{n-2} + \cdots + d_1P_0 + 1, \quad 1 \leq n \leq k - 1, \\
P_n &= d_nP_{n-1} + d_{n-1}P_{n-2} + \cdots + d_{n+2-k}P_{n+1-k} + P_{n-k}, \quad n \geq k.
\end{align*}
\]

For \( k = 2 \), observe that \( P_n/Q_n \) is the \( n \)-th convergent to the continued fraction expansion \([0; 1 + d_1, d_2, d_3, \ldots] \).

**Proposition 3.4** For all \( n \geq 0 \), \( |s_n|a_1 = Q_n - P_n \).

**PROOF.** Induction on \( n \). □

### 4 Generalized singular words

Recall the standard Sturmian word \( c_\alpha \) of slope \( \alpha = [0; 1+d_1, d_2, d_3, \ldots] \), \( d_1 \geq 1 \). Melançon [19] (also see [4]) introduced the singular words \( (w_n)_{n \geq 1} \) of \( c_\alpha \) defined by

\[ w_n = \begin{cases} 
    as_nb^{-1} & \text{if } n \text{ is odd,} \\
    bs_na^{-1} & \text{if } n \text{ is even,}
\end{cases} \]

with the convention \( w_{-2} = \varepsilon \), \( w_{-1} = a \), \( w_0 = b \). It is easy to show that the set of factors of \( c_\alpha \) of length \( |s_n| \) is given by

\[ \Omega_{|s_n|}(c_\alpha) = \mathcal{C}(s_n) \cup \{w_n\}. \]
Indeed, using (2) and (6), one finds that $s$ standard episturmian word (See [19,4,12] for instance.) Also note that in this 2-letter case $D$ also note that in this 2-letter case $\nu L_n$ ab (resp. $s_n = u_{L_n} ab$) if $n$ is odd (resp. even).

Singular words are profoundly useful in studying properties of factors of $c_s$ (e.g., [4,12,11,18,19,23]). It is for this very reason that we now generalize these words for the standard episturmian word $s$. Firstly, however, we prove some basic results concerning the words $s_n$ and $u_{L_n}$, as detailed in the next section.

4.1 Useful results

For each $n \geq 0$, set $D_n := u_{L_{n+1}}$. Observe that, for any $m \geq 1$,

$$|D_m| = (d_{m+1} \cdot 1)|s_m| + \sum_{j=0}^{m-1} d_{j+1}|s_j|. \quad (10)$$

Indeed, using (2) and (6), one finds that

$$D_m = u_{L_{m+1}} = h_{L_{m+1}} h_{L_{m+1}} \cdots h_{L_0} = h_{L_{m+1}} h_{L_{m+1}} \cdots h_{L_0} = s_{d_{m+1}} \cdot s_{d_{m+1}} \cdots s_{d_1} s_1. \quad (11)$$

Also note that $D_0 = a_1^{d_1-1}$ since $D_0 = u_{d_1} = h_{d_1} h_{d_1} \cdots h_1 h_0 = h_0^{d_1-1}$. For technical reasons, we shall set $D_{-j} := a_{k+1-j}$ and $|D_{-j}| = -1$ for $1 \leq j \leq k$.

**Proposition 4.1** Let $1 \leq i \leq k$. For all $n \geq 1 - k$, $a_i$ is the last letter of $s_n$ if $n \equiv i - 1 \pmod{k}$.

**PROOF.** Since we have $s_{1-k} = a_2$, $s_{2-k} = a_3$, $\ldots$, $s_{-1} = a_k$, $s_0 = a_1$, the result follows immediately from the definition of the words $s_n$ (see (5)). \qed

**Proposition 4.2** For all $n \geq 0$, $s_{n+1} D_{n-k+1} = s_n D_n$, and hence $|D_n| - |D_{n-k+1}| = |s_{n+1}| - |s_n|$.

**PROOF.** The claim holds for $0 \leq n \leq k-2$ since $s_{n+1} D_{n-k+1} = s_n s_{d_{n+1}} \cdots s_{d_1} a_{n+2} a_{n+2} = s_n D_n$, and for $n \geq k-1$, $s_{n+1} D_{n-k+1} = s_{d_{n+1}} s_{d_{n+1}} \cdots s_{d_{n-k+1}} s_{d_{n-k+1}} = s_n D_n$. \qed

**Proposition 4.3** For all $n \geq 1$, $|s_n| > |D_{n-1}|$.

**PROOF.** We proceed by induction on $n$. The result is clearly true for $n = 1$ since $|s_1| = |a_1^{d_1} a_2| = |D_0 a_1 a_2| = |D_0| + 2$. Now assume the result holds for some $n \geq 2$. Then, using Proposition 4.2,

$$|s_{n+1}| = |s_n| + |D_n| - |D_{n-k+1}| > |D_{n-1}| + |D_n| - |D_{n-k+1}| \geq |D_n|,$$

since $|D_{n-k+1}| \leq |D_{n-1}|$. \qed

Recall that the words $D_n$ and $s_n$ are prefixes of $s$ for all $n \in \mathbb{N}$. Thus, according to Proposition 4.3, the palindromes $D_0$, $D_1$, $\ldots$, $D_{n-1}$ are prefixes of $s_n$. In fact,
maximal index $i$ such that $D_i$ is a proper prefix of $s_n$ is $i = n - 1$, which is evident from the following result.

**Proposition 4.4** For all $n \geq 0$, $D_n = s_n^{d_n+1}D_{n-k}$.

**PROOF.** Firstly, $D_0 = a_1^{d_1-1} = s_0^{d_1}a_1^{-1} = s_0^{d_1}D_k$ and, for $1 \leq n \leq k - 1$, we have

$$D_n = s_n^{d_n+1-1}s_{n-1}^{d_{n-1}} \cdots s_0^{d_0} = s_n^{d_n+1-1}s_{n-1}^{a_n^{-1}}D_{n-k}.$$ 

Now take $n \geq k$. Then

$$D_n = s_n^{d_n+1-1}s_{n-1}^{d_{n-1}} \cdots s_{n-k}^{d_{n-k}}s_{n-k}D_{n-k} = s_n^{d_n+1-1}s_nD_{n-k} = s_n^{d_n+1}D_{n-k}.$$ 

**Proposition 4.5** For all $n \geq 0$, $s_n = D_{n-k}\bar{s}_nD_{n-k}^{-1}$.

**PROOF.** We proceed by induction on $n$. For $n = 0$, $D_{-k}\bar{s}_0D_0^{-1} = a_1^{-1}a_1 = a_1 = s_0$. Assume the result holds for some $n \geq 1$. Then, using Proposition 4.2,

$$s_{n+1} = s_nD_nD_{n-k+1}^{-1} = D_{n-k}\bar{s}_nD_{n-k}^{-1}D_{n-k+1}^{-1}.$$ 

Therefore, invoking Proposition 4.4 and (5), for $1 \leq n \leq k - 2$, we have

$$s_{n+1} = D_{n-k}\bar{s}_n(s_n)^{d_n+1}D_{n-k+1}^{-1} = D_{n-k+1}a_{n+2}a_{n+1}^{-1}D_{n-k+1}^{-1}.$$ 

And, for $n \geq k - 1$,

$$s_{n+1} = D_{n-k}\bar{s}_n(s_n)^{d_n+1}D_{n-k+1}^{-1} = D_{n-k+1}\bar{s}_nD_{n-k+1}^{-1},$$

as required. □

**Remark 4.6** This result shows, in particular, that $\bar{s}_n = D_{n-k}^{-1}s_nD_{n-k}$, i.e., $\bar{s}_n$ is the $|D_{n-k}|$-th conjugate of $s_n$ for each $n \geq k$. (For $0 \leq n \leq k - 1$, $\bar{s}_n$ is the $(|s_n| - 1)$-st conjugate of $s_n$ since $\bar{s}_n = a_{n+1}s_na_{n+1}^{-1}$.) The following two corollaries are direct results of the above proposition.

**Corollary 4.7** For any $n \geq 0$, the word $\bar{s}_nD_{n-k}$ is a palindrome. In particular, let $U_n = D_{n-k}$ and $V_n = \bar{s}_nD_{n-k}^{-1}$. Then $s_n = U_nV_n$ is the unique factorization of $s_n$ as a product of two palindromes.

**PROOF.** From Proposition 4.5, we have $s_n = D_{n-k}\bar{s}_nD_{n-k}^{-1} = U_nV_n$, and whence $D_{n-k}^{-1}s_n = \bar{s}_nD_{n-k}^{-1}$. It is therefore clear that $\bar{s}_nD_{n-k}^{-1}$ is a palindrome. The uniqueness
of the factorization \( s_n = U_n V_n \) is immediate from the primitivity of \( s_n \), which follows from Lemma 2.1(i), together with Proposition 3.2. (Recall that since \( s_n \) is primitive, there are exactly \( |s_n| \) different conjugates of \( s_n \).) \( \square \)

**Corollary 4.8** For all \( n \geq 0 \), \( s_n = D_n s_n D_n^{-1} \).

**PROOF.** Propositions 4.4 and 4.5. \( \square \)

**Notation:** Now, for each \( n \in \mathbb{N} \), we define the words \( G_{n,r} \) by

\[
s_n = D_{n-r} G_{n,r}, \quad 1 \leq r \leq k - 1.
\]

**Example 4.9** In the case of Sturmian words \( c_\alpha, r = 1 \) and \( s_n = D_{n-1} G_{n,1} = u_{L_n} G_{n,1} \), for all \( n \geq 1 \), where \( G_{n,1} = ab \) or \( ba \), according to \( n \) odd or even, respectively.

**Example 4.10** Recall that when all \( d_i = 1 \), \( s \) is the Tribonacci sequence over \( \{a_1, a_2, a_3\} \equiv \{a, b, c\} \). For \( n = 4 \), we have \( s_n = s_4 = abacabaabacab, D_2 = aba, D_3 = abacab \), and hence

\[
G_{4,1} = abacab \text{ and } G_{4,2} = cabaabacab.
\]

**Note.** Since \( D_{n-r} = a_{k+1+n-r}^{-1} \) for \( 0 \leq n < r \), we also set

\[
G_{n,r} = a_{k+1+n-r}s_n \quad \text{for } 0 \leq n < r. \quad (12)
\]

**Proposition 4.11** For all \( n \geq 1 \), \( s_n s_{n-1} G_{n-1,k-1}^{-1} = s_{n-1} s_n G_{n-1}^{-1} \).

**PROOF.** It is easily checked that the result holds for \( 1 \leq n \leq k - 1 \), since

\[
s_n s_{n-1} G_{n-1,k-1}^{-1} = s_n D_{n-k} = s_n a_{n+1}^{-1},
\]

and

\[
s_{n-1} s_n G_{n-1}^{-1} = s_{n-1} D_{n-1} = s_n a_{n+1}^{-1}.
\]

Now take \( n \geq k \). Then, using (11), we have

\[
s_n s_{n-1} G_{n-1,k-1}^{-1} = s_n D_{n-k}
\]

\[
= (s_n s_{n-1}^{d_{n-1}} s_n^{d_{n-2}} s_{n-k+2} s_n^{d_{n-k}})^{d_{n-k+1}} s_n^{d_{n-k}} s_{n-k-1}^{d_{n-k}} \cdots s_0^{d_1} d_1
\]

\[
= s_{n-1} (s_n s_{n-1}^{d_{n-1}} s_n^{d_{n-2}} s_{n-k+2} s_n^{d_{n-k}})^{d_{n-k+1}} s_n^{d_{n-k}} s_{n-k-1}^{d_{n-k}} \cdots s_0^{d_1} d_1
\]

\[
= s_{n-1} D_{n-1}
\]

\[
= s_{n-1} s_n G_{n-1}^{-1}. \quad \square
\]

**Remark 4.12** Recall Example 3.1. For \( c_\alpha \) with \( \alpha = [0; 1 + d_1, d_2, d_3 \ldots] \), it is well-known that, for all \( n \geq 2 \), \( s_n s_{n-1} (xy)^{-1} = s_{n-1} s_n (yx)^{-1} \), where \( x, y \in \{a, b\}, x \neq y \), and \( xy \prec_\alpha s_{n-1} \). This is known as the Near-Commutative Property of the words \( s_n \) and \( s_{n-1} \). Because \( s_n s_{n-1} (xy)^{-1} = s_n D_{n-2} \) and \( s_{n-1} s_n (yx)^{-1} = s_{n-1} D_{n-1} \), Proposition 4.11 is merely an extension of this property to standard episturmian words \( s \). It is also worthwhile noting that Proposition 4.11 shows that \( s_n \) is a prefix of \( s_{n-1} s_n \).
Hereafter, we set $d_{-j} = 0$ for $j \geq 0$.

Proposition 4.2 implies that $|s_{n+1}| - |D_n| = |s_n| - |D_{n-k+1}|$, and hence $|G_{n+1,1}| = |G_{n,k-1}|$. In fact, we have the following:

**Proposition 4.13**  For all $n \geq 1$, $G_{n,1} = \tilde{G}_{n-1,k-1}$.

**PROOF.** One can write

$$G_{n,1} = D_{n-1}^{-1}s_n = D_{n-1}^{-1}s_{n-1}^d s_{n-2}^{d_{n-1}} \cdots s_{n-k+1}^{d_{n-k+2}} s_{n-k}^{d_{n-k+1}}$$

where $s_{n-r+1}$ is a prefix of $s_n$. Hence, one can write

$$G_{n,r} = G_{n-r+1,1}s_{n-r+1}^{-1}s_n = \tilde{G}_{n-r,k-1}s_{n-r+1}^{-1}s_n$$

by applying Proposition 4.13.

Now, one easily deduces from Proposition 4.1 that $a_m \prec_p \tilde{s}_{n-r}$ if $n \equiv m+r-1 \pmod{k}$, and thus $a_m \prec_p \tilde{G}_{n-r,k-1} \prec_p \tilde{s}_{n-r}$ if $n \equiv m+r-1 \pmod{k}$.

(ii) For $0 \leq n < r$, $G_{n,r} = a_{k+1+n-r}s_n$ and, for each $n \geq r$, we have $G_{n,r} \prec_s s_n$. Hence, $a_m \prec_s G_{n,r}$ if $n \equiv m-1 \pmod{k}$, by Proposition 4.1. □

### 4.2 Singular $n$-words of the $r$-th kind

By definition of the words $(s_n)_{n \geq 1-k}$ (see (5)) and the fact that $s = \lim_{n \to \infty}s_n$, one deduces that, for any $n \geq 0$, $s$ can be written as a concatenation of blocks of the form $s_n$, $s_{n-1}$, …, $s_{n-k+1}$, i.e.,

$$s = [(s_{n}^{d_{n+1}}s_{n-1}^{d_{n}} \cdots s_{n-k+2}^{d_{n-k+1}}s_{n-k+1})^{d_{n+2}}s_{n}^{d_{n+1}} \cdots s_{n-k+3}^{d_{n-k+4}}s_{n-k+2}^{d_{n-k+3}}]^{d_{n+3}}$$

$$\cdots (s_{n}^{d_{n+1}}s_{n-1}^{d_{n}} \cdots s_{n-k+2}^{d_{n-k+1}}s_{n-k+1})^{d_{n+2}}s_{n}^{d_{n+1}} \cdots s_{n-k+4}^{d_{n-k+5}}s_{n-k+3}^{d_{n-k+4}} \cdots .$$

We shall call this unique decomposition the $n$-partition of $s$. This will be a useful tool in our subsequent analysis of powers of words occurring in $s$ (Section 6, to follow).
Note. Uniqueness of the factorization (14) is proved inductively. The initial case $n = 0$ is trivial. For $n \geq 1$, the factorization of $s_n$ in terms of the $s_{n-i}$ given by (5) is unique because the $s_{n-i}$ end with different letters (by Proposition 4.1). So it is clear that every $(n+1)$-partition of $s$ gives rise to an $n$-partition, in which the positions of $s_{n-k+1}$ blocks uniquely determine the positions of $s_{n+1}$ blocks in the original $(n+1)$-partition (since $s_{n+1} = s_{n+1}^{d_{n+1}}s_{n-1}^{d_n} \cdots s_{n-k+3}s_{n-k+1}$). Accordingly, uniqueness of the $n$-partition implies uniqueness of the $(n+1)$-partition.

Remark 4.15 Since each factor of $s$ has exactly $k$ different return words (see Section 2.2.2), two consecutive $s_{n+1-i}$ blocks ($1 \leq i \leq k$) of the $n$-partition are separated by a word $V$, of which there are $k$ different possibilities. From now on, it is advisable to keep this observation in mind.

Lemma 4.16 Let $1 \leq r \leq k - 1$. For any $n \in \mathbb{N}^+$, a factor $u$ of length $|s_n|$ of $s$ is a factor of at least one of the following words:

- $C_{j}(s_n)$, $0 \leq j \leq |s_n| - 1$;
- $s_{n-r+1}^{d_{n-r+1}} \cdots s_{n-k+2}s_{n-k}s_{n-1}^{d_n} \cdots s_{n-r+1}s_{n-r}^r$ if $n \geq r$;
- $a_{n+1}s_{n+1}a_{n-r+1}^r$ if $n < r$.

Note. The word $s_{n-r+1}^{d_{n-r+1}} \cdots s_{n-k+2}s_{n-k}s_{n-1}^{d_n} \cdots s_{n-r+1}s_{n-r}^r$ $(1 \leq r \leq k - 1)$ has length $|s_n|$.

PROOF. In the $n$-partition of $s$, one observes that two consecutive $s_n$ blocks make the following $k$ different appearances:

$$s_n s_n \quad \text{and} \quad \underbrace{s_n s_{n-1}^{d_n} \cdots s_{n-r+1}s_{n-r}^r}_{(s)} , \quad 1 \leq r \leq k - 1.$$ 

Evidently, any factor of length $|s_n|$ of $s$ is a factor of one of the above $k$ different words.

Now, factors of length $|s_n|$ of $s_n s_n$ are simply conjugates of $s_n$. Furthermore, for $n \geq r$, the first $|s_{n-r+1}^{d_{n-r+1}} \cdots s_{n-k+2}s_{n-k}s_{n-1}^{d_n} \cdots s_{n-r+1}s_{n-r}^r|$ factors of length $|s_n|$ of $(*)$ are again just conjugates of $s_n$.

The remaining factors of length $|s_n|$ of $(*)$ are factors of

$$s_{n-r+1}^{d_{n-r+1}} \cdots s_{n-k+2}s_{n-k}s_{n-1}^{d_n} \cdots s_{n-r+1}s_{n-r}^r.$$ 

For $n < r$, one can write $(*)$ as $s_n s_{n-1}^{d_n} \cdots a_{n-r+k+1}s_n = s_n s_n a_{n+1}^{-1} a_{n-r+k+1}s_n$, of which the first $|s_n| - 1$ factors of length $|s_n|$ are conjugates of $s_n$, and the other factors of length $|s_n|$ are factors of $a_{n+1}s_n a_{n+1}^{-1} a_{n-r+k+1}s_n$. □

Lemma 4.17 For any $n \geq 1$, $\sum_{j=1}^{k-1} |D_{n-j}| = |s_n| - k$.

PROOF. Induction on $n$ and Proposition 4.2. □

Lemma 4.18 Let $1 \leq r \leq k - 1$. For any $n \geq r$, we have

$$s_{n-r}^{d_{n-r+1}} \cdots s_{n-k+1}s_{n-k}^{d_n} \cdots s_{n-r+1}s_{n-r} = D_{n-r} \bar{G}_{n,r},$$

and for $1 \leq n < r$, $a_{n+1}s_n a_{n+1}^{-1} a_{n-r+k+1} = \bar{G}_{n,r}$.  

12
PROOF. For $1 \leq n < r$, one can write $\tilde{G}_{n,r} = \tilde{s}_n a_{n-r+k+1} = a_{n+1} s_n a_{n+1} a_{n-r+k+1}$, by Remark 4.6. Now take $n \geq r$. Then, using Corollary 4.8 and Proposition 4.4,

$$D_{n-r} \tilde{G}_{n,r} = D_{n-r} \tilde{s}_n D_{n-r}^{-1}$$

$$= D_{n-r} D_{n-r}^{-1} s_n D_{n-r}^{-1}$$

$$= D_{n-r} D_{n-r}^{-1} s_n D_{n-r}^{-1} s_{n-r+2}$$

$$= D_{n-r} D_{n-r}^{-1} s_n s_{n-r+2}$$

$$= s_{n-r+1} s_{n-r+2} s_{n-r+1} s_{n-r+2}.$$  

\[\Box\]

Whence, it is now plain to see that each word $\tilde{G}_{n,r} s_n = \tilde{G}_{n,r} D_{n-r} G_{n,r}$ is a factor of $s$. We will now partition the set of factors of length $|s_n|$ of $s$ into $k$ disjoint classes.

**Theorem 4.19** Let $1 \leq r \leq k - 1$. For any $n \in \mathbb{N}^+$, the set of factors of length $|s_n|$ of $s$ can be partitioned into the following $k$ disjoint classes:

- $\Omega^0_n := \mathcal{C}(s_n) = \{C_j(s_n) : 0 \leq j \leq |s_n| - 1\}$;
- $\Omega^r_n := \{w \in \mathcal{A}_n^* : |w| = |s_n| \text{ and } w < x^{-1} \tilde{G}_{n,r} D_{n-r} G_{n,r} x^{-1}\}$, where $x$ is the last letter of $G_{n,r}$.

That is, $\Omega_{s_n}(s) = \Omega^0_n \dot{\cup} \Omega^1_n \dot{\cup} \cdots \dot{\cup} \Omega^{k-1}_n$.

**PROOF.** First observe that Lemma 2.1(i), coupled with Proposition 3.2, implies that each $s_n$ is primitive, and hence $|\Omega^0_n| = |s_n|$. Also note that $\tilde{\Omega}^0_n := \{\tilde{w} : w \in \Omega^0_n\} = \Omega^0_n$; i.e., $\Omega^0_n$ is closed under reversal, which is deduced from Corollary 4.7.

We shall use Lemma 4.16 to partition $\Omega_{s_n}(s)$ into $k$ disjoint classes; the first being $\Omega^0_n = \mathcal{C}(s_n)$. Now consider the factors of length $|s_n|$ of the words

$$s_{n-r+1} s_{n-k+1} s_{n-k} s_{n-1} \cdots s_{n-r+1} s_{n-r} s_n \quad (n \geq r). \quad (15)$$

Since (15) can be written as $D_{n-r} \tilde{G}_{n,r} D_{n-r} G_{n,r}$ (by Lemma 4.18), the first $|D_{n-r}| + 1$ factors of length $|s_n| = |D_{n-r} G_{n,r}|$ are conjugates of $\tilde{s}_n$ (and hence of $s_n$) and the last factor is just $s_n$. Hence, all other factors of length $|s_n|$ of (15) are factors of $x^{-1} \tilde{G}_{n,r} D_{n-r} G_{n,r} x^{-1}$, where $x$ is the last letter of $G_{n,r}$. Moreover, $D_{n-r}$ appears exactly once (and at a different position) in each word in

$$\Omega^r_n := \{w \in \mathcal{A}_k^* : |w| = |s_n| \text{ and } w < x^{-1} \tilde{G}_{n,r} D_{n-r} G_{n,r} x^{-1}\};$$

whence $|\Omega^r_n| = |G_{n,r}| - 1$. Since the letter just before $D_{n-r}$ (equivalently, the last letter of $\tilde{G}_{n,r}$) in the word $x^{-1} \tilde{G}_{n,r} D_{n-r} G_{n,r} x^{-1}$ is different for each $r \in [1, k - 1]$, it is evident that $\Omega^0_n, \Omega^1_n, \ldots, \Omega^{k-1}_n$ are pairwise disjoint.

Now, for $1 \leq n < r$, other than words in the sets $\Omega^0_n, \Omega^1_n, \ldots, \Omega^{k-1}_n$, the remaining factors of length $|s_n|$ of $s$ are factors of

$$a_{n+1} s_n a_{n+1} a_{n-r+k+1} s_n = \tilde{s}_n a_{n-r+k+1} s_n = a_{n+1} s_n a_{n+1} a_{n-r+k+1} s_n \quad (16)$$

(see Lemma 4.16). The first factor of length $|s_n|$ of the word (16) is $\tilde{s}_n$ (i.e., the $(|s_n| - 1)$-st conjugate of $s_n$) and the last is just $s_n$. All other factors of length $|s_n|$ of (16) are factors
of

\[a_{n+1}^{-1} \bar{s}_n a_{n-r+k+1} s_n a_n^{-1} = a_{n+1}^{-1} \bar{G}_{n,r} D_{n-r} G_{n,r} a_{n+1}^{-1} - a_{n+1}^{-1} \bar{G}_{n,r} D_{n-r} G_{n,r} a_{n+1}^{-1} \cdot \]

Defining \( \Omega_n^r := \{ w \in \mathcal{A}_k^n \mid |w| = |s_n| \text{ and } w < a_{n+1}^{-1} \bar{G}_{n,r} D_{n-r} G_{n,r} a_{n+1}^{-1} \} \), one can check that \( |\Omega_n^r| = |G_{n,r}| - 1 \) and \( \Omega_n^0, \Omega_n^1, \ldots, \Omega_n^{n-1} \) are pairwise disjoint.

It remains to show \( \bigcup_{j=0}^{k-1} \Omega_n^j = \Omega_{[s_n]}(s) \) for all \( n \geq 1 \). Indeed, \( |\Omega_{[s_n]}(s)| = (k - 1)|s_n| + 1 \) (from the complexity function for \( k \)-strict standard episturmian words), and we have

\[
\sum_{j=0}^{k-1} |\Omega_n^j| = |s_n| + \sum_{j=1}^{k-1} (|G_{n,j}| - 1) = |s_n| + \sum_{j=1}^{k-1} (|s_n| - |D_{n-j}| - 1)
\]

\[
= k|s_n| - k + 1 - \sum_{j=1}^{k-1} |D_{n-j}|
\]

\[
= k|s_n| - k + 1 - (|s_n| - k) \quad \text{(by Lemma 4.17)}
\]

\[
= (k - 1)|s_n| + 1. \quad \square
\]

Let us remark that the sets \( \Omega_n^r \) are closed under reversal since \( x^{-1} \bar{G}_{n,r} D_{n-r} G_{n,r} x^{-1} \) is a palindrome; that is \( \Omega_n^r := \{ \bar{w} : w \in \Omega_n^r \} = \Omega_n^r \). We shall call the factors of \( s \) in \( \Omega_n^r \) the singular \( n \)-words of the \( r \)-th kind. Such words will play a key role in our study of powers of words occurring in \( s \).

Evidently, for Sturmian words \( c_\alpha \), \( \Omega_n^1 = \{ w_n \} \) and we have \( \Omega_{[s_n]}(c_\alpha) = \mathcal{C}(s_n) \cup \{ w_n \} \), as before.

5 Index

A word of the form \( w = (uv)^n u \) is written as \( w = z^r \), where \( z = uv \) and \( r := n + |u|/|z| \). The rational number \( r \) is called the exponent of \( z \), and \( w \) is said to be a fractional power.

Now suppose \( x \) is an infinite word. For any \( w < x \), the index of \( w \) in \( x \) is given by the number

\[ \text{ind}(w) = \sup \{ r \in \mathbb{Q} \mid w^r < x \}, \]

if such a number exists; otherwise, \( w \) is said to have infinite index in \( x \). Furthermore, the greatest number \( r \) such that \( w^r \) is a prefix of \( x \) is called the prefix index of \( w \) in \( x \). Obviously, the prefix index is zero if the first letter of \( w \) differs from that of \( x \), and it is infinite if and only if \( x \) is purely periodic.

For all \( n \geq 0 \), define the words

\[ t_n := D_{n-k+1} G_{n+1,k-1} \quad \text{and} \quad r_n := s_{n-1} D_{n-1} = s_{n-1}^{d_{n-1}} s_{n-2}^{d_{n-2}} \cdots s_1^{d_1} s_0^{d_0}. \]

Note. By convention, \( r_0 = a_k a_k^{-1} = \varepsilon \), and \( t_n = a_{n+1}^{-1} a_{n+3} s_{n+1} \) for \( 0 \leq n \leq k - 2 \).

The next two results extend those of Berstel [2].

**Lemma 5.1** For all \( n \geq 1 \), the word \( r_{n+1} \) is the greatest fractional power of \( s_n \) that is a prefix of \( s \), and the prefix index of \( s_n \) in \( s \) is \( 1 + d_{n+1} + |D_{n-k}|/|s_n| \).
PROOF. First we take \( n \geq k \). Observe that the longest common prefix shared by the words \( s_n \) and \( t_n \) is

\[
D_{n-k+1} = s_{n-k+1}^{d_{n-k+2}^{-1}}r_{n-k+1},
\]

since

\[
s_n = D_{n-k+1}G_{n,k-1} \quad \text{and} \quad t_n = D_{n-k+1}G_{n+1,k-1} = D_{n-k+1}\tilde{G}_{n+2,1}
\]

where \( G_{n,k-1} \) and \( G_{n+1,k-1} \) do not share a common first letter, by Proposition 4.14. Clearly, \( s_{n+1}s_n \prec_p s \), and we have

\[
s_{n+1}s_n = s_n^{d_{n+1}}s_{n-1}^{d_n} \cdots s_{n-k+3}^{d_{n-k+2}}s_{n-k+2}s_{n-k+1}s_n
\]

\[
= s_n^{d_{n+1}+1}(s_{n-k}^{d_{n-k+2}^{-1}}s_{n-k})^{-1}D_{n-k+1}G_{n,k-1}
\]

\[
= s_n^{d_{n+1}+1}(s_{n-k}^{d_{n-k+1}^{-1}}s_{n-k-1} \cdots s_{1}^{d_1}d_0)G_{n,k-1} \quad \text{(by (11))}
\]

\[
= s_n^{d_{n+1}+1}D_{n-k}G_{n,k-1}
\]

\[
= s_n^{d_{n+1}+1}t_{n-1}.
\]

Hence, \( s_n^{d_{n+1}+1} \) is a prefix of \( s \). Also observe that the longest common prefix of \( t_{n-1} \) and \( s_n \) is \( D_{n-k} \) since

\[
t_{n-1} = D_{n-k}G_{n,k-1} \quad \text{and} \quad s_n = D_{n-k}G_{n,k}
\]

where \( G_{n,k-1} \) and \( G_{n,k} \) have different first letters, by Proposition 4.14. Further, from (18) and Proposition 4.4, we have

\[
s_{n+1}s_n = s_n^{d_{n+1}+1}t_{n-1} = s_n^{d_{n+1}+1}D_{n-k}G_{n,k-1} = s_nD_{n-k}G_{n,k-1} = r_{n+1}G_{n,k-1}.
\]

Thus, the greatest fractional power of \( s_n \) that is a prefix of \( s \) is \( r_{n+1} \) with

\[
|r_{n+1}| = |s_nD_n| = |s_n^{d_{n+1}+1}D_{n-k}| = (d_{n+1}+1)|s_n| + |D_{n-k}|;
\]

whence the prefix index of \( s_n \) in \( s \) is \( 1 + d_{n+1} + |D_{n-k}| / |s_n| \).

Similarly, for \( 1 \leq n \leq k-1 \), we have

\[
s_{n+1}s_n = s_n^{d_{n+1}}s_{n-1}^{d_n} \cdots s_0^{d_0}a_{n+2}s_n
\]

\[
= s_n^{d_{n+1}+1}a_{n+1}^{-1}a_{n+2}s_n
\]

\[
= s_n^{d_{n+1}+1}D_{n-k}G_{n,k-1}
\]

\[
= r_{n+1}G_{n,k-1}.
\]

Therefore, the greatest fractional power of \( s_n \) (= \( D_{n-k}G_{n-1,k-1} \)) that is a prefix of \( s_{n+1}s_n \prec_p s \) is \( r_{n+1} \), where \( |r_{n+1}| = (d_{n+1}+1)|s_n| + |D_{n-k}| = (d_{n+1}+1)|s_n| - 1 \). That is, the prefix index of \( s_n \) in \( s \) is \( 1 + d_{n+1} - 1 / |s_n| \) for \( 1 \leq n \leq k-1 \). □

**Lemma 5.2** For all \( n \geq 1 \), the index of \( s_n \) as a factor of \( s \) is at least \( 2 + d_{n+1} + |D_{n-k}| / |s_n| \), and hence \( s \) contains cubes.

We will show later that the index of \( s_n \) is exactly \( 2 + d_{n+1} + |D_{n-k}| / |s_n| \).
PROOF. Setting \( e = 1 + d_{n+1} + |D_{n-k}|/|s_n| \), we will show that \( s_{n+k+2} \) contains a power of \( s_n \) of exponent \( 1 + e \). Certainly, using Proposition 4.11, one can write

\[
\begin{align*}
\frac{d_{n,k} s_{n+k+2}^{-1}}{d_{n,k+1} s_{n+k+1} s_{n+k} D_{n+k}} & = \frac{d_{n,k+1} s_{n+k+1} G_{n+k+1, k+1} G_{n+k, k-1} D_{n+k}}{d_{n,k+1} s_{n+k} D_{n+k} G_{n+k, k-1} D_{n+k}} \\
\end{align*}
\]

The suffix \( s_{n+k} D_{n+k} G_{n+k, k-1} D_{n+k} D_{n+2} \) contains the exponent \( 1 + e \) of \( s_n \). More precisely, \( s_{n+k} \) ends with \( s_n \), and \( D_{n+k} G_{n+k, k-1} \) shares a prefix of length \( |D_{n+k}| \) with \( s_{n+k+1} \). Thus, since \( r_{n+1} \) is a prefix of \( s \) of length \( |r_{n+1}| = |s_n| + |D_n| < |D_{n+k}| \), we have \( s_n r_{n+1} < s_{n+k} D_{n+k} < s_{n+k+2} \). □

6 Powers occurring in \( s \)

For each \( m, l \in \mathbb{N} \) with \( l \geq 2 \), let us define the following set of words:

\[
\mathcal{P}(m; l) := \{ w \in \mathcal{A}_k^* : |w| = m, \ w^l \prec s \},
\]

where \( s \) is the \( k \)-strict standard episturmian word over \( \mathcal{A}_k = \{a_1, a_2, \ldots, a_k\} \) with directive word \( \Delta \) given by (4). Also, let \( p(m; l) := |\mathcal{P}(m; l)| \).

The next theorem is a generalization of Theorem 1 in [6]. It gives all the lengths \( m \) such that there is a non-trivial power of a word of length \( m \) in \( s \). Firstly, let us define the following \( k \) sets of lengths for fixed \( n \in \mathbb{N}^+ \):

\[
\begin{align*}
\mathcal{D}_1(n) := \{ r |s_n| : 1 \leq r \leq d_{n-1} \}, \\
\mathcal{D}_i(n) := \{ s_n^r s_{n-1} \cdots s_{n+3-i} s_{n+1-i} : 1 \leq r \leq d_{n+1}, \ 2 \leq i \leq k-1 \}, \\
\mathcal{D}_k(n) := \{ s_n^r s_{n-1} \cdots s_{n+3-k} s_{n+1-k} : 1 \leq r \leq d_{n+1} - 1 \}.
\end{align*}
\]

Theorem 6.1. Let \( m, n \in \mathbb{N}^+ \) be such that \( |s_n| \leq m < |s_{n+1}| \) and suppose \( m \not\in \bigcup_{i=1}^k \mathcal{D}_i(n) \). Then \( p(m; l) = 0 \) for all \( l \geq 2 \).

Remark 6.2. Put simply, the above theorem states that if a word \( w \) has a non-trivial integer power in \( s \), then \( |w| \in \bigcup_{i=1}^k \mathcal{D}_i(n) \) for some \( n \). For instance, if \( k = 3 \), we have

\[
\bigcup_{i=1}^3 \mathcal{D}_i(n) = \{|s_n|^r, s_n^r s_{n-1} : 1 \leq r \leq d_{n+1} \} \cup \{|s_n^r s_{n-1} s_{n-2} : 1 \leq r \leq d_{n+1} - 1 \}.
\]

In the particular case of the Tribonacci sequence, Theorem 6.1 implies that if \( w^l \) is a factor, then \( |w| \in \{||s_n|, |s_n| + |s_{n-1}| \} \) for some \( n \), where the lengths \( (|s_i|)_{i \geq 0} \) are the Tribonacci numbers: \( T_0 = 1, T_1 = 2, T_2 = 4, T_i = T_{i-1} + T_{i-2} + T_{i-3}, i \geq 3 \).
The proof of Theorem 6.1 requires several lemmas. Let us first observe that in the n-partition of s (see (14)) to the left of each s_n block, there is an s_{n+1-j} block for some j \in [1, k]. Also note that each s_{n+1-j} is a prefix of s_n. Furthermore, to the left of each s_{n+1-i} block is another s_{n+1-i} block or an s_{n+2-i} block, for each i \in [2, k].

Lemma 6.3 Let n \in \mathbb{N}^+. Consider a word w \prec s of the form w = us_nv for some words u, v \in A_n^*, u \neq \varepsilon.

(i) If w = u_1 u_2, where u_1 \prec s s_{n+1-i} for some i \in [1, k] and u_2 \prec_p s_n, then u_1 = u.

(ii) If w = u_1 s_{n+1-i} u_2 for some i \in [2, k], where u_1 \prec_s s_{n+2-i} and u_2 \prec_p s_n, then u_1 = u or u_1 s_{n+1-i} = u.

(iii) If w = u_1 s_{n+1-i} u_2 for some i \in [2, k-1], where u_1 \prec_s s_{n+1-i} and u_2 \prec_p s_n, then u_1 = u or u_1 s_{n+1-i} = u.

**PROOF.** (i) Other than the case when u_1 = u, u_2 = s_n and v = \varepsilon, the only other possibility is:

\[
\begin{array}{ccc}
  u & s_n & v \\
  u_1 & & \end{array}
\]

(Note that u_1 \prec_s s_{n+1-i} for some i \in [1, k], and therefore |u_1| \leq |s_{n+1-i}| \leq |s_n|.)

In this case, using the figure, we write u_1 = uu', s_n = u'v', u_2 = v'v for some u', v' (u' \neq \varepsilon). As v' is a prefix of s_n, we have s_n = v'v'' for some v'', thus u' and v'' are conjugate. So there exist e, f and non-negative integers p, q such that v' = (ef)^p e, u' = (ef)^q, and v'' = (ef)^q with ef primitive. Hence s_n = (ef)^{p+q} e. As u' is a suffix of s_{n+1-i} which is a prefix of s_n, we must have, by primitivity of ef, s_{n+1-i} = (ef)^r, and then r = 1. But u' is non-empty, so u' = ef = s_{n+1-i}, and it follows that u = \varepsilon; a contradiction.

(ii) Let i \in [2, k] be a fixed integer. Since u_1 \prec_s s_{n+2-i} and u_2 \prec_p s_n, we have |u_1| \leq |s_{n+2-i}| \leq |s_n| and |u_2| \leq |s_n|. Accordingly, there exist only three possibilities (other than u_1 = u or u_1 s_{n+1-i} = u), and these are:

\[
\begin{array}{ccc}
  u & s_n & v \\
  u_1 & s_{n+1-i} & u_2 \\
\end{array}
\]

or

\[
\begin{array}{ccc}
  u & s_n & v \\
  u_1 & s_{n+1-i} & u_2 \\
\end{array}
\]

or

\[
\begin{array}{ccc}
  u & s_n & v \\
  u_1 & s_{n+1-i} & u_2 \\
\end{array}
\]

In the first instance, us_nv = u_1 s_{n+1-i} u_2 = uu's_{n+1-i}v'v, where u_2 = v'v with v' \prec_s s_n and u_1 = uu' with u' \prec_p s_n. That is, s_n = u's_{n+1-i}v', where u' \prec_p s_n and v' \prec_s s_n, and u_1 = uu' \prec_s s_{n+2-i}. Therefore, u' \prec_s s_{n+2-i}, and hence the word s_{n+1-i} must be preceded
by the last letter of \( s_{n+2-i} \). However, since \( u' \) is also a prefix of \( s_n = s_{n-1}^{d_{n-1}} \cdots s_{n+1-k}^{d_{n+1-k-1}} s_{n-k}^{d_{n-k}} \), where \( s_{n-1}, \ldots, s_{n+1-k}, s_{n-k} \) do not share a common last letter (by Proposition 4.1), one is forced to presume that \( u' = s_{n-1}^{d_{n-1}} s_{n-2}^{d_{n-2}} \cdots s_{n+2-i}^{d_{n+2-i}} \) (resp. \( u' = \varepsilon \)) when \( i \in [3, k] \) (resp. \( i = 2 \)). This contradicts the fact that \( 1 \leq |u'| < |s_{n+2-i}| \).

In the second instance, we have \( u s_n v = u_1 s_{n+1-i} u_2 = u u' s_{n+1-i} u_2 \), where \( u_1 = uu' \) with \( u' \prec_p s_n \) and \( u_2 \prec_p s_n \). Consider the word \( w' := w u_2^{-1} = u s_n v u_2^{-1} = u s_n v' \), i.e., \( u' = u s_n v' = u_1 s_{n+1-i} v \), where \( v' \prec_p v \) and \( v' \prec s s_{n+1-i} \). Since \( u_1 \prec_s s s_{n+2-i} \) and \( s_{n+1-i} \prec_p s_n \), it follows from assertion (i) that \( u_1 = u \) and hence \( s_{n+1-i} = s v' \), which is absurd unless \( i = 1 \) and \( v' = \varepsilon \). But \( i > 1 \), so this situation is impossible.

Lastly, \( u s_n v = u_1 s_{n+1-i} u_2 = u_1 s_{n+1-i} v' v \), where \( u_2 = v' v \prec_p s_n \) with \( v' \prec s s_n \). Consider the word \( w' := u_1^{-1} w = u_1^{-1} u s_n v = u' s_n v \), i.e., \( w' = u s_n v = s_{n+1-i} u_2 \), where \( u' \prec s u \) and \( u' \prec_p s_{n+1-i} \). Since \( u_2 \prec_p s_n \), one obtains, as an immediate consequence of claim (i), \( u' = s_{n+1-i} \), \( v = \varepsilon \), and hence \( u = u_1 s_{n+1-i} \); a contradiction since \( |u_1 s_{n+1-i}| > |u_1| \).

One can prove assertion (iii) in a similar manner. \( \square \)

**Lemma 6.4** Let \( c \in A_k \) and \( n \in \mathbb{N} \) be fixed. Consider an occurrence of \( c s_n \) in \( s \). Then the letter \( c \) is the last letter of a block \( s_{n+1-i} \) of the \( n \)-partition of \( s \), for some \( i \in [1, k] \), and the integer \( i \) (equiv. the block \( s_{n+1-i} \)) is uniquely determined by \( c \). In particular, in every occurrence of \( s_{n+1-i} s_n \) in \( s \), the word \( s_{n+1-i} \) is a block in the \( n \)-partition of \( s \).

That is, occurrences of words \( w \) containing \( c s_n \) \((c \in A_k)\) must be aligned to the \( n \)-partition of \( s \).

**Proof.** This assertion follows from Lemma 6.3. The case \( n = 0 \) is trivial, and for \( n \geq 1 \), observe from Lemma 6.3 that the given \( s_n \) is either an \( s_n \) block in the \( n \)-partition of \( s \) or has an \( s_{n+1-j} \) block of the \( n \)-partition as a prefix, for some \( j \in [2, k] \). In the first case, to the left of \( s_n \) there is an \( s_{n+1-l} \) block, for some \( l \in [1, k] \). Whereas, in the second case, there is an \( s_{n+1-j} \) or \( s_{n+2-j} \) block (of the \( n \)-partition) to the left of \( s_n \). That is, \( s_n \) is preceded by an \( s_{n+1-i} \) block of the \( n \)-partition for some \( i \in [1, k] \). Since the last letters of \( s_n, s_{n-1}, \ldots, s_{n+1-k} \) are mutually distinct (by Proposition 4.1), it is clear that \( i \) (and hence \( s_{n+1-i} \)) is uniquely determined by the letter \( c \). \( \square \)

We can now determine the exact index of \( s_n \) in \( s \).

**Lemma 6.5** For any \( n \geq 1 \), the word of maximal length that is a factor of both \( s \) and the infinite sequence \( (s_n)^\omega := s_n s_n s_n \cdots \) is \( s_{n+1}^{d_{n+1}} \cdots D_{n-k} \), i.e., \( \text{ind}(s_n) = 2 + d_{n+1} + |D_{n-k}| / |s_n| \).

**Proof.** According to Lemma 6.4, any occurrence of \( s_n^p \) \((p \geq 2)\) must be aligned to the \( n \)-partition of \( s \). By inspection of the \( n \)-partition of \( s \) (see (14)), it is not hard to see that between two successive \( s_{n+1-k} \) blocks there is a word possessing one of the following \( k \) forms:

\[
\begin{align*}
  & s_n^{d_n} s_{n+1-k}^{d_{n+1-k}} \cdots D_{n-k}^{d_{n-k}+1} s_{n-k}^{d_{n-k}}, & q \in \{0, 1\}, \\
  & s_n^{d_n+1} s_{n+1-k}^{d_{n+1-k}} \cdots s_{n+2-i}^{d_{n+2-i}} s_{n+1-i}^{d_{n+1-i}} s_n^{d_n} \cdots s_{n+2-k}^{d_{n+2-k}}, & i \in [2, k-1].
\end{align*}
\]
Thus, the alignment property implies that an occurrence of \( s_n^p \) \((p \geq 2)\) is either a prefix of
\[
s_n^{d_{n+1} + r} s_n^{d_n} s_{n+1}^{-1} \cdots s_{n+3-k} s_{n+1-k} z_1
\]
for some integer \( r \leq 1 \) and suitable \( z_1 \), or a prefix of
\[
s_n^{d_{n+1} + r} s_n^{d_{n+1} - i} s_{n+1}^{-1-i} s_n^{d_{n} + d_{n+1} - i-k} s_{n+1-k} z_2
\]
for some \( i \in [2, k-1] \), \( r \leq d_{n+1} \) and suitable \( z_2 \).

Now, suppose \( s_n^p \) is a prefix of the word (19). Since \( s_{n-1}^{d_{n-1}} \cdots s_{n+2-k} s_{n+1-k} s_n \) is not a prefix of \( s_n s_n \) (in fact, it is the word \( s_n (s_{n+1 - k} s_{n-k})^{-1} s_n = s_n t_{n-1} \), \( s_n^p \) must also be a prefix of
\[
s_n^{d_{n+1} + r} s_n^{d_{n} - i} s_{n+1}^{-1-i} s_n^{d_{n} + d_{n+1} - i-k} s_{n+1-k} s_n = s_n^{d_{n+1} + r+1} t_{n-1}.
\]
As in the proof of Lemma 5.1, one can show that the prefix index of \( s_n \) in the word (21) is \( d_{n+1} + r + 1 + |D_{n-k}|/|s_n| \), which is at most \( d_{n+1} + 2 + |D_{n-k}|/|s_n| \). Furthermore, in the word (20), it is clear that the prefix index of \( s_n \) is less than for (19) (since \( s_{n-1}^{d_{n-1}} \cdots s_{n+2-k} s_{n+1-k} s_n \) has length less than the word \( s_n^{d_{n} - i} s_{n+1}^{-1-i} s_n^{d_{n} + d_{n+1} - i-k} s_{n+1-k} s_n \) and is not a prefix of \( s_n s_n \)). Whence, it has been shown that \( \text{ind}(s_n) \leq d_{n+1} + 2 + |D_{n-k}|/|s_n| \), and so the result is now an easy consequence of Lemma 5.2 (which gives \( \text{ind}(s_n) \geq d_{n+1} + 2 + |D_{n-k}|/|s_n| \)). \( \square \)

The following analogue of Lemma 3.5 in [5] is required in order to prove Theorem 6.1.

**Lemma 6.6** Let \( n \in \mathbb{N}^+ \) and suppose \( u < s \) with \( |s_n| \leq |u| < |s_{n+1}| \). Then the following assertions hold.

1. For all \( i \in [1, k] \), if \( u \) starts at position \( l \) in some \( s_{n+1-i} \) block in the \( n \)-partition of \( s \) and also starts at position \( m \) in some factor \( s_{n+1-i} \) of \( s \), then \( l = m \).
2. For all \( i \in [1, k-1] \), if \( u \) can start at position \( l \) in \( s_{n+1-i} \) and at position \( m \) in \( s_{n-i} \), then \( l = m \).

**Proof.** By inspection of the \( n \)-partition of \( s \), notice that, for \( 1 \leq i \leq k-1 \), an \( s_{n+1-i} \) block is followed by either an \( s_{n+1-i} \) block, an \( s_n \) block, or an \( s_{n-i} \) block. Furthermore, an \( s_{n+1-k} \) block is always followed by an \( s_n \) block.

Let \( u_{n+1-i} \) be the prefix of \( u \) of length \( |s_{n+1-i}| \).

1. Let \( 1 \leq i \leq k \) and consider an occurrence of \( u \) that starts in an \( s_{n+1-i} \) block of the \( n \)-partition of \( s \). If this \( s_{n+1-i} \) block is followed by an \( s_{n+1-i} \) block, then \( u_{n+1-i} \) is a conjugate of \( s_{n+1-i} \) as \( u_{n+1-i} \prec s_{n+1-i} s_{n+1-i} \) and \( |u_{n+1-i}| = |s_{n+1-i}| \). Similarly, if this \( s_{n+1-i} \) block is followed by an \( s_n \) block, then \( u_{n+1-i} \) is a conjugate of \( s_{n+1-i} \) since \( s_{n+1-i} \prec s_n \). And, if this \( s_{n+1-i} \) block is followed by an \( s_{n-i} \) block, then again \( u_{n+1-i} \) is a conjugate of \( s_{n+1-i} \). Indeed, in the \((n+k-i)\)-partition of \( s \), \( s_{n+1-i} \) is always followed by an \( s_{n+k-i} \) block, which has \( s_{n+1-i} \) as a prefix; whence \( u_{n+1-i} \prec s_{n+1-i} s_{n+1-i} \). So, in any case, \( u_{n+1-i} \) is a conjugate of \( s_{n+1-i} \), and the result follows from the fact that the conjugates of \( s_{n+1-i} \) are distinct.

2. Let \( 1 \leq i \leq k-1 \). Suppose the word \( u \) has occurrences starting in \( s_{n+1-i} \) blocks as well as \( s_{n-i} \) blocks in the \( n \)-partition of \( s \). (Note that this implies \( n \geq i \).) First consider
an occurrence of \( u \) beginning in a block of the form \( s_{n-i} \) of the \( n \)-partition. As an \( s_{n-i} \) block is always followed by an \( s_{n+k-i-1} \) block in the \( (n + k - i - 1) \)-partition of \( s \) and \( s_{n+1-i} \prec s_{n+k-i-1} \), we have

\[
 u_{n+1-i} \prec s_{n-i} s_{n+1-i} = s_{n-i} d_{n-i+1}^{d_{n-i}} \cdots s_{n+2-i-k}^{d_{n+3-i-k}} s_{n+1-i-k}.
\]

Thus, in light of Lemma 6.4, we have the following fact:

\[
 cs_{n-i} \not\prec u_{n+1-i} \quad \text{where } c \in \mathcal{A}_k \text{ and } c \prec s_{n+1-i-k}.
\] (22)

Consider an occurrence of \( u \) starting in an \( s_{n+1-i} \) block of the \( n \)-partition, which can be factorized as

\[
 s_{n+1-i} = s_{n-i}^{d_{n-i+1}} s_{n-i-1}^{d_{n-i}} \cdots s_{n+2-i-k}^{d_{n+3-i-k}} s_{n+1-i-k}.
\] (23)

We distinguish two cases, below.

**Case 1:** The word \( u \) begins in the left-most \( s_{n-i} \) block in (23) when \( d_{n-i+1} \geq 2 \). In this case, \( u_{n-i} \) is a conjugate of \( s_{n-i} \) and hence, as deduced in (1), the starting position of \( u \) in this \( s_{n-i} \) block must coincide with its starting position in any occurrence of \( s_{n-i} \) in the \( n \)-partition of \( s \).

**Case 2:** The word \( u \) does not start in the left-most \( s_{n-i} \) block in (23). The block to the right of \( s_{n+1-i} \) in the \( n \)-partition is either another \( s_{n+1-i} \), or an \( s_{n-i} \), or an \( s_n \). In any case, \( s_{n-i} \) is a prefix of this block to the right of \( s_{n+1-i} \), which implies \( u_{n+1-i} \) contains an occurrence of \( s_{n+1-i-k} s_{n-i} \). This contradicts (22). \( \square \)

**Proof of Theorem 6.1** Clearly, \( p(m; l_1) \geq p(m; l_2) \) if \( l_1 \leq l_2 \). Thus, it suffices to show that for \( m \not\in \bigcup_{i=1}^k \mathcal{D}_i(n) \), we have

\[
p(m; 2) = 0,
\] (24)

i.e., there are no squares of words of length \( m \) in \( s \).

Suppose (24) does not hold for some \( m \) satisfying

\[
m \not\in \bigcup_{i=1}^k \mathcal{D}_i(n),
\] (25)

and let \( u \) be a word of length \( m \) with \( |s_n| \leq m < |s_{n+1}| \) such that \( u^2 \prec s \). For convenience, we shall write \( u^2 = u^{(1)} u^{(2)} \) to allow us to refer to the two separate occurrences of \( u \). Let \( 1 \leq i, j \leq k \). Obviously, \( u^{(1)} \) starts at position \( q \), say, in some \( s_{n+1-i} \) block of the \( n \)-partition of \( s \). Further, by Lemma 6.6, \( u^{(2)} \) also starts in some \( s_{n+1-j} \) block of the \( n \)-partition of \( s \) at position \( q \). From the proof of Lemma 6.5, recall that two consecutive \( s_{n+1-k} \) blocks in the \( n \)-partition of \( s \) are separated by a word of one of the following \( k \) forms:

\[
s_n^{d_{n+1}} s_{n-1}^{d_{n}} \cdots s_{n+2-k}^{d_{n+3-k}}, \quad r \in \{0, 1\},
\]
or

\[
s_n^{d_{n+1}} s_{n-1}^{d_{n}} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}^{d_{n+1}} s_{n-1}^{d_{n}} \cdots s_{n+2-k}^{d_{n+3-k}}, \quad i \in [2, k - 1].
\]
If we also keep in mind that \(|s_n| \leq |u| < |s_{n+1}|\), then using Lemma 6.6 we see that the possible lengths \(|u|\) of \(u\) are:

\[
|s^r_n| \quad \text{and} \quad |s^r_ns_{n-1} \cdots s_{n-i+2}s_{n-i}|\]

where \(1 \leq i \leq k-1\) and \(1 \leq r \leq d_{n+1}\) (with \(r \neq d_{n+1}\) if \(i = k-1\) as \(|u| < |s_{n+1}|\)). Therefore, \(m\) does not satisfy (25); a contradiction. \(\Box\)

The next five propositions, which have some interest in themselves, are needed in the next two sections where we shall prove our main results concerning squares, cubes, and higher powers in \(s\).

**Notation:** Given \(l \in \mathbb{N}\) and \(w \in A_k\), denote by \(\text{Pref}_l(w)\) the prefix of \(w\) of length \(l\) if \(|w| > l\), \(w\) otherwise. Likewise, denote by \(\text{Suff}_l(w)\) the suffix of \(w\) of length \(l\) if \(|w| > l\), \(w\) otherwise.

Recall that \(\Omega^r_n\) denotes the set of singular \(n\)-words of the \(r\)-th kind \((1 \leq r \leq k-1)\), as defined in Theorem 4.19.

**Proposition 6.7** Let \(n \in \mathbb{N}^+\). Suppose \(w \in \Omega^1_{n+1-i}\) for some \(i \in [1, k-1]\) and let \(v = \text{Pref}_l(w)\) where \(1 \leq l \leq |G_{n+1-i,1}| - 1\). Then the word \(vs_{n+1-i}\) occurs at position \(p\) in \(s\) if and only if the \(n\)-partition of \(s\) contains an \(s_n\) starting at position \(p + l\) and an \(s_{n-i}\) ending at position \(p + l - 1\). In particular, \(w\) occurs at exactly those positions where \(vs_{n+1-i}\) occurs in \(s\).

**PROOF.** Let \(i \in [1, k-1]\) be fixed and let \(1 \leq l \leq |G_{n+1-i,1}| - 1\).

First note that \(|w| = |s_{n+1-i}|\) and \(w \preceq x^{-1}G_{n+1-i,1}D_{n-i}G_{n+1-i,1}x^{-1}\), where \(x \in A_k\), by definition of \(\Omega^1_{n+1-i}\). Since \(|D_{n-i}G_{n+1-i,1}| = |s_{n+1-i}|\), the word \(v = \text{Pref}_l(w)\) is a suffix of \(x^{-1}G_{n+1-i,1}\), which, in turn, is a suffix of \(s_{n-i}\) as \(G_{n+1-i,1} = G_{n-i,k-1}\).

Now, by Lemma 6.6, the word \(s_{n+1-i}\) can only occur at the starting positions of blocks (in the \(n\)-partition) of the form \(s_n, s_{n-1}, \ldots, s_{n-k}\), all of which have different last letters (by Proposition 4.1). In particular, each \(s_{n-j}\) block \((0 \leq j \leq k-1, j \neq i)\) of the \(n\)-partition of \(s\) has a different last letter to \(s_{n-i}\) (and hence \(v\)). One should note, however, that an \(s_{n-i}\) block of the \(n\)-partition is never followed by an \(s_{n+1-i}\) block (except if \(i = 1\), in which case we do have certain \(s_n\) blocks preceded by \(s_{n-1}\) blocks). Also observe that if \(z = \text{Suff}_l(s_{n+1-i})\), then an \(s_{n-i}\) block of the \(n\)-partition is only ever followed by \(s_{n+1-i}z^{-1} = D_{n-i}G_{n+1-i,1}z^{-1}\) if it is followed by an \(s_n\) block of the \(n\)-partition. Taking all of this into account, one deduces that the word \(vs_{n+1-i}\) occurs only at positions in \(s\) where an \(s_{n-i}\) block of the \(n\)-partition is followed by an \(s_n\) block, which has \(s_{n+1-i}\) as a prefix. This completes the proof of the first assertion.

As for the second assertion, recall that \(w\) begins with the word \(v\) which is a non-empty suffix of \(x^{-1}G_{n+1-i,1} = x^{-1}G_{n-i,k-1} \preceq s_{n-i}\). Consequently, \(w\) occurs at every \((|s_{n-i}| - l + 1)\)-position of an \(s_{n-i}\) block that is followed by an \(s_n\) block in the \(n\)-partition of \(s\), i.e., \(w\) occurs where the prefix \(vs_{n+1-i}\) of \(vs_n\) occurs in \(s\). By Lemma 6.6, the only other position where \(w\) may occur (besides where \(vs_{n+1-i}\) occurs) is in the \((|s_{n-i}| - l + 1)\)-position of an \(s_n\) block that is preceded by an \(s_{n-i}\) block. Now, to the right of this type.
of $s_n$ block (in the $n$-partition) there appears another $s_n$ block or an $s_{n-1}$ block. The fact that $s_{n+1-i}s_{n-i} \prec_p s_{n}s_{n-1} \prec_p s_n s_n$ implies that $w$ ends with the prefix of $s_{n-i}$ of length $|s_{n-i}| - l$. More precisely, $w$ ends with the word

$$D_{n+1-i-k}z_1$$

where $z_1$ is a non-empty prefix of $\tilde{G}_{n+1-i,1}$ of length $|z_1| = |G_{n+1-i,1}| - l$. On the other hand, by definition of $w$, we have that $w$ ends with

$$D_{n-i}z_2$$

where $z_2$ is a non-empty prefix of $G_{n+1-i,1}$ of length $|z_2| = |G_{n+1-i,1}| - l$. It is impossible for both situations to occur, so we conclude that $w$ occurs at exactly those positions where $vs_{n+1-i}$ occurs. $\square$

Notation: For $n \geq 1$, denote by $\mathbb{P}_n$ the set of all formal positions of $s_{n+1-i}^1s_{n-1}$ in the $(n-1)$-partition of $s$.

Proposition 6.8 For any $n \in \mathbb{N}^+$, the set of all positions of $D_n$ in $s$ is $\mathbb{P}_n$.

PROOF. We proceed by induction on $n$. For $n = 1$, $D_n = D_1 = s_1^{d_2-1}s_0^{d_1}$, and hence $D_1$ occurs at exactly those places in $s$ where $s_1^{d_2-1}s_0 = (a_1^{d_1}a_2^{d_2-1}a_1$ occurs in the 0-partition of $s$. We claim that there is a one-to-one correspondence from the set of all positions of $D_n$ in the $(n-1)$-partition of $s$ to the set of all positions of $D_{n+1}$ in the $n$-partition of $s$ (see (14)). Assume that $\mathbb{P}_n$ gives all of the occurrences of $D_n$ in the $(n-1)$-partition of $s$. Since $D_{n+1} = s_{n+1}^{d_{n+2}-1}s_nD_n = D_n\tilde{s}_n(\tilde{s}_{n+1})^{d_{n+2}-1}$, $D_{n+1}$ occurs at any place in $\mathbb{P}_{n+1}$. Conversely, since each occurrence of $D_{n+1}$ in (14) naturally gives rise to an occurrence of $D_n$ in the $(n-1)$-partition of $s$, the word $D_{n+1}$ must occur in $s$ at exactly those places given by $\mathbb{P}_{n+1}$. $\square$

Consider two distinct occurrences of a factor $w$ in $s$, say

$$s = uwv = u'wv', \quad |u'| > |u|,$$

where $v, v' \in A_k^*$. These two occurrences of $w$ in $s$ are said to be positively separated (or disjoint) if $|u'| > |uw|$, in which case $u' = uwz$ for some $z \in A_k^*$, and hence $s = uwzwv'$. 

Proposition 6.9 For any $n \in \mathbb{N}^+$, successive occurrences of a singular word $w \in \bigcup_{j=1}^{k-1} \Omega_n^j$ in $s$ are positively separated.

PROOF. Let $1 \leq r \leq k - 1$. For $1 \leq n \leq r$, observe that

$$\Omega_n^r = \{w \in A_k^* : |w| = |s_n| \text{ and } w \neq s_{n-1}D_{n-1}s_{n-1} \},$$

where $D_{n-r} = a_{n-r+k+1}$ for $1 \leq n < r$. It is left to the reader to verify that consecutive occurrences of a word $w \in \Omega_n^r$ ($1 \leq n \leq r$) are positively separated in $s$.

Now take $n \geq r + 1$ and suppose $w \in \Omega_n^r$. Then $D_{n-r}$ will occur in $w$. By Proposition 6.8, the word $D_{n-r}$ occurs at exactly those places where $s_{n-r+1}^{d_{n-r+1}-1}s_{n-r-1}$ occurs in the
(n − r − 1)-partition of s. First note that the letter just before $D_{n-r}$ in $w$ is the last letter of $G_{n,r}$, which is the first letter $G_{n,r}$, and hence the last letter of $s_{n−r−k}$ (by Propositions 4.1 and 4.14). On the other hand, in the word $w$, the letter just after $D_{n-r}$ is the first letter of $G_{n,r}$. Since there are $k$ different return words of $s_{n−r−1}^{d_n−r+1} s_{n−r−1}$ in $s$, there exist $k$ different possibilities for occurrences of $s_{n−r−1}^{d_n−r+1} s_{n−r−1}$ in the $(n − r − 1)$-partition of $s$; namely:

\begin{equation}
(s_{n-r}^{d_n-r+1} s_{n-r-1}) s_{n-r-1}^{d_n-r-1} s_{n-r-2}^{d_n-r-2} \cdots s_{n-r-k+1}^{d_n-r-k+2} s_{n-r-k}^{d_n-r-k+1} s_{n-r-1}^{d_n-r-1} s_{n-r-1}^{d_n-r-1} = D_{n-r} D_{n-r} D_{n-r} D_{n-r-1} D_{n-r} \end{equation}

(1)

\begin{equation}
D_{n-r} D_{n-r-k} D_{n-r} D_{n-r-1} D_{n-r} = D_{n-r} (s_{n-r}^{d_n-r+1} D_{n-r-1}) = D_{n-r} (s_{n-r}^{d_n-r+1} G_{n-r,1})
\end{equation}

(2)

\begin{equation}
D_{n-r} D_{n-r-l} (s_{n-r}^{d_n-r+1} s_{n-r-1}) = \begin{cases} D_{n-r} G_{n-r,l} s_{n-r}^{d_n-r+1} - 2 s_{n-r-1} & \text{if } d_{n-r+l} \geq 2, \\ D_{n-r} G_{n-r-1,l-1} & \text{if } d_{n-r+l} = 1, \end{cases}
\end{equation}

where $2 \leq l \leq k - 1$;

\begin{equation}
(s_{n-r}^{d_n-r+1} s_{n-r-1}) (s_{n-r}^{d_n-r+1} s_{n-r-1}) s_{n-r-1}^{d_n-r-1} s_{n-r-2}^{d_n-r-2} \cdots s_{n-r-k+1}^{d_n-r-k+3} s_{n-r-k+1} = \begin{cases} D_{n-r} s_{n-r-1}^{d_n-r-1} (s_{n-r}^{d_n-r+1} G_{n-r,k-1}) & \text{if } d_{n-r+l} \geq 2, \\ D_{n-r} G_{n-r-1,k-2} & \text{if } d_{n-r+l} = 1. \end{cases}
\end{equation}

(3)

Thus, if $d_{n-r+l} \geq 2$, the word $D_{n-r}$ is followed by either $s_{n-r}$, $G_{n-r,l}$, or $s_{n-r-1}$, of which only $s_{n-r}$ has the same first letter as $G_{n,r}$. Similarly, if $d_{n-r+l} = 1$, the word $D_{n-r}$ is followed by either $G_{n-r,1}$, $G_{n-r-1,l-1}$, or $G_{n-r-1,k-2}$, of which only $G_{n-r,1}$ has the same first letter as $G_{n,r}$. Therefore, only in case (1) will we have $D_{n-r}$ followed by the first letter of $G_{n,r}$. Accordingly, one deduces that any occurrence of $w$ in $s$ corresponds to a formal occurrence of the word

\[ s_{n-r-k}^{d_n-r+1} s_{n-r-1}^{d_n-r-1} s_{n-r-2}^{d_n-r-2} \cdots s_{n-r-k+1}^{d_n-r-k+2} s_{n-r-k}^{d_n-r-k+1} s_{n-r-1}^{d_n-r-1} s_{n-r-1}^{d_n-r-1} \]

in the $(n − r − 1)$-partition of $s$. Hence, we conclude that occurrences of $w$ are positively separated in $s$ since a word of the above form is positively separated in the $(n − r − 1)$-partition.

\[ \square \]

The next proposition follows from Lemma 6.5, and Propositions 6.7 and 6.9.
Proposition 6.10 Let \( n \in \mathbb{N}^+ \) and suppose \( u \prec s \) with \( |u| = |s_n| \). Then \( u^2 \prec s \) if and only if \( u \in C(s_n) \). In particular, if \( u \) is a singular word of any kind of \( s \), then \( u^2 \not\prec s \). Moreover, for any \( n \geq k - 1 \), if \( u^2 \prec s \) with \( |s_n| \leq |u| < |s_{n+1}| \), then \( u \) does not contain a singular word from the set \( \Omega_{n+2-k}^1 \).

PROOF. As \( s_n^{d_n+1+2}D_{n-k} \prec s \) (see Lemma 6.5), the square of any conjugate of \( s_n \) is a factor of \( s \). (Note that \( s_n^{d_n+1+2}D_{n-k} = s_n^{d_n+1+2}a_n^{-1} \) for \( 1 \leq n \leq k - 1 \).) Now recall that the set of all factors of \( s \) of length \(|s_n|\) is the disjoint union of the sets \( C(s_n) \) and \( \cup_{j=1}^{i} \Omega_j \). Consequently, the first two assertions are deduced from Proposition 6.9.

For the last statement, let \( n \geq k - 1 \) and suppose \( u^2 = u^{(1)}u^{(2)} \) is an occurrence of \( u^2 \) in \( s \), where \(|s_n| \leq |u| < |s_{n+1}|\). Also assume \( w \in \Omega_{n+2-k}^1 \) and \( w \prec u \). Clearly, \( w \) occurs in both \( u^{(1)} \) and \( u^{(2)} \) at the same position. By Proposition 6.7 (with \( i = k - 1 \)), different occurrences of \( w \) correspond to different occurrences of \( s_{n+1-k} \)-blocks in the \( n \)-partition of \( s \) (as an \( s_{n+1-k} \)-block is always followed by an \( s_n \)-block). Between two consecutive \( s_{n+1-k} \)-blocks in the \( n \)-partition, there is a word taking one of the following forms:

\[
s_r s_n^{d_n+1} s_n^{-1} \cdots s_n^{d_n+3-k}, \quad r \in \{0, 1\},
\]
or

\[
s_n^{d_n+1} s_n^{-1} \cdots s_n^{d_n+3-i} s_n^{d_{n+1}} s_n^{-1} \cdots s_n^{d_n+3-k}, \quad i \in [2, k-1].
\]

Therefore, the distance between consecutive occurrences of \( s_{n+1-k} \)-blocks in the \( n \)-partition of \( s \) is

\[
|s_{n+1}| \quad (r = 0), \quad |s_{n+1}| + |s_n| \quad (r = 1), \quad \text{or} \quad |s_{n+1}| - (|D_{n+1-i}| - |D_{n+1-k}|) + |s_{n+1}|,
\]

and all of these distances are at least \(|s_{n+1}|\), which implies \(|u| \geq |s_{n+1}|\); a contradiction. \( \square \)

More generally, we have the following proposition.

Proposition 6.11 Let \( n \in \mathbb{N}^+ \) and suppose \( u^2 \prec s \) with \(|s_n| \leq |u| < |s_{n+1}|\). Then \( u \) does not contain a singular word from the set \( \Omega_{n+1-i}^1 \) for any \( i \in [1, k-1] \).

PROOF. The case when \( i = k - 1 \) is proved in Proposition 6.10, so take \( i \in [1, k-2] \). Let \( u^2 = u^{(1)}u^{(2)} \) be an occurrence of \( u^2 \) in \( s \), where \(|s_n| \leq |u| < |s_{n+1}| \). Assume \( w \in \Omega_{n+1-i}^1 \) for some \( i \in [1, k-2] \), and \( w \prec u \). Clearly, \( w \) occurs in both \( u^{(1)} \) and \( u^{(2)} \) at the same position. By Proposition 6.7, different occurrences of \( w \) correspond to different occurrences of \( s_{n-i} \)-blocks that are followed by \( s_n \) blocks in the \( n \)-partition of \( s \). By inspection of the \( n \)-partition (see (14)), the word of minimal length that separates two such \( s_{n-i} \)-blocks is

\[
s_n^{d_n+1} s_n^{d_n} \cdots s_n^{d_n+3-k} s_n^{d_{n+1}} s_n^{d_n} \cdots s_n^{d_{n+2-i}}.
\]

That is, the minimal distance between two consecutive occurrences of an \( s_{n-i} \)-block (with each appearance followed by an \( s_n \) block) is \(|s_{n+1}| + |s_n| + |D_n| - |D_{n-i}| > |s_{n+1}| + |s_n|\), which implies \(|u| > |s_{n+1}| + |s_n|\); a contradiction. \( \square \)
6.1 Squares

The next two main theorems concern squares of factors of $s$ of length $m < d_1 + 1 = |s_1|$ and length $m \geq |s_1|$, respectively.

A letter $a$ in a finite or infinite word $w$ is said to be separating for $w$ if any factor of length 2 of $w$ contains the letter $a$. For example, $a$ is separating for the infinite word $(aaba)^\omega$. If $a$ is separating for an infinite word $x$, then it is clearly separating for any factor of $x$. According to [8, Lemma 4], since the standard episturmian word $s$ begins with $a_1$, the letter $a_1$ is separating for $s$ and its factors.

**Theorem 6.12** For $1 \leq r \leq d_1$, we have

$$ p(r; 2) = \begin{cases} 1 & \text{if } r \leq (d_1 + 1)/2, \\ 0 & \text{if } r > (d_1 + 1)/2. \end{cases} $$

In particular, $\mathcal{P}(r; 2) = \{(a_1^r)^2\}$ for $r \leq (d_1 + 1)/2$, and $\mathcal{P}(r; 2) = \emptyset$ for $r > (d_1 + 1)/2$.

**PROOF.** Consider a factor $u$ of $s$ with $|u| = r \leq d_1$. As $a_1$ is separating for $s$ and $a_1$ occurs in runs of length $d_1$ or $d_1 + 1$ (inspect the 0-partition of $s$), we have that $u$ is either $a_1^r$ or a conjugate of $a_1^{r-1}a_j$ for some $j$, $1 < j \leq k$. Further, it is evident that there are no squares of words conjugate to $a_1^{r-1}a_j$, $1 < j \leq k$. And, using the same reasoning for words $u$ of the form $a_1^r$, one determines that $s$ contains the square of $u$ if and only if $2r \leq d_1 + 1$, in which case there exists exactly one square of each such factor of length $r$ of $s$; namely $(a_1^r)^2$. $\Box$

Let $w$ be a factor of $s$ with $|w| \in \bigcup_{i=1}^k D_i(n)$ for some $n$. Roughly speaking, the next theorem shows that if $w^2$ is a factor of $s$, then $w$ is a conjugate of a finite product of blocks from the set $\{s_n, s_{n+1}, \ldots, s_{n+k}\}$, depending on $|w|$ and $d_{n+1}$. For example, if $|w| = r|s_n|$ for some $r$ with $1 \leq r < 1 + d_{n+1}/2$, then $w^2 < s$ if and only if $w$ is one of the first $|s_n|$ conjugates of $s_n$.

**Theorem 6.13** Let $n$, $r \in \mathbb{N}^+$.

(i) For $1 \leq r \leq d_{n+1}$,

$$ p(|s_n^r|; 2) = \begin{cases} |s_n| & \text{if } 1 \leq r < 1 + d_{n+1}/2, \\ |D_{n-k}| + 1 & \text{if } d_{n+1} \text{ is even and } r = 1 + d_{n+1}/2, \\ 0 & \text{if } 1 + d_{n+1}/2 < r \leq d_{n+1}. \end{cases} \quad (26) $$

In particular,

$$ \mathcal{P}(|s_n^r|; 2) = \begin{cases} \{C_j(s_n^r) : 0 \leq j \leq |s_n| - 1\} & \text{if } 1 \leq r < 1 + d_{n+1}/2, \\ \{C_j(s_n^r) : 0 \leq j \leq |D_{n-k}|\} & \text{if } d_{n+1} \text{ is even and } r = 1 + d_{n+1}/2, \\ \emptyset & \text{if } 1 + d_{n+1}/2 < r \leq d_{n+1}. \end{cases} \quad (27) $$

25
(ii) For $1 \leq r \leq d_{n+1}$ and $i \in [2, k]$ (with $r \neq d_{n+1}$ if $i = k$), we have

$$p(|s_n^r s_n \cdots s_{n+3-i} n+1-i|; 2) = |D_{n+1-i}| + 1. \quad (28)$$

In particular,

$$\mathcal{P}(|s_n^r s_n \cdots s_{n+3-i} n+1-i|; 2) = \{C_j(s_n^r s_n \cdots s_{n+3-i} n+1-i) : 0 \leq j \leq |D_{n+1-i}|\}. \quad (29)$$

**Remark 6.14** For standard Sturmian words $c_n$, we have $s_n = D_{n-1} x y$, where $x, y \in \{a, b\}$ (x $\neq$ y), and hence $|D_{n-1}| = q_n - 2$ for all $n \geq 1$. Accordingly, Theorem 6.13 agrees with Theorem 3 in [6] for the case of a 2-letter alphabet.

The proof of Theorem 6.13 requires the following three lemmas.

**Lemma 6.15** Let $n \in \mathbb{N}^+$ and set $u_i := s_n^r s_n \cdots s_{n+3-i} n+1-i$ for each $i \in [2, k]$ and $1 \leq r \leq d_{n+1} - 1$. Then, for all $i \in [2, k]$, we have

$$(s_n^d s_{n-1} \cdots s_{n+i-1} n+1-i)^2 D_{n+1-i} \prec s, \quad (30)$$

and

$$u_i D_{n+1-i} \prec (s_n^d s_{n-1} \cdots s_{n+i-1} n+1-i)^2. \quad (31)$$

**PROOF.** Let us first note that, for $i = k$, $(s_n^d s_{n-1} \cdots s_{n+i-1} n+1-i)^2 D_{n+1-k} = s_{n+1} D_{n+1-k}$ is a factor of $s$ (by Lemma 5.1). Now, for $i \in [2, k-1]$, by inspection of the $n$-partition of $s$, the word

$$s_n^d s_{n-1} \cdots s_{n+i-1} n+1-i \in \cdots s_{n+i-1} n+1-i \cdots s_{n+i} n+1$$

is a factor of $s$ (where $D_{n+1-k}$ is a prefix of $s_n$ for $n \geq k - 1$, and $D_{n+1-k}^{-1} = a_n+2$ for $1 \leq n \leq k - 2$). Thus, assertion (30) is proved.

As for the second assertion (31), one can write

$$(s_n^d s_{n-1} \cdots s_{n+i-1} n+1-i)^2$$

$$= s_n^d u_i^d s_{n+i-1} n+1-i \cdots s_{n+i} n+1-i$$

$$= s_n^d u_i^d D_{n+i-1} D_{n+i-1}^{-1} s_{n+i} n+1-i \cdots s_{n+i} n+1-i$$

which yields the result since $D_{n-k}$ is a prefix of $s_n$ and $s_{n-1}$ for $n \geq k$, and $D_{n-k}^{-1} = a_n+1$ for $1 \leq n \leq k - 1$. \Box
Lemma 6.16 Let \( n \in \mathbb{N}^+ \) and let \( u^2 = u^{(1)}u^{(2)} \) be an occurrence of \( u^2 \) in \( s \), where \( |s_n| \leq |u| < |s_{n+1}| \).

(i) For all \( n \geq 1 \), if \( |u| = |s_n^r| \) with \( 1 \leq r \leq d_{n+1} \), then \( u^{(1)} \) begins in an \( s_n \) block of the \( n \)-partition of \( s \). Moreover, \( u^2 \) is a factor of \( s_{d_{n+1}+2}s_{n+1}v^{-1} = s_{n+1}D_{n-k} \), where \( |v| = |s_n| - |D_{n-k}| \).

(ii) Let \( i \in [2, k-1] \). For all \( n \geq i - 1 \), if \( |u| = |s_n^r s_{n+1}\cdots s_{n+2-i} s_{n+1-i}| \) with \( 1 \leq r \leq d_{n+1} \), then \( u^{(1)} \) starts in an \( s_n \) block and contains an \( s_{n+1-i} \) block that is followed by an \( s_n \) block in the \( n \)-partition of \( s \). Moreover, \( u^2 \) is a factor of \((s_n^r s_{n+1}\cdots s_{n+2-i} s_{n+1-i})^2D_{n+1-i} \), which is a factor of

\[
(s_n^{d_{n+1}}s_{n+1}^d\cdots s_{n+2-i}^d s_{n+1-i})^2D_{n+1-i}.
\]

(iii) For all \( n \geq k-1 \), if \( |u| = |s_n^r s_{n+1}\cdots s_{n+3-k} s_{n+k-1}| \) with \( 1 \leq r \leq d_{n+1} - 1 \), then \( u^{(1)} \) starts in an \( s_n \) block and contains an \( s_{n+k-1} \) block of the \( n \)-partition of \( s \). Moreover, \( u^2 \) is a factor of \((s_n^r s_{n+1}\cdots s_{n+2-k} s_{n+1-k})^2D_{n+1-k} \), which is a factor of

\[
s_{n+1}^2 = (s_{n+1}^d s_{n+2-i}^d s_{n+1-i}^d)^2.
\]

**PROOF.** (i) By similar arguments to those used in the proof of Theorem 6.1, the first claim is obtained from the fact that \( |u| = r|s_n| \) with \( 1 \leq r \leq d_{n+1} \), together with Lemma 6.6. For the second claim, one uses the fact that an \( s_n \) block in which \( u^{(1)} \) starts is followed by the word \( s_n^p s_{n+1}^d\cdots s_{n+3-k} s_{n+k-1} s_{n+3-i} s_{n+3-i} s_{n+k-1} \) for some \( i \in [2, k] \) and \( 0 \leq p \leq d_{n+1} \). Hence, we have

\[
s_{n+1}^d = s_n^d(s_n^{d_{n+2-i}}s_{n+1-i}^d\cdots s_{n+k-1}^d)^{-1}s_{n+1-k}^dG_{n+1-k}\]

where \( |v| = |s_n| - |D_{n-k}| \). Therefore, the assertion holds provided \( u^{(2)} \) does not contain the word \( s_{n-1}s_{n-1}^d G_{n+1}^{-1} \) for some non-empty proper prefix \( w \) of \( G_{n-1} \). Indeed, if \( s_{n-1}s_{n-1}^d G_{n+1}^{-1} < u^{(2)} \), then \( s_{n-1}D_{n-1}w = D_{n-k}G_{n-1}D_{n-1}w \) is a factor of \( u^{(2)} \), where \( w \prec_p G_{n-1} \). But this situation is absurd (by Proposition 6.11) since this word contains a singular \( n \)-word of the first kind.

(ii) From Lemma 6.6, we can argue (as in the proof of Theorem 6.1) that \( u^{(1)} \) begins in an \( s_{n+1-i} \) block that is followed by an \( s_n \) block in the \( n \)-partition, or contains an \( s_{n+1-i} \) block that is followed by an \( s_n \) block. However, in the first case, we see that \( u \) would contain a singular word from the set \( \Omega^1_{n+2-i} \), since \( u \) would contain \( s_{n+1-i} s_{n+2-i} \), which has \( s_{n+1-i} s_{n+2-i} \) as a prefix (see Proposition 6.7). This contradicts Proposition 6.11, and so only the second case can occur. By reasoning as above and using the fact
that \(|u| = |s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}|\), \(u^{(1)}\) must start in the left-most \(s_n\) block in the word
\[s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i} s_n,\]
which appears in the \(n\)-partition of \(s\). Since \(D_{n+1-i} \prec_p s_{n+2-i} \prec_p s_n\), \(u^{(1)}\) ends within the first \(|D_{n+1-i}|\) letters of the \(s_n\) block to the right of the \(s_{n+1-i}\) block. Otherwise, \(u\) would contain a singular word \(w \in \Omega_{n+2-i}\) which contradicts Proposition 6.11. To the left of the \(s_{n+1-i}\) block, there is the word \(s_n^{d_{n+1}} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i} k\) and, in view of the fact that \(|u| = |s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}|\) with \(1 \leq r \leq d_{n+1}\), one deduces that there exists a \(p \in \mathbb{N}\) with \(0 \leq p \leq |D_{n+1-i}|\) such that \(u^{(1)}\) starts at position \(p\) in \(s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i} s_n\). This implies \(u^2 \prec v^2 D_{n+1-i}\) where \(v := s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}\). It remains to show that \(v^2 D_{n+1-i}\) is contained in
\[(s_{n+1}^{d_{n+1}} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i})^2 D_{n+1-i}\]
which, in turn, is a factor of \(s\). Indeed, this fact is easily deduced from Lemma 6.15.

(iii) The proof of this assertion is similar to that of (ii), but with \(i = k\) and \(1 \leq r \leq d_{n+1} - 1\). The details are left to the reader. □

**Lemma 6.17** For all \(n, r \in \mathbb{N}^+\) and \(i \in [2, k]\), the word \(v := s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}\) is primitive.

**PROOF.** Suppose on the contrary that the given word \(v\) is not primitive, i.e., suppose \(v = u^p\) for some non-empty word \(u\) and integer \(p \geq 2\). Then \(|v|_a = p |u|_a\) for each letter \(a_j \in A_k\), i.e., \(p\) divides \(|v|_a\) for each \(a_j \in A_k\). In particular, \(p\) divides
\[|v|_{a_1} = r(Q_n - P_n) + d_n(Q_n-1 - P_{n-1}) + \cdots + (Q_{n+1-i} - P_{n+1-i}) = |v| - (r P_n + d_n P_{n-1} + \cdots + d_{n+3-i} P_{n+2-i} + P_{n+1-i});\]
by Proposition 3.4. Thus, \(p\) must also divide \(r P_n + d_n P_{n-1} + \cdots + d_{n+3-i} P_{n+2-i} + P_{n+1-i}\). But \(\gcd(|v|, r P_n + d_n P_{n-1} + \cdots + d_{n+3-i} P_{n+2-i} + P_{n+1-i}) = 1\), which yields a contradiction; whence \(p = 1\), and therefore \(v\) is primitive. □

**Proof of Theorem 6.13** We simply prove that (27) and (29) hold as the elements of the respective sets are mutually distinct (by Lemma 6.17), which implies (26) and (28).

(i) As shown previously, for each \(i \in [2, k]\), \(v := s_n^{d_{n+1}} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i} s_n C_{n-i-1}^{-1} = s_n^{d_{n+1}+2} D_{n-k}\) is a factor of \(s\). Thus, by Lemma 6.16(i), it suffices to find all of the squares of words \(u\) with \(|u| = r |s|_a\) \((1 \leq r \leq d_{n+1})\) that occur in the word \(v\). In fact, one need only consider occurrences of \(u^2\) starting in the left-most \(s_n\) block of \(v\), and the result easily follows.

(ii) Let \(i \in [2, k-1]\). By Lemma 6.15, the word \((s_n^{d_{n+1}} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i})^2\) is a factor of \(s\) and it contains the word
\[v := (s_n^{r} s_{n-1}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i})^2 D_{n+1-i};\]

28
for any \( r \) with \( 1 \leq r \leq d_{n+1} \). Therefore, \( u^2 \) is a factor of \( s \) for each word \( u \) given by
\[
u := C_j(s_n^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}), \quad 0 \leq j \leq |D_{n+1-i}|.
\]

Conversely, if \( u^2 \preceq s \) with \( |u| = |s_n^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}| \) for some \( 1 \leq r \leq d_{n+1} \), then \( u^2 \preceq v \), by Lemma 6.16(ii). And, since \( |v| = 2|u| + |D_{n+1-i}| \), we must have \( u = C_j(s_n^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}) \) for some \( j \) with \( 0 \leq j \leq |D_{n+1-i}| \).

The case \( i = k \) is proved similarly, using Lemma 6.16(iii). \( \square \)

6.2 Cubes and higher powers

Our subsequent analysis of cubes and higher powers occurring in \( s \) is now an easy task due to the above consideration of squares. Extending Theorem 6.13 (see Theorem 6.19 below), only requires the following lemma, together with arguments used in the proof of Theorem 6.13.

Lemma 6.18 Let \( n \in \mathbb{N}^+ \) and suppose \( u^3 \preceq s \) with \( |s_n| \leq |u| < |s_{n+1}| \). Then \( u^3 \) does not contain a singular word from the set \( \Omega_{n+1-i}^1 \) for any \( i \in [1, k-1] \).

PROOF. Suppose on the contrary that \( u^3 = u^{(1)}(u^{(2)}u^{(3)}) \) contains a singular word \( w \in \Omega_{n+1-i}^1 \) for some \( i \in [1, k-1] \). By Proposition 6.11, \( w \) is not a factor of \( u^{(3)} \), and therefore every occurrence of \( w \) must begin in \( u^{(1)} \) or \( u^{(2)} \), both of which are followed by \( u \) again. Accordingly, there exists a \( p \in \mathbb{N} \) such that \( w \) starts at position \( p \) in both \( u^{(1)} \) and \( u^{(2)} \). Reasoning, as in the proofs of Propositions 6.10 and 6.11, yields the contradiction \( |u| \geq |s_{n+1}| \). \( \square \)

Theorem 6.19 Let \( n, r, l \in \mathbb{N}^+, l \geq 3 \).

(i) For \( 1 \leq r \leq d_{n+1} \),
\[
p(|s_n^r|; l) = \begin{cases} |s_n| & \text{if } 1 \leq r < (d_{n+1} + 2)/l, \\
|D_{n-k}| + 1 & \text{if } r = (d_{n+1} + 2)/l, \\
0 & \text{if } (d_{n+1} + 2)/l < r \leq d_{n+1}. \end{cases}
\]

In particular,
\[
P(|s_n^r|; l) = \begin{cases} \{C_j(s_n^r) : 0 \leq j \leq |s_n| - 1\} & \text{if } 1 \leq r < (d_{n+1} + 2)/l, \\
\{C_j(s_n^r) : 0 \leq j \leq |D_{n-k}|\} & \text{if } r = (d_{n+1} + 2)/l, \\
\emptyset & \text{if } (d_{n+1} + 2)/l < r \leq d_{n+1}. \end{cases}
\]

(ii) For \( 1 \leq r \leq d_{n+1} \) and \( i \in [2, k] \) (with \( r \neq d_{n+1} \) if \( i = k \)), we have
\[
p(|s_n^{r} s_{n+1-i}^{d_n} \cdots s_{n+2-i}^{d_{n+3-i}} s_{n+1-i}|; l) = 0.
\]

PROOF. (i) Suppose \( u \preceq s \) with \( |u| = r|s_n| \) for some \( r \), \( 1 \leq r \leq d_{n+1} \), and consider an occurrence of \( u^l = u^{(1)}u^{(2)} \cdots u^{(l)} \) in \( s \), \( l \geq 3 \). By Lemma 6.16(i), \( u^{(1)}u^{(2)} \) begins in an \( s_n \)
block of the $n$-partition of $s$, and by Lemma 6.18, $u^l$ does not contain a singular $n$-word of the first kind. So, as in the proof of Theorem 6.13(i), one infers that $u^l$ is contained in the word $v := s_n^{d_{n+1}+2}D_{n-k} \ll s$, and the rest of the proof follows in much the same fashion.

(ii) In this case, assume $u^3 = u^{(1)}u^{(2)}u^{(3)}$ occurs in $s$. By (ii) and (iii) of Lemma 6.16, $u^{(1)}$ begins in an $s_n$ block and contains an $s_{n+1-i}$ block that is followed by an $s_n$ block in the $n$-partition of $s$. Accordingly, $u^3$ contains the word $s_{n+1-i}s_n$, and hence contains a singular word $w \in \Omega_{n+2-i}^1$, which contradicts Lemma 6.18. □

6.3 Examples

**Example 6.20** Let us demonstrate Theorems 6.13 and 6.19 with an explicit example. Consider the standard episturmian word $s$ over $A_3 = \{a_1, a_2, a_3\} \equiv \{a, b, c\}$ with periodic directive word $(abca)^\omega$. We have $(d_n)_{n \geq 1} = (1, 1, 2, 2, 1, 2)$, and hence

- $s_1 = ab$, $|s_1| = 2$;
- $s_2 = abac$, $|s_2| = 4$;
- $s_3 = abacabacaba, |s_3| = 11$;
- $s_4 = abacabacaabacababaabacabab, |s_4| = 32$;
- $s_5 = abacabacaabacababaabacabacabacaabacabacabacabacaab, |s_5| = 58$.

Also, $D_0 = \varepsilon$, $D_1 = a$, $D_2 = abacaba$, and $D_3 = abacabacaabacabaca$.

We shall simply consider squares and cubes of words of length $m$ occurring in $s$ with $|s_3| \leq m < |s_6|$. By Theorem 6.1, we need only consider lengths $m$ in the set

\[ S := \{|s_3|, 2|s_3|, |s_4|, |s_5|, 2|s_5|, |s_3s_2|, |s_3^2s_2|, |s_3^2s_2s_1|, |s_4s_3|, |s_5s_4|, |s_5^2s_4|, |s_5s_4s_3|\}. \]

According to Theorem 6.13(i), for $3 \leq n \leq 5$, we have

- $p(|s_3|; 2) = |s_3| = 11$,
- $p(2|s_3|; 2) = |D_0| + 1 = 1$,
- $p(|s_4|; 2) = |s_4| = 32$,
- $p(|s_5|; 2) = |s_5| = 58$,
- $p(2|s_5|; 2) = |D_2| + 1 = 8$.

Also, part (ii) of Theorem 6.13 gives

- $p(|s_3s_2|; 2) = |D_2| + 1 = 8$,
- $p(|s_3^2s_2|; 2) = |D_2| + 1 = 8$,
- $p(|s_3^2s_2s_1|; 2) = |D_1| + 1 = 2$,
- $p(|s_4s_3|; 2) = |D_3| + 1 = 23$,
- $p(|s_5s_4|; 2) = |D_4| + 1 = 34$,
- $p(|s_5^2s_4|; 2) = |D_4| + 1 = 34$,
- $p(|s_5s_4s_3|; 2) = |D_3| + 1 = 23$.
Furthermore, from Theorem 6.19, one has

\[ p(|s_3|; 3) = |s_3| = 11, \]
\[ p(2|s_3|; 3) = 0, \]
\[ p(|s_4|; 3) = |D_1| + 1 = 2, \]
\[ p(|s_5|; 3) = |s_5| = 58, \]
\[ p(2|s_5|; 3) = 0, \]

and \( p(m; 3) = 0 \) for all other lengths \( m \in S \).

For instance, the sole factor of \( s \) of length \( 2|s_3| = 22 \) that has a square in \( s \) is
\[ s_3^2 = abacabacabaabacabacaba, \]
and the eight squares of length \( 2|s_3s_2| = 30 \) are the squares of the first eight conjugates of
\[ s_3s_2 = abacabacabaacacaba; \]
namely
\[ (abacabacabaacaba)^2, \ (bacabacabaabaca)^2, \ldots, \ (cabaabacabacaba)^2. \]

The only factors of length \( |s_4| = 32 \) that have a cube in \( s \) are the first two conjugates of \( s_4 \), i.e.,
\[ s_4^3 ≺ s \quad \text{and} \quad (C_1(s_4))^3 = (a^{-1}s_4a)^3 ≺ s. \]

**Example 6.21** The \( k \)-bonacci word is the standard episturmian word \( η_k \in A_k^\omega \) with directive word \( (a_1a_2\cdots a_k)^\omega \). Since all \( d_i = 1 \), we have \( s_n = s_{n-1}s_{n-2}\cdots s_{n-k} \) for all \( n \geq 1 \) (and the lengths \( |s_n| \) are the \( k \)-bonacci numbers). Thus, for fixed \( n \in \mathbb{N}^+ \) and \( l \geq 2 \), if \( w^l ≺ η_k \) with \( |s_n| \leq |w| < |s_{n+1}| \), then we necessarily have \( |w| = |s_n| + |s_{n-1}| + \cdots + |s_{n+1-i}| \) for some \( i \in [1, k-1] \) (by Theorem 6.1). The preceding main theorems reveal that
\[ \mathcal{P}(1; 2) = \{a_1\}, \ \mathcal{P}(|s_n|; 2) = \mathcal{C}(s_n) = \Omega_n^0 \quad \text{and} \quad \mathcal{P}(|s_n|; 3) = \{C_j(s_n) : 0 \leq j \leq |D_{n-k}|\}. \]

Furthermore, for each \( i \in [2, k-1] \), we have
\[ \mathcal{P}(|s_ns_{n-1}\cdots s_{n+1-i}|; 2) = \{C_j(s_ns_{n-1}\cdots s_{n+1-i}) : 0 \leq j \leq |D_{n+1-i}|\}. \]

All other \( \mathcal{P}(|w|; l) = \emptyset, \ l \geq 2 \). In particular, \( k \)-bonacci words are 4-power free.

### 7 Concluding remarks

Using the results of Section 6, it is possible to determine the exact number of distinct squares in each building block \( s_n \), which extends Fraenkel and Simpson’s result [10] concerning squares in the finite Fibonacci words. Such work forms part of the present author’s PhD thesis [13, Chapters 6 and 7].

Theorems 6.12, 6.13 and 6.19 also suffice to describe all integer powers occurring in any (episturmian) word \( t \in A_k^\omega \) that is equivalent to \( s \). (See [15, Theorem 3.10] for a definition of such \( t \).) The problem of determining all integer powers occurring in general standard episturmian words (with not all \( d_i \) necessarily positive) remains open.
Acknowledgements

The author would like to thank the anonymous referee for their helpful comments and suggestions.

References


