THE COMPLEXITY OF FIBONACCI-LIKE KNEADING SEQUENCES

H. BRUIN AND O. VOLKOVA

ABSTRACT. The Fibonacci(-like) unimodal maps that have been studied in recent years give rise to a zero-entropy minimal subshift on two symbols, generated by the kneading sequence. In this paper we computed the word-complexity of such subshifts exactly.

1. INTRODUCTION

Topological entropy was introduced in the 1960s [1] as a way of classifying dynamical systems (as a topological invariant) and measuring their complexity. When studied from a symbolic viewpoint, the topological entropy indicates the exponential growthrate of the number of symbolic codes $p(n)$ that describe trajectories of length n in some alphabet. Also if the rate is 0 (e.g. substitution systems, piecewise isometries or polygonal billiards), $p(n)$ remains a useful measure of the complexity of the system, see [11] for a survey.

Let us fix our alphabet $\{0,1\}$. A (one-sided) *subshift* Σ is a shift-invariant, closed (in product topology) subset of $\{0,1\}^{\mathbb{N}}$. In this paper, we will only consider *minimal* subshifts, i.e. $\Sigma = \overline{\{\sigma^n(s)\}_n}$ for each $s \in \Sigma$, where σ denotes the right-shift. The *language* $\mathcal{L}$ of Σ is the collection of all finite words w (including the empty word ϵ) which are subwords of some (and hence all) $s \in \Sigma$. The *complexity* of the language $\mathcal{L}$ is the function

$$p(n) = \#\{w \in \mathcal{L} : \ |w| = n\},$$

where $|w|$ indicates the length of the word.

It is well-known [13] that $p(n) \geq n+1$ unless Σ consists of a single cycle of periodic strings. The complexity $p(n) = n+1$ is obtained by the Sturmian sequences which describe, among other things, the behavior of circle rotations. *Sublinear complexity* (i.e. $p(n) \leq Cn$ for some $C > 1$) is shared by substitution subshifts [21, 9, 2]. For instance, the Fibonacci substitution sequence $1011010110110\dots$ generated by the substitution $1 \to 10, \ 0 \to 1$ has complexity $p(n) = n+1$, because it happens to coincide with the Sturmian sequence describing the golden ratio circle rotation. Mossé et al. [16, 20] has worked out methods to compute the complexity in the

2000 *Mathematics Subject Classification.* Primary: 68R15 - Secondary: 37E05, 58F12 .

Key words and phrases. complexity, subshift, unimodal map, Fibonacci map.

The research was supported by a visitors grant of the London Mathematical Society.

case of substitutions of constant length, and J. Cassaigne, e.g. [8], has given more general methods, relying on the counting of left and right-special words. But for many cases to compute the complexity exactly remains an unsolved problem.

In this paper we study the complexity of a different class of subshifts, which stem from interval maps with specific combinatorial properties. Let $f : [0,1] \to [0,1]$ be a unimodal map, e.g. $f(x) = ax(1-x)$. The map has a unique critical point $c = \frac{1}{2}$, and $f|[0,c)$ is increasing and $f|(c,1]$ is decreasing. Let us call J a *branch* of f^n if it is a maximal closed interval on which f^n is monotone. The branch is called *central* if $c \in \partial J$. Because f is assumed to be symmetric, the image $D_n := f^n(J)$ is the same for both central branches. We call n a *cutting time* if $D_n \ni c$, and we write them in increasing order as $\{S_k\}$, starting with $S_0 = 1$. The combinatorial properties of f are completely determined by its cutting times. The maps that we are interested in satisfy the relation

$$S_k - S_{k-1} = \max\{1, S_{k-d}\} \text{ for a fixed } d \geq 1. \tag{1}$$

For $d = 1$, $S_k = 2^k$, and the corresponding map is the Feigenbaum map. For $d = 2$, the S_k are the Fibonacci numbers, and the corresponding map is known as *Fibonacci map*, [15]. For any $d \geq 1$, there exists a unimodal map f_d satisfying (1). The critical point c is recurrent, and its omega-limit set $\omega(c)$ is a minimal Cantor set. Fibonacci(-like) maps have drawn attention in the past decade because of their exceptional measure-theoretic properties, see [18, 6] and the general reference on unimodal maps [19]. For each d, there are unimodal polynomials satisfying (1) such that $\omega(c)$ is an attracting Cantor set [6, 4]. In [7], the spectral properties of $f|\omega(c)$ were investigated; it was shown that $f|\omega(c)$ is isomorphic to a $d-1$-dimensional torus rotation for $d = 2, 3, 4$, whereas for $d \geq 5$, $f|\omega(c)$ is weakly mixing. (For $d = 1$ (i.e. the Feigenbaum map) f acts on $\omega(c)$ as a dyadic adding machine.) In addition, it was shown that $f|\omega(c)$ is an almost one-to-one[1] factor of an (adic) enumeration system and for $d \geq 2$ also an almost one-to-one factor of a subshift on d symbols generated by $1 \to 1d,\ 2 \to 1,\ \dots, d \to d-1$. See [5] for generalizations.

For a symbolic approach we use standard kneading theory. For $x \in [0,1]$, define the *itinerary* $e(x) = e_1(x)e_2(x)\dots$ by

$$e_k(x) = \begin{cases} 1 & \text{if } f^k(x) > c, \\ 0 \text{ and } 1 & \text{if } f^k(x) = c, \\ 0 & \text{if } f^k(x) < c. \end{cases}$$

The itinerary of c, denoted as $K = e_1e_2e_3\dots$ is called the *kneading sequence*. For example, for the Fibonacci map is

$$K = 100111011001010011100100111011001 1\dots \tag{2}$$

The cutting times can be retrieved from K because they satisfy:

$$S_0 = 1 \text{ and } S_{k+1} = \min\{i > S_k : e_i \neq e_{i-S_k}\} \text{ for } k \geq 0.$$

[1]This means that a dense set of points in the factor space has only one preimage under the factor map.

The subshift corresponding to K is $\Sigma_d = \overline{\{\sigma^n(K)\}_{n\geq 1}}$. We call (Σ_d, σ) the *Fibonacci kneading shift*, or *Fibonacci-like kneading shift* for $d \geq 3$, in order to distinguish it from the Fibonacci substitution shift (based on substitution $1 \to 10,\ 0 \to 1$) and the Fibonacci subshift of finite type (with transition matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.) Note that (Σ_d, σ) is an almost one-to-one extension of $(\omega(c), f)$: the critical point and all its preimages have two itineraries, but all other points in $\omega(c)$ have only one.

The purpose of this paper is to compute the complexity of Σ_d for each $d \geq 1$.

Theorem 1.1. *The complexity of the Fibonacci kneading subshift satisfies:* $p(1) = 2$, $p(2) = 4$, $p(3) = 7$, *and if* $l \geq 4$, *then:*

$$p(l) = \begin{cases} 4l - S_{k-1} - 2 & \text{if } S_k \leq l < S_k + S_{k-3} \quad k \text{ is even,} \\ 4l - S_{k-1} - 3 & \text{if } S_k < l \leq S_k + S_{k-3} \quad k \text{ is odd,} \\ 3l + S_{k-1} - 2 & \text{if } S_k + S_{k-3} \leq l < S_{k+1} \quad k \text{ is even,} \\ & \text{if } S_k + S_{k-3} < l \leq S_{k+1} \quad k \text{ is odd.} \end{cases}$$

The complexity of the case $d = 1$, i.e. the Feigenbaum map was computed earlier by Rauzy [22]. We include it for completeness.

Theorem 1.2. *The complexity of the Feigenbaum subshift satisfies:* $p(1) = 2$, $p(2) = 3$, *and if* $l \geq 3$ *and for* k *such that* $2^k \leq l < 2^{k+1}$:

$$p(l) = \begin{cases} 2l - 2^{k-1} & \text{for } 2^k \leq l < 2^k + 2^{k-1}, \\ l + 2^k & \text{for } 2^k + 2^{k-1} \leq l < 2^{k+1}. \end{cases}$$

Remark: The Feigenbaum kneading sequence can be constructed in many other ways. It is the fixed point of the substitution $1 \to 10,\ 0 \to 11$, see [2], as well as a Toeplitz sequence. It also appears in studies on Beatty sequence, cf. [23]. The complexity of all possible itineraries (i.e. not restricted to $\omega(c)$) is known, see [3, Exercise 9.3.10].

Theorems 1.2 and 1.1 are special cases of the following result.

Theorem 1.3. *If* Σ_d *is the subshift corresponding to* $S_k - S_{k-1} = S_{k-d}$, *then if* $l \geq S_{d+1}$ *the complexity function satisfies*

$$p(l+1) - p(l) = \begin{cases} 2d & \text{for } S_{k-d+2} + t \leq l \leq S_{k-d+2} + S_{k-3d+3} + t - 1, \\ 2d - 1 & \text{for } S_{k-d+2} + S_{k-3d+3} + t \leq l \leq S_{k-d+1} + t - 1. \end{cases}$$

for $t \equiv k \bmod d$, *where* $k \bmod d$ *denotes the remainder in* $\{0, 1, \ldots, d-1\}$ *under division by* d.

Remark: Being a factor of a substitution shift (see above), results of Durand [10] show that any Fibonacci-like unimodal map has sublinear complexity (and is uniquely ergodic). The above theorem gives the exact complexity. Note that the number of subsequent l's where $p(l+1) - p(l) = 2d$ is S_k for some k.

Corollary 1.1. *The system Σ_d has sublinear complexity; more precisely: $2d-1 < \liminf_l p(l)/l < \limsup_l p(l)/l < 2d$.*

Therefore, these subshifts are no counter-example to the question raised in the Ph.D. thesis of Heinis [14]. There are known subshifts such that $\alpha := \lim_{n\to\infty} p(n)/n$ exists and is integer. Heinis showed that there is no subshift for $\alpha \in (1,2)$, leaving open the problem for non-integer values of α greater than 2.

Corollary 1.2. *If $d < \tilde{d}$, then $\Sigma_{\tilde{d}}$ is not a continuous factor of Σ_d.*

A word $w \in \mathcal{L}$ is called *right-special* if both *successors* $w0$ and $w1$ belong to $\mathcal{L}$. It is well-known that $p(n+1) - p(n)$ is precisely the number of right-special words of length n.

Theorem 1.4. *Let Σ_d be the Fibonacci-like subshift and $B_m = e_1 \dots e_m$ the initial m-word of the Fibonacci-like kneading sequence. Then the word w of length $l \geq S_{d+1}$ is right-special if w is a suffix of the word:*

$$\begin{cases} B_{S_{k+t}-1} & \text{for } S_{k-1} \leq l < S_k, \\ e_{S_{k+d}+1-t'} \dots e_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1} & \text{for } S_{k-d+2} + t \leq l \leq S_{k-d+2} + S_k + t' - 1 \end{cases}$$

where $t \equiv k \bmod d$ and $t' \equiv (k-1) \bmod d$.

Acknowledgement: We would like to thank the referees for valuable suggestions and for drawing our attention to some papers in the bibliography.

2. The lower bound for the number of right-special words

Let Σ_d be the subshift associated to the unimodal maps with cutting times satisfying $S_0 = 1$ and

$$S_k - S_{k-1} = \max\{1, S_{k-d}\}$$

Let $K = K_d = e_1 e_2 e_3 \dots$ be the kneading sequence and $B_m = e_1 \dots e_{m-1} e_m$ the prefix of K of length m. Let also $B'_m = e_1 \dots e_{m-1} e'_m$ be the same prefix with the last symbol changed to $e'_m := 1 - e_m$.

Lemma 2.1. *The kneading sequence can be constructed by the rule*

$$B_{S_d} = 100\dots0 \text{ and } B_{S_k} = B_{S_{k-1}} B'_{S_{k-d}} \text{ for } k > d.$$

Proof. Because $1, 2, \dots, d+1$ are all cutting times, $c_1 > 0$ and $D_i = [c_i, c_1]$ with $c_i < c$ for $i = 2, \dots, d+1$. Hence K starts with $e_1 e_2 \dots e_{d+1} = 100\dots0$. For the induction step, assume that $B_{S_{k-1}}$ for $k > d$ is given, and $D_{S_k} \ni c$. Since $S_k = S_{k-1} + S_{k-d}$ is the next cutting time, $f^{S_{k-d}}$ is monotone on $(c, c_{S_{k-1}})$. Therefore, $D_{S_{k-1}+i} \not\ni c$ for $i = 1, \dots, S_{k-d} - 1$, and hence $e_{S_{k-1}+i} = e_i$. However, $D_{S_k} \ni c$, so $e_{S_k} \neq e_{S_{k-d}}$. □

Lemma 2.2. *If* $t \equiv k \bmod d$*, then*

$$e_{S_k-t+1} \dots e_{S_k} = e_{S_{k+1}-t+1} \dots e_{S_{k+1}}$$

and

$$e_{S_k-t} \neq e_{S_{k+1}-t}.$$

Proof. Using a decomposition rule from Lemma 2.1 several times we get:

$$\begin{aligned} B_{S_k} &= B_{S_{k-1}} B'_{S_{k-d}} \\ &= B_{S_{k-1}} B_{S_{k-d-1}} B_{S_{k-2d}} \\ &= \dots B_{S_{k-(n-1)d-1}} B^{(\prime)}_{S_{k-nd}} \\ &= \dots B^{(\prime)}_{S_t} \end{aligned}$$

where $B^{(\prime)}_{S_{k-nd}}$ denotes $B_{S_{k-nd}}$ if n is even and $B'_{S_{k-nd}}$ otherwise.

$$\begin{aligned} B_{S_{k+1}} &= B_{S_k} B'_{S_{k+1-d}} \\ &= B_{S_k} B_{S_{k-d}} B_{S_{k+1-2d}} \\ &= \dots B_{S_{k-(n-1)d}} B^{(\prime)}_{S_{k+1-nd}} \\ &= \dots B^{(\prime)}_{S_{t+1}} \end{aligned}$$

It is easy to see that for $t = 0$: $B_{S_0} = 1$, $B_{S_1} = 10$, $B'_{S_0} = 0$, $B'_{S_1} = 11$ and therefore $e_{S_k-t} \neq e_{S_{k+1}-t}$. For $1 \le t < d$, Lemma 2.1 gives

$$B_{S_t} = 1\underbrace{00\dots0}_{t \text{ zeroes}},\; B_{S_{t+1}} = 1\underbrace{00\dots0}_{t+1 \text{ zeroes}} \text{ and } B'_{S_t} = 1\underbrace{00\dots0}_{t-1 \text{ zeroes}}1,\; B'_{S_{t+1}} = 1\underbrace{00\dots0}_{t \text{ zeroes}}1.$$

Therefore $e_{S_t-t+1} \dots e_{S_t} = e_{S_{t+1}-t+1} \dots e_{S_{t+1}}$ and $e_{S_t-t} \neq e_{S_{t+1}-t}$. □

In this section we present candidates of words w with two successors.

Proposition 2.1. (Case A) *Given* $l \ge d$ *and* k *such that* $S_{k-1} \le l < S_k$*, then the* l*-suffixes of* $B_{S_k-1}, B_{S_{k+1}-1}, \dots, B_{S_{k+d-1}-1}$ *are right-special and all different.*

Proof. Take $i \in \{k, \dots, k+d-1\}$. Clearly $we_{S_i} := e_{S_i-l} \dots e_{S_i}$ is the suffix of B_{S_i}. By Lemma 2.1, $B_{S_{i+d}} = B_{S_{i+d-1}} B'_{S_i}$ has suffix we'_{S_i}. Therefore both $w0$ and $w1$ appear in $\mathcal{L}$, proving that w is right-special.

The words w found this way are all different, because they have different suffixes, see Lemma 2.2.

Since B_{S_i-1} is a suffix of $B_{S_{i+d}-1}$, there are no more right-special words of this form. □

Proposition 2.2. (Case B) *Given* $l \ge d$ *and* k *such that*

$$S_{k-d+2} + t \le l \le S_k + S_{k-d+2} + t' - 1, \tag{3}$$

for $t \equiv k \bmod d$ *and* $t' \equiv (k-1) \bmod d$*. Then the* l*-suffix of*

$$e_{S_{k+d}+1-t'} e_{S_{k+d}+2-t'} \dots e_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1}$$

is right-special. Moreover, if two different values of k *satisfy* (3)*, then the corresponding* l*-suffixes are different.*

Proof. Using Lemma 2.1 repeatedly we get that

$$\begin{aligned} B_{S_{k+2d}+S_{k-d+2}} &= B_{S_{k+2d}} B_{S_{k-d+2}} \\ &= B_{S_{k+2d-1}} B'_{S_{k+d}} B_{S_{k-d+2}} \\ &= B_{S_{k+2d-1}} B_{S_{k+d-1}} B_{S_k} B_{S_{k-d+2}}, \end{aligned}$$

and

$$\begin{aligned} B_{S_{k+d+1}+S_{k-2d+2}} &= B_{S_{k+d+1}} B_{S_{k-2d+2}} \\ &= B_{S_{k+d}} B'_{S_{k+1}} B_{S_{k-2d+2}} \\ &= B_{S_{k+d}} B_{S_k} B_{S_{k-d+1}} B_{S_{k-2d+2}} \\ &= B_{S_{k+d}} B_{S_k} B'_{S_{k-d+2}}. \end{aligned}$$

By Lemma 2.2, $e_{S_{k+d-1}-t'+1} \dots e_{S_{k+d-1}} = e_{S_{k+d}-t'+1} \dots e_{S_{k+d}}$ for $t' \equiv (k+d-1) \bmod d$ and $e_{S_{k+d-1}-t'} \neq e_{S_{k+d}-t'}$. Therefore, if w is a suffix of $e_{S_{k+d}-t'+1} \dots e_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1}$, then both $w0$ and $w1$ appear in $\mathcal{L}$, so w is right-special. By Lemma 2.2, these suffixes w are different for different values of $k \bmod d$ when $|w| \geq d$.

In order to find out if these words are different from the right-special words in Case A, we decompose $B_{S_{k+2}-1} = B_{S_{k+1}} B_{S_{k-d+2}-1}$. By Lemma 2.2, $e_{S_{k+1}-t+1} \dots e_{S_{k+1}} B_{S_{k-d+2}-1}$ is the longest suffix of $B_{S_{k+2}-1}$ that is identical to a suffix of $B_{S_k} B_{S_{k-d+2}-1}$ for $t \equiv k \bmod d$. Therefore, for Case B to be disjoint from Case A, the length of the suffix should be at least $S_{k-d+2} + t$. □

Proof of Theorem 1.4 . This is a direct combinations of Propositions 2.1 and 2.2. Case A gives the suffixes of $B_{S_{k+t}-1}$ and case B is responsible for the suffixes of $e_{S_{k+d}+1-t'} \dots e_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1}$ □

3. The upper bound for the number of right-special words

Proposition 3.1. *Let* Σ_d *be the subshift for* $S_k = S_{k-1} + S_{k-d}$*. Then for* n *sufficiently large, there are at most* $2d$ *right-special words of length* n*.*

Given an n-word w, the *n-cylinder set* I_w is the set of points x whose itinerary $i(x)$ starts with w. Each cylinder set is an open interval[2], say $I_w = (c_{-a}, c_{-b})$ for some $0 < a < b < n$, where c_{-a} and c_{-b} indicate the appropriate points in $f^{-a}(c)$ and $f^{-b}(c)$ respectively. The n-cylinder sets partition the interval $[0,1]$, but $w \in \mathcal{L}_d$

[2]Except when I_w is adjacent to the boundary of $[0,1]$.

only if I_w intersects orb(c). More specifically, $w \in \mathcal{L}_d$ is a right-special word if I_w contains a point $c_{-n} \in f^{-n}(c)$ and both components of $I_w \setminus \{c_{-n}\}$ intersect orb(c).

Remark: From this observation, it follows that $\#\{w : w \text{ is right-special of length } n\} \geq \#\{f^{-n}(c) \cap \omega(c)\}$. In fact, $\#\{f^{-n}(c) \cap \omega(c)\} = d$ for each n sufficiently large (this follows from the construction of adic transformations factoring over $(\omega(c), f)$, see [7]). The corresponding words w are precisely the d right-special words of Case A, see Proposition 2.1.

A point in $z \in f^{-n}(c)$ is called a *closest precritical point* if $f^n|(c, z)$ is monotone. If z is a closest precritical point, then n is a cutting time. Indeed, if J is a central branch of f^n, then $J \ni z$ and $c = f^n(z) \in f^n(J)$. Let z_k denote the closest precritical points with $n = S_k$, where the context should make clear if z_k is to the left or to the right of c.

Lemma 3.1. *If $I_w = (c_{-a}, c_{-b})$ is an n-cylinder and $0 < a < b < n$, then $b - a$ is a cutting time. If w is right-special, then $c_{-n} \in I_w$ for some $c_{-n} \in f^{-n}(c)$, and both $n - b$ and $n - a$ are cutting times.*

Proof. The interval $(c, c_{a-b}) = f^a(I_w)$ is contained in the central branch of f^{b-a}. Because $c = f^{b-a}(c_{a-b}) \in f^{b-a}(J)$, $b - a$ is a cutting time.

If w is right-special, then both $w0$ and $w1$ are allowed words, and hence $f^n(I_w)$ must intersect both $[0, c)$ and $(c, 1]$. Hence $f^{-n}(c) \cap I_w \neq \emptyset$. Both components of $I_w \setminus \{c_{-n}\}$ are $n+1$-cylinders, so by the above arguments, both $n - a$ and $n - b$ are cutting times. □

Lemma 3.2. *Recall that the image-closure of the central branch of f^n is $D_n = [c_n, c_{\beta(n)}]$ for $\beta(n) = n - \max\{S_k < n\}$. For every n, $D_{\beta(n)} \supset D_n$.*

Proof. This was proven by induction in [7, Lemma 5]. □

In the next lemma, we will use the notation $Q(k) = \max\{0, k - d\}$, so $S_k - S_{k-1} = S_{Q(k)}$.

Lemma 3.3. *The point $c_{S_k} \in (z_{Q(k+1)}, z_{Q(k+1)-1})$, and if r is such that $S_r < S_r + S_k < S_{r+1}$, then $z_{Q(k+1)} \in D_{S_r+S_k}$ if and only if $Q(r+1) = k+1$. (In this case, $c_{S_k+S_r} \in (z_{Q(k+1)+1}, z_{Q(k+1)})$.*

Proof. Let l be minimal such that $z_l \in (c, c_{S_k}) \subset D_{S_k}$. Then $S_k + S_l = S_{k+1}$ is the first cutting time after S_k. But this means that $S_l = S_{k+1} - S_k = S_{Q(k+1)}$. This proves the first statement.

For the second statement, notice that $D_{S_k+S_r} = f^{S_r}(c, c_{S_k})$, and $c \notin D_{S_r+S_k}$. Moreover, by Lemma 3.2 $D_{S_r+S_k} \subset D_{S_k}$. If $z_{Q(k+1)} \in D_{S_r+S_k}$, then $S_r + S_k + S_{Q(k+1)} = S_r + S_{k+1}$ is the first cutting time after S_r. In other words, $Q(r+1) = k+1$. In this case, $f^{S_{Q(k+1)}}(z_{Q(k+1)+1}) = z_{Q(Q(k+1)+1)}$, whereas $c_{S_{r+1}} \in (z_{Q(r+2)}, z_{Q(r+2)-1})$. Because $Q(Q(k+1)+1) = Q(Q^2(r+1)+1) < Q(r+2)$, $z_{Q(k+1)+1} \notin D_{S_r+S_k}$.

Now for the reverse implication of the "if and only if" statement, assume that $Q(r+1) = k+1$. Then $S_{r+1} = S_r + S_{Q(r+1)} = S_r + S_{k+1}$, so $S_{r+1} - (S_r + S_k) = S_r + S_{k+1} - (S_r + S_k) = S_{Q(k+1)}$. Therefore $D_{S_r+S_k}$ must contain a point in $f^{-S_{Q(k+1)}}(c)$. But $f^{-S_{Q(k+1)}}(c) \cap (z_{Q(k+1)}, z_{Q(k+1)-1}) = \emptyset$ because of the construction of closest precritical points. Therefore $D_{S_r+S_k} \ni z_{Q(k+1)}$.

Note that because Q is onto and is strictly increasing for $k \geq d+1$, there is one and only one r such that $Q(r+1) = k+1$. □

A word $w \in \Sigma_d$ of length n is right-special only if the corresponding cylinder set I_w has the property that $c_{-n} \in I_w$ and both components of $I_w \setminus \{c_{-n}\}$ contain a point from orb(c), say c_L and c_R respectively. In Figure 1, we drew the possibilities for the configurations of the corresponding sets D_L and D_R.

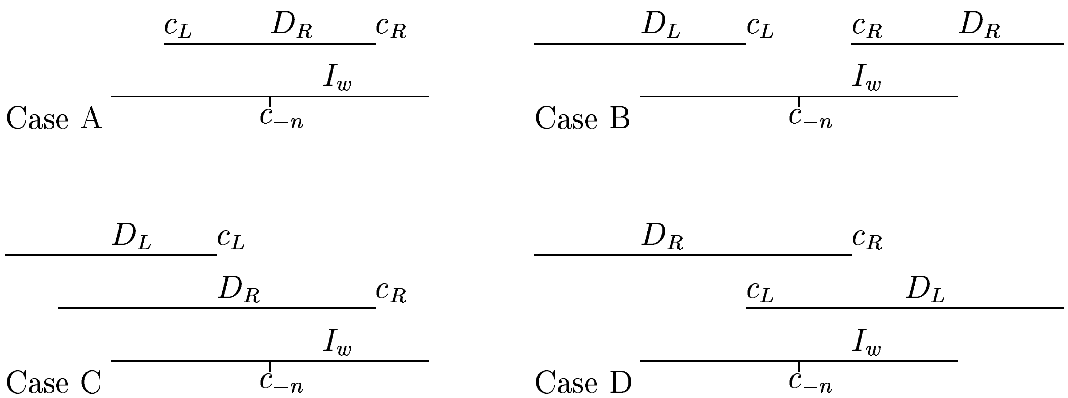

FIGURE 1. Configurations of D_L and D_R with respect to I_w.

Lemma 3.4. *Let $I_w = (c_{-a}, c_{-b})$ be an n-cylinder set containing c_{-n}. If c_j is such that $c_j \in I_w$ and $D_j \ni c_{-n}$, then there exists i such that c_j is a boundary point of D_i and $c_{-n} \in D_i \subset I_w$.*

Proof. If $D_j \subset I_w$, then we can take $i = j$ and there is nothing to prove. Otherwise, D_j contains a boundary point of I_w, say $c_{-a} \in D_j$. It follows that $D_{j+a} \ni c$ and $c_{a-n} \in (c, c_{j+a})$, and there are no closest precritical points of lower order between c_{j+a} and c_{a-n}. Obviously, $j+a$ is a cutting time, say S_k, so $c_{a-n} = z_{Q(k+1)}$. By Lemma 3.3, there exists r such that $c_{a-n} \in D_{S_r+S_k} \subset (c, c_{S_k}]$. But then $D_{S_r+S_k-a} = D_{S_r+j} \ni c_{-n}$; this is the required interval. □

Corollary 3.1. *Cases C and D reduce to Case A in Figure 1.*

Proof. In Case C, the previous lemma gives an interval $D_j \subset D_R$ such that $D_i \ni c_{-n}$ and the endpoint $c_{\beta(i)}$ of D_i equals c_R. Hence the interval D_i is in Case A. A similar argument works for Case D. □

Proof of Proposition 3.1. The above discussion showed that whenever w is a right-special word, there are D_L and D_R as in Figure 1, Case A or B. (They correspond to Cases A and B in the previous section.

Case A: Since $c_{-n} \in D_R \subset I_w$, f^n maps D_R monotonically onto $D_{R+n} \ni c$. Hence $S_{k-1} < R < R+n = S_k$ for some cutting time S_k, and $w = e_{S_k - n} \dots e_{S_k - 1}$ is the suffix of $B_{S_k - 1}$. Because the suffix of length n are the same for B_{S_k-1} and $B_{S_{Q(k)}-1} = B_{S_{k-d}-1}$ for each k (provided $n < S_{Q(k)} - S_{Q(k)-1}$), there are at most d different words w of this type. In the previous section, we saw that there are exactly d different words w of this type.

Case B: Let us assume that $a < b$ and $D_R \ni c_{-b}$. (Otherwise we interchange the role of D_R and D_L.) Assume also that c_R is closest to c_{-n} among all c_i between c_{-n} and c_{-b} such that $D_i \ni c_{-b}$. This means that $D_{\beta(R)} \supset I_w$. Indeed, $D_{\beta(R)} = [c_{\beta^2(R)}, c_{\beta(R)}]$ and if $c_{\beta^2(R)} \in (c_{-n}, c_{-b})$, then c_R was not closest to c_{-n}, contradiction the above assumption. If $c_{\beta^2(R)} \in (c_{-a}, c_{-n})$, then, due to Lemma 3.3, we can reduce this case to Case A.

Apply the iterate f^a. Then the picture is as Figure 2. Rename $R' = R + a$, and

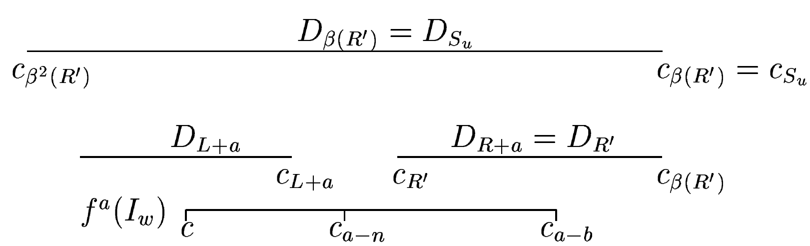

FIGURE 2. Case B after applying iterate f^a.

note that $\beta(R') = \beta(R) + a$. Because $D_{\beta(R)} \ni c_{-a}$, $D_{\beta(R')} \ni c$, so $\beta(R') =: S_u$ is a cutting time, with $b - a = S_{Q(u+1)}$ and $n - a = S_{Q(u+1)+1}$. This follows because c_{a-b} and c_{a-n} are adjacent closest precritical points. By Lemma 3.3, $R' = S_u + S_t$ where $Q(t+1) = u + 1$. It follows that

$$n - b = n - a - (b - a) = S_{Q(u+1)+1} - S_{Q(u+1)} = S_{Q(Q(u+1)+1)} = S_{Q(Q^2(t+1)+1)}.$$

Since $R + b = R' + (b - a) = S_u + S_t + S_{Q(u+1)} = S_t + S_{u+1} = S_{t+1}$, we have

$$w = e_R \dots e_{R+b} e_{R+b+1} \dots e_{R+n-1} = e_{S_{t+1}-b+1} \dots e_{S_{t+1}} e_1 \dots e_{S_{Q(Q^2(t+1)+1)}-1}.$$

Using Lemma 2.1 repeatedly, we get

$$\begin{aligned} B_{S_{t+1}+S_{Q(Q^2(t+1)+1)}-1} &= B_{S_{t+1}} B_{S_{Q(Q^2(t+1)+1)}-1} \\ &= B_{S_t} B'_{S_{Q(t+1)}} B_{S_{Q(Q^2(t+1)+1)}-1} \\ &= B_{S_t} B_{S_{Q(t+1)-1}} B_{S_{Q^2(t+1)}} B_{S_{Q(Q^2(t+1)+1)}-1} \\ &= B_{S_t} B_{S_{Q(t+1)-1}} B_{S_{Q^2(t+1)+1}-1} \\ &= B_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1}, \end{aligned}$$

for $k = Q(t+1) - 1 = u$. Therefore w is a suffix of $B_{S_{k+d}} B_{S_k} B_{S_{k-d+2}-1}$, and $n = |w| > n - a = S_{k-d+1}$. The maximal possible length is given in Proposition 2.2. □

4. The remaining proofs

Proof of Theorem 1.3. Proposition 2.1 gives d right-special words of length l observing Case A. By Proposition 2.2, a right-special word of Case B "appears" for $l = S_{k-d+2} + (k \bmod d)$, and "disappears" again at $l = S_k + S_{k-d+2} + (k-1 \bmod d)$. Write $t = (k \bmod d)$. The smallest number $\geq S_{k-d+2} + t$ of the form $S_{\tilde{k}} + S_{\tilde{k}-d+2} + ((\tilde{k}-1) \bmod d)$ is $S_{k-d+1} + S_{k-2d+3} + (k \bmod d) = S_{k-d+2} + S_{k-3d+3} + (k \bmod d)$. Therefore the number of right-special l-words is maximal $(= 2d)$ if $S_{k-d+2} + t \leq l \leq S_{k-d+2} + S_{k-3d+3} + t - 1$ for $t \equiv k \bmod d$. □

Proof of Theorems 1.1 and 1.2. Theorem 1.2 follows directly from Theorem 1.3. For Theorem 1.1, Theorem 1.3 implies that if $S_k + S_{k-3} \leq l < S_{k+1}$ (or $S_k + S_{k-3} < l \leq S_{k+1}$ depending on whether k is even or odd), there have been $k-2$ blocks of length $S_0, S_1, \dots, S_{k-3}$ where $p(i+1) - p(i) = 4$. It follows by induction that $\sum_{j=0}^{k-3} S_j = S_{k-1} - 2$. For other values of i $(4 \leq i < l)$, $p(i+1) - p(i) = 3$. Therefore $p(l) = 3l + S_{k-1} + C$, and a single check shows that the constant $C = -2$. A similar argument gives $p(l)$ for other values of l. □

Proof of Corollary 1.1. The cutting times satisfying $S_k - S_{k-1} = S_{k-d}$ increase exponentially. Between $S_{k-d+2} + (k \bmod d)$ and $S_{k-d+3} + (k+1 \bmod d)$ there is a block of length S_{k-3d+3} where $p(i+1) - p(i) = 2d$ and a block of length $\approx S_{k-2d+3} - S_{k-3d+3} = S_{k-2d+2}$ where $p(i+1) - p(i) = 2d-1$. The length of these blocks are comparable to S_{k-d+2}. Therefore

$$\begin{aligned} \liminf_l \frac{p(l)}{l} &= \lim_k \frac{p(S_{k-d+2} + (k \bmod d))}{S_{k-d+2} + (k \bmod d)} \\ &< \lim_k \frac{p(S_{k-d+2} + S_{k-3d+3} + (k \bmod d) - 1)}{S_{k-d+2} + S_{k-3d+3} + (k \bmod d) - 1} = \limsup_l \frac{p(l)}{l}, \end{aligned}$$

and $2d-1 < \liminf_l p(l)/l < \limsup_l p(l)/l < 2d$. □

Proof of Corollary 1.2. It is well-known (see e.g. [17]) that if $\varphi : \Sigma_d \to \Sigma_{\tilde{d}}$ is a semiconjugacy, then φ is generated by a sliding block code $\varphi(x)_i = \Phi(x_i \dots x_{i+N})$ for some N, independently of x. Therefore (cf. [12] and [2, Corollary 3.1.1]), each l-word $w \in \mathcal{L}_{\tilde{d}}$ is uniquely determined by an $l+N$-word $v \in \mathcal{L}_d$:

$$w_1 \dots w_l = \Phi(v_1 \dots v_{1+N}) \dots \Phi(v_l \dots v_{l+N}).$$

It follows that $p_{\Sigma_{\tilde{d}}}(l) \leq p_{\Sigma_d}(l+N)$ for all l, contradicting Theorem 1.3. □

Remark: Recall that for the Fibonacci map f, $(\omega(c), f)$ is a factor of the Fibonacci substitution shift Σ_{sub}. Yet the complexity of Σ_{sub} is $p(l) = l+1$ (it is a Sturmian subshift), whereas the Fibonacci kneading subshift Σ_2 has complexity $p(l) \geq 3l$ for l sufficiently large. This shows that the factor map $\pi : \Sigma_{sub} \to \omega(c)$ does not extend to a continuous factor map $\tilde{\pi} : \Sigma_{sub} \to \Sigma_2$. Indeed, as c has two itineraries $i(c) = \{0K(f)m,\ 1K(f)\}$ in Σ_2, defining $\tilde{\pi} = i \circ \pi_{sub}$ makes it double-valued. One can remedy this by giving c only one itinerary, say $0K(f)$, but then $i \circ \pi_{sub}$ is not continuous anymore.

References

[1] R. Adler, A. Konheim, M. McAndrew, *Topological entropy,* Trans. Amer. Math. Soc. **114** (1965) 309–319.
[2] P. Arnoux, V. Berthé, S. Ferenczi, S. Ito, C. Mauduit, M. Mori, J. Peyrière, A. Siegel, J.-I. Tamura, Z.-Y. Wen, *Substitutions in dynamics, arithmetics and combinatorics,* Lect. Notes. Math. **1794** (2002).
[3] K. Brucks, H. Bruin, *Topics from one-dimensional dynamics,* London Mathematical Society, Student Texts **62** Cambridge University Press 2004.
[4] H. Bruin, *Topological conditions for the existence of absorbing Cantor sets,* Trans. Amer. Math. Soc. **350** (1998) 2229–2263.
[5] H. Bruin, *Minimal Cantor systems and unimodal maps,* J. Difference Eq. Appl. **9** (2003) 305–318.
[6] H. Bruin, G. Keller, T. Nowicki and S. van Strien, *Wild Cantor attractors exist,* Ann. of Math. **143** (1996) 97-130.
[7] H. Bruin, G. Keller and M. St. Pierre, *Adding machines and wild attractors,* Ergod. Th. and Dyn. Sys. **17** (1997) 1267-1287.
[8] J. Cassaigne, *Complexité et facteurs spéciaux,* Bull. Math. Simon Stevin, **4** (1997) 67-88.
[9] F. Durand, *A characterization of substitutive sequences using return words,* Discrete Math. **179** (1998) 89-101.
[10] F. Durand, *Linearly recurrent subshifts have a finite number of non-periodic subshift factors,* Ergod. Th. & Dynam. Sys. **20** (2000), 1061–1078.
[11] S. Ferenczi, *Complexity of sequences of dynamical systems,* Combinatorics and number theory (Tiruchirappalli, 1996), Discrete Math. **206** (1999) 145–154.
[12] S. Ferenczi, *Rank and symbolic complexity,* Ergod. Th.& Dyn. Sys. **16** (1996) 663–682.
[13] G. Hedlund, M. Morse, *Symbolic dynamics II. Sturmian trajectories,* Amer. J. Math. **62** (1940) 1-42.

[14] A. Heinis, *Arithmetics and combinatorics of words of low complexity,* Ph.D. Thesis, Leiden 2003.
[15] F. Hofbauer, G. Keller, *Some remarks on recent results about S-unimodal maps,* Ann. Inst Henri Poincaré **53** (1990) 413–425.
[16] S. Jaeger, R. Lima, B. Mossé, *Symbolic analysis of finite words: the complexity function* Bull. Braz. Math. Soc. **34** (2003) 457–477.
[17] D. Lind, B. Marcus, *An introduction to symbolic dynamics and coding,* Cambridge Univ. Press 1995.
[18] M. Lyubich, J. Milnor, *The Fibonacci unimodal map,* J. Amer. Math. Soc. **6** (1993) 425-457.
[19] W. de Melo, S. van Strien, *One-dimensional dynamics,* Springer, Berlin Heidelberg New York, (1993).
[20] B. Mossé, *Reconnaissabilité des substitutions et complexité des suites automatiques,* Bull. Soc. Math. France **124** (1996), 329-346
[21] J.-J. Pansiot, *Complexité des facteurs des mots infinis engendrés par morphismes itérés,* in ICALP 1984, Lecture Notes in Computer Science **172** Springer (1984) 380-389.
[22] G. Rauzy, *Suites a valeurs dans un alphabet fini,* Seminaire de Theorie des Nombres de Bordeaux **25** (1982-1983).
[23] R. Tijdeman, *Fraenkel's conjecture for six sequences,* Discrete Math. **222** (2000) 223–234.

Department of Mathematics and Statistics, University of Surrey, Guildford GU2 7XH United Kingdom

E-mail address: H.Bruin@surrey.ac.uk

Institute of Mathematics, National Academy of Sciences of Ukraine, Tereshchenkivska street 3, 01601 Kiev, Ukraine

E-mail address: oxana@imath.kiev.ua