SOME EXTREMALITIES OF THE BINARY FIBONACCI SEQUENCE

JEAN-PAUL ALLOUCHE

Abstract. The binary Fibonacci sequence is the sequence of 0's and 1's obtained by starting from 0 and iterating in parallel the substitution rules $0 \to 01$, $1 \to 0$ infinitely many times: the first construction steps successively yield the binary strings 0, 01, 010, 01001, ... (whose lengths are $1, 2, 3, 5, ...$). This sequence is in some sense one of the "simplest" non- periodic sequences. It can also be obtained by playing billiard on a square. In this survey we describe some "extremal" properties of the binary Fibonacci sequence and of similar sequences (the *Sturmian* sequences): in particular we recall unexpected inequalities involving these sequences and their shifted sequences.

1. Introduction

Starting from 0 and iteratively applying the substitution rules $0 \to 01$, $1 \to 0$ yields

$$0, \quad 01, \quad 010, \quad 01001, \quad 01001010, \quad \ldots$$

Note that the lengths of the successive strings of 0's and 1's above are $1, 2, 3, 5, 8, \ldots$ where the reader might recognize their favorite sequence. Also note that this sequence of strings of 0's and 1's gets "closer and closer" to an infinite sequence

$$F = 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots$$

which is thus called the *binary Fibonacci sequence*.

Is F a periodic sequence? If not, how close is F to being periodic? Are their other ways to generate this sequence?

Letting S denote the map that erases the first term of an infinite sequence, and S^k the map that erases the first k letters of an infinite sequence one has

$$\begin{array}{lcl} S^0(F) & = & 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^1(F) & = & 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^2(F) & = & 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^3(F) & = & 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ \ldots & & \end{array}$$

It is then immediate to see that $S^1(F)$ and $S^2(F)$ are smaller than $1F$ and larger than $0F$ for the lexicographical order[1] Is this also true for $S^k(F)$ for all $k \geq 1$? Are their other sequences with these properties? We will survey these questions and see how they are related to the *Sturmian sequences*. We will also allude to similar but different "extremal" properties of a non-Sturmian sequence, namely the Thue-Morse sequence.

[1] A binary sequence is lexicographically smaller than another binary sequence if they begin with a same –possibly empty– prefix followed by a 0 for the first sequence and by a 1 for the second sequence.

2. Some definitions

In this section we will give a few definitions. More details can be found in the books [9, 7, 4].

Definition 2.1. *An* alphabet $\mathcal{A}$ *is a finite set. If the cardinality of* $\mathcal{A}$ *is equal to* 2*, in particular if* $\mathcal{A} = \{0, 1\}$*,* $\mathcal{A}$ *is called a* binary *alphabet. The elements of an alphabet are called* letters*. A* word *on the alphabet* $\mathcal{A}$ *is a finite string of letters. We let* $\mathcal{A}^*$ *denote the set of all words on* $\mathcal{A}$ *(including the empty word* $\emptyset$ *with no letter). The total number of letters of a word* w *in* $\mathcal{A}^*$ *is called its length and denoted by* $|w|$*. In particular* $|\emptyset| = 0$*.*

We define on the set of words $\mathcal{A}^*$ an operation, the concatenation "." of two words, which consists of writing the two words one after the other, e.g., $011.10 = 01110$. This operation is clearly associative. The empty word $\emptyset$ is clearly a unit. Furthermore if v and w are two words in $\mathcal{A}^*$, then $|v.w| = |v| + |w|$. We have thus proved the following proposition.

Proposition 2.2. *Let* $\mathcal{A}$ *be an alphabet. Then* $(\mathcal{A}^*, .)$ *is a monoid, called the* free monoid *generated by* $\mathcal{A}$*. The map* $w \to |w|$ *is a homomorphism from* $(\mathcal{A}, .)$ *onto* $(\mathbb{N}, +)$ *(where* $\mathbb{N}$ *is the set of all nonnegative integers* $\{0, 1, 2, \ldots\}$*).*

As usual, having a structure (monoids) we look for the functions (homomorphisms of monoids) that preserve the structure. In other words we look for maps such that the image of the concatenation of two words is the concatenation of their images. We write a formal definition.

Definition 2.3. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be two alphabets. A* morphism g *from* $\mathcal{A}^*$ *to* $\mathcal{B}^*$ *is a map from* $\mathcal{A}^*$ *to* $\mathcal{B}^*$ *such that for any words* v *and* w *belonging to* $\mathcal{A}^*$*, one has* $g(v.w) = g(v).g(w)$*. [We let "." denote the concatenation both in* $\mathcal{A}^*$ *and* $\mathcal{B}^*$*.] A morphism* g *is said to be* non-erasing *if for any word* $w \in \mathcal{A}^*$ *such that* $w \neq \emptyset$*, then* $g(w) \neq \emptyset$*.*

Remark 2.4. *Of course a morphism* g *is completely defined by its values on the alphabet* $\mathcal{A}$*. Namely a word is equal to the concatenation of its letters (e.g., in* $\{0, 1\}^*$ *one has* $011 = 0.1.1$*).*

Example 2.5. *Let* $\mathcal{A} = \mathcal{B} = \{0, 1\}$*. Define* $f : \mathcal{A} \to \mathcal{A}^*$ *by* $f(0) = 01$*,* $f(1) = 0$*. Then* f *defines a morphism* $\mathcal{A}^* \to \mathcal{A}^*$*: e.g.,* $f(011) = f(0.1.1) = f(0).f(1).f(1) = 01.0.0 = 0100$*.*

Now we want to have a notion of convergence for a sequence of words on a given alphabet. To this effect we first introduce the set $\mathcal{A}^\infty$ which is the set of all (infinite) sequences $(a_n)_{n\geq 1}$ with values in $\mathcal{A}$. Two sequences are "close" if they coincide for a "large" initial range of indexes. A (finite) word can be considered as an infinite sequence by completing it with an infinite "tail" of symbols all equal to a "new" letter. More formally:

Definition 2.6. *Let* $\mathcal{A}$ *be an alphabet. Let* $\diamond$ *be a letter that does not belong to* $\mathcal{A}$*. Then any word* w *in* $\mathcal{A}^*$ *is identified with the sequence* $w \diamond \diamond \diamond \cdots$ *in* $(\mathcal{A} \cup \{\diamond\})^\infty$*.The set* $\mathcal{A}^\infty \cup \mathcal{A}^*$ *is identified with the subset of* $(\mathcal{A} \cup \{\diamond\})^\infty$ *that consists of sequences that either do not take the value* $\diamond$*, or whose values are eventually (i.e., from some index on) equal to* $\diamond$*. A* distance δ *is defined on* $(\mathcal{A} \cup \{\diamond\})^\infty$ *by: if* $X = (x_n)_{n\geq 1}$ *and* $Y = (y_n)_{n\geq 1} \in (\mathcal{A} \cup \{\diamond\})^\infty$*, then*

$$\delta(X, Y) = (1 + \inf\{n \geq 1;\ x_n \neq y_n\})^{-1},$$

where, if $X = Y$ *then* $\inf\{n \geq 1; x_n \neq y_n\} = +\infty$*, hence* $\delta(X, Y) = 0$*.*

In the sequel we will identify $\mathcal{A}^*$ with the set of infinite sequences in $(\mathcal{A} \cup \{\diamond\})^\infty$ that are eventually equal to $\diamond$.

Now coming back to Example 2.5, what happens when iterating f? Letting $f^{(k)}$ denote the k-th iterate of f (so that $f^{(0)}$ is just the identity map id) and applying $f^{(k)}$ to 0 yields

$$\begin{array}{ll} f^{(0)}(0) & = id(0) = 0 \\ f^{(1)}(0) & = f(0) = 01 \\ f^{(2)}(0) & = f(f(0)) = f(01) = 010 \\ f^{(3)}(0) & = f(f(f(0))) = f(f(01)) = f(010) = 01001 \\ & \ldots \end{array}$$

It looks like (and it can be proved) that the sequence of words $f^{(n)}(0)$, i.e., the sequence of words $0, 01, 010, 01001, \ldots$ converges (for the distance defined above) to the infinite word $F = 01001010010\ldots$. Furthermore, extending f to infinite words "by continuity" (i.e., $f(a_0a_1a_2\ldots) = f(a_0)f(a_1)f(a_2)\ldots$) one sees that the infinite word F is a fixed point of (the extended morphism) f, namely $f(F) = F$. This is an example of a general situation that is described in the following theorem.

Theorem 2.7. *Let $\mathcal{A}$ be an alphabet. Let g be a nonerasing morphism $\mathcal{A}^* \to \mathcal{A}^*$, such that there exists a letter $a \in \mathcal{A}$ with $g(a) = az$ where z is a non-empty word in $\mathcal{A}^*$. Then $g^{(k)}(a)$ converges to an infinite sequence which is a fixed point of the extension by continuity of morphism g. This infinite sequence is called the* iterated fixed point *of g beginning with a.*

3. The binary Fibonacci sequence

The sequence in Example 2.5 above is obtained as the limit of the sequence of words $0, 01, 010, 01001, \ldots$, whose lengths are $1, 2, 3, 5, \ldots$. This justifies the following definition.

Definition 3.1. *Let $\mathcal{A} = \{0, 1\}$. Let f be the morphism defined on $\{0, 1\}$ by $f(0) = 01$, $f(1) = 0$. The* binary Fibonacci sequence *F is the iterative fixed point of the morphism f beginning with 0. The first terms of F are $F = 01001010010\ldots$*

The binary Fibonacci sequence can also be defined as follows (see, e.g., [9, 4]).

Proposition 3.2. *Let F be the binary Fibonacci sequence. It has the following properties.*

(i) *F is the limit of the sequence of words U_n defined by $U_0 = 0$, $U_1 = 01$, and for every $n \geq 0$, $U_{n+2} = U_{n+1}U_n$.*

(ii) *Let $\varphi = (1 + \sqrt{5})/2$ be the golden ratio, so that $1/\varphi^2 = 0.381966014...$. For $n \geq 1$, let $x_n = \lfloor n/\varphi^2 \rfloor$, where $\lfloor y \rfloor$ is the integer part of the real number y. This yields the sequence $0, 0, 1, 1, 1, 2, 2, 3, \ldots$. The first difference of this sequence is equal to F.*

Proof. (partial) Here we only prove (i). It suffices to prove that $U_{n+1} = f^{(n)}(0) = f^{(n)}(U_0)$. Thus it suffices to prove that for all n one has $U_{n+1} = f(U_n)$. This is done by induction on n. For $n = 0$: $U_1 = 01 = f(0) = f(U_0)$. For $n = 1$: $U_2 = U_1U_0 = 010 = f(01) = f(U_1)$. Suppose that $U_{k+1} = f(U_k)$ for all $k \leq n$ for some $n \geq 1$. Then $U_{n+2} = U_{n+1}U_n = f(U_n)f(U_{n-1})$ by the induction hypothesis. But $f(U_n)f(U_{n-1}) = f(U_nU_{n-1}) = f(U_{n+1})$ and we are done. $\square$

Remark 3.3. *The first property given above for the binary Fibonacci sequence F is one more reason to call it Fibonacci.*

Remark 3.4. *The property given in Proposition 3.2 (ii) can be used to construct the sequence F as follows: the subset $\mathbb{N}^2$ of the lattice $\mathbb{Z}^2$ generates a square grid; take the half straight line of slope $1/\varphi$ starting from the origin and look at the places where it crosses the individual grid lines; mark a 0 if the grid line is vertical and a 1 if the grid line is horizontal; the sequence of 0's and 1's that is obtained is equal to F. Furthermore by "folding" this construction onto the* *unit square, one sees that F can be generated by playing (perfect) billiard on a square, starting from the bottom left corner with a slope equal to $1/\varphi$.*

Now we state two lemmas that will prove useful in the sequel.

Lemma 3.5. *The binary Fibonacci sequence is not eventually periodic.*

Proof. If a sequence $(v_n)_{n\geq 1}$ is eventually periodic, then for any value a taken by the sequence, the frequency of this value (i.e., the limit when N goes to infinity of $\sharp\{n \leq N,\ v_n = a\}/N$) exists and is rational. To prove that F is not eventually periodic, it thus suffices to use the definition of F given in Proposition 3.2 (i) and to prove that the number of 1's occurring in the word U_n divided by the length of U_n tends to an irrational value when n goes to infinity. Define $F_0 = 0$, $F_1 = 1$ and, for $n \geq 0$, $F_{n+2} = F_{n+1} + F_n$ (the F_n's are the Fibonacci numbers). It is easily proved by induction that the number of 1's occurring in U_n is equal to F_n, while the length of U_n is equal to F_{n+2}. Since F_n/F_{n+2} tends to $1/\varphi^2$ when n goes to infinity, we are done. $\square$

The proof of the next lemma is classical. Actually the statement – but not the proof – holds for all "Sturmian" sequences (see Definition 4.1).

Lemma 3.6. *If z is a block of consecutive letters occurring in F, then the blocks $0z0$ and $1z1$ cannot both occur in F.*

Proof. The lemma is proved by induction on the length of z. For $|z| = 0$ and $|z| = 1$ the statement is true since it is clear from the definition of F that the block 11 does not occur in F; neither do the blocks 000 and 111. Now let z be a word of length ≥ 2 occurring in F such that the statement of the lemma is true for all words of length $< |z|$ occurring in F, and such that $0z0$ and $1z1$ occur in F. Since $1z1$ occurs in F, the word z (of length ≥ 2) must begin and end with a 0 (the block 11 does not occur in F). Let $z = 0x0$, where x is possibly the empty word. Since $0z0 = 00x00$ occurs in F, we have that x is not empty (0000 does not occur in F); furthermore x is not reduced to a single letter (if $x = 0$, then $00x00 = 00000$ which cannot occur in F; if $x = 1$, then $1z1 = 10101$ which cannot occur in F, since it would be the image by f of 000 which does not occur in $F = f(F)$). Thus x, which must begin and end with a 1, is equal to $1y1$ for some word y. Now since $0x = 01y1$ occurs in $F = f(F) = f(0)f(1)f(0)f(0)\dots$ with $f(0) = 01$ and $f(1) = 0$, the only way this can happen is that there exists some block w in F such that $0x = 01y1 = f(w)$. Since $0z0 = 00x00$ is necessarily followed in F by a 1, the word $0z01 = 00x001$ occurs in F. But $00x001 = 0f(w)001$; thus we must have $00x001 = 0f(w)001 = f(1w10)$. In particular $1w1$ occurs in F. Since $1z1 = 10x01$ occurs in F and must be preceded by a 0, and since $010x01 = f(0w0)$, we finally have that both $0w0$ and $1w1$ occur in F. But $|w| \leq |f(w)| = |0x| < |z|$, contradicting the induction hypothesis. $\square$

We will see that, in some precise sense, the binary Fibonacci sequence (which is not periodic, as we have seen) is as close as possible to being periodic.

Definition 3.7. *The* block-complexity $(p_v(k))_{k\geq 1}$ *of a sequence $(v_n)_{n\geq 1}$ on an alphabet $\mathcal{A}$ is defined by: $p_v(k)$ is the number of distinct blocks of consecutive letters of length k occurring in the sequence $(v_n)_{n\geq 1}$.*

Remark 3.8. *For the block-complexity of $(v_n)_{n\geq 1}$ we clearly have: $1 \leq p_v(k) \leq (\sharp\mathcal{A})^k$ for all $k \geq 1$. A constant sequence has $p(k) = 1$ for all k, while any "random" sequence is expected to have all possible words on $\mathcal{A}$ occurring, thus to have $p(k) = (\sharp\mathcal{A})^k$. This gives a reason for calling p_v the (block-)complexity of the sequence v. The following classical proposition shows that periodic sequences are the sequences with the least possible complexity.*

Proposition 3.9 (Morse-Hedlund [10]). *Let $(v_n)_{n\geq 1}$ be a sequence on an alphabet $\mathcal{A}$ such that $\sharp\mathcal{A} \geq 2$. Let $(p_v(k))_{k\geq 1}$ be its block-complexity. Then the following properties are equivalent.*

(i) *The sequence $(v_n)_{n\geq 1}$ is eventually periodic (i.e., periodic from some index on);*
(ii) *$(p_v(k))_{k\geq 1}$ is eventually constant;*
(iii) *$(p_v(k))_{k\geq 1}$ is bounded;*
(iv) *There exists an integer $k_0 \geq 1$ such that $p_v(k_0) \leq k_0$;*
(v) *There exists a non-negative integer m such that $p_v(m) = p_v(m+1)$.*

Proof. Since the block-complexity p_v is integer-valued and non-decreasing, it is either eventually constant or it goes to infinity: this easily implies that (ii) and (iii) are equivalent and that they imply (iv). Also if p_v is increasing, one has for any $k \geq 1$ the inequality $p_v(k+1) > p_v(k)$, hence $p_v(k+1) \geq 1 + p_v(k)$; this implies by induction that for any $k \geq 1$ one has $p_v(k) \geq k - 1 + p_v(1) \geq k+1$. Thus, by contraposition, (iv) implies (v). Since the block-complexity of an eventually periodic sequence is clearly bounded, it remains to prove that (v) implies that the sequence v is eventually periodic.

Now we start from a sequence $(v_n)_{n\geq 1}$ such that its complexity p_v has the property: there exists an integer $m \geq 1$ such that $p_v(m+1) = p_v(m)$. This implies that any block of length m occurring in the sequence $(v_n)_{n\geq 1}$ can be extended on the right to a block of length $m+1$ occurring in the sequence in exactly one way. But this block of length $m+1$ can be extended to the right in only one way (think of its suffix of length m). Iterating this remark proves that given a word of length m occurring in the sequence, the sequence of letters that follow it is uniquely determined. In other words, given a block w of length m in the sequence, there exists a sequence $r(w)$ depending only on w such that the initial sequence ends with $wr(w)$. Looking at all the blocks of length m occurring in the initial sequence, there must be two of them occurring at two distinct places but being identical (recall that $\mathcal{A}$ is finite), say blocks w' and w'', but with $w' = w''$. They must be followed by the same infinite sequence $r(w') = r(w'')$: hence the initial sequence is eventually periodic. $\square$

This proposition clearly implies the following corollary.

Corollary 3.10. *Let $(v_n)_{n\geq 1}$ be a sequence on the alphabet $\mathcal{A}$ that is not eventually periodic. Let p_v be its block-complexity. Then, for all $k \geq 1$, one has $p_v(k) \geq k+1$.*

Thus, in the sense of block-complexity, a non-eventually periodic sequence that would satisfy $p_v(k) = k+1$ for all $k \geq 1$ would be as close to eventually periodic as possible. The binary Fibonacci sequence has this property.

Proposition 3.11. *Let F be the binary Fibonacci sequence. Let p_F be its block-complexity. Then, for all $k \geq 1$ one has $p_F(k) = k+1$.*

Proof. Since F is not eventually periodic from Lemma 3.5 (more precisely from the proof of (iv) $\Rightarrow$ (ii) in Lemma 3.5), we have $p_F(k) \geq k+1$ for all $k \geq 1$ from Corollary 3.10.

To prove that $p_F(k) \leq k+1$ for all $k \geq 1$, it suffices to prove that for all $k \geq 1$ one has $p_F(k+1) - p_F(k) \leq 1$. It thus suffices to prove that there is at most one block of each length occurring in F that can be extended (to the right) in two ways to a block occurring in F. In other words we would like to prove that for any length ℓ, there is at most one word w of length ℓ occurring in F such that both words $w0$ and $w1$ occur in F. Suppose that both w' and w'' have length ℓ and that $w'0$, $w'1$, $w''0$ and $w''1$ occur in F. If $w' \neq w''$, let y be their longest common suffix. Then there are letters preceding y in w' and in w'', and they must be distinct. Say $w' = x'0y$ and $w'' = x''1y$, where x' and x'' are possibly empty words. Thus the words $0y0$, $0y1$, $1y0$, and $1y1$ occur in F which is impossible from Lemma 3.6. $\square$

4. Sturmian sequences

A natural question is whether the properties of the binary Fibonacci sequence are specific, or whether they are shared by other sequences. We begin with the following definition.

Definition 4.1. *A sequence on an alphabet $\mathcal{A}$ is called a* Sturmian sequence *if its block-complexity p satisfies: $p(k) = k + 1$ for all $k \geq 1$.*

Example 4.2. *The binary Fibonacci sequence is a Sturmian sequence defined on the alphabet $\{0, 1\}$.*

Remark 4.3. *For a Sturmian sequence one has $p(1) = 2$, thus a Sturmian sequence is defined on an alphabet of cardinality 2.*

The following theorem is due to Morse and Hedlund and to Coven and Hedlund.

Theorem 4.4. [10, 5] *A sequence $(z_n)_{n\geq 1}$ defined on $\mathcal{A} = \{0, 1\}$ is Sturmian if and only if there exists an irrational number $\alpha \in (0, 1)$ and a real number β such that*
either $z_n = \lfloor (n + 1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor$ for all $n \geq 0$,
or $z_n = \lceil (n + 1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil$ for all $n \geq 0$,
where $\lfloor y \rfloor$ is the largest integer less than or equal to y and $\lceil y \rceil$ is the least integer larger than or equal to y.

We formulate one more definition.

Definition 4.5. *If $(z_n)_{n\geq 1}$ is a Sturmian sequence, let α and β be as in Theorem 4.4 above. The irrational α is called the* slope *of $(z_n)_{n\geq 1}$ while β is called the* intercept *of $(z_n)_{n\geq 1}$. If $\beta = 0$ the sequence $(z_n)_{n\geq 1}$ is called a* characteristic Sturmian sequence.

Example 4.6. *The Fibonacci sequence is a characteristic Sturmian sequence. (By Proposition 3.2 (ii), the nth term of the Fibonacci sequence is $\lfloor (n + 1)/\varphi^2 \rfloor - \lfloor n/\varphi^2 \rfloor$.)*

Remark 4.7. *There is (at least) one property of the binary Fibonacci sequence that most Sturmian or characteristic Sturmian sequences do not have, namely that F is an iterative fixed point of a binary morphism: the iterative fixed points of binary morphisms form a countable set, while the set of (characteristic) Sturmian sequences is not countable (consider the slopes).*

This section shows that Sturmian sequences share in particular two properties of the binary Fibonacci sequence, namely the one described in Theorem 4.4 and the fact (Definition 4.1) that they are as close as possible to periodic sequences in the sense of block complexity. The next section will give another general result involving Sturmian sequences and characteristic Sturmian sequences.

5. Inequalities involving Sturmian and characteristic Sturmian sequences

We begin with a definition.

Definition 5.1. *Let $X = (x_n)_{n\geq 1}$ and $Y = (y_n)_{n\geq 1}$ be two sequences with values in $\{0, 1\}$. Then X is said to be smaller in the* lexicographical order *than Y if there exists $n_0 \geq 1$ such that $x_n = y_n$ for all $n \leq n_0$ and $x_{n_0+1} = 0$ while $y_{n_0+1} = 1$. We note $X < Y$.*

Remark 5.2. *It is not difficult to see that the relation $X \leq Y$ if either $X = Y$ or $X < Y$ (in the lexicographical order), is a total order on the binary sequences.*

Recall from the introduction that the kth shift S^k of a sequence $(x_n)_{n\geq 1}$ is the sequence $(x_{n+k})_{n\geq 1}$ and that the first shifted sequences of the binary Fibonacci sequence are respectively

$$\begin{array}{lcl} S^0(F) & = & 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^1(F) & = & 1\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^2(F) & = & 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ S^3(F) & = & 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ \ldots \\ \ldots & & \end{array}$$

It is immediate that $0F \leq S^j(F) \leq 1F$ for $j = 1, 2, 3$. This property is actually true for all j, and even more: it is true for all characteristic sequences. If $U = (u_n)_{n\geq 1}$ is a binary sequence and a belongs to $\{0, 1\}$, we let aU denote the sequence $U' = (u'_n)_{n\geq 1}$ defined by $u'_1 = a$ and $u'_n = u_{n-1}$ for all $n \geq 2$. Then we have the following two theorems.

Theorem 5.3. *A non-eventually periodic sequence $U = (u_n)_{n\geq 1}$ on $\{0, 1\}$ is a characteristic Sturmian sequence if and only if, for all $k \geq 0$,*

$$0U < S^k(U) < 1U.$$

Furthermore, we have $0U = \inf\{S^k(U),\ k \geq 0\}$ and $1U = \sup\{S^k(U),\ k \geq 0\}$.

Theorem 5.4. *A non-eventually periodic sequence $V = (v_n)_{n\geq 1}$ on $\{0, 1\}$ is Sturmian if and only if there exists a sequence $U = (u_n)_{n\geq 1}$ on $\{0, 1\}$ such that $0U \leq S^k(V) \leq 1U$ for all $k \geq 0$. Moreover, U is the unique characteristic Sturmian sequence with the same slope as V, and we have $0U = \inf\{S^k(V),\ k \geq 0\}$ and $1U = \sup\{S^k(V),\ k \geq 0\}$.*

Remark 5.5. *Both theorems above were totally or partially discovered several times (see comments and references in* [3, Section 5]*).*

The two theorems above characterize Sturmian and characteristic Sturmian sequences. One can ask whether other sequences or families of sequences satisfy conditions of the same kind. We will see in the next section that this is indeed the case.

6. Other extremal properties of infinite sequences

In a 1983 paper Cosnard and the author, studying the iteration of unimodal functions from the unit interval to itself, introduced the set Γ defined by:

$$\Gamma = \{A \in \{0, 1\}^\infty,\ \forall k \geq 0,\ \overline{A} \leq S^k(A) \leq A\}$$

where $\overline{A}$ is the sequence obtained by exchanging 0's and 1's in A (see [2], also see [1]). They proved several properties of the set Γ, in particular a kind of fractal property of this set. It happens that the least non-periodic element in this set is the sequence $S^1(M) = 11010011001\ldots$, where M is the famous *Thue-Morse sequence* which can be defined as the iterative fixed point of the morphism $0 \to 01,\ 1 \to 10$. Other sequences similar to $S^1(M)$ and called q-mirror sequences play a specific rôle in the set Γ.

Curiously enough, a set $\widetilde{\Gamma}$ almost identical to Γ was studied independently in the paper [6] by Erdős, Joó and Komornik in 1990. This set is defined by:

$$\widetilde{\Gamma} = \{A \in \{0, 1\}^\infty,\ \forall k \geq 0,\ \overline{A} < S^k(A) < A\}.$$

It permits the characterization of the real numbers $\beta \in (1, 2)$ which are *univoque*, i.e., such that there is only one expansion of 1 as $1 = \sum_{k\geq 1} a_k \beta^{-k}$ with $a_k \in \{0, 1\}$.

Pairs of sequences satisfying similar inequalities occur, e.g., in a 1990 paper by Hubbard and Sparrow about *Lorenz maps* (see [8]). The authors consider *allowed pairs* (V, W) of distinct binary sequences in $\{0, 1\}^\omega$: they satisfy

$$W \leq S^k(W) < V \quad \text{and} \quad W < S^k(V) \leq V \quad \text{for all } k \geq 0.$$

Note that characteristic Sturmian sequences Z essentially correspond to $W = 0Z$ and $V = 1Z$, while the sequences Z in Γ and $\widetilde{\Gamma}$ essentially corespond to $V = Z$ and $W = \overline{Z}$.

Acknowledgments. We heartily thank C. Ballot, V. Bosser and D. Simon for their nice organization of the *Seventeenth International Conference on Fibonacci Numbers and Their Applications*, and the referee and Christian Ballot for very careful and useful readings of this paper.

References

[1] J.-P. Allouche, *Théorie des nombres et automates*, Thèse d'État, Université Bordeaux I, 1983, available at the URL `https://tel.archives-ouvertes.fr/tel-00343206/document`

[2] J.-P. Allouche, M. Cosnard, *Itérations de fonctions unimodales et suites engendrées par automates*, C. R. Acad. Sci. Paris, Série I **296** (1983), 159–162.

[3] J.-P. Allouche, A. Glen, *Extremal properties of (epi)Sturmian sequences and distribution modulo* 1, Enseign. Math. **56** (2010), 365–401.

[4] J.-P. Allouche, J. Shallit, *Automatic Sequences. Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[5] E. M. Coven, G. A. Hedlund, *Sequences with minimal block growth*, Math. Systems Theory **7** (1973), 138–153.

[6] P. Erdős, I. Joó, V. Komornik, *Characterization of the unique expansions* $1 = \sum q^{-n_i}$ *and related problems*, Bull. Soc. Math. France **118** (1990), 377–390.

[7] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Mathematics **1794**, Springer-Verlag, Berlin, 2002.

[8] J. H. Hubbard, C. T. Sparrow, *The classification of topologically expansive Lorenz maps*, Comm. Pure Appl. Math. **43** (1990), 431–443.

[9] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.

[10] M. Morse, G. A. Hedlund, *Symbolic dynamics II. Sturmian trajectories*, Amer. J. Math. **62** (1940), 1–42.

MSC2010: 11B39, 68R15, 11B85, 37B10, 11A63.

CNRS, Institut de Mathématiques de Jussieu-PRG, Équipe Combinatoire et Optimisation, Université Pierre et Marie Curie, Case 247, 4 Place Jussieu, F-75252 Paris Cedex 05 France

E-mail address: `jean-paul.allouche@imj-prg.fr`