Synchronizing delay for binary uniform morphisms

Karel Klouda\textsuperscript{a}, Kateřina Medková\textsuperscript{b}

\textsuperscript{a} Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00, Prague 6
\textsuperscript{b} Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Blýskalova 7, 115 19, Prague 1

Abstract

Circular D0L-systems are those with finite synchronizing delay. We introduce a tool called graph of overhangs which can be used to find the minimal value of synchronizing delay of a given D0L-system. By studying the graphs of overhangs, a general upper bound on the minimal value of a synchronizing delay of a circular D0L-system with a binary uniform morphism is given.

Keywords: D0L-system, circularity, synchronizing delay

2000 MSC: 68R15

1. Introduction

Circular codes are a classical notion studied in theory of codes \cite{1}. A set $X$ of finite words is a code if each word in $X^+$ has a unique decomposition into words from $X$. If we slightly modify the requirement of uniqueness, we get the definition of a circular code: $X$ is a circular code if each word in $X^+$ written in a circle has a unique decomposition into words from $X$.

An analogue to codes in the family of D0L-systems are D0L-systems that are injective on the set of all factors of their languages. Circularity is defined as slightly relaxed injectivity: a D0L-system is circular if long enough factors of its language have a unique preimage (under the respective morphism) in the language except for some prefix and suffix bounded in length by some constant. This constant is called a synchronizing delay and it is studied in this paper.

In the case of D0L-systems circularity is connected with repetitiveness. As stated in \cite{2}, a non-circular D0L-system is repetitive, i.e., for each $k \in \mathbb{N}$ there exists a word $v$ such that $v^k$ is a factor of the language. In fact, if a D0L-system is not pushy (which is always true if the morphism is uniform and the language is infinite), then circularity is equivalent to non-repetitiveness \cite{3}.

As explained by Cassaigne in \cite{4}, knowledge of the value of the synchronizing delay can be very helpful when analysing the structure of bispecial factors in languages of D0L-systems. This idea was further developed by one of the authors in \cite{5}, where an algorithm for generating all bispecial factors is given. This algorithm works for...
circular and non-pushy D0L-systems and its computational complexity depends on the value of the synchronizing delay. This fact and the absence of any known bound on the value of synchronizing delay is the main motivation of the present work.

Unfortunately, it seems it is not easy to find such a bound. Therefore we focus on the simplest case: a circular D0L-system with binary $k$-uniform morphism with $k \geq 2$. Using the notion of the graph of overhangs introduced in Subsection 2.2 we prove the following result. The details of the proof are given in Section 3.

Theorem 1. If the morphism $\varphi$ of a circular D0L-system $\{(a, b), \varphi, a\}$ is $k$-uniform, then the minimum value of its synchronizing delay, denoted by $Z_{\text{min}}$, is bounded as follows:

(i) $Z_{\text{min}} \leq 8$ if $k = 2$,

(ii) $Z_{\text{min}} \leq k^2 + 3k - 4$ if $k$ is an odd prime number,

(iii) $Z_{\text{min}} \leq k^2 \left( \frac{k}{d} - 1 \right) + 5k - 4$ otherwise,

where number $d$ is the least divisor of $k$ greater than 1.

2. Preliminaries

A finite set of symbols is an alphabet, denoted by $\mathcal{A}$. The set of all finite words over $\mathcal{A}$ is denoted by $\mathcal{A}^*$, the empty word is $\varepsilon$ and $\mathcal{A}^+ = \mathcal{A}^* \setminus \{\varepsilon\}$. If a word $u \in \mathcal{A}^*$ is a concatenation of three words $x, y$ and $z$ from $\mathcal{A}^*$, i.e., $u = xyz$, the word $x$ is a prefix of $u$, $y$ its factor and $z$ a suffix. We put $x^{-1}u = yz$ and $ux^{-1} = xy$. The length of the word $u$ equals the number of letters in $u$ and is denoted by $|u|$; $|u|_{w}$ denotes the number of occurrences of a letter $a$ in $u$.

A mapping $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ is a morphism if for every $v, u \in \mathcal{A}^*$ we have $\varphi(vu) = \varphi(v)\varphi(u)$. A triplet $G = (\mathcal{A}, \varphi, w)$ is a D0L-system, if $\varphi$ is a morphism on $\mathcal{A}$ and $w \in \mathcal{A}^+$. The word $w$ is called an axiom. The language of $G$ is the set $L(G) = \{\varphi^n(w) : n \in \mathbb{N}\}$. The set of all factors of elements of $L(G)$ is denoted by $S(L(G))$. The alphabet is always considered to be the minimal alphabet necessary, i.e., $\mathcal{A} \cap S(L(G)) = \mathcal{A}$.

A D0L-system $G = (\mathcal{A}, \varphi, w)$ is injective on $S(L(G))$ if for every $u, v \in S(L(G))$, $\varphi(u) = \varphi(v)$ implies that $u = v$. It is clear that if $\varphi$ is injective, then $G$ is injective. If $\varphi$ is non-erasing, i.e., $\varphi(a) \neq \varepsilon$ for all $a \in \mathcal{A}$, then $G$ is a propagating D0L-system, shortly PD0L-system.

Given a D0L-system $G = (\mathcal{A}, \varphi, w)$, we say that a letter $a$ is bounded if the set $\{\varphi^n(a) : n \in \mathbb{N}\}$ is finite. If a letter is not bounded, it is unbounded. The system $G$ is pushy if $S(L(G))$ contains infinitely many factors containing bounded letters only.

A D0L-system $G$ is repetitive if for any $k \in \mathbb{N}$ there is a non-empty word $v$ such that $v^k$ is a factor from $S(L(G))$. By [5], any repetitive D0L-system is strongly repetitive, i.e., there is a non-empty word $v$ such that $v^k$ is a factor for all $k \in \mathbb{N}$. We say that $G$ is unboundedly repetitive if there is $v$ containing at least one unbounded letter such that $v^k$ is a factor for all $k \in \mathbb{N}$. 

2
2.1. Circular D0L-systems

In [4], a circular D0L-system is defined using the notion of synchronizing point (see Section 3.2 in [4] for details). We give here an equivalent definition employing the notion of interpretation.

**Definition 2.** Let \( G = (\mathcal{A}, \varphi, w) \) be a PD0L-system and \( u \in S(L(G)) \). A triplet \((p, v, s)\), where \( p, s \in \mathcal{A}^* \) and \( v \in S(L(G)) \), is an interpretation of the word \( u \) if \( \varphi(v) = pus \).

**Definition 3.** Let \( G = (\mathcal{A}, \varphi, w) \) be a PD0L-system. We say that two interpretations \((p, v, s)\) and \((p', v', s')\) of a word \( u \in S(L(G)) \) are synchronized at position \( k \) if there exist indices \( i \) and \( j \) such that

\[
\varphi(v_i \cdots v_j) = pu_1 \cdots u_k \quad \text{and} \quad \varphi(v'_i \cdots v'_{j}) = p'u_1 \cdots u_k
\]

with \( v = v_1 \cdots v_n \in \mathcal{A}^* \), \( v' = v'_1 \cdots v'_m \in \mathcal{A}^* \) and \( u = u_1 \cdots u_l \in \mathcal{A}^l \) (if \( k = 0 \), we put \( u_1 \cdots u_l = \varepsilon \)). Two interpretations that are not synchronized at any position are called non-synchronized.

We say that a word \( u \in S(L(G)) \) has a synchronizing point at position \( k \) with \( 0 \leq k \leq |u| \) if all its interpretations are pairwise synchronized at position \( k \).

**Definition 4.** Let \( G = (\mathcal{A}, \varphi, w) \) be a PD0L-system injective on \( S(L(G)) \). We say that \( G \) is circular if there is a positive integer \( Z \), called a synchronizing delay, such that any \( u \) from \( S(L(G)) \) longer than \( Z \) has a synchronizing point. The minimal constant \( Z \) with this property is denoted by \( Z_{\text{min}} \).

By the results from [2, 3], non-circular systems are repetitive (and by [6] also strongly repetitive). In fact, a DOL-system injective on \( S(L(G)) \) is not circular if and only if it is unboundedly repetitive [4], i.e., there exists \( v \) containing an unbounded letter such that \( v^k \in S(L(G)) \) for all \( k \in \mathbb{N} \). Since this property can be checked by a simple algorithm [7], we can easily verify whether a given DOL-system injective on \( S(L(G)) \) is circular or not.

The notion of circularity is inspired by the notion of circular code:

**Definition 5.** A subset \( X \) of \( \mathcal{A}^* \) is called a code over alphabet \( \mathcal{A} \) if for any word \( v \in X^* \) there are uniquely given a number \( n \) and words \( x_1, x_2, \ldots, x_n \) from \( X \) so that \( v = x_1 x_2 \cdots x_n \).

The set \( X \) is a circular code over \( \mathcal{A} \) if for all \( n, m \geq 1 \), \( x_1, \ldots, x_n, y_1, \ldots, y_m \in X \), \( p \in \mathcal{A}^* \) and \( s \in \mathcal{A}^* \) it holds that:

\[
(sx_2x_3 \cdots x_nps = y_1y_2 \cdots y_m \text{ and } x_1 = \varepsilon) \Longrightarrow (n = m, p = \varepsilon \text{ and } x_i = y_i \forall i = 1, \ldots, n).
\]

2.2. Graph of overhangs

Here we introduce the basic tool we use to prove the main result of this paper: graphs of overhangs. To understand the motivation of the definition let us consider a D0L-system \( G = (\mathcal{A}, \varphi, w) \) and a word \( u \) from \( S(L(G)) \) that has two non-synchronized interpretations \((p_1, x, s_1)\) and \((p_2, y, s_2)\) with \( x = x_1 x_2 x_3 x_4, x_i \in \mathcal{A} \) and \( y = y_1 y_2 y_3 y_4 \).

Let \( \varphi(x_1) = u_i, \varphi(x_2) = u_i, \varphi(x_3) = u_i, \varphi(x_4) = u_i \) and \( \varphi(y_1) = u_j, \varphi(y_2) = u_j, \varphi(y_3) = u_j, \varphi(y_4) = u_j \) (see the top line in Figure[1]). This structure can be decomposed
into overhangs (see Definition 6 and the second line in Figure 1). Each overhang has left and right overlapping words, in the figure denoted by \( s_1, s_2, s_3 \) and \( s_4 \). The structure of these overhangs can be captured as a graph (see Definition 8 and the bottom line in Figure 1): for instance the first overhang is connected with an directed edge with the second overhang since their right and left overlapping words are equal to \( s_2 \).

Figure 1: The construction of a graph of overhangs, the numbers \( n_1, n_2 \) and \( n_3 \) are the lengths of overlaps of \( u_{i_1} \) and \( u_{i_2} \), \( u_{i_3} \) and \( u_{i_4} u_{i_5} \), and \( u_{i_6} u_{i_7} \) and \( u_{i_8} \), respectively.

**Definition 6.** Given a PD0L-system \((\mathcal{A}, \varphi, w)\), let \( X = \{\varphi(a) \mid a \in \mathcal{A}\} \). An ordered triplet \((u_1 u_2 \cdots u_m, v_1 v_2 \cdots v_n, |x|)\), where \( x \) is a non-empty word over \( \mathcal{A} \) and \( u_i \) and \( v_j \) are elements of \( X \) for all \( 1 \leq i \leq m \) and \( 1 \leq j \leq n \), is called an overhang if the following conditions are satisfied:

(i) \( x \) is a suffix of \( u_1 u_2 \cdots u_m \) but not of \( u_2 u_3 \cdots u_m \).

(ii) \( x \) is a prefix of \( v_1 v_2 \cdots v_n \) but not of \( v_1 v_2 \cdots v_{n-1} \).

(iii) either \( x \neq u_1 u_2 \cdots u_m \) or \( x \neq v_1 v_2 \cdots v_n \).

(iv) \(|v_1 \cdots v_{n-1}| < |x(u_2 \cdots u_m)^{-1}|\).

Here again we put \( v_1 \cdots v_{n-1} = \varepsilon \) if \( n = 1 \) and \( u_2 \cdots u_m = \varepsilon \) if \( m = 1 \).

The word \( x \) is called the common factor of the overhang and the words \( u_1 u_2 \cdots u_m x^{-1} \) and \( x^{-1} v_1 v_2 \cdots v_n \) are left and right overhang, respectively.

**Example 7.** Consider the PD0L-system \( G = ((0, 1, 2), \varphi, 0) \) with \( \varphi \) given by \( 0 \rightarrow 011, 1 \rightarrow 1120, 2 \rightarrow 120 \). The set \( X \) in the definition above reads \( \{011, 1120, 120\} \).

To find all possible overhangs, we have to go through all elements \((u, v)\) of \( X \times X \) and look for a non-empty word \( x \) which is both a suffix of \( u \) and prefix of \( v \) and for \( z \in X^* \) such that \((uz, v, |x|)\) or \((u, zv, |x|)\) is an overhang. After doing so we find
Definition 8. Given a PD0L-system $G = (A, \varphi, w)$. A graph of overhangs $GO_G$ for the PD0L-system $G$ is given by this rule: words $s_1$ and $s_2$ are vertices of $GO_G$ connected with a directed edge $(s_1, s_2)$ labelled by $(u_1 u_2 \cdots u_m, v_1 v_2 \cdots v_n, |x|)$ if the triplet $(u_1 u_2 \cdots u_m, v_1 v_2 \cdots v_n, |x|)$ is an overhang such that $s_1$ is its left and $s_2$ its right overhang.

Example 9 (continued). The graph of overhangs for the PD0L-system from Example 7 is depicted in Figure 2.

Figure 2: The graph of overhangs for Example 7.

To demonstrate usefulness of graphs of overhangs for the study of synchronizing delay, we first consider the simpler case of circular codes.

Lemma 10. Given a PD0L-system $G = (A, \varphi, w)$. The set $X = \{ \varphi(a) \mid a \in A \}$ is a code if and only if the graph of overhangs $GO_G$ does not contain a cycle containing the empty word as a vertex.

The set $X$ is a circular code if and only if its graph of overhangs does not contain any cycle.

Proof. We prove the second part of the statement, the first part can be proved analogously.

Assume that the graph of overhangs contains a cycle

$$s_1 \xrightarrow{(v_1 s_1|x)} s_2 \xrightarrow{(v_2 s_2|x)} \cdots \xrightarrow{(v_l s_l|x)} s_1.$$

For the word $y = s_1 x_1 s_2 \ldots s_l x_l$ we have

$$y = s_1 x_1 s_2 \ldots s_l x_l = v_1 v_2 \cdots v_l = s_1 u_1 u_2 \ldots u_l (s_1)^{-1},$$

where $v_i, u_j \in X^*$. It follows that $X$ is not a circular code by definition.

If $X$ is not a circular code, there is a word $z \in X^*$ such that

$$z = v_1 v_2 \cdots v_n = s u_1 u_2 \ldots u_m (s)^{-1},$$

where $v_i, u_j \in X$ and $s$ is a suffix of $u_m$. These two decompositions of $z$ give an analogue of two non-synchronized interpretation of $z$. Using the same procedure as in Figure 1.
we can find numbers $1 \leq n_1 < \cdots < n_{k-1} < n$ and $1 \leq m_1 < \cdots < m_{k-1} < m$ and words $x_i$ and $s_i$, $1 \leq i \leq k-1$, so that we get the following cycle in the graph of overhangs:

$$s \xrightarrow{(v_1, \ldots, v_{n_1}, u_1, \ldots, u_{m_1}, |x_1|)} s_2 \cdots s_{k-1} \xrightarrow{(v_{n_{k-1}}, \ldots, v_{n_1}, u_{m_{k-1}}, \ldots, u_{m_1}, |x_1|)} s.$$

To explain the connection between circular codes and circular D0L-systems, consider a D0L-system \( G = (\mathcal{A}, \varphi, w) \) that is not circular. It implies that for any \( L \in \mathbb{N} \) there must be a word longer than \( L \) with two non-synchronized interpretations. These two non-synchronized interpretations can be decomposed into overhangs as in Figure 1. If \( L \) is big enough, there must be two overhangs so that the left overhang of one of them equals the right overhang of the other. Therefore they can be glued together to form a cycle which means that the set \( X = \{ \varphi(a) \mid a \in \mathcal{A} \} \) is not a circular code.

By Lemma 10 and by the previous paragraph, if a D0L-system \( G \) is not circular, then \( GOG \) must contain an infinite walk and this happens if and only if \( GOG \) contains a cycle. However, existence of a cycle in \( GOG \) is not a sufficient condition as follows from the following example.

**Example 11.** It is known that Thue-Morse D0L-system \( G_{TM} = ((0,1), \varphi_{TM}, 0) \), where \( \varphi_{TM}(0) = 01 \) and \( \varphi_{TM}(1) = 10 \), is circular with minimal synchronizing delay \( Z_{\min} = 3 \). However, its graph of overhangs contains two cycles, see Figure 3. These two cycles correspond to words \((01)^j\) or \((10)^j\) for \( j = 1, 2, 3, \ldots \), but such words are in \( S(L(G_{TM})) \) only for \( j \leq 2 \).

![Figure 3: Graph of overhangs of the D0L-system \( G_{TM} = ((0,1), \varphi_{TM}, 0) \).](image)

**Definition 12.** Let \( \ell > 1 \) and

$$s_1 \xrightarrow{(v_1, \ldots, v_{n_1}, |x_1|)} s_2 \xrightarrow{(v_{n_1}, \ldots, v_1, u_1, \ldots, u_{m_1}, |x_1|)} s_3 \cdots s_\ell \xrightarrow{(v_1, \ldots, v_{n_1}, u_1, \ldots, u_{m_1}, |x_1|)} s_{\ell+1}$$

be a walk in the graph of overhangs of a PD0L-system \( G = (\mathcal{A}, \varphi, w) \). We call this walk \( G \)-admissible if both the words \( v_1 \cdots v_{\ell} \) and \( u_1 \cdots u_{\ell} \) are in \( S(L(G)) \) and \( u_i, v_i \in \{ \varphi(a) \mid a \in \mathcal{A} \} \) for all \( 1 \leq i \leq \ell \).

The number \( |(s_1)^{-1}v_1 \cdots v_\ell| \) is called the word-length of the walk.

With these definitions, we can state a sufficient and necessary condition for \( G \) being non-circular: \( G \) is non-circular if for any \( L \in \mathbb{N} \) there is a \( G \)-admissible walk in \( GOG \) of word-length greater than \( L \). Moreover, if \( G \) is circular, then the word-longest \( G \)-admissible walk is tightly connected with the longest non-synchronized factor and so with the constant \( Z_{\min} \).
Lemma 13. \textit{Let }G\text{ be a circular PD0L-system. Let }L_{\text{max}}\text{ denote the maximum }L\text{ such that there is a }G\text{-admissible walk in }GO_G\text{ of word-length }L.\text{ It holds that }

\[ L_{\text{max}} \leq Z_{\text{min}} \leq L_{\text{max}} + 2M - 3, \]

where \( M = \max_{a \in A} |\varphi(a)|. \)

\textit{Proof.} Let the word-longest walk in \( GO_G \) be denoted by

\[ p = s_1 \overset{(v_1, u_1, x_1)}{\rightarrow} s_2 \overset{(v_2, u_2, x_2)}{\rightarrow} \ldots s_{\ell} \overset{(v_\ell, u_\ell, x_\ell)}{\rightarrow} s_{\ell+1}. \]

By the construction of \( GO_G \) we have that \((s_1)^{-1}v_1 \cdots v_\ell\) is non-synchronized. Since \( Z_{\text{min}} \) equals the length of the longest non-synchronized factor from \( S(L(G)) \), it holds that \(|(s_1)^{-1}v_1 \cdots v_\ell| = L_{\text{max}} \leq Z_{\text{min}}. \) To obtain an upper bound, consider the longest possible word \( x \) such that \( xu_1 \cdots u_\ell(s_{\ell+1})^{-1} \) is a non-synchronized factor in \( S(L(G)) \). It is easy to see that if \(|x| \geq M - 1\), then we can find a word-longer walk in \( GO_G \) than \( p \). Therefore \(|x| \leq M - 2. \) Analogously we prove that if \((s_1)^{-1}v_1 \cdots v_\ell\) is non-synchronized, than \(|y| \leq M - 1 \) (see Figure 4).

![Figure 4: The bold line denotes the longest possible prolongation of a non-synchronized factor corresponding to the word-longest walk in }GO_G\text{.}]

In the following sections we use the last lemma to get an upper bound on \( Z_{\text{min}} \) by finding the word-longest possible walk in \( GO_G \) for \( G \) with uniform binary morphisms.

3. Proof of the main result

There is no known estimate on the (minimal) synchronizing delay of a PD0L-system. Here we restrict ourselves to the simplest class of morphisms: uniform morphisms defined on a binary alphabet. We further assume that the axiom of the D0L-system is a letter; considering a general case would not bring much novelty but technical difficulties.

\textbf{Definition 14.} A morphism \( \varphi \) on the binary alphabet \{a, b\} is \textit{k-uniform} for some integer \( k \geq 2 \) if \(|\varphi(a)| = |\varphi(b)| = k. \)
Because of the simple structure of DOL-systems with a \( k \)-uniform morphism, we can find an explicit list of morphisms that are not circular. We state a proof here as it is quite simple using the results from \([7]\). However, a complete characterisation of repetitive (and so non-circular) DOL-systems with binary morphisms was done in \([8]\). Partial characterisation of circular DOL-systems with a binary \( k \)-uniform morphism was also given in \([9, 10]\).

**Lemma 15.** A DOL-system \( G = (\{a, b\}, \varphi, a) \) with \( k \)-uniform morphism \( \varphi \) is not circular if and only if one of the following conditions is satisfied:

(i) \( \varphi(a) = \varphi(b) \),

(ii) \( \varphi(a) = a^k \) or \( \varphi(b) = b^k \),

(iii) \( \varphi(a) = b^k \) and \( \varphi(b) = a^k \),

(iv) \( k = 2m + 1, m \geq 1, \) and \( \varphi(a) = (ab)^m a, \varphi(b) = (ba)^m b, \)

(v) \( k = 2m + 1, m \geq 1, \) and \( \varphi(a) = (ba)^m b, \varphi(b) = (ab)^m a. \)

**Proof.** As explained above, circular DOL-systems are not unboundedly repetitive. Since the DOL-system \( G \) cannot contain a bounded letter (as it is \( k \)-uniform with \( k > 1 \)), it is unboundedly repetitive if and only if it is repetitive. Hence, we need to find all \( \varphi \) for which the DOL-system is non-injective or repetitive. Clearly, a \( k \)-uniform morphism \( \varphi \) is not injective if and only if \( \varphi(a) = \varphi(b) \) (condition (i)). It follows from \([7]\) that \( G \) is repetitive if and only if there is a primitive word \( u \) (i.e., a word for which \( u^j \ell \) implies \( \ell = 1 \)) in \( S(L(G)) \) such that \( \varphi^j(u) = u' \) for some integers \( j \geq 1 \) and \( \ell \geq 2 \). Moreover, the factor \( u \) cannot contain any unbounded letter twice and \( j \) is bounded by the number of letters in the alphabet: hence we must have \( u \in \{a, b, ba, ab\} \) and \( j \in \{1, 2\} \). Analysing all these eight possible cases we obtain conditions (ii) – (v).

3.1. List of all possible components of graphs of overhangs

In general, even for a simple-looking morphism over a small alphabet the corresponding graph of overhangs can be of very complex structure. However, in the case of (binary) uniform morphisms, these graphs are always divided into components containing relatively small number of vertices; for a binary uniform morphism this number is at most four as follows from these two simple observations:

**Lemma 16.** Given a DOL-system \( G = (\mathcal{A}, \varphi, w) \) with a \( k \)-uniform morphism \( \varphi \). For the graph of overhangs of \( G \) the following holds:

(i) A label of any edge is of the form of \( (\varphi(a), \varphi(b), \ell) \) with \( a, b \in \mathcal{A} \) and \( 0 < \ell < k \).

(ii) If there is an edge from a vertex \( s_1 \) to a vertex \( s_2 \) with label \( (\varphi(a), \varphi(b), \ell) \), then the words \( s_1 \) and \( s_2 \) are of the same length \( k - \ell \).

Applied on the case of binary morphism we obtain this:
Corollary 17. Given a D0L-system $G = (\mathcal{A}, \varphi, w)$ with an injective $k$-uniform morphism $\varphi$ defined over the binary alphabet $\mathcal{A} = \{a, b\}$. The graph of overhangs $GO_G$ consists of weakly connected components each containing at most four vertices. Labels of these vertices are all of the same length $k - \ell$, $0 < \ell < k$, and the edges are labelled by four possible labels: $(\varphi(a), \varphi(a), \ell), (\varphi(a), \varphi(b), \ell), (\varphi(b), \varphi(a), \ell), (\varphi(b), \varphi(b), \ell)$. 

Proof. It is a direct consequence of the previous lemma that $GO_G$ is divided into weakly connected components and also that within each component all possible labels of its edges are four overhangs $(\varphi(a), \varphi(a), \ell), (\varphi(a), \varphi(b), \ell), (\varphi(b), \varphi(a), \ell), (\varphi(b), \varphi(b), \ell)$. Further, the left overhangs of $(\varphi(a), \varphi(a), \ell)$ and $(\varphi(a), \varphi(b), \ell)$ are the same words. The same is true for $(\varphi(b), \varphi(a), \ell)$ and $(\varphi(b), \varphi(b), \ell)$. Similarly, the right overhangs are equal words for $(\varphi(a), \varphi(a), \ell)$ and $(\varphi(b), \varphi(a), \ell)$ and for $(\varphi(a), \varphi(b), \ell)$ and $(\varphi(b), \varphi(b), \ell)$ (see Figure 5). It follows that each component has at most four vertices. 

![Figure 5: An illustration of a weakly connected component of a graph of overhangs for a binary $k$-uniform morphism.](image)

As explained above, we are interested in the word-longest $G$-admissible walks (see Definition 12) in the graphs of overhangs. The previous corollary says that it is sufficient to look for such walks only in components (in what follows by components we always mean the weakly connected components from Corollary 17) with at most four vertices. The component in Figure 5 can be considered as maximal, since all other components that can appear in the graph of overhangs for an injective binary $k$-uniform morphism arise by merging some vertices of it and by omitting some edges. In other words, all possible components are subgraphs of the graphs depicted in Figures 5, 6, 7, and 8 (labels are denoted by $\varphi(x)$ instead of $(\varphi(x), \varphi(y), \ell)$ in order to save some space).

3.2. Infeasible subgraphs

Here we prove that the components of the graph of overhangs of circular D0L-systems with a binary $k$-uniform morphism cannot contain the following subgraphs: two loops on one vertex (Figure 9a), a vertex with a loop that is at the same time a vertex of a cycle over two vertices (Figure 9b) and two vertices with a loop that are connected with an edge (Figure 9c). All the proofs are quite similar: we use the following three auxiliary lemmas to get a contradiction with circularity; namely, we show that these subgraphs appear in the graphs only for non-circular D0L-systems, i.e., only for morphisms listed in Lemma 15.

Lemma 18 (Lyndon, Schützenberger [11]). Let $x, y \in \mathcal{A}^+$. The following three conditions are equivalent:

\[\begin{array}{c}
\text{Lemma 18 (Lyndon, Schützenberger [11]). Let } x, y \in \mathcal{A}^+. \text{ The following three conditions are equivalent:}
\end{array}\]
Figure 6: All possible components with three vertices arise by omitting edges in one of these six graphs.

(i) $xy = yx$;
(ii) There exist integers $i, j > 0$ such that $x^i = y^j$;
(iii) There exist $z \in \mathcal{A}^*$ and integers $p, q > 0$ such that $x = z^p$ and $y = z^q$.

**Lemma 19** (Krieger [12]). Let $x, y, z, t \in \mathcal{A}^+$. Assume the following equalities hold:

(i) $xy = zt$ (equivalently, $ty = zx$);
(ii) $yx = xz$.

Then $y = z$ and $x = t$.

**Lemma 20** (Krieger [12]). Let $x, y, z, t \in \mathcal{A}^+$. Assume the following equalities hold:

(i) $xy = yt$;
(ii) $tz = zx$.

Then there exists $u \in \mathcal{A}^*, v \in \mathcal{A}^*$ and integers $i \geq 1$ and $j, m \geq 0$ such that $x = (uv)^i$, $t = (yu)^j$, $y = (uv)^j u$, $z = (vu)^m v$. If in addition $|v| = |z|$, then either $v = \epsilon$ and $m = j + 1$, or $|u| = |v|$ and $m = j$.

**Lemma 21.** Let $G = (\mathcal{A}, \varphi, a)$ be a circular D0L-system with a binary $k$-uniform morphism $\varphi$ and $GO_G$ its graph of overhangs. Then $GO_G$ does not contain two loops on a common vertex (see Figure 9(a)) as its subgraph.

**Proof.** Assume that there is a component containing a vertex $s$ such that $|s| = k - \ell$ and there are two loops starting and ending in $s$. We have six possible pairs of labels for the two edges of the loops (see Figure 8). We consider only the following three cases since the others can be handled analogously.
First assume that the two loops on \( s \) are labelled with \((\varphi(a), \varphi(a), \ell)\) and \((\varphi(b), \varphi(a), \ell)\). We must have \( \varphi(a) = sy = ys = zs \) and \( \varphi(b) = sz \). It follows that \( z = y \) and therefore \( \varphi(a) = \varphi(b) \) and \( G \) is not circular.

Now assume the two labels read \((\varphi(a), \varphi(a), \ell)\) and \((\varphi(b), \varphi(b), \ell)\). We get \( \varphi(a) = sy = ys = zs \). By Lemma \([20]\) applied to \( x = t = s \) we must have \( s = (uv)^j = (vu)^j, y = (uv)^j u \) and \( z = (vu)^j v \) with integers \( i \geq 1, j, m \geq 0 \). Since \( vuv = vu \), Lemma \([18]\)(iii) says that both \( u \) and \( v \) are powers of the same word and hence we have again \( \varphi(a) = \varphi(b) \).

Finally, assume the labels are \((\varphi(a), \varphi(b), \ell)\) and \((\varphi(b), \varphi(a), \ell)\). It holds that \( \varphi(a) = \varphi(b) \).
Theorem 22. Let \( G = (\mathcal{A}, \varphi, a) \) be a circular D0L-system with a binary \( k \)-uniform morphism \( \varphi \) and \( GO_G \) its graph of overhangs. Then \( GO_G \) does not contain a cycle on two vertices with a loop on one of them (see Figure 10) as its subgraph.

Proof. The proof is quite similar to the previous one: we again assume that the structure can be found in the graph on vertices \( s_1 \) and \( s_2 \). The words \( s_1 \) and \( s_2 \) must be of the same length which is denoted by \( k - t \). As follows from the Figure 7(f) and g), the labels of the cycle are either \((\varphi(a), \varphi(a), \ell)\) and \((\varphi(b), \varphi(b), \ell)\) or \((\varphi(a), \varphi(b), \ell)\) and \((\varphi(b), \varphi(a), \ell)\).

Assume the former case is true, the latter one can be handled analogously. We get \( \varphi(a) = s_1y = ys_2, \varphi(b) = s_2z = zs_1 \) and \( |y| = |z| \). Due to Lemma 20 we can write \( \varphi(a) = (uv)^{m+1}u, \varphi(b) = (vu)^{m+1}v, \) \( s_1 = (uv)^i \) and \( s_2 = (vu)^i \) for some \( u \in \mathcal{A}^* \), \( v \in \mathcal{A}^* \), and integers \( i \geq 1 \) and \( j, m \geq 0 \). If \( v = \epsilon \) and \( m = j + 1 \), then \( \varphi(a) = \varphi(b) \) which contradicts circularity, therefore we must have \( \varphi(a) = (uv)^{m+1}u, \varphi(b) = (vu)^{m+1}v \) with \( u \) and \( v \) non-empty and of the same length.

Assume that the loop is on the vertex \( s_1 \) (the other case is again analogous), the label of the loop must read \((\varphi(a), \varphi(a), \ell)\). Such an overhang can exist only if \( uv = vu \). This is again a contradiction since \( \varphi(a) = \varphi(b) \).

Lemma 23. Let \( G = (\mathcal{A}, \varphi, a) \) be a circular D0L-system with a binary \( k \)-uniform morphism \( \varphi \) and \( GO_G \) its graph of overhangs. Then \( GO_G \) does not contain two vertices with a loop that are also connected with an edge (see Figure 9(c)) as its subgraph.

Proof. This structure can appear in \( GO_G \) only as a subgraph of the graphs f) and g) in Figure 7.

Assume the loops are on the vertices \( s_1 \) with label \((\varphi(a), \varphi(a), \ell)\) and on \( s_2 \) with label \((\varphi(b), \varphi(b), \ell)\) and the edge from \( s_1 \) to \( s_2 \) is labelled by \((\varphi(a), \varphi(b), \ell)\) (the other case can be obtained just by exchanging the letters \( a \) and \( b \)). It must hold that \( \varphi(a) = s_1x = xs_1 = s_1y \) and \( \varphi(b) = yzs_2 = s_2z = zs_2 \). It follows that \( x = y = z \) and so \( \varphi(a) = \varphi(b) \) by Lemma 18(iii). Hence \( G \) is not injective and so non-circular.

Assume now the loops are on the vertex \( s_1 \) with label \((\varphi(a), \varphi(b), \ell)\) and on the vertex \( s_2 \) with label \((\varphi(b), \varphi(a), \ell)\) and the edge from \( s_1 \) to \( s_2 \) is labelled with \((\varphi(a), \varphi(a), \ell)\) (the other case can be obtained just by exchanging the letters \( a \) and \( b \)). It holds that \( \varphi(a) = s_1x = s_1y = ys_2 = zs_2 \) and \( \varphi(b) = xs_1 = s_2z \). It follows that \( x = y = z \) and by Lemma 19 applied for \( x = t \) we have \( \varphi(a) = \varphi(b) \).

3.3. Word-longest walks

If we put all the previous results together, we can say that any walk in the graph of overhangs of a circular D0L-system \( G \) with a binary \( k \)-uniform morphism is of one of the shapes depicted in Figure 10. As we are looking for the word-longest \( G \)-admissible walk, it suffices to focus on walks in graphs a) and g) in Figure 10. However, the word-length of a \( G \)-admissible walk depends on how many times such a walk goes along the edge of the loop (in case a)) or along the both edges of the cycle (case g)). Denote the
maximal number of those transitions along the edges by $R_1$ and $R_2$, respectively. We recall that these numbers must be finite for circular D0L-systems.

Denote the three vertices in graph a) in Figure 10 by $s_1, s_2, s_3$ ($s_2$ is the vertex with the loop) and the common factors of the labels of the respective edges by $x_1, x_2, x_3$. We set $R_1$ to the maximum $R \in \mathbb{N}$ such that the walk $s_2 \rightarrow s_2 \rightarrow \cdots \rightarrow s_2$ $R$ arrows is $G$-admissible for some circular D0L-system $G$ with a binary $k$-uniform morphism.

The word-length of any $G$-admissible walk in any subgraph of graph a) in Figure 10 is bounded by $|x_1(s_2s_2)^R|s_2s_3| = |x_1| + R_1k + k \leq R_1k + 2k - 1$.

Similarly, we denote the vertices of the cycle in graph g) in Figure 10 by $s_1$ and $s_2$ and the common factors of the labels by $x_1$ and $x_2$. We set $R_2$ to the maximum $R \in \mathbb{N}$ such that the walk $s_1 \rightarrow s_2 \rightarrow \cdots \rightarrow s_2 \rightarrow s_1$ $2R$ arrows is $G$-admissible for some circular D0L-system $G$ with a binary $k$-uniform morphism. The word-length of any $G$-admissible walk in any subgraph of graph g) in Figure 10 is bounded by $|(s_1s_2)^R|s_1s_1| = 2kR_2 + |x_1| \leq 2kR_2 + k - 1$ (we have considered the walk that goes through the cycle $(R_2 + 1/2)$ times).

By Lemma 13, the minimum synchronizing delay $Z_{min}$ is equal to or less than the number $\max[R_1k + 4k - 4, 2kR_2 + 3k - 4]$. It remains to find an upper bound on $R_1$ and $R_2$. We start with $R_1$: the label of the loop in graph a) in Figure 10 is one of these four overhangs: $(\varphi(a), \varphi(a), \ell), (\varphi(a), \varphi(b), \ell), (\varphi(b), \varphi(b), \ell)$ and $(\varphi(b), \varphi(a), \ell)$. As the last two cases are analogous to the first or the second one, we focus only on labels $(\varphi(a), \varphi(a), \ell)$ and $(\varphi(a), \varphi(b), \ell)$: in the former case $R_1 = R_a$ and in the latter case $R_1 = \min\{R_a, R_b\}$, where $R_a = \max\{\ell \in \mathbb{N} : a^\ell \in S(L(G))\}$ and $R_b = \max\{\ell \in \mathbb{N} : b^\ell \in S(L(G))\}$.

**Lemma 24.** Let $G = \{A, \varphi, a\}$ be a circular D0L-system with a binary $k$-uniform morphism $\varphi$ and $GO_G$ its graph of overhangs containing a vertex $s$ with a loop labelled with $(\varphi(a), \varphi(a), \ell)$. Then the following holds:

1. $R_a \leq k - 1$ if $k$ is a prime number,
2. \( R_a \leq k\left(\frac{k}{2} - 1\right) + 1 \) otherwise.

where number \( d \) is the least divisor of \( k \) greater than 1. These bounds are attained for some DOL-system.

**Proof.** We find a morphism \( \varphi \) so that the value of \( R_a \) for the resulting DOL-system is maximum. We have that \( \varphi(a) = xy = yx \) and \( \varphi(b) \) is arbitrary. By Lemma [18] there exists a word \( z \) such that \( \varphi(a) = z^m \) for some \( z \in \mathcal{A}^+ \) and integer \( m \geq 2 \). Since \( G \) is circular, one of the following is true by Lemma [15]:

(i) \( \varphi(a) = z^m, \varphi(b) = a^k \), and \( |z|_a > 0, |z|_b > 0 \). This is possible only if \( k \) is not prime.

(ii) \( \varphi(a) = z^m, \varphi(b) = \ell \), where \( |z|_a > 0, |z|_b > 0, |w|_a > 0 \) and \( |w|_b > 0 \). This is possible only if \( k \) is not prime, too.

(iii) \( \varphi(a) = b^q, \varphi(b) = w \), where \( |w|_a > 0 \) and \( |w|_b > 0 \).

We analyse these three cases separately.

Case (i): Let \( p, q \) denote the maximum numbers such that \( a^p \) is a prefix of \( \varphi(a) \) and \( a^q \) is suffix of \( \varphi(a) \). We have that \( R_a = R_b + p + q \). Since \( p + q < |z| \leq \frac{k}{2} \), the maximum value of \( R_a \) is attained if we maximize the value of \( R_b \) (regardless of the value of \( p + q \)). The letter \( b \) is present only in \( \varphi(a) \), thus \( R_b \) is given by the maximal power of \( b \) in the word \( zz \) (note that \( z \) contains the letter \( a \)). So \( R_b \) is maximum possible if \( m = d \), where \( d \) is the least divisor of \( k \) greater than 1. We get \( |z| = \frac{k}{2} \) and so the maximum possible value of \( R_b \) is \( \frac{k}{2} - 1 \). It follows that \( p + q \leq 1 \) and \( R_a \leq k\left(\frac{k}{2} - 1\right) + 1 \). This bound is attained for \( \varphi(a) = (ab^{k-1})^d, \varphi(b) = a^k \).

Case (ii): We have \( |z|_a \leq \frac{k}{2} - 1 \) and \( |\varphi(b)|_a \leq k - 1 \), where \( d \) is again the least divisor of \( k \) greater than 1. The value of \( R_a \) is equal to the maximum value of \( \ell \) such that \( a^\ell \) is a factor of \( \varphi(a)\varphi(a), \varphi(a)\varphi(b), \varphi(b)\varphi(a) \) or \( \varphi(b)\varphi(b) \). It is easy to see that \( R_a \leq k - 1 + \frac{k}{2} - 1 = \frac{3k - 4}{2} - 2 \).

Case (iii): The letter \( a \) is present only in the word \( \varphi(b) = w \), thus the maximum power of \( a \) in \( S(L(G)) \) is equal to the maximum power of \( a \) that appears in the word \( ww \). Since \( |w|_b > 0 \), we get \( R_a \leq k - 1 \). This bound is attained for \( \varphi(a) = b^q, \varphi(b) = b^{d-k-1} \).

If \( k \) is a prime number, the proof is finished with \( R_a \leq k - 1 \). Let \( k \) be not prime and \( d \) its least divisor greater than one. It is an easy exercise to show that

\[
k - 1 < k \leq \frac{k(d + 1)}{d} - 2 < k\left(\frac{k}{d} - 1\right) + 1,
\]

which concludes the proof. \( \square \)

**Lemma 25.** Let \( G = [\mathcal{A}, \varphi, a] \) be a circular DOL-system with a binary \( k \)-uniform morphism \( \varphi \) and \( GO_{\ell} \) its graph of overhangs containing a vertex \( s \) with a loop labelled with \( (\varphi(a), \varphi(b), \ell) \). Then for \( R_1 = \max\{R_a, R_b\} \) the following holds:

1. \( R_1 \leq k \) if \( k \) is even.

2. \( R_1 \leq k - 1 \) if \( k \) is odd.
These bounds are attained for some D0L-system.

Proof. It follows from the assumptions that \( \varphi(a) = xy, \varphi(b) = yx \), where \( x, y \) are non-empty words, and \( |\varphi(a)|_a = |\varphi(b)|_a > 0 \) and \( |\varphi(a)|_b = |\varphi(b)|_b > 0 \).

To find the largest powers of \( a \) and \( b \) it suffices to go through this four factors of length \( 2k \): \( \varphi(a)\varphi(a), \varphi(a)\varphi(b), \varphi(b)\varphi(a) \) and \( \varphi(b)\varphi(b) \). We must have \( R_a + R_b \leq 2k \).

Since \( R_1 = \min(R_a, R_b) \), its maximal possible value for even \( k \) is attained if \( R_a = R_b = k \), if \( k \) is odd, the maximum value of \( R_1 \) is \( k - 1 \). These maximum values of \( R_1 \) are attained for the morphism \( \varphi(a) = a^{\frac{k}{2}}b^{\frac{k}{2}}, \varphi(b) = b^{\frac{k}{2}}a^{\frac{k}{2}} \).

It remains to consider the cycle (i.e., case g) in Figure [10]. The labels of the cycle are either \( (\varphi(a), \varphi(a), \ell) \) and \( (\varphi(b), \varphi(b), \ell) \) or \( (\varphi(a), \varphi(b), \ell) \) and \( (\varphi(b), \varphi(a), \ell) \). We consider only the former case since the latter one is analogous. Let \( R_{ab} = \max(\ell \in \mathbb{N} : (ab)\ell \in S(L(G))) \) and \( R_{ba} = \max(\ell \in \mathbb{N} : (ba)\ell \in S(L(G))) \), then \( R_2 = \min(R_{ab}, R_{ba}) \).

Lemma 26. Let \( G = (A, \varphi, \alpha) \) be a circular D0L-system with binary \( k \)-uniform morphism \( \varphi \) and \( GO_G \) its graph of overhangs containing a cycle with labels \( (\varphi(a), \varphi(a), \ell) \) and \( (\varphi(b), \varphi(b), \ell) \). Then it holds that \( R_2 \leq \frac{k^2}{2} \).

Proof. As explained in the proof of Lemma [22] we know that \( \varphi(a) = (uv)^{i+1}u, \varphi(b) = (vu)^{i+1}v \) for some words \( u, v \in A^* \), \( |u| = |v| \), \( u \neq v \) and integers \( i \geq 1, j \geq 0 \). By Lemma [13] one of the words \( \varphi(a) \) or \( \varphi(b) \) must contain \( aa \) or \( bb \) as a factor (otherwise the D0L-system \( G \) is not circular). Thus, some of the factors \( u, v, uv, vu \) contain \( aa \) or \( bb \). In order to find the maximum number of repetitions of \( ab \) and \( ba \) in \( S(L(G)) \), we can restrict ourselves to words \( \varphi(a)\varphi(a), \varphi(a)\varphi(b), \varphi(b)\varphi(a) \) and \( \varphi(b)\varphi(b) \). Obviously numbers \( R_{ab} \) and \( R_{ba} \) will be largest possible, if \( u \) and \( v \) are longest possible, i.e., \( \varphi(a) = uvu, \varphi(b) = vuv \). Since \( R_2 = \min(R_{ab}, R_{ba}) \), we get \( R_2 \leq \frac{k^2}{2} \).

As explained at the beginning of this subsection the minimum synchronizing delay \( Z_{\text{min}} \) is equal to or less than the number \( \max(R_1k + 4k - 4, 2kR_2 + 3k - 4) \). If \( k = 2 \), then the maximum value of \( R_1 \) is 2 (by Lemma [25]) and \( R_2 = 0 \) (as a cycle cannot appear in \( GO_G \)), it follows that \( Z_{\text{min}} \leq 2k + 4k - 4 = 8 \).

If \( k \) is an odd prime number, then the maximum value of \( R_1 \) is \( k - 1 \). Since \( R_2 \) is less than or equal to \( (k - 2)/2 \) we have

\[
Z_{\text{min}} \leq \max(R_1k + 4k - 4, 2kR_2 + 3k - 4) = \max(k^2 + 3k - 4, k^2 + k - 4) = k^2 + 3k - 4.
\]

Finally, if \( k \) is not prime and not equal to 2, the maximal value of \( R_1 \) equals

\[
\max \left( k \left( \frac{k}{d} - 1 \right) + 1, k \right) = k \left( \frac{k}{d} - 1 \right) + 1,
\]

where \( d \) is the smallest divisor of \( k \) greater than one. Since \( R_2 \) is still less than or equal to \( (k - 2)/2 \), we have that

\[
Z_{\text{min}} \leq R_1k + 4k - 4 = k^2 \left( \frac{k}{d} - 1 \right) + 5k - 4.
\]

This concludes the proof of Theorem[1]
4. Conclusion

The upper bounds in Theorem 1 can be slightly improved: for instance if \( k \) is a prime number, we have proved that any word containing more than \( k^2 \left( \frac{k}{d} - 1 \right) + \frac{k(d-1)}{d} + 1 \) letters has a synchronizing point. This bound is attained for the morphism \( \varphi(a) = (ab^{d-1})^d \), \( \varphi(b) = a^d \). However, the proofs of these improved bounds are very technical and difficult to follow, therefore we do not state them here. The proof techniques presented in this paper could be used for uniform morphisms over three or more letter alphabets. Unfortunately, the higher the number of letters is, the higher the number of subgraphs of the respective graph of overhangs to be considered. In order to obtain some reasonable bound on minimal synchronizing delay for a general D0L-system, some other proof techniques must be used.

5. Acknowledgment

We acknowledge financial support by the Czech Science Foundation, grant GAČR 13-35273P (the second author), and by the Grant Agency of the Czech Technical University in Prague, grant SGS11/162/OHK4/3T/14 (the first author).

References

[1] J. Berstel, D. Perrin, C. Reutenauer, Codes and Automata, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2009.
[2] F. Mignosi, P. Séébold, If a D0L language is \( k \)-power free then it is circular, in: ICALP ’93: Proceedings of the 20th International Colloquim on Automata, Languages and Programming, Spriger-Verlag, London, UK, 1993, pp. 507–518.
[3] K. Klouda, Š. Starosta, Characterization of circular D0L-systems, submitted, arXiv:1401.0038.
[4] J. Cassaigne, An algorithm to test if a given circular HDOL-language avoids a pattern, in: IFIP World Computer Congress’94, Elsevier, 1994, pp. 459–464.
[5] K. Klouda, Bispecial factors in circular non-pushy D0L languages, Theoret. Comput. Sci. 445 (2012) 63–74, doi:10.1016/j.tcs.2012.05.007
[6] A. Ehrenfeucht, G. Rozenberg, Repetition of subwords in D0L languages, Inform. Comput. 53 (1983) 13–35.
[7] K. Klouda, Š. Starosta, An algorithm enumerating all infinite repetitions in a D0L-system, J. Discrete Algorithms 33 (2015) 130–138.
[8] Y. Kobayashi, F. Otto, P. Séébold, A complete characterization of repetitive morphisms over the two-letter alphabet, in: T. Jiang, D. Lee (Eds.), Computing and Combinatorics, Vol. 1276 of LNCS, Springer Berlin Heidelberg, 1997, pp. 393–402.
[9] A. E. Frid, Applying a uniform marked morphism to a word, Discrete Math. Theoret. Comput. Sci. 3 (1999) 125–140.

[10] A. E. Frid, On uniform dol words, in: STACS’98, LNCS 1373, Springer, 1998, pp. 544–554.

[11] R. C. Lyndon, M. P. Schützenberger, The equation $a^m = b^n c^p$ in free group., Michigan Math. J. 9 (1962) 289–298.

[12] D. Krieger, On critical exponents in fixed points of $k$-uniform binary morphisms, RAIRO-Theoret. Inf. Appl. 43 (2009) 41–68.