Linearly recurrent systems

V. Berthé*
CNRS, IRIF, Université Paris Diderot

Exercises

Exercise 1. Prove that a clopen set is a finite union of cylinders.

Exercise 2. Build an example of an infinite word that is recurrent but not uniformly recurrent.

Exercise 3. Prove that an infinite word u is recurrent if and only if the shift map S is onto on $\overline{\mathcal{O}(u)}$.

Exercise 4. Prove that every substitution σ on the alphabet $\mathcal{A}$ defines a continuous map from $\mathcal{A}^{\mathbb{N}}$ to $\mathcal{A}^{\mathbb{N}}$. Give an example of a substitution for which σ is one-to-one, and an example of a substitution for which σ is not one-to-one.

Exercise 5 (Recognizability)**.** Let $\sigma : \mathcal{A}^* \to \mathcal{B}^*$ be a non-erasing morphism (the image of any non-empty word is non-empty). Let $x \in \mathcal{A}^{\mathbb{Z}}$, and let

$$C_\sigma(x) = \{|\sigma(x_{[0,\ell)})| : \ell \geq 0\} \cup \{-|\sigma(x_{[\ell,0)})| : \ell < 0\}.$$

A non-erasing morphism σ is said to be recognizable on x if there exists ℓ such that, for each $m \in C_\sigma(x)$, $m' \in \mathbb{Z}$, $y_{[m-\ell,m+\ell)} = y_{[m'-\ell,m'+\ell)}$ implies that $m' \in C_\sigma(x)$, with $y = \sigma(x)$.

Is the Thue-Morse substitution $\sigma\colon a \mapsto ab$, $b \mapsto ba$ recognizable on $\sigma^\infty(a)$? Same question for the Fibonacci substitution $\tau\colon a \mapsto ab$, $b \mapsto a$.

Let σ be a primitive morphism that is recognizable on some x and injective on letters. Let $X = \overline{\mathcal{O}(x)}$ and $Y = \cup_{k\in\mathbb{Z}} S^k\sigma(X)$, where S is the shift map.

Prove that the set $\sigma(X)$ is a clopen subset of Y, that the map $\sigma : X \to \sigma(X)$ is a homeomorphism, and that the collection $\mathcal{P} = \{S^k\sigma([a]) : a \in \mathcal{A}, 0 \leq k < |\sigma(a)|\}$ is a clopen partition of Y.

Let us note that recognizability has been proved by Mossé (1992) for primitive substitutions, and by Bezuglyi, Kwiatkowski and Medynets (2009) for any aperiodic substitution.

What can be said for the substitution $\sigma\colon 0 \mapsto 010$, $1 \mapsto 10$ in the case of one-sided words in $\mathcal{A}^{\mathbb{N}}$ and one-side recognizability?

*Many thanks to S. Barbieri for his careful reading and suggestions.

Exercise 6. Most infinite words we shall consider have an at most linear factor complexity function ($\forall n \in \mathbb{N}$, $p_u(n) \leq Cn$). In this case, even if the shift is not injective, this exercise shows that it can be made invertible up to a set which is at most countable.

Let u be a recurrent infinite word such that its factor complexity satisfies $p_u(n) \leq Cn$ for all n and some constant C. Prove that there exists a finite set F such that, if D is the (at most countable, and shift-invariant) set $\bigcup_{n\in\mathbb{Z}} S^n F$, the shift S is one-to-one from $X_u \setminus D$ to $X_u \setminus D$.

We can note that if u is recurrent and not eventually periodic, then the set X_u is not countable, and $X_u \setminus D$ is not empty.

Exercise 7. For any $s \geq 1$, construct an example of an infinite word u with factor complexity $p_u(n) = s + n$ for n large enough. Same question with u uniformly recurrent.

Exercise 8 (Factor complexity)**.** Let $u \in \mathcal{A}^{\mathbb{N}}$ with the cardinality of $\mathcal{A}$ being finite. Prove that if there exists n such that $p_u(n+1) = p_u(n)$, then u is eventually periodic. (A sequence is said eventually periodic if it is periodic from a certain index on.) Deduce that that u is eventually periodic if and only if there exists $C > 0$ such that $p_u(n) \leq Cn$ for all n.

What happens in the case of a sequence defined over $\mathbb{Z}$?

Let $X \subset \mathcal{A}^{\mathbb{N}}$ be a subshift. Do the infinite words in X all have the same factor complexity?

Let u be an infinite word such that the set of its factors is not equal to $\mathcal{A}^*$. Prove that there exists a real number α, with $1 \leq \alpha < \mathrm{Card} A$, such that $p_u(n) = O(\alpha^n)$.

Does there exist an infinite word with factor complexity $[\log(n)]$?

What is the factor complexity of a coding of an irrational rotation on $\mathbb{R}/\mathbb{Z}$ by two intervals?

Give an exemple of a two-dimensional word in $\mathcal{A}^{\mathbb{Z}^2}$ having at least one period and an unbounded rectangular factor complexity, where the rectangular factor complexity counts the number of rectangular factors of a given size.

What is the rectangular factor complexity of the two-dimensional word u in $\{0,1\}^{\mathbb{Z}^2}$ defined by: $u_{m,n} = 1$ if $m\sqrt{2} + n\pi \in [0, 1/2)$ modulo 1, for all $(m,n) \in \mathbb{Z}^2$?

Exercise 9 (Sturmian words)**.** A Sturmian word is an infinite word u with factor complexity function satisfying $p_u(n) = n + 1$ for all n. Prove that a Sturmian word is recurrent, and even uniformly recurrent. Give an example of a bi-infinite word u with factor complexity function satisfying $p_u(n) = n + 1$, for all n, that is not recurrent.

A Sturmian word can also be described as a binary coding of a rotation. Prove that the language $\mathcal{L}_u$ associated with a Sturmian word u is palindromic, that is, the reverse word $w_n \cdots w_1$ of any word $w_1 \cdots w_n$ in $\mathcal{L}_u$ also belongs to $\mathcal{L}_u$.

Prove that the set of factors of a Sturmian word u is balanced, that is, for any factors v, w of u of the same length, one has $||v|_1 - |w|_1| \leq 1$.

Exercise 10. An infinite word $u \in \mathcal{A}^{\mathbb{N}}$ is said to be C-balanced if for any factors v, w of u of the same length, one has $||v|_i - |w|_i| \leq C$, for any $i \in \mathcal{A}$. Prove that a C-balanced infinite word admits uniform frequencies for letters.

Give an example of a 2-balanced infinite word.

Exercise 11 (Factor complexity of the Fibonacci word)**.** The Fibonacci word is the fixed point of $\sigma \colon 0 \mapsto 01,\ 1 \mapsto 1$.

1. Prove that every factor w of the Fibonacci word can be uniquely written as follows:
$$w = a\sigma(v)b,$$
where v is a factor of the Fibonacci word, $a \in \{\varepsilon, 1\}$, and $b = 0$, if the last letter of w is 0, and $b = \varepsilon$, otherwise.

2. Prove that if w is a left special factor distinct from the empty word, then there exists a unique left special factor v such that $w = \sigma(v)b$, where $b = 0$, if the last letter of w is 0, and $b = \varepsilon$, otherwise. Deduce the general form of the left special factors.

3. Prove that the Fibonacci word is not ultimately periodic.

4. Prove that the complexity function of the Fibonacci word equals $n+1$ for every n.

Exercise 12. Is the word *abaabaabababaab* a factor of some Sturmian word?

Exercise 13. Let σ be the substitution defined on $\{0,1\}$ by $\sigma \colon 0 \mapsto 001$, $1 \mapsto 1$. Prove that the fixed point starting with 0 of σ has quadratic factor complexity.

Exercise 14 (Arnoux-Rauzy words)**.** Let $\mathcal{A} = \{1, 2, \ldots, d\}$. The set of Arnoux-Rauzy substitutions is defined as $\mathcal{S}_{AR} = \{\mu_i \mid i \in \mathcal{A}\}$ where

$$\mu_i :\ i \mapsto i,\ j \mapsto ij \text{ for } j \in \mathcal{A} \setminus \{i\}\,.$$

An *Arnoux-Rauzy word* is an infinite word $u \in \mathcal{A}^{\mathbb{N}}$ whose set of factors coincides with the set of factors of a sequence of the form

$$\lim_{n\to\infty} \mu_{i_0}\mu_{i_1}\cdots\mu_{i_n}(1),$$

where the sequence $(i_n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}}$ is such that every letter in $\mathcal{A}$ occurs infinitely often in $(i_n)_{n\geq 0}$. The sequence $(\mu_{i_n})_n$ is called the directive sequence of $u = \lim_{n\to\infty} \mu_{i_0}\mu_{i_1}\cdots\mu_{i_n}(1)$.

Prove that one recovers Sturmian words in the case $d = 2$.

Prove that for every sequence $(i_n)_n$ and for any $i = 1, \cdots, d$, then $\lim_{n\to\infty} \mu_{i_0}\mu_{i_1}\cdots\mu_{i_n}(i)$ exists.

Compare the words $\lim_{n\to\infty} \mu_{i_0}\mu_{i_1}\cdots\mu_{i_n}(i)$, for $i = 1, \cdots, d$.

Prove that two distinct directive sequences produce distinct words.

Consider now the set of substitutions $\mathcal{S}'_{AR} = \{\mu_i \mid i \in \mathcal{A}\}$ where

$$\mu'_i :\ i \mapsto i,\ j \mapsto ji \text{ for } j \in \mathcal{A} \setminus \{i\}\,.$$

Let $(i_n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}}$ be such that every letter in $\mathcal{A}$ occurs infinitely often in $(i_n)_{n\geq 0}$. Prove that $\lim_{n\to\infty} \mu'_{i_0}\mu'_{i_1}\cdots\mu'_{i_n}(1)$ exists, and that $\lim_{n\to\infty} \mu_{i_0}\mu_{i_1}\cdots\mu_{i_n}(1)$, and $\lim_{n\to\infty} \mu'_{i_0}\mu'_{i_1}\cdots\mu'_{i_n}(1)$, have the same set of factors.

Prove that the fixed point starting with 1 of the Tribonacci substitution $\sigma\colon 1 \mapsto 12,\ 2 \mapsto 13,\ 3 \mapsto 1$ is an Arnoux-Rauzy word.

Prove that every prefix w of an Arnoux-Rauzy word u satisfies $iw \in \mathcal{L}_u$ for all $i \in \mathcal{A}$. Prove that Arnoux-Rauzy words have factor complexity $2n+1$ and are uniformly recurrent.

Prove that an Arnoux-Rauzy word is linearly recurrent if and only if it has bounded strong partial quotients, that is, each substitution of $\mathcal{S}_{AR}$ occurs in its directive sequence $(\mu_{i_n})_n$ with bounded gaps.

Exercise 15 (Episturmian words)**.** The reversal of a word $w_1\cdots w_n$ is the word $w_n\cdots w_1$. A factor w of an infinite word u is said to be special it there exists at least two distinct letters a and b such that wa and wb are factors of u. An infinite word is said to be episturmian if the set of its factors is closed under reversal and has at most one right special factor of each length. Prove that Arnoux-Rauzy words are episturmian. Prove that an episturmian infinite word is uniformly recurrent and has uniform frequencies.

Exercise 16 (Chacon word)**.** The Chacon morphism σ is defined over the alphabet $\{0,1\}$ by $\sigma : 0 \mapsto 0010,\ 1 \mapsto 1$. Observe that the substitution σ is not primitive. Prove that the Chacon word $x = \sigma^{\omega}(0)$ begins with the following sequence of words $(b_n)_{n\geq 0}$:

$$b_0 = 0, \text{ and } \ \forall n \in \mathbb{N},\ b_{n+1} = b_n b_n 1 b_n.$$

Deduce that the Chacon word x is uniformly recurrent. Deduce also that the Chacon word has uniform frequencies. Deduce that $X_x = \overline{\mathcal{O}(x)}$ is uniquely ergodic and minimal.

Exercise 17 (Rauzy graph of words)**.** The Rauzy graph Γ_n of words of length n of an infinite word u on a finite alphabet $\mathcal{A}$ is an oriented graph. Its vertices are the factors of length n of u and its edges are defined as follows: there is an edge from U to V if V follows U in the infinite word u, i.e., if there exists a word W and two letters x and y such that $U = xW$, $V = Wy$ and xWy is a factor of the u. There are $p_u(n+1)$ edges and $p_u(n)$ vertices, where $p_u(n)$ denotes the factor complexity function of u.

Prove that the graphs of words Γ_n of u are all connected. Prove that the infinite word u is recurrent if and only if the graphs of words Γ_n are all strongly connected.

Let U be a vertex of the graph Γ_n. Denote by U^+ the number of edges of Γ_n with origin U and by U^- the number of edges of Γ_n with end vertex U. In other words, U^+ (respectively U^-) counts the number of right (respectively left) extensions of U in u.

What are the possible shapes of a graph of words Γ_n for a Sturmian word? Deduce that Sturmian words are uniformly recurrent. Same question for Arnoux-Rauzy words.

We now restrict ourselves to recurrent infinite words for which the frequencies of factors do exist. Let

U and V be two vertices linked by an edge such that $U^+ = 1$ and $V^- = 1$. What can be said on the frequencies of the two factors U and V?

Prove that for a recurrent infinite word u of factor complexity function $p_u(n)$, the frequencies of factors of length n take at most $3(p(n+1) - p(n))$ values.

Prove that the frequencies of factors of length n take at most $p(n+1) - p(n) + r_n + l_n$ values, where r_n (respectively l_n) denotes the number of factors having more than one right (respectively left) extension. Deduce that if u is a recurrent with at most linear factor complexity, then the frequencies of its factors of given length take a finite number of values.

Exercise 18 (Topological conjugacy)**.** A topological dynamical system (Y, T) is a topological factor of (X, S) if there exists a continuous map π from X onto Y which conjugates the maps S and T, i.e., $\pi \circ S = T \circ \pi$. If π is moreover injective, then it is said to be a topological conjugacy. Observe that π^{-1} is continuous and that it also conjugates T and S.

Let u, v be two infinite words with values in finite alphabets, and let X_u, X_v denote respectively the associated symbolic dynamical systems.

1. Suppose that (X_v, S) is a topological factor of (X_u, S), where S denotes the shift. Let ϕ denote the conjugation map from X_u onto X_v. Prove that the map ϕ satisfies the following: there exists a positive integer q such that for every i, the coordinate of index i of $\phi(x)$ depends only on $(x_i, \ldots, x_{i+q})$.

   More generally, Curtis-Hedlund-Lyndon theorem states that every factor map is given by a sliding block code (that is, a map defined by a local rule).

2. Deduce that if (X_u, S) and (X_v, S) are topologically conjugate, then they have the same topological entropy, and if (X_v, S) is a topological factor of (X_u, S), then $H_{top}(v) \leq H_{top}(u)$.

3. Prove that if (X, T) and (Y, T) are conjugate topologically symbolic systems, then there exists a constant c such that, for all $n > c$,

$$p_X(n - c) \leq p_Y(n) \leq p_X(n + c).$$

   Prove also that the boundedness of first-order and second-order differences of the complexity function is preserved.

Prove that topological conjugacy preserves unique ergodicity.

Exercise 19. Give an example of a non-minimal non-uniquely ergodic subshift (X, S) such that any infinite word x in X has uniform frequencies.

Exercise 20. Let u be an infinite word with uniform factor frequencies, that is, for any factor w, the sequence $\frac{|u_k \ldots u_{k+n}|_w}{n+1}$ tends to a limit f_w, uniformly in k.

For every word w, consider the shift-invariant measure μ defined by $\mu([w]) = f_w$. Prove that $\frac{1}{N}\sum_{n=0}^{N-1} g(S^n x) \to \int g\, d\mu$ for every x in the subshift $X_u = \overline{\mathcal{O}(u)}$ and for every continuous function g. Deduce that the subshift (X_u, S) is uniquely ergodic.

Exercise 21. Prove that every sequence in a linearly recurrent subshift is linearly recurrent with the same constant. Prove that a linearly recurrent infinite word has uniform frequencies.

Exercise 22. Let G be a compact metric group. Let $T : G \to G$, $x \mapsto ax$ be a rotation of G. Prove that (G, T) is minimal if and only if $\{a^n,\ n \in \mathbb{N}\}$ is dense in X. Prove that (G, T) is minimal if and only if it is uniquely ergodic.