Sturmian words with balanced construction

Hanna Uscka-Wehlou

Uppsala University, Department of Mathematics
Box 480, SE-751 06 Uppsala, Sweden
hania@wehlou.com, http://hania.wehlou.com

Abstract. In this paper we define Sturmian words with balanced construction. We formulate a fixed-point theorem for Sturmian words and analyze the set of all fixed points. The inspiration for this work came from the Kolakoski word and the general idea of self-reading sequences by Păun and Salomaa. The basis for this article is the author's earlier research on the influence of the continued fraction elements in the expansion of $a \in]0,1[\setminus \mathbf{Q}$ on the construction of runs for the upper mechanical word with slope a and intercept 0.

Keywords: upper mechanical word; irrational slope; Sturmian word; continued fraction; hierarchy of runs; fixed point; self-reading sequence; Kolakoski word; Freeman chain code.

1 Introduction

Word theory has grown very intensively during the last century. The theory has found numerous applications in computer science, which has stimulated its fast development. Both mathematicians and theoretical computer scientists have been working on problems connected with word theory and related domains. A very good illustration of the results of this work and of the variety of domains and subjects word theory is connected to, is presented in Pytheas Fogg (2002) [16], Lothaire (2002) [13], Allouche and Shallit (2003) [1], Perrin and Pin (2004) [15], Karhumäki (2004) [10], Berthé, Ferenczi and Zamboni (2005) [3], and Berstel et al. (2008) [2].

This paper about binary words is inspired mainly by ideas of three persons: William G. Kolakoski, Herbert Freeman and Azriel Rosenfeld.

Self-reading sequences have been examined by a lot of researchers. Some general definitions of those can be found in [9, 14]. William G. Kolakoski has described probably the most famous self-reading sequence, very well known to the community of theoretical computer scientists; see [12] and [16, p. 93]. The Kolakoski word is defined as one of the two fixed points of the run-length encoding Δ; see [5, 6]. These words are identical with their own run-length encoding sequences. The one beginning with 2 is: $K = 2211212212211211221211212211211212212211212212\cdots$. Brlek et al. have studied some generalizations of the Kolakoski word to an arbitrary alphabet, which got the name of *smooth words*; see [6] and references there.

A simple example of a self-reading sequence is the Morse sequence u which begins with a and is defined as the fixed point of the *Morse substitution* σ defined over the alphabet $\{a,b\}$ by $\sigma(a) = ab$, $\sigma(b) = ba$, thus $u = abbabaabbaababbabaababbaabbabaabbaababbaabbabaababbabaabbaabab\ldots$; see also [16, p. 7]. Another simple self-reading sequence is the Fibonacci word defined as the fixed point w beginning with 1 of the substitution $\varphi(1) = 10,\ \varphi(0) = 1$; see also [16, p. 7]. We show on Figure 1 how to construct w. The arrows pointing downwards show how we use the definition of the substitution φ, the arrows pointing upwards show how to use the fixed-point condition $w = \varphi(w)$. Because $\varphi(w)$ is being formed faster than w, we get in each step enough information to be able to construct w.

Generally, the *characteristic words* of irrational numbers with purely periodic continued fraction (CF) expansion (i.e., some quadratic surds) are also fixed points of corresponding substitutions, as has been shown in the paper by Shallit (1991) [18]. These fixed points, and, in particular, the Fibonacci word, are Sturmian. They are also examples of self-reading sequences.

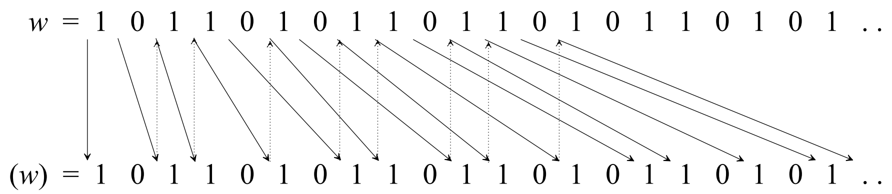


Fig. 1. The Fibonacci sequence as the fixed point w beginning with 1 of the substitution $\varphi(1) = 10$, $\varphi(0) = 1$.

The *cutting sequence* of grid lines by the half-line $y = ax$ for $a \in]0,1[\setminus \mathbf{Q}$ and $x > 0$ (i.e., the line passes through no lattice points) is one of binary representations of $y = ax$, where 0 denotes a vertical grid crossing and 1 a horizontal one. Such a sequence for straight lines with irrational slopes is Sturmian [16, p. 143]. There is a close relationship between the cutting sequence of $y = ax$ for $a \in]0,1[\setminus \mathbf{Q}$ and the line's *chain code* (which is the same, or the same up to the transformation "replace 10 by 1", as the characteristic word with slope a, depending on whether the line is *naive* or *standard*). Herbert Freeman (1970) [8, p. 260] observed that in the chain code of a digital straight line "successive occurrences of the element occurring singly are as uniformly spaced as possible". This property has been formalized and has got the name of *balance property*; see [25]. The self-similarity properties formulated by Bruckstein (1991) [7] form a quantitative expression of this uniformity principle.

Azriel Rosenfeld described in his paper from 1974 [17] the run-hierarchical structure of digital lines. On each level k (for $k \geq 2$) we have runs$_k$ which are composed of a single occurring run$_{k-1}$ (long L_{k-1} or short S_{k-1}) and a maximal sequence of runs$_{k-1}$ (short S_{k-1} or long L_{k-1}, respectively) following after this single one or preceding it. On some levels the long runs are the most frequent (coming in sequences), while on other levels the short runs are the mainly occurring ones.

In Uscka-Wehlou (2008) [22] we presented a CF-based description of upper mechanical words, which reflects the run-hierarchical structure of words. The present idea is to create a *run-construction encoding* operator, by analogy to the run-length encoding operator. The latter is very well known and was used for coding the Thue–Morse word by Brlek in 1988 [5] and the former is a new concept, defined for the first time in the present paper (Definition 6). We will look for the fixed points of the run-construction encoding operator. For them even the constructional distribution is uniform, in the way as described by Freeman. In the main theorem of this paper (Theorem 4) we show that every infinite sequence of positive natural numbers such that all the elements indexed by numbers greater than 1 are greater than 1 generates exactly one fixed point of the run-construction encoding operator. All of them are self-generating sequences, identical with their own run-construction encoding sequences, by analogy with the Kolakoski word. In the second half of this paper we present a number of examples. We also examine the set of all fixed points (Theorem 5) and formulate a number of questions and combinatorial problems for further research (on p. 8 after Proposition 3, and in Section 6).

2 A continued-fraction-based description of upper mechanical words

In [22] we presented a recursive description by CFs of upper mechanical words. Let us recall the definiton of those; cf. Lothaire (2002) [13, p. 53].

Definition 1. *Given two real numbers a and r with $0 \leq a \leq 1$, we define two infinite words $s(a,r), s'(a,r)\colon \mathbf{N} \to \{0,1\}$ by $s_n(a,r) = \lfloor a(n+1)+r \rfloor - \lfloor an+r \rfloor$ and $s'_n(a,r) = \lceil a(n+1)+r \rceil - \lceil an+r \rceil$. The word $s(a,r)$ is the* lower mechanical word *and $s'(a,r)$ is the* upper mechanical word *with* slope a *and* intercept r. *A lower or upper mechanical word is* irrational *or* rational *according as its slope is irrational or rational.*

In the present paper we deal with the special case when $a \in \,]0,1[$ is irrational and $r = 0$. In this case we will denote the lower and upper mechanical words by $s(a)$ and $s'(a)$ respectively. We have $s_0(a) = \lfloor a \rfloor = 0$ and $s'_0(a) = \lceil a \rceil = 1$ and, because $\lceil x \rceil - \lfloor x \rfloor = 1$ for irrational x and $\lceil x \rceil - \lfloor x \rfloor = 0$ only for integers, we have

$$s(a) = 0c(a), \quad s'(a) = 1c(a) \tag{1}$$

(meaning 0, resp. 1 concatenated to $c(a)$). The word $c(a)$ is called the *characteristic word* of a. For each $a \in \,]0,1[\, \setminus \mathbf{Q}$, the characteristic word associated with a is thus the following infinite word $c(a)\colon \mathbf{N}^+ \to \{0,1\}$:

$$c_n(a) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor = \lceil a(n+1) \rceil - \lceil an \rceil, \quad n \in \mathbf{N}^+. \tag{2}$$

It is well known that the equality of characteristic words gives the equality of corresponding slopes, i.e., for any $a, a' \in \,]0,1[\, \setminus \mathbf{Q}$, if $c(a) = c(a')$, then $a = a'$; cf. Lothaire (2002) [13, p. 62, Lemma 2.1.21].

We assume that, for each $a \in \,]0,1[\, \setminus \mathbf{Q}$, its simple CF expansion is given, expressed as $a = [0; a_1, a_2, a_3, \ldots]$, and we know the positive integers a_i for all $i \in \mathbf{N}^+$. These are called the *elements* (or *partial quotients*) of the CF. Let us recall that

$$[a_0; a_1, a_2, a_3, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}\ .$$

In our case, when $a \in \,]0,1[\, \setminus \mathbf{Q}$, we have $a_0 = \lfloor a \rfloor = 0$ and the sequence of the CF elements $(a_1, a_2, \ldots)$ is infinite. For more information about CFs see Khinchin (1997) [11].

Our CF description of upper mechanical words from [22] was based on our earlier one by digitization parameters from [19] and the following *index jump function*, introduced by the author in [20].

Definition 2. *For each $a \in \,]0,1[\, \setminus \mathbf{Q}$, the* index jump function $i_a\colon \mathbf{N}^+ \to \mathbf{N}^+$ *is defined by $i_a(1) = 1$, $i_a(2) = 2$, and $i_a(k+1) = i_a(k) + 1 + \delta_1(a_{i_a(k)})$ for $k \geq 2$, where $\delta_1(x) = \begin{cases} 1, x = 1 \\ 0, x \neq 1 \end{cases}$, and a_j for $j \in \mathbf{N}^+$ are the CF elements of a.*

The index jump function is a renumbering which avoids elements following directly after some 1's in the CF expansion (in particular, it avoids every second element in the sequences of consecutive 1's with index greater than 1); see also [21].

In [22], upper mechanical words were described according to the hierarchy of runs on all levels, as introduced by Azriel Rosenfeld (1974) [17, p. 1265]. A run of the first level is a maximal sequence 10^m, meaning the letter 1 followed by $m \geq 0$ letters 0. For a given slope, there are only two possible run lengths, runs with the smaller length we call *short* runs (S_1) and runs with the largest length we call *long* runs (L_1). The same holds for the other levels: a run of level n is a maximal sequence of runs of level $n-1$, i.e., $S^k_{n-1} L_{n-1}$, $S_{n-1} L^k_{n-1}$, $L_{n-1} S^k_{n-1}$ or $L^k_{n-1} S_{n-1}$ and the cardinality-wise run length of run_n, denoted by $\|\mathrm{run}_n\|$, is the number (here $k+1$) of runs_{n-1} forming it. We denote by $|w|$ the binary-word length of a 0-1 word w, i.e., the total number of its letters. The following theorem shows how exactly the run-hierarchical structure of $s'(a)$ for each $a \in \,]0,1[\, \setminus \mathbf{Q}$ depends on the CF elements of a. Because of (1) and (2), this gives also a description of lower mechanical and characteristic words.

Theorem 1 ([22]; a CF description of upper mechanical words). *Let $a \in \,]0,1[\, \setminus \mathbf{Q}$ and $a = [0; a_1, a_2, \ldots]$. For $s'(a)$ as in Definition 1 we have $s'(a) = \lim_{k\to\infty} P_k$, where $P_1 = S_1 =$*

10^{a_1-1}, $L_1 = 10^{a_1}$, *and, for* $k \geq 2$,

$$P_k = \begin{cases} L_k = S_{k-1}^{a_{i_a(k)}} L_{k-1} & \text{if} \quad a_{i_a(k)} \neq 1 \quad \text{and} \quad i_a(k) \quad \text{is even} \\ S_k = S_{k-1} L_{k-1}^{a_{i_a(k)+1}} & \text{if} \quad a_{i_a(k)} = 1 \quad \text{and} \quad i_a(k) \quad \text{is even} \\ S_k = L_{k-1} S_{k-1}^{-1+a_{i_a(k)}} & \text{if} \quad a_{i_a(k)} \neq 1 \quad \text{and} \quad i_a(k) \quad \text{is odd} \\ L_k = L_{k-1}^{1+a_{i_a(k)+1}} S_{k-1} & \text{if} \quad a_{i_a(k)} = 1 \quad \text{and} \quad i_a(k) \quad \text{is odd,} \end{cases} \tag{3}$$

where the function i_a *is defined in Definition 2. The meaning of the symbols is the following: for* $k \geq 1$, P_k *is the* ***P****refix number* k, S_k *is the* ***S****hort* run_k *and* L_k *is the* ***L****ong* run_k*. To make the recursive formula (3) complete, we add that for each* $k \geq 2$, *if* $P_k = S_k$, *then* L_k *is defined in the same way as* S_k, *with the only difference that the exponent defined by* $a_{i_a(k)}$ *(or by* $a_{i_a(k)+1}$*) is increased by* 1*. If* $P_k = L_k$, *then* S_k *is defined in the same way as* L_k, *with the only difference that the exponent defined by* $a_{i_a(k)}$ *(or by* $a_{i_a(k)+1}$*) is decreased by* 1*.*

The value of the index jump function for each natural $k \geq 2$ describes the index of the CF element which determines the most frequent run on level $k-1$ (denoted main_{k-1}), which we can formulate as the following corollary. The corollary also describes the cardinality-wise run length on each digitization level and shows how to conclude about the kind of the prefix P_{k-1} as obtained in (3) (long L_{k-1} or short S_{k-1}) from the parity of $i_a(k)$.

Corollary 1. *Let* $a \in \,]0,1[\, \setminus \mathbf{Q}$ *and* $a = [0; a_1, a_2, a_3, \ldots]$*. If* $s'(a)$ *is the upper mechanical word with slope* a *and intercept* 0 *as defined in Definition 1, then, in the run-hierarchic structure of* $s'(a)$ *we have for each* $k \geq 2$

- $a_{i_a(k)} \geq 2 \quad \Rightarrow \quad \text{main}_{k-1} = S_{k-1}, \qquad a_{i_a(k)} = 1 \quad \Rightarrow \quad \text{main}_{k-1} = L_{k-1},$
- $i_a(k)$ *is odd* $\quad \Rightarrow \quad P_{k-1} = L_{k-1}, \qquad i_a(k)$ *is even* $\quad \Rightarrow \quad P_{k-1} = S_{k-1},$

where i_a *is the corresponding index jump function. Moreover, the cardinality-wise run length on each level is the following:* $\|S_n\| = b_n, \|L_n\| = b_n + 1$, *where*

$$b_1 = a_1 \quad \text{and, for} \quad n \geq 2, \quad b_n = \begin{cases} a_{i_a(n)}, & a_{i_a(n)} \neq 1 \\ 1 + a_{i_a(n)+1}, & a_{i_a(n)} = 1 \ . \end{cases} \tag{4}$$

Corollary 1 follows immediately from Theorem 1.

Let us recall the concept of the *sequence of length specification* which was first introduced by the author in [23] (Definition 3 there).

Definition 3. *For any irrational* $a = [0; a_1, a_2, \ldots]$, *the sequence* $(b_n)_{n \in \mathbf{N}^+} = (\|S_n\|)_{n \in \mathbf{N}^+}$ *of short run lengths on all levels in the run-hierarchical construction of the upper mechanical word* $s'(a)$ *with slope* a *and intercept* 0*, will be called the* sequence of length specification*.*

It is clear from (4), that for each $a \in \,]0,1[\, \setminus \mathbf{Q}$, the corresponding sequence of length specification $(b_n)_{n \in \mathbf{N}^+}$ fulfills $b_1 \in \mathbf{N}^+$ and, for each $n \geq 2$, $b_n \geq 2$. In [23] we also showed that each sequence fulfilling these condition is the sequence of length specification for some slopes and the cardinality of the set of these slopes is of the continuum. For a fixed index jump function (i.e., a sequence of values $(d_n)_{n \in \mathbf{N}^+}$ such that $d_1 = 1, d_2 = 2$ and, for all $k \geq 2$ $d_k \in \mathbf{N}^+$ and $d_{k+1} - d_k = 1$ or $d_{k+1} - d_k = 2$) there exists exactly one slope with $(b_n)_{n \in \mathbf{N}^+}$ as sequence of length specification [23, 24].

3 The constructional word

In this section we will define (Definition 4) a new binary word associated with the upper mechanical word $s'(a)$ for $a \in]0,1[\setminus \mathbf{Q}$ and we will call it the *constructional word*. It follows from Definition 2 that, for any $a \in]0,1[\setminus \mathbf{Q}$ and $n \geq 2$

$$a_{i_a(n)} = 1 \quad \Leftrightarrow \quad i_a(n+1) = i_a(n) + 2 \qquad \text{and} \qquad a_{i_a(n)} \geq 2 \quad \Leftrightarrow \quad i_a(n+1) = i_a(n) + 1. \tag{5}$$

The sequence $(i_a(n))_{n \in \mathbf{N}^+}$ is thus strictly increasing and the difference between each two consecutive elements of this sequence is equal to 1 or to 2. This gives us an idea of defining a new two-letter word associated with a. This word will be called the *constructional word* and it will code the structure of $s'(a)$ in terms of long and short runs on all the levels, according to Corollary 1 and (5).

Definition 4. *Let $a \in]0,1[\setminus \mathbf{Q}$. The* constructional word *of a is $\gamma = \gamma(a)$, defined by*

$$\gamma_n = i_a(n+2) - i_a(n+1) - 1$$

for $n \in \mathbf{N}^+$, where i_a is the index jump function defined in Definition 2.

It follows from (5) that the constructional word for all $a \in]0,1[\setminus \mathbf{Q}$ is a 0-1 word, and, for all $n \in \mathbf{N}^+$, $\gamma_n = 1 \Leftrightarrow a_{i_a(n+1)} = 1$ and $\gamma_n = 0 \Leftrightarrow a_{i_a(n+1)} \geq 2$. This gives us the following proposition.

Proposition 1. *For each $a \in]0,1[\setminus \mathbf{Q}$ and for each $n \in \mathbf{N}^+$ we have $\gamma_n = \delta_1(a_{i_a(n+1)})$, where i_a is the index jump function defined in Definition 2.*

Corollary 1 shows clearly why γ got **the name of constructional word**. The elements $a_{i_a(k)}$ for $k \geq 2$ of the CF expansion of the slope $a \in]0,1[\setminus \mathbf{Q}$ determine the **construction** of runs$_k$ as sets of short and long runs$_{k-1}$. The indices $k \in \mathbf{N}^+$ numbering letters of γ equal to 1 are the same as the indices of the digitization levels with the most frequent long runs$_k$ (L_k).

Example 1. If the slope a is $e - 2 = [0; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, \ldots, 1, 1, 2n, 1, 1, \ldots]$, then the index jump function i_a is formed as follows:

$$\begin{array}{rl} a = & [0; \overset{b_1}{1}, \overset{b_2}{2}, \overbrace{\underline{1}, 1}^{b_3}, \overset{b_4}{4}, \overbrace{\underline{1}, 1}^{b_5}, \overset{b_6}{6}, \overbrace{\underline{1}, 1}^{b_7}, \overset{b_8}{8}, \overbrace{\underline{1}, 1}^{b_9}, \overset{b_{10}}{10}, \overbrace{\underline{1}, 1}^{b_{11}}, \ldots] \\ & \quad \downarrow \ \downarrow \ \downarrow \quad \ \downarrow \ \downarrow \quad \ \downarrow \ \downarrow \quad \ \downarrow \ \downarrow \quad \ \downarrow \ \downarrow \quad \ \ldots \\ (i_a(k))_{k \in \mathbf{N}^+} = & (\ 1, 2, \ 3, \quad 5, \ 6, \quad 8, \ 9, \quad 11, 12, \quad 14, 15, \quad \ldots). \end{array}$$

In the last row we presented the first eleven elements of the sequence of the values of the index jump function for this a, so $(i_a(k))_{1 \leq k \leq 11}$. The sequence of length specification for $a = e - 2$ is $(b_n)_{n \in \mathbf{N}^+} = (1, 2, 2, 4, 2, 6, 2, 8, 2, 10, 2, \ldots 2, 2n, 2, \ldots)$. The constructional word is $\gamma(e-2) = (01)^\omega$. On odd-numbered levels k short runs (S_k) are the most frequent runs, while on even-numbered levels k long runs (L_k) dominate. The run-hierarchical structure of the digital line $y = (e-2)x$ was thoroughly discussed in the author's paper [21, p. 2252, Example 14].

Definition 4 describes how to form the word γ for $a \in]0,1[\setminus \mathbf{Q}$, in terms of the corresponding function i_a. The following proposition is a kind of converse to this definition. It says, how to find the function i_a, given the constructional word of a.

Proposition 2. *If $a \in]0,1[\setminus \mathbf{Q}$ and $\gamma = \gamma(a)$ is the constructional word associated with a, then we have for $n \geq 3$*

$$i_a(n) = n + \sum_{j=1}^{n-2} \gamma_j. \tag{6}$$

Proof. By induction, from Definitions 2 and 4. □

We know, from the author's papers [23, 24], that $(i_a(n))_{n\in\mathbf{N}^+}$ and $(b_n)_{n\in\mathbf{N}^+}$ determine the slope $a \in\,]0,1[\, \setminus \mathbf{Q}$. Because of Definition 4 and Proposition 2 we know that also $(\gamma_n)_{n\in\mathbf{N}^+}$ and $(b_n)_{n\in\mathbf{N}^+}$ determine the slope $a \in\,]0,1[\, \setminus \mathbf{Q}$.

4 Introduction to Sturmian words

In this section we provide a very brief introduction to Sturmian words, based on Lothaire (2002) [13].

Let $\mathcal{A}$ be a set of symbols usually called the *alphabet*. We denote by $\mathcal{A}^\star$ (in some papers denoted by $\mathcal{A}^{(\mathbf{N})}$) the set of all finite words over $\mathcal{A}$ (i.e., finite sequences of elements from $\mathcal{A}$) and by ε the empty word. We denote by $\mathcal{A}^\omega$ ($\mathcal{A}^{\mathbf{N}}$) the set of (right) infinite words (i.e., sequences of symbols in $\mathcal{A}$ indexed by non-negative integers). In this paper we consider only right infinite words.

A finite word w is a *factor* of a (finite or infinite) word x if there exist words u (finite) and y such that $x = uwy$. **Sturmian words** are defined as infinite words which have exactly $n+1$ different factors of length n for every natural n. In particular, they have 2 factors of length 1, which means that each Sturmian word is constructed of exactly 2 letters, which we can call 0 and 1, thus $\mathcal{A} = \{0,1\}$.

A word $x \in \mathcal{A}^\omega$ is *periodic* if it is of the form $x = z^\omega$ for some $z \in \mathcal{A}^\star \setminus \{\varepsilon\}$, *eventually periodic* if it is of the form $x = yz^\omega$ for some $y, z \in \mathcal{A}^\star \setminus \{\varepsilon\}$, and *aperiodic* if it is not eventually periodic; cf. Lothaire (2002) [13, p. 9]. We need the following definition to formulate a theorem which shows equivalent characterizations of Sturmian words (Theorem 2).

Definition 5 (Lothaire 2002:48). *For binary words with letters* 0 *and* 1 *we define the following.*

- *The* height *of a finite word x is the number $h(x)$ of letters equal to* 1 *in x.*
- *Given two finite words x and y of the same length, their* balance *is* $\delta(x,y) = |h(x) - h(y)|$.
- *A set of finite words X is* balanced *if* $(\ x, y \in X\ \wedge\ |x| = |y|\ )\ \Rightarrow\ \delta(x,y) \le 1$.
- *An infinite word is itself* balanced *if the set of its factors (thus, finite words) is balanced.*

Theorem 2 (Lothaire 2002:57). *Let s be an infinite word. We have the following equivalence:*
s is Sturmian $\Leftrightarrow$ s is balanced and aperiodic $\Leftrightarrow$ s is irrational (lower or upper) mechanical.

5 A fixed-point theorem for Sturmian words. Self-generating run construction.

In Sections 2 and 3 we described two words over a two-letter alphabet $\{0,1\}$ associated with an irrational positive slope $a < 1$. The first of them, the upper mechanical word, is Sturmian (Theorem 2), the second one, the constructional word, can obviously be any 0-1 word. One could try to describe the slopes $a \in\,]0,1[\, \setminus \mathbf{Q}$, for which the levels with the most frequent run being long (or, dually, short) are uniformly distributed (for such a we will call $s'(a)$ *words with balanced construction*). And an even more demanding condition would be: find these a for which $\gamma(a) = c(a)$. For these a, $s'(a) = 1c(a)$ will be called *word with self-balanced construction*, because the distribution of the levels with the most frequent run being long (equivalently: the distribution of pairs (a_l, a_{l+1}) of CF elements of a such that $a_l = 1$, $l \ge 2$, and a_l is not immediately preceded by an odd number of consecutive CF elements equal to 1 and with indices greater than 1) is the same as the distribution of the letter 1 in the characteristic word $c(a)$; c.f. Proposition 4 on page 8 and the discussion there.

We consider $\{0,1\}^\omega$, the set of all right infinite two-letter words composed of 0's and 1's and let $\mathcal{UM}_0 \subset \{0,1\}^\omega$ be the subset of all upper mechanical words with positive irrational slopes less than 1 and with intercept 0 (which are Sturmian according to Theorem 2).

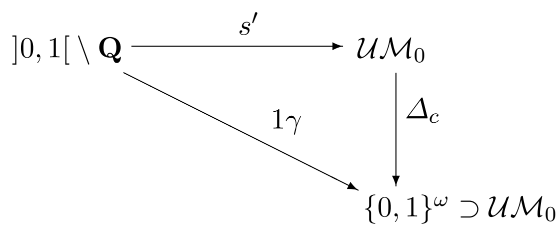
Definitions 1 and 4 give us two mappings from $]0,1[\setminus \mathbf{Q}$ to $\{0,1\}^\omega$. The first one maps each $a \in]0,1[\setminus \mathbf{Q}$ to the associated upper mechanical word $s'(a) = 1c(a)$:

$$s' \colon]0,1[\setminus \mathbf{Q} \longrightarrow \mathcal{UM}_0 \subset \{0,1\}^\omega,$$

the second one maps each $a \in]0,1[\setminus \mathbf{Q}$ to the associated constructional word $\gamma(a)$ concatenated with prefix 1:

$$1\gamma \colon]0,1[\setminus \mathbf{Q} \longrightarrow \{0,1\}^\omega, \quad (1\gamma)(a) = 1\gamma(a).$$

Definition 6. *The* run-construction encoding operator $\Delta_c \colon \mathcal{UM}_0 \longrightarrow \{0,1\}^\omega$ *is defined as* $\Delta_c = (1\gamma) \circ (s')^{-1}$.



The mapping is well defined (Lemma 2.1.21 from Lothaire 2002:62 mentioned in Section 2). We can also describe this operator by analogy with the run-length encoding operator as in [6]:

$$\Delta_c(s'(a))(0) = 1, \quad \Delta_c(s'(a))(n) = \delta_1(a_{i_a(n+1)}) \quad \text{for} \quad n \in \mathbf{N}^+$$

which, according to Corollary 1, can be written in the following, more illustrative way:

$$\Delta_c(s'(a))(n) = \begin{cases} 0, & S_n \text{ is the most frequent run on level } n \\ 1, & L_n \text{ is the most frequent run on level } n \end{cases} \quad \text{for} \quad n \in \mathbf{N}^+.$$

Definition 7. *Let* $a \in]0,1[\setminus \mathbf{Q}$. *The upper mechanical word* $s'(a)$ *has*

- balanced construction *if its constructional word* $\gamma(a)$ *is a characteristic word* $c(\alpha)$ *(not necessarily with irrational slope) for some* α.
- Sturmian-balanced construction *if* $\gamma(a)$ *is a characteristic word* $c(\alpha)$ *for some* $\alpha \in]0,1[\setminus \mathbf{Q}$.
- self-balanced construction *if* $1\gamma(a) = \Delta_c(1c(a)) = 1c(a)$, *i.e., its constructional word is equal to its characteristic word, i.e.,* $s'(a)$ *is a fixed point of* Δ_c.

Clearly: self-balanced construction $\Rightarrow$ Sturmian-balanced construction $\Rightarrow$ balanced construction.

Example 2. The words $s'(a)$ with $a = [0; a_1, a_2, a_3, \ldots]$, where $a_k \geq 2$ for all $k \geq 2$, have **balanced construction**. We have $i_a(k) = k$ for all $k \in \mathbf{N}^+$ and $a_{i_a(k)} \geq 2$ for all $k \geq 2$. This means that the constructional word $\gamma = \gamma(a)$ is defined by $\gamma_n = 0$ for all $n \in \mathbf{N}^+$, which is the characteristic word with slope 0. This also means that no upper mechanical word with dominating short run on all digitization levels can be a fixed point of Δ_c.

Example 3. The words $s'(a)$ with $a = [0; a_1, 1, a_3, 1, a_5, 1, a_7, \ldots]$, where $a_{2k-1} \in \mathbf{N}^+$ for all $k \in \mathbf{N}^+$, have **balanced construction**. We have $i_a(1) = 1$ and $i_a(k) = 2k-2$ for $k \geq 2$, and $a_{i_a(k)} = 1$ for all $k \geq 2$. This means that the constructional word $\gamma = \gamma(a)$ is defined by $\gamma_n = 1$ for all $n \in \mathbf{N}^+$, which is the characteristic word with slope 1. This also means that no upper mechanical word with dominating long run on all digitization levels can be a fixed point of Δ_c.

Let us recall the following theorem, which is a merge of Lagrange's theorem from 1770 with Euler's theorem from 1737; see [4, pp. 66–71].

Theorem 3 (Euler, Lagrange). *Quadratic surds (i.e., algebraic numbers of the second degree), and only they, are represented by periodic or eventually periodic CFs.*

Example 4. A generalization of Example 3: For each $k \in \mathbf{N}^+ \setminus \{1\}$ and for each infinite matrix $A = [a_{ij}]_{i\in\mathbf{N}^+, j\in[1,k]_\mathbf{Z}}$, where $a_{i1} \in \mathbf{N}^+$ and $a_{ij} \in \mathbf{N}^+ \setminus \{1\}$ for $i \in \mathbf{N}^+$ and $j \in [2,k]_\mathbf{Z}$, the upper mechanical words $s'(a)$ with slopes $a = [0; a_{11}, \ldots, a_{1k}, 1, a_{21}, \ldots, a_{2k}, 1, a_{31}, \ldots, a_{3k}, 1, a_{41}, \ldots, a_{4k}, 1, a_{51}, \ldots]$ have **balanced construction**. We have $a_{i_a(nk+1)} = a_{nk+n} = 1$ for all $n \in \mathbf{N}^+$. The constructional words of all these slopes for a fixed k are $0^{k-1}10^{k-1}10^{k-1}1\ldots$. They correspond to the word with slope $\frac{1}{k}$. If all the rows of the matrix A are identical, the CF expansion is periodic and a is quadratic irrational.

Example 5. A generalization of Example 3: For each $k \in \mathbf{N}^+ \setminus \{1\}$ and each pair of infinite matrices $[a_{ij}]_{i\in\mathbf{N}^+, j\in[1,k]_\mathbf{Z}}$ and $[a'_{ij}]_{i\in\mathbf{N}^+, j\in[1,k+1]_\mathbf{Z}}$ such that $a_{i1}, a'_{i1} \in \mathbf{N}^+$ and $a_{ij}, a'_{is} \in \mathbf{N}^+ \setminus \{1\}$ for all indices $i \in \mathbf{N}^+$, $j \in [2,k]_\mathbf{Z}$ and $s \in [2, k+1]_\mathbf{Z}$, the upper mechanical words $s'(a)$ with slopes $[0; a_{11}, \ldots, a_{1k}, 1, a'_{11}, \ldots, a'_{1,k+1}, 1, a_{21}, \ldots, a_{2k}, 1, a'_{21}, \ldots, a'_{2,k+1}, 1, a_{31}, \ldots]$ have **balanced construction**. The constructional words of all these $s'(a)$ for fixed k are $0^{k-1}10^k10^{k-1}10^k10^{k-1}1\ldots$. They correspond to the upper mechanical words $s'(a)$ with slopes $a = \frac{2}{2k+1}$.

Proposition 3. *There exist no quadratic surds which are slopes to upper mechanical words with Sturmian-balanced construction.*

Proof. Let $a \in \,]0,1[\, \setminus \mathbf{Q}$ be any quadratic surd. If there are no 1's in the CF expansion of a, then $\gamma(a) = 000\cdots$, which is the characteristic word with slope 0, which is rational. If there is a 1 in the CF expansion of a, then, according to Theorem 3, either this 1 is only in the beginning of the CF (if we have eventual periodicity) or is repeated periodically, which will lead to a characteristic word of a rational number, if any. □

Quadratic surds with purely periodic CF expansion are slopes of fixed points of corresponding substitutions as defined in Shallit (1991) [18]. It follows from Proposition 3, that no quadratic surds can be slopes of fixed points of Δ_c. No quadratic surds can have Sturmian-balanced construction, but some of them have balanced construction. It would be an interesting combinatorial exercise to describe all the quadratic surds with balanced construction (give a necessary and sufficient condition on the CF expansion of slopes) and, generally, to give a necessary and sufficient condition on the elements of the CF expansion of $a = [0; a_1, a_2, \ldots]$ to generate an upper mechanical word $s'(a)$ with balanced construction, Sturmian-balanced construction, self-balanced construction. Proposition 1 and Definition 7 give us the following characterization of the CF expansion of the slopes of fixed points of Δ_c. Let $a = [0; a_1, a_2, \ldots]$. A pair (a_l, a_{l+1}) of CF elements in the expansion of a will be called an *essential pair* if $a_l = 1$, $l \geq 2$, and the element $a_l = 1$ is immediately preceded by an even number, i.e., $0, 2, 4, \ldots$, of consecutive 1's with index greater than 1, i.e., $\exists\ k \in \mathbf{N}$, $[0; a_1, a_2, \ldots] = [0; a_1, a_2, \ldots, a_{l-2k-1}, \underbrace{1, 1, \ldots, 1, 1}_{2k}, a_l, a_{l+1}, \ldots]$ and, if $l - 2k - 1 \geq 2$, then $a_{l-2k-1} \geq 2$; cf. *essential 1's* in [23, 24].

Proposition 4. *If $a = [0; a_1, a_2, \ldots]$, then $s'(a)$ is a fixed point of Δ_c iff $c_n(a) = \delta_1(a_{i_a(n+1)})$ for all $n \in \mathbf{N}^+$, where $c(a)$ is the corresponding characteristic word.*

This means that for a fixed point $s'(a) = 1c(a)$, each essential pair in the CF expansion of a reflects in the letter 1 on the corresponding place in $c(a)$, while the letters 0 of $c(a)$ appear on places corresponding to the places of remaining (i.e., no members of essential pairs) CF elements a_k $(k \geq 2)$ of a.

An upper mechanical word with slope $a \in]0,1[\setminus \mathbf{Q}$ and intercept 0 is a word with self-balanced construction (a fixed point of Δ_c) if its construction according to the hierarchy of runs and the arrangement of 0's and 1's in the word itself are made according to the same rules. The levels with dominating long runs are uniformly distributed like the 1's in the original characteristic word corresponding to the upper mechanical word. The following theorem is the main result of this paper.

Theorem 4 (main result). *Let $(b_n)_{n\in\mathbf{N}^+}$ be any sequence of natural numbers such that $b_1 \in \mathbf{N}^+$ and $b_n \geq 2$ for all $n \geq 2$. There exists exactly one fixed point of Δ_c with $(b_n)_{n\in\mathbf{N}^+}$ as the sequence of its length specification as defined in Definition 3.*

Proof. We will show how to find the fixed point $w = s'(a)$ corresponding to given $(b_n)_{n\in\mathbf{N}^+}$. The uniqueness will follow from the construction. In our reasoning we will use the following rules:

R1. Fixed point condition: for each $k \in \mathbf{N}$, $\mathrm{pref}_{k+1}(w) = 1\gamma_1 \cdots \gamma_k$, where $\mathrm{pref}_{k+1}(w)$ denotes the $k+1$ letters long prefix of the upper mechanical word $w = s'(a)$ we are looking for, and $\gamma = \gamma(a)$.

R2. $(\gamma_1, \gamma_2, \ldots, \gamma_k)$ determines $(i_a(1), i_a(2), \ldots, i_a(k+2))$ according to (6)

R3. $(\gamma_1, \gamma_2, \ldots, \gamma_k)$ and $(b_1, b_2, \ldots, b_{k+1})$ determine $(a_1, a_2, \ldots, a_{i_a(k+1)})$ according to **R2**, Proposition 1 and (4) in the following way. For $j = 1, 2, \ldots, k$

$$\gamma_j = 1 \quad \Rightarrow \quad \left[a_{i_a(j+1)} = 1 \ \wedge \ a_{i_a(j+1)+1} = b_{j+1} - 1\right], \qquad \gamma_j = 0 \quad \Rightarrow \quad a_{i_a(j+1)} = b_{j+1}$$

R4. According to **R2** and **R3**, $(\gamma_1, \ldots, \gamma_k)$ and $(b_1, \ldots, b_{k+1})$ determine (uniquely!) the prefix P_{k+1} (as in the run-hierarchical description (3)) of the upper mechanical word w we are looking for.

One can see that we need to describe a way of finding $(\gamma_1, \gamma_2, \ldots)$ to be able to reconstruct the fixed point (according to the condition **R1**) with the length specification $(b_1, b_2, b_3, \ldots)$. Because we do have whole $(b_1, b_2, b_3, \ldots)$, **R1**–**R4** imply that it is enough to show that for any $k \in \mathbf{N}^+$ we have $|P_{k+1}| > k+1 = |1\gamma_1 \cdots \gamma_k|$, i.e., that the prefixes produced of $(\gamma_1, \gamma_2, \ldots, \gamma_k)$ and $(b_1, b_2, b_3, \ldots)$ are on each step of the construction long enough to supply us with more information about the constructional word, which enables us to continue the construction using **R1**. So we have to prove $|P_{k+1}| > k+1$ for each $k \in \mathbf{N}^+$ (which is actually a severe understatement; see Corollary 1 in [22] and the last but one column in the table in Example 6).

Let us first suppose that $b_1 \geq 2$. We know from Theorem 1, that P_1 is short and $|P_1| = b_1 \geq 2 > 1$, so, because of the recursive construction (3), we get by easy induction $|P_{k+1}| \geq 2^{k+1} > k+1$.

If $b_1 = 1$, we get from Theorem 1 $|P_1| = 1$, which does not look well, because $|P_1| < 2$. To continue our construction, we have to get our information about γ_1 from somewhere else than P_1 and **R1**. Because the first run of level 1 is always short, we know that $s'(a) = 1c(a) = 11\ldots$, thus, **R1** gives us $\gamma_1 = 1$. This implies (rule **R3**) that $a_2 = a_{i_a(1+1)} = 1$ (and $a_3 = b_2 - 1$) and we get the following prefix of $s'(a)$: $P_2 = S_2 = S_1 L_1^{a_3} = 1(10)^{b_2-1}$ from (3) (the second row of the formula, because $i_a(2) = 2$ and $a_2 = 1$). Now we have already $|P_2| \geq 3 > 1+1$ for any b_2 and again, we obtain by induction $|P_{k+1}| \geq 2^{k-1} \cdot 3 > k+1$ for $k \geq 1$, which completes the proof. □

The speed of finding the fixed point grows together with b_1, but we have shown that even in case $b_1 = 1$ we can both get started and go on with our construction. Let us take the length specification $b_1 = 1$ and $b_n = 2$ for $n \geq 2$. This gives the slowest possible process of finding of the slope of the fixed point, but still, even in this worst case, it is possible to construct the unique fixed point:

Example 6. We will find the fixed point of Δ_c with the length specification $(1, 2, 2, 2, \ldots)$. We are thus looking for such $a \in]0,1[\setminus \mathbf{Q}$ that $\mathrm{pref}_{k+1}(s'(a)) = 1\gamma_1 \cdots \gamma_k$ for each $k \in \mathbf{N}$, and $(1, 2, 2, 2, 2, \ldots)$ is the corresponding sequence of length specification. At the starting point we only know that the first letter of $s'(a)$ is, by definition, 1. In each of the following steps we get $P_2, P_3, \ldots$ (step n gives us P_{n+1}).

The facts that $b_1 = 1$ and that the first run of level 1 is short, gives us only the information, that $s'(a) = 1c(a) = 11\ldots$ thus, because $\mathrm{pref}_2(s'(a)) = 1\gamma_1$, we get $\gamma_1 = 1$, which implies (rule **R3**) that $a_2 = a_{i_a(1+1)} = 1$ (and $a_3 = b_2 - 1 = 1$) and we get the following prefix of $s'(a)$: $P_2 = S_2 = S_1L_1 = 110$ from (3) (the second row of the formula, because $i_a(2) = 2$ and $a_2 = 1$). We have moreover $i_a(3) = 3 + \gamma_1 = 4$ (even number).

Further, because $110 = \mathrm{pref}_3(s'(a)) = 1\gamma_1\gamma_2$, we get $\gamma_2 = 0$, which means that $a_{i_a(2+1)} = a_4 = b_3 = 2$ and $i_a(4) = 4 + \gamma_1 + \gamma_2 = 5$. We get $P_3 = L_3 = S_2^2L_2 = 11011011010$ (from the first row of (3), because $a_{i_a(3)} \neq 1$ and $i_a(3)$ is even). This gives us, because of **R1**, $\gamma_3 = 1$, $\gamma_4 = 1$, $\gamma_5 = 0$, $\gamma_6 = 1$, $\gamma_7 = 1$, $\gamma_8 = 0$, $\gamma_9 = 1$, $\gamma_{10} = 0$, which, according to **R4**, allows us to get $P_4, \cdots, P_{11}$. Prefixes P_2, P_3 and P_4 are illustrated on Figure 2. One can see the analogy to Figure 1.

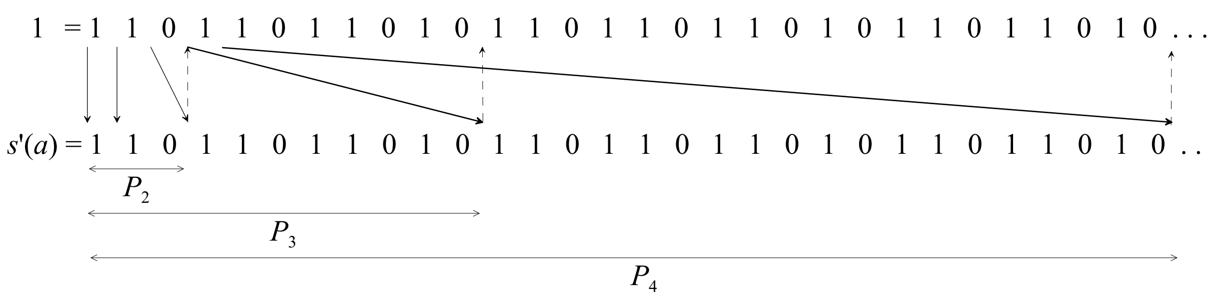


Fig. 2. The prefixes P_2, P_3 and P_4 of the fixed point of Δ_c with the length specification $(b_n)_{n\in\mathbf{N}^+} = (1, 2, 2, 2, \ldots)$.

We can summarise the data we have until now in the following table. In the next to last column, $|P_{k+1}|$ denotes the binary-word length of prefix P_{k+1} (total number of 0's and 1's forming it).

given	k	$i_a(k+1)$	$a_{i_a(k+1)}$	b_{k+1}	S_{k+1}	L_{k+1}	gives P_{k+1}	$\|P_{k+1}\|$	gives
$\gamma_1 = 1$	1	2	$= 1$	2	S_1L_1	$S_1L_1^2$	S_1L_1	3	γ_2
$\gamma_2 = 0$	2	4	$\neq 1$	2	S_2L_2	$S_2^2L_2$	$S_2^2L_2$	11	$\gamma_3, \ldots, \gamma_{10}$
$\gamma_3 = 1$	3	5	$= 1$	2	L_3S_3	$L_3^2S_3$	$L_3^2S_3$	30	$\gamma_{11}, \ldots, \gamma_{29}$
$\gamma_4 = 1$	4	7	$= 1$	2	L_4S_4	$L_4^2S_4$	$L_4^2S_4$	79	$\gamma_{30}, \ldots, \gamma_{78}$
$\gamma_5 = 0$	5	9	$\neq 1$	2	L_5S_5	$L_5S_5^2$	L_5S_5	128	$\gamma_{79}, \ldots, \gamma_{127}$
$\gamma_6 = 1$	6	10	$= 1$	2	S_6L_6	$S_6L_6^2$	S_6L_6	305	$\gamma_{128}, \ldots, \gamma_{304}$
$\gamma_7 = 1$	7	12	$= 1$	2	S_7L_7	$S_7L_7^2$	S_7L_7	787	$\gamma_{305}, \ldots, \gamma_{786}$
$\gamma_8 = 0$	8	14	$\neq 1$	2	S_8L_8	$S_8^2L_8$	$S_8^2L_8$	2843	$\gamma_{787}, \ldots, \gamma_{2842}$
$\gamma_9 = 1$	9	15	$= 1$	2	L_9S_9	$L_9^2S_9$	$L_9^2S_9$	7742	$\gamma_{2843}, \ldots, \gamma_{7741}$
$\gamma_{10} = 0$	10	17	$\neq 1$	2	$L_{10}S_{10}$	$L_{10}S_{10}^2$	$L_{10}S_{10}$	12641	$\gamma_{7742}, \ldots, \gamma_{12640}$

We proceed in this way. The fixed point $s'(a) = 1c(a)$ is

$$\underbrace{\overbrace{110}^{S_2}\,\overbrace{110}^{S_2}\,\overbrace{11010}^{L_2}\,\overbrace{11011011010}^{L_3}\,\overbrace{11011010}^{S_3}}_{L_4}\,\underbrace{110110110101101101101011011010}_{L_4}\,\underbrace{1101101101011011010}_{S_4}\ldots$$

so the constructional word $\gamma(a) = 101101101011011011010\ldots$, which gives the following slope a:

$$[0;1,\underbrace{1,1}_{a_{i_a(2)}=1},2,\underbrace{1,1}_{a_{i_a(4)}=1},\underbrace{1,1}_{a_{i_a(5)}=1},2,\underbrace{1,1}_{a_{i_a(7)}=1},\underbrace{1,1}_{a_{i_a(8)}=1},2,1,1,2,1,1,1,1,2,1,1,1,1,2,1,1,1,1,2,1,1,2,\ldots].$$

Let us analyze the set of all fixed points of Δ_c.

Theorem 5. *Let* $\mathrm{Fix}(\Delta_c) \subset \mathcal{UM}_0$ *denote the set of all fixed points of* Δ_c*. Then:*

1. $\mathrm{Fix}(\Delta_c) \subset s'(]0, \frac{2}{3}[\setminus \mathbf{Q})$*; numbers* 0 *and* $\frac{2}{3}$ *are accumulation points of* $(s')^{-1}(\mathrm{Fix}(\Delta_c))$*.*
2. $\mathrm{card}(\mathrm{Fix}(\Delta_c))$ *is equal to that of the continuum.*

Proof. It is clear that we can go as near as we want towards 0. If we take $b_1 \to \infty$ then $[0; b_1, b_2, \ldots] \to 0$.

It remains to be shown that we cannot have a fixed point with slope larger than $\frac{2}{3}$. We will look for maximal a such that $1c(a)$ is a fixed point of Δ_c. First, to get as large slope as possible, we have to have $b_1 = 1$ (otherwise the slope is less than $\frac{1}{2}$). So, we proceed as in Example 6, $s'(a) = 11\ldots$, thus $\gamma_1 = 1$, so $a_2 = 1$, which means that $a_{i_a(2)} = 1$, so $i_a(3) = 4$ (a_3 is not of the form $a_{i_a(k)}$ for any $k \in \mathbf{N}^+$). The maximal possible slope of a fixed point begins with $[0; 1, 1, \ldots]$. We are absolutely free in the choice of the next element, because it does not affect the constructional word, as it is not a value of the index jump function. So, to make the slope maximal, we choose 1, because $[a_0; a_1, a_2, \ldots] < [a'_0; a'_1, a'_2, \ldots]$ iff $(a_0, -a_1, a_2, -a_3, a_4, -a_5, \ldots) \overset{\text{lexic.}}{<} (a'_0, -a'_1, a'_2, -a'_3, a'_4, -a'_5, \ldots)$, where the second inequality is according to the lexicographical order on sequences. Taking $b_3 \to \infty$ (thus, making the slope as large as possible), we get the limit value of $\frac{2}{3}$, because $[0; 1, 1, 1, b_3, \ldots] \to \frac{2}{3}$. We can also illustrate the solution with the following table:

given	k	$i_a(k+1)$	$a_{i_a(k+1)}$	b_{k+1}	S_{k+1}	L_{k+1}	gives P_{k+1}
$\gamma_1 = 1$	1	2	$= 1$	2	$S_1 L_1$	$S_1 L_1^2$	$S_1 L_1$
$\gamma_2 = 0$	2	4	$\neq 1$	b_3	$S_2^{b_3-1} L_2$	$S_2^{b_3} L_2$	$S_2^{b_3} L_2$

so the largest slopes of fixed points have the form $[0; 1, 1, 1, b_3, \ldots]$ and tend to $\frac{2}{3}$ when $b_3 \to \infty$.

$$s'(a) = S_2^{b_3} L_2 \ldots = (110)^{b_3}(11010)\ldots \overset{b_3 \to \infty}{\to} s'\left(\frac{2}{3}\right).$$

To prove the second statement of the theorem, we only need to recall that, according to Theorem 4, each sequence of length specification generates exactly one fixed point and each fixed point has its length specification (the same). The set of all fixed points has thus the same cardinality as the set of all sequences of length specification, which is the same as this of $\mathbf{N}^{\mathbf{N}}$. □

6 Conclusions and open problems

In this paper we have defined a run-construction encoding operator by analogy to the well-known run-length encoding operator and we formulated and proved a fixed-point theorem for Sturmian words. We also presented some combinatorial problems concerning quadratic surds (on p. 8, after Proposition 3). Some questions and problems arise also in connection with the run-construction encoding operator and the set of its fixed points. Theorem 5 gives us some answers. It states that the cardinality of the set of all fixed points is equal to that of the continuum (which follows from the main theorem of this paper, Theorem 4) and that no slopes of fixed points are larger than $\frac{2}{3}$. No fixed point of substitutions as described in Shallit (1991) [18] can be a fixed point of the run-construction encoding operator. Proposition 3 states that no quadratic surds can be slopes of fixed points of the operator.

There are still some problems to be solved. For example:

- Is the set of slopes of all fixed points of the run-construction encoding operator, i.e. the set $(s')^{-1}(\mathrm{Fix}(\Delta_c))$, dense in $]0, \frac{2}{3}[\setminus\mathbf{Q}$? Does it have accumulation points different from 0 and $\frac{2}{3}$?
- What kind of irrational numbers are the slopes of fixed points? Are they all transcendental?
- An algorithm finding fixed points related to the equivalence classes defined by sequences of length specification $(b_1, b_2, \ldots)$ could be written.
- How can we use the fixed points in digital geometry?

Acknowledgments. I am grateful to Christer Kiselman for comments on earlier versions of the manuscript.

References

1. Allouche, J.-P.; Shallit, J.: *Automatic Sequences*, Cambridge Univ. Press (2003).
2. Berstel, J.; Lauve, A.; Reutenauer, Ch.; Saliola, F.: *Combinatorics on Words: Christoffel Words and Repetition in Words*, CRM monograph series 27, Université de Montréal and American Mathematical Society (2008).
3. Berthé, V.; Ferenczi, S.; Zamboni, L. Q.: Interactions between Dynamics, Arithmetics and Combinatorics: the Good, the Bad, and the Ugly. *Contemporary Mathematics*, **385**, pp. 333–364, 2005.
4. Beskin, N.M.: *Fascinating Fractions*. Mir Publishers, Moscow (1986) (Revised from the 1980 Russian edition).
5. Brlek, S.: Enumeration of factors in the Thue–Morse word. *Discrete Applied Mathematics* **24** (1989) 83–96.
6. Brlek, S.; Jamet, D.; Paquin, G.: Smooth words on 2-letter alphabets having same parity. *Theoret. Comput. Sci.* **393** (2008) 166–181.
7. Bruckstein, A. M.: Self-Similarity Properties of Digitized Straight Lines, *Contemporary Mathematics* Vol. **119**, 1991.
8. Freeman, H.: Boundary encoding and processing. In *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld (Eds.), 241–266. New York and London: Academic Press 1970.
9. Istrate, G.: A note on self-reading sequences. *Discrete Applied Mathematics* **50** (1994) 201–203.
10. Karhumäki, J.: Combinatorics on words: A new challenging topic. In M. Abel, ed., *Proceedings of FinEst*, pp. 64–79. Estonian Mathematical Society, Tartu, 2004.
11. Khinchin, A. Ya.: *Continued Fractions*. Dover Publications, third edition (1997).
12. Kolakoski, W.: Self generating runs, Problem 5304, *Amer. Math. Monthly* **72** (1965) 674; Solution: *Amer. Math. Monthly* **73** (1966) 681–682.
13. Lothaire, M.: *Algebraic Combinatorics on Words*, Cambridge Univ. Press (2002).
14. Păun, G.; Salomaa, A.: Self-Reading Sequences. *Amer. Math. Monthly* **103** (1996) 166–168.
15. Perrin, D.; Pin, J.-E.: *Infinite Words; Automata, Semigroups, Logic and Games*. Pure and Applied Mathematics Vol. 141, Elsevier, 2004.
16. Pytheas Fogg, N.: *Substitutions in Dynamics, Arithmetics and Combinatorics*. Lecture Notes in Math. **1794**, Springer Verlag (2002).
17. Rosenfeld, A.: Digital straight line segments. *IEEE Transactions on Computers* **c-32**, No. 12, 1264–1269 (1974).
18. Shallit, J.: Characteristic Words as Fixed Points of Homomorphisms. Univ. of Waterloo, Dept. of Computer Science, Tech. Report CS-91-72, 1991.
19. Uscka-Wehlou, H.: Digital lines with irrational slopes. *Theoretical Computer Science* **377** (2007) 157–169.
20. Uscka-Wehlou, H.: Continued Fractions and Digital Lines with Irrational Slopes. In D. Coeurjolly et al. (Eds.): DGCI 2008, LNCS **4992**, pp. 93–104, 2008.
21. Uscka-Wehlou, H.: Run-hierarchical structure of digital lines with irrational slopes in terms of continued fractions and the Gauss map. *Pattern Recognition* **42** (2009) 2247–2254.
22. Uscka-Wehlou, H.: A Run-hierarchical Description, by Continued Fractions, of Upper Mechanical Words with Irrational Slopes. In *Proceedings of 12th Mons Theoretical Computer Science Days (Mons, Belgium), 27–30 August 2008*. Preprint: `http://wehlou.com/hania/files/uu/mons08rev.pdf`.
23. Uscka-Wehlou, H.: Two Equivalence Relations on Digital Lines with Irrational Slopes. A Continued Fraction Approach to Upper Mechanical Words. In press in *Theoretical Computer Science*, doi: 10.1016/j.tcs.2009.04.026 (2009).
24. Uscka-Wehlou, H.: Continued fractions, Fibonacci numbers, and some classes of irrational numbers. Submitted to a journal, 2009.
25. Vuillon, L.: Balanced words. *Bull. Belg. Math. Soc.* 10 (2003), 787–805.