Combinatorial and Dynamical study of substitutions around the Theorem of Cobham

Fabien Durand
Centro de Modelamiento Matemático
UMR 2071 UCHILE–CNRS
Blanco Encalada 2120, Piso 7
Casilla 170, Correo 3
Santiago, Chile
e-mail : fdurand@dim.uchile.cl

April 9, 2002

1 Introduction and motivations

Given a subset E of $\mathbb{N} = \{0, 1, 2, \cdots\}$ can we find an elementary algorithm which accepts the elements of E and rejects those that do not belong to E ?

By "elementary algorithm" we mean a finite state automaton. This question originates from the work of Büchi (1960, [Bu]). Cobham gave two answers to this question. In 1969 he proved that the existence of such an algorithm deeply depends on the numeration base, more precisely :

First Cobham's Theorem. *Let p and q be two multiplicatively independent integers greater or equal to 2. Then, a set $E \subset \mathbb{N}$ is both p-recognizable and q-recognizable if and only if E is the finite union of arithmetic progressions.*

Where "p-recognizable" means that there exists an automaton which accepts exactly the language consisting in the expansions in base p of the elements of E. For example, we will see that the set $\{2^n; n \in \mathbb{N}\}$ is 2-recognizable and, as it is not a finite union of arithmetic progressions, it can not be 3-recognizable. Cobham's Theorem also says that the set $\{2n; n \in \mathbb{N}\}$ is p-recognizable for any $p \in \mathbb{N}$. But it does not tell us anything about the structure of recognizable sets of integers. For example, we can not deduce whether the set of prime numbers is recognizable or not ? In fact it is not (see [Co2] for the details) using the second answer of Cobham which gives a complete description of their structure :

Second Cobham's Theorem. *A set $E \subset \mathbb{N}$ is p-recognizable if and only if its characteristic sequence $\mathsf{x} \in \{0, 1\}^{\mathbb{N}}$ ($\mathsf{x}_i = 1$ if and only if $i \in E$) is the image by a letter to letter morphism of a fixed point of a substitution of constant length.*

From this result we remark there are few recognizable sets, the family of recognizable sets is numerable.
The original proof of the First Cobham's Theorem was considered as "almost elementary" but highly technical. In 1974, Eilenberg suggested in [Ei] to find some more readable proofs. Hansel gave some ideas in this matter [Ha1] . Later, in 1993, Michaux and Villemaire [MV] found a new proof of this Theorem using the formalism of the first order logic. In the same time other characterizations were given in terms of congruences with finite index [Ei] and in terms of algebraic power series [CKMR]. Due to all these characterizations, many generalizations of the First Cobham Theorem can be stated (see [BHMV] for a very nice overview). Semenov [Se] gave the first one extending it to recognizable subsets of $\mathbb{N}^n$. The most prolific generalization, in the sense that a lot of work has been done in this direction [BH1, BH2, BP, Du3, Fab1, Fab2, Ha2, Sh], propose to obtain the same kind of results for non-standard numeration systems such as the numeration system given by the Fibonacci sequence $U_0 = 1, U_1 = 2$ and $U_{n+2} = U_{n+1}+U_n$. In [Du3] it is proved that such a result can be obtained for a large class of linear numeration systems (including the Fibonacci numeration system), more precisely :

Theorem 1 (systnum) *Let U and V be two Bertrand numeration systems, α and β be two multiplicatively independent β-numbers such that $L(U) = L(\alpha)$ and $L(V) = L(\beta)$, and E a subset of $\mathbb{N}$. If E is U-recognizable and V-recognizable then E is a finite union of arithmetic progressions.*

We will see that this result is a corollary of more general results concerning substitutions. This is a result of Fabre [Fab2] (Theorem 49) which enables us to apply substitution results to non-standard numeration systems. It is analogous to the Second Cobham's Theorem but for non-standard numeration systems. In [Bes] the author found a weaker result but with completely different methods using first order logic (the characteristic polynomial of the linear numeration systems he considered have to be the minimal polynomial of some Pisot numbers).
Let us describe the way it will be proved in this paper. First we need to remark that a set $E \subset \mathbb{N}$ is a finite union of arithmetic progressions if and only if its characteristic sequence is ultimately periodic. Hence the First Cobham Theorem can be formulated in the following equivalent way : *Let p and q be two multiplicatively independent integers greater or equal to 2. Let A be a finite alphabet and $\mathsf{x} \in A^{\mathbb{N}}$. Then the sequence x is generated by both a substitution of length p and a substitution of length q if and only if it is ultimately periodic.*
To a substitution σ is associated an integral square matrix $M \neq 0$ which has non-negative coefficients. It is known (see [LM] for instance) that such a matrix has a real eigenvalue α which is greater than the modulus of all others eigenvalues. It is usually called the dominant eigenvalue of M. If x is the image by a letter to letter morphism of a fixed point of σ then we will say that x is α-substitutive. An easy computation shows that if σ is of length p then $\alpha = p$. Furthermore if a sequence is generated by a substitution of length p then it is p-substitutive. Note that the converse is not true. This suggests the following conjecture stated by G. Hansel.

Conjecture. *Let α and β be two multiplicatively independent Perron numbers. Let A be a* *finite alphabet. Let* x *be a sequence of* $A^{\mathbb{N}}$. *Then,* x *is both* α*-substitutive and* β*-substitutive if and only if* x *is ultimately periodic.*

This is true when we restrict to primitive substitutive sequences [Du2] (Theorem 30) and also in some non-primitive case (Theorem 46). This is from Theorem 46 we will get Theorem 1. Moreover there is an analogous result for substitutive subshifts. More precisely we have :

Theorem 2 *Let* (X, T_X) *and* (Y, T_Y) *be two subshifts generated by respectively an* α*-substitutive primitive sequence* x *and a* β*-substitutive primitive sequence* y*, where* α *and* β *are multiplicatively independent. Then,* (X, T_X) *and* (Y, T_Y) *are isomorphic if and only if there are periodic with the same period.*

The main notion of this paper is the notion of *return word*. The following theorem ([Du1, HZ1]) shows the relevance of this notion for substitution. It makes evident the "self-induced" structure of substitutions and provide a very useful tool to study substitutions, namely the *derivative sequences.*

Theorem 3 (caracterisation) *Let* x *be a uniformly recurrent sequence. The following are equivalent*

1. x *is a substitutive sequence ;*
2. *the set of its u-derivative sequences is finite, u being a word of* $L(\mathsf{x})$ *;*
3. *the set of its u-derivative sequences is finite, u being a prefix of* x.

In terms of subshifts this theorem says that subshifts generated by primitive substitutions only have a finite number of induced systems on cylinders.
The paper is organized as follows. Section 2 is devoted to the definitions concerning sequences and subshifts and to some results. We recall Perron's Theorem which will be used quite often. In Section 3 we prove Theorem 3 and make some comments about substitutive subshifts. We prove Theorem 2 in Section 4. We also recall a nice result of Holton and Zamboni saying that, for the unique shift-invariant measure, the measures of cylinders of a minimal substitution subshift lie in a finite union of arithmetic progressions and we point out it gives a new proof of Theorem 2. Numeration systems appear in Section 5 where Theorem 1 is proved.

2 Definitions and background

2.1 Words, sequences and morphisms

Words and sequences. We call *alphabet* a finite set of elements called *letters*. Let A be an alphabet, a *word* on A is an element of the free monoïd generated by A, denoted by A^*, i.e. a finite sequence (possibly empty) of letters. Let $x = x_0 x_1 \cdots x_{n-1}$ be a word, its *length* is n and is denoted by $|x|$. The *empty-word* is denoted by ϵ, $|\epsilon| = 0$. The set of non-empty words on A is denoted by A^+. If $J = [i, j]$ is an interval of $\mathbb{N} = \{0, 1 \cdots\}$ then x_J denote

the word $x_i x_{i+1} \cdots x_j$ and is called a *factor* of x. We say that x_J is a prefix of x when $i = 0$ and a suffix when $j = n - 1$. If u is a factor of x, we call *occurrence* of u in x every integer i such that $x_{[i,i+|u|-1]} = u$. Let u and v be two words, we denote by $L_u(v)$ the number of occurrences of u in v.
The elements of $A^{\mathbb{N}}$ are called *sequences*. For a sequence $\mathsf{x} = (\mathsf{x}_n; n \in \mathbb{N}) = \mathsf{x}_0\mathsf{x}_1 \cdots$ we use the notation x_J and the terms "occurrence" and "factor" exactly as for a word. The set of factors of length n of x is written $L_n(\mathsf{x})$, and the set of factors of x, or *language* of x, is represented by $L(\mathsf{x})$; $L(\mathsf{x}) = \cup_{n \in \mathbb{N}} L_n(\mathsf{x})$. The sequence x is *periodic* if it is the infinite concatenation of a word v. A *gap* of a factor u of x is an integer g which is the difference between two successive occurrences of u in x. We say that x is *uniformly recurrent* if each factor has bounded gaps. We have the following property which proof is left to the reader.

Proposition 4 *If a sequence is ultimately periodic and uniformly recurrent then it is periodic.*

A way to evaluate the complexity of a sequence x is the so called *symbolic complexity*. It is the integer function $p_{\mathsf{x}} : \mathbb{N} \to \mathbb{N}$ where $p_{\mathsf{x}}(n)$ is the number of different words of length n. The following classical result about complexity asserts that if x is not ultimately periodic then its complexity is at least $n + 1$ (see [HM]).

Proposition 5 *Let* x *be a sequence. If for some* n $p_{\mathsf{x}}(n) \leq n$ *then* x *is ultimately periodic.*

Morphisms. Let A, B and C be three alphabets. A *morphism* τ is a map from A to B^*. Such a map induces by concatenation a map from A^* to B^*. If $\tau(A)$ is included in B^+, it induces a map from $A^{\mathbb{N}}$ to $B^{\mathbb{N}}$. All these maps are written τ also.
To a morphism τ, from A to B^*, is naturally associated the matrix $M_\tau = (m_{i,j})_{i \in B, j \in A}$ where $m_{i,j}$ is the number of occurrences of i in the word $\tau(j)$. To the composition of morphisms corresponds the multiplication of matrices. For example, let $\tau_1 : B \to C^*$, $\tau_2 : A \to B^*$ and $\tau_3 : A \to C^*$ be three morphisms such that $\tau_1 \circ \tau_2 = \tau_3$ (we will quite often forget the composition sign), then we have the following equality: $M_{\tau_1} M_{\tau_2} = M_{\tau_3}$. In particular if τ is a morphism from A to A^* we have $M_{\tau^n} = M_\tau^n$ for all non-negative integers n.

2.2 Dynamical systems and subshifts

By a *dynamical system* we mean a pair (X, S) where X is a compact metric space and S a continuous map from X onto itself. We say that it is a *Cantor system* if X is a Cantor space. That is, X has a countable basis of its topology which consists of closed and open sets and does not have isolated points. The system (X, S) is *minimal* whenever X and the empty set are the only S-invariant closed subsets of X. We say that a minimal system (X, S) is *periodic* whenever X is finite. We say it is p-periodic if $\text{Card}(X) = p$.
Let (X, S) and (Y, T) be two dynamical systems. We say that (Y, T) is a *factor* of (X, S) if there is a continuous and onto map $\phi : X \to Y$ such that $\phi S = T \phi$ (ϕ is called *factor map*). If ϕ is one-to-one we say that ϕ is an *isomorphism* and that (X, S) and (Y, T) are *isomorphic*.

Let (X,S) be a minimal Cantor system and $U \subset X$ be a clopen set. Minimality implies that $\min\{n > 0 \ ; \ S^n(x) \in U\}$ exists for all $x \in U$. Hence, for all $x \in U$ we can define $S_U : U \to U$, the *induced transformation on* U, by

$$S_U(x) = S^{r_U(x)}(x), \text{ where } r_U(x) = \min\{n > 0 \ ; \ S^n(x) \in U\}.$$

The pair (U, S_U) is a minimal Cantor system, we say that (U, S_U) is the *induced system of* (X,S) *with respect to* U.
In this paper we deal with Cantor systems called *subshifts*. Let A be an alphabet. We endow $A^{\mathbb{N}}$ with the topology defined by the metric

$$d(\mathsf{x},\mathsf{y}) = \frac{1}{2^n} \ \text{ with } \ n = \text{Inf}\{k; \mathsf{x}_k \neq \mathsf{y}_k\},$$

where $\mathsf{x} = (\mathsf{x}_n; n \in \mathbb{N})$ and $\mathsf{y} = (\mathsf{y}_n; n \in \mathbb{N})$ are two elements of $A^{\mathbb{N}}$. By a *subshift* on A we shall mean a couple $(X, T_{/X})$ where X is a closed T-invariant ($T(X) = X$) subset of $A^{\mathbb{N}}$ and T is the *shift transformation*

$$\begin{array}{rccl} T & : & A^{\mathbb{N}} & \to & A^{\mathbb{N}} \\ & & (\mathsf{x}_n; n \in \mathbb{N}) & \mapsto & (\mathsf{x}_{n+1}; n \in \mathbb{N}). \end{array}$$

We call language of X the set $L(X) = \{\mathsf{x}_{[i,j]}; \mathsf{x} \in X, i \leq j\}$. Let u and v be two words of A^*. The set

$$[u]_X = \{\mathsf{x} \in X; \mathsf{x}_{[|u|)} = u\}$$

is called *cylinder*. The family of these sets is a base of the induced topology on X. When it will not create confusion we will write $[u]$ and T instead of $[u]_X$ and $T_{/X}$.
Let x be a sequence on A and $\Omega(\mathsf{x})$ be the set $\{\mathsf{y} \in A^{\mathbb{N}}; \mathsf{y}_{[i,j]} \in L(\mathsf{x}), \forall \ [i,j] \subset \mathbb{N}\}$. It is clear that $(\Omega(\mathsf{x}), T)$ is a subshift. We say that $(\Omega(\mathsf{x}), T)$ is the subshift generated by x. When x is a sequence we have $\Omega(\mathsf{x}) = \overline{\{T^n\mathsf{x}; n \in \mathbb{N}\}}$. Let (X,T) be a subshift on A, the following are equivalent:

i) (X,T) is minimal.
ii) For all $\mathsf{x} \in X$ we have $X = \Omega(\mathsf{x})$.
iii) For all $\mathsf{x} \in X$ we have $L(X) = L(\mathsf{x})$.

We also have that $(\Omega(\mathsf{x}), T)$ is minimal if and only if x is uniformly recurrent.

2.3 Return words and derivatives of a sequence

For the rest of the section x is a uniformly recurrent sequence on the alphabet A and (X,T) is the minimal subshift it generates. We recall that all sequences in X are uniformly recurrent. Let u be a non-empty word of $L(X)$.

Definition 6 *A word w on A is a* return word *to u in x if there exist two consecutive occurrences j,k of u in x such that $w = \mathsf{x}_{[j,k)}$.*

The set of return words to u is denoted by $\mathcal{R}_u(\mathsf{x})$. It is immediate to check that a word $w \in A^+$ is a return word if and only if:

$\imath$) $uwv \in L(\mathsf{x})$ (i.e. uwv is a factor of x);
$\imath\imath$) u is a prefix of wu;
$\imath\imath\imath$) the word wu contains only two occurrences of u.

Remarks.

1. As x is uniformly recurrent, the difference between two consecutive occurrences of u in x is bounded, and the set $\mathcal{R}_u(\mathsf{x})$ of return words to u is finite.

2. The statement $\imath\imath$) cannot be simplified: it is not equivalent to *u is a prefix of w*. For example, if aaa is a factor of x then the word a is a return word to aa.

3. From this characterization, it follows that the set of return words to u is the same for all $y \in X$, hence we set $\mathcal{R}_u(X) = \mathcal{R}_u(\mathsf{x})$.

If it is clear from the context, we write $\mathcal{R}_u$ instead of $\mathcal{R}_u(\mathsf{x})$.

Lemma 7 *If u is prefix of some word $v \in L(\mathsf{x})$ then the return words to v are concatenation of return words to u, i.e. $\mathcal{R}_v \subset \mathcal{R}_u^+$.*

Lemma 8 *If x is not periodic. Then*

$$m_n = \text{Inf}\{|v|; v \in \mathcal{R}_{\mathsf{x}[0,n]}\} \to +\infty \ \textit{when} \ n \to +\infty.$$

Proof. We have seen (Proposition 7) that $\mathcal{R}_{\mathsf{x}[0,n+1]}$ is included in $\mathcal{R}^+_{\mathsf{x}[0,n]}$, then $m_n \leq m_{n+1}$. Let suppose $(m_n)_{n\in\mathbb{N}}$ stationary at the rank n_0: there exist an integer k and, for every $n > n_0$, a word v_n where $|v_n| = k$ and v_n is an element of $\mathcal{R}_{\mathsf{x}[0,n]}$. If $n \geq k$ then $\mathsf{x}[0,n]$ is a prefix of $v_n\mathsf{x}[0,n]$. Therefore, for all integers j such that $0 \leq j \leq n-k$, we deduce $\mathsf{x}[j] = \mathsf{x}[k+j]$. It follows that x is periodic with period k. This completes the proof. □

Definition 9 *Let A be an alphabet. A finite subset $\mathcal{C}$ of A^+ is a* code *if every word $u \in A^+$ admits at most one decomposition in a concatenation of elements of $\mathcal{C}$.*

Lemma 10 ([Du1]) *The set $\mathcal{R}_u$ is a code.*

It will be convenient to label the return words to u with respect to x. Put

$$R_u(\mathsf{x}) = \{1, \ldots, \text{Card}(\mathcal{R}_u)\}$$

and let $\Theta_{\mathsf{x},u} : R_u(\mathsf{x}) \to \mathcal{R}_u$ be the bijection defined as follows: let $\mathcal{R}_u$ be ordered according to the rank of first occurrence in x, and $\Theta_{\mathsf{x},u}(k)$ defined to be the k^{th} element of $\mathcal{R}_u$ for this order.
We consider $R_u(\mathsf{x})$ as an alphabet, and $\phi_{\mathsf{x},u.v}$ as a map from $R_u(\mathsf{x})$ to A^+. Clearly, for all $\mathsf{y} \in X$ we have $R_u(\mathsf{y}) = R_u(\mathsf{x})$ but for some $\mathsf{y} \in X$ we have $\Theta_{\mathsf{x},u}$ is different from $\Theta_{\mathsf{y},u}$. Nevertheless when it will not create confusion we will write $R_u = R_u(\mathsf{x})$ and $\Theta_u = \Theta_{\mathsf{x},u}$. The Lemma 10 can be stated as follows:

Corollary 11 *$\Theta_{x,u} : R_u^+ \to A^+$ and $\Theta_{x,u} : R_u^{\mathbb{N}} \to A^{\mathbb{N}}$ are one to one.*

This corollary implies that if y is a concatenation of return word to u then the equation $\Theta_{x,u}(z) = y$ admits a unique solution. This remark is important for the next definition and for some proofs in the next sections.
Let $i \geq 0$ be the smallest integer such that v is a prefix of $y = T^i(x)$. The sequence y is a concatenation of return words, thus we can define:

Definition 12 *Let $i \geq 0$ be the smallest integer such that u is a prefix of $y = T^i(x)$. The u-*derivative *of x is the unique sequence $\mathcal{D}_u(x)$ on the alphabet R_u such that:*

$$\Theta_{x,u}(\mathcal{D}_u(x)) = y \ .$$

We remark that $\Theta_{x,u} = \Theta_{y,u}$ and $\mathcal{D}_u(x) = \mathcal{D}_u(y)$. For example if $x = aabbaabbbbabaaaabbaa...$ then $\mathcal{D}_{aa}(x) = ABCCA....$

Lemma 13 ([Du1]) *Let v be a non-empty prefix of $\mathcal{D}_u(x)$ and $w = \Theta_{x,u}(v)u$. Then w is a prefix of x and $\mathcal{D}_v(\mathcal{D}_u(x)) = \mathcal{D}_w(x)$.*

2.4 Topological interpretation

The notion of derivative sequence is the combinatorial analogue to the notion of induced system, as we explain now.
Let $u \in L(X)$ and $x \in [u]$. We set $\Theta_u = \Theta_{x,u}$. The cylinder sets $[wu]$ for $w \in \mathcal{R}_u$ are obviously pairwise disjoint; they are included in the cylinder set $[u]$ by the property $ii)$ of return words. Let $y \in [u]$, and n be the smallest positive occurrence of u in y; then $w = y_{[0,n)}$ is a return word, and $y \in [wu]$. Thus $\{[wu];\ w \in \mathcal{R}_u\}$ is a partition of $[u]$. Moreover, if $w \in \mathcal{R}_u$ and $y \in [wu]$, the first return time of y to $[u]$ is $|w|$ by the property $iii)$ of return words. It follows that

$$\mathcal{Q} = \{T^j[wu]\ w \in \mathcal{R}_u \text{ and } 0 \leq j < |w|\}$$

is a Kakutani–Rohlin partition of X, with base $[u]$. Using the bijection Θ_u, this partition can also be written:

$$\mathcal{Q} = \{T^j[\Theta_u(k)u];\ k \in R_u \text{ and } 0 \leq j < |\Theta_u(k)|\} \ .$$

Let S be the shift on $R_u^{\mathbb{N}}$ and Y the subshift spanned by $\mathcal{D}_u(x)$.

Proposition 14 *Θ_u is an isomorphism of (Y, S) onto the system induced by (X, T) on the cylinder set $[u]$.*

Proof. We know that $\Theta_u : R_u^{\mathbb{N}} \to A^{\mathbb{N}}$ is one to one. As

$$\Theta_u(\mathcal{D}_u(x)) = x \text{ and, for all } y \in \mathcal{R}_u^{\mathbb{N}},\ \Theta_u(Sy) = T^{|\Theta_u(y_0)|}\,\Theta_u(y)$$

we get $\Theta_u(S^n\mathcal{D}_u(\mathsf{x})) \in X$ for all n, thus $\Theta_u(Y) \subset X$. By definition of Θ_u, we have $\Theta_u(Y) \subset [u]$.
Let $\mathsf{z} \in [u]$. There exists a sequence (n_i) of integers such that $T^{n_i}\mathsf{x} \to \mathsf{z}$; as $[u]$ is open in X, for i large enough $T^{n_i}\mathsf{x} \in [u]$, and n_i is an occurrence of u in x; it follows that $T^{n_i}\mathsf{x} = \Theta_u(S^{k_i}\mathcal{D}_u(\mathsf{x}))$ for some k_i, and $T^{n_i}\mathsf{x} \in \Theta_u(Y)$; finally we get $\mathsf{z} \in \Theta_u(Y)$ and $\Theta_u(Y) = [u]$.
Let $\mathsf{y} \in Y$ and $\mathsf{z} = \Theta_u(\mathsf{y})$; the first return time of z to $[u]$ is $n = |\Theta_u(\mathsf{y}_0)|$; thus the image of z by the first return time transformation is $T^n\mathsf{z} = \Theta_u(S\mathsf{y})$, and the lemma is proved. □

2.5 Perron's Theorem

The Perron's Theorem will be central in the following sections. Let A be a square matrix, we set $\rho(A) = \max\{|\lambda|;\ \lambda \text{ eigenvalue of } A\}$. We say a matrix is primitive whenever there exists $n \in \mathbb{N}$ such that A^n has strictly positive entries.

Theorem 15 (Perron's Theorem) *Let A be a primitive $n \times n$ matrix. Then*

1) $\rho(A) > 0$;
2) $\rho(A)$ is an eigenvalue of A;
3) There is $x \in \mathbb{R}^n$ with strictly positive coefficients such that $Ax = \rho(A)x$;
4) If $y \in \mathbb{R}^n$ is a strictly positive eigenvector then it is a multiple of x;
5) $\rho(A)$ is an algebraically (and hence geometrically) simple eigenvalue of A;
6) $|\lambda| < \rho(A)$ for every eigenvalue $\lambda \neq \rho(A)$;
7) $[\rho(A)^{-1}A]^m \to L$ as $m \to +\infty$, where $L \equiv xy^T$, $Ax = \rho(A)x$, $A^Ty = \rho(A)y$, $x > 0$, $y > 0$, and $x^Ty = 1$.

For non-negative matrices, Perron's Theorem does not remain valid but we have the following.

Theorem 16 *Let A be a non-negative $n \times n$ matrix. Then $\rho(A)$ is a eigenvalue of A and there exists a vector x with non-negative entries, $x \neq 0$, such that $Ax = \rho(A)x$.*

For more details see [HJ] or [LM]. In both case we call $\rho(A)$ the *dominant eigenvalue of A*. A real number $\lambda \geq 1$ is a *Perron number* if it is an algebraic integer that strictly dominates all its other algebraic conjugates. It is clear that the dominant eigenvalue of a primitive matrix is a Perron number, and in fact from Paragraph 4.4 in [LM] the dominant eigenvalue of a non-negative matrix is also a Perron number. Lind proved the converse is true (see [LM] for details).

Theorem 17 *Let λ be a real number. Then, λ is a Perron number if and only if there is a primitive integral matrix A such that $\rho(A) = \lambda$.*

3 A characterization of substitutive sequences using return words

For more informations about substitutions we refer the reader to [Qu], [DHS] and [Ho]. This section is devoted to the proof of Theorem 3. Moreover at the end we make some remarks about induced systems.

3.1 Substitutive sequences

Definition 18 *A* substitution *on the alphabet A is a morphism $\sigma : A \to A^*$ satisfying :*

1. *There exists $a \in A$ such that a is the first letter of $\sigma(a)$;*
2. *For all $b \in A$, $\lim_{n\to+\infty} |\sigma^n(b)| = +\infty$.*

In some papers (see [Pan] for example) the condition 2) is not required to be a substitution and our definition corresponds to what Pansiot call *growing substitutions.*
It is classical that $(\sigma^n(aa\cdots); n \in \mathbb{N})$ converges in $A^{\mathbb{N}}$ to a sequence x. The substitution σ being continuous on $A^{\mathbb{N}}$ this sequence is a *fixed point* of σ, i.e. $\sigma(\mathsf{x}) = \mathsf{x}$.
Whenever the matrix associated to τ is primitive we say that τ is a *primitive substitution.* It is equivalent to the fact that there exists n such that for all a and b in A, a has an occurrence in $\sigma^n(b)$. It is a *substitution of constant length* p if for all $a \in A$ the length of $\sigma(a)$ is p.
Let B be an other alphabet, we say that a morphism ϕ from A to B^* is a *letter to letter morphism* when $\phi(A)$ is a subset of B. Then the sequence $\phi(\mathsf{x})$ is called *substitutive*, and *primitive substitutive* if τ is primitive. The matrix of τ is non-negative and consequently (Theorem 16) has a dominant eigenvalue α. We will say that it is the dominant eigenvalue of τ and that $\phi(\mathsf{x})$ is α*-substitutive.*

Proposition 19 *Let A and B be two alphabets. Let $\mathsf{x} \in A^{\mathbb{N}}$ be an α-substitutive sequence and ϕ be a morphism from A to B^+. Then there exists $n \in \mathbb{N}$ such that the sequence $\phi(\mathsf{x})$ is α^n-substitutive. Moreover, if x is α-substitutive primitive then there exists $n \in \mathbb{N}$ such that the sequence $\phi(\mathsf{x})$ is α^n-substitutive primitive.*

Proof. There exist a substitution ζ with fixed point y and a letter to letter morphism ρ from C to A^* such that $\mathsf{x} = \rho(\mathsf{y})$. Setting $\phi = \varphi \circ \rho$, we have $\varphi(\mathsf{x}) = \phi(\mathsf{y})$.
Let $D = \{(c,k); c \in C \text{ and } 0 \leq k \leq |\phi(c)| - 1\}$ and $\psi : C \to D^*$ the morphism defined by:

$$\psi(c) = (c,0)\dots(c,|\phi(c)|-1).$$

There an integer n such that $|\zeta^n(c)| \geq |\phi(c)|$ for all c in C.
Let τ be the morphism from D to D^* defined by:

$$\begin{array}{rrcll} & \tau((c,k)) & = & \psi(\zeta^n(c)_{[k,k]}) & \text{if } 0 \leq k < |\phi(c)|-1, \\ \text{and} & \tau((c,|\phi(c)|-1)) & = & \psi(\zeta^n(c)_{[|\phi(c)|-1,|\zeta^n(c)|-1]}) & \text{otherwise.} \end{array}$$

For all c in C we have

$$\begin{aligned} \tau(\psi(c)) = \ & \tau((c,0)\cdots(c,|\phi(c)|-1)) = \psi(\zeta^n(c)_{[0,0]})\cdots\psi(\zeta^n(c)_{[|\zeta^n(c)|-1,|\zeta^n(c)|-1]}) \\ & = \psi(\zeta^n(c)), \end{aligned}$$

hence $\tau(\psi(y)) = \psi(\zeta^n(y)) = \psi(y)$. In this way $\psi(y)$ is the fixed point of the substitution τ which begins with $(e,0)$ and $\tau\psi = \psi\zeta^n$. From this last equality we observe that α^n is the dominant eigenvalue of τ.

Let χ be the letter to letter morphism from D to B defined by $\chi((c,k)) = \phi(c)_{[k,k]}$ for all (c,k) in D. For all c in C we obtain

$$\chi(\psi(c)) = \chi((c,0)\cdots(c,|\phi(c)|-1)) = \phi(c),$$

and then $\chi(\psi(y)) = \phi(y)$. Consequently $\varphi(\mathsf{x})$ is α^n-substitutive.
The relation $\tau\psi = \psi\zeta^n$ implies that for all $k \in \mathbb{N}$ we have $\tau^k\psi = \psi\zeta^{kn}$. Hence if ζ is primitive then τ is primitive too, and, if x is α-substitutive primitive then $\varphi(\mathsf{x})$ is α^n-substitutive primitive. □

Example. Let $A = \{a,b,c\}$ and $\sigma : A \to A^*$ the substitution defined by $\sigma(a) = aba$, $\sigma(b) = cb$ and $\sigma(c) = bcc$. Its dominant eigenvalue is $(1+\sqrt{5})/2$ and σ is not primitive. Hence its fixed point $\mathsf{x} = abacbababcc...$ is a $\left((1+\sqrt{5})/2\right)$-substitutive sequence.
Let $\phi : A \to \{a,b\}^*$ be the morphism defined by $\phi(a) = ab$, $\phi(b) = a$ and $\phi(c) = bbb$. Using the construction in the proof of Proposition 19 we will obtain a substitution $\tau : C \to C^*$, with fixed point z, and a letter to letter morphism $\varphi : C \to \{a,b\}^*$ such that $\varphi(\mathsf{z}) = \phi(\mathsf{x}) = \mathsf{y}$.
Let $C = \{(a,1),(a,2),(b,1),(c,1),(c,2),(c,3)\}$. The morphisms ψ, τ and φ are defined by :

$$\psi(a) = (a,1)(a,2), \ \ \psi(b) = (b,1), \ \ \psi(c) = (c,1)(c,2)(c,3),$$

$$\begin{array}{lllllll}
\tau((a,1)) & = & \psi(\sigma(a)_{[1]}) & = & \psi(a) & = & (a,1)(a,2), \\
\tau((a,2)) & = & \psi(\sigma(a)_{[2]}) & = & \psi(ba) & = & (b,1)(a,1)(a,2), \\
\tau((b,1)) & = & \psi(\sigma(b)_{[1]}) & = & \psi(cb) & = & (c,1)(c,2)(c,3)(b,1), \\
\tau((c,1)) & = & \psi(\sigma(c)_{[1]}) & = & \psi(b) & = & (b,1)(b,2)(b,3), \\
\tau((c,2)) & = & \psi(\sigma(c)_{[2]}) & = & \psi(b) & = & (b,1)(b,2)(b,3), \\
\tau((c,3)) & = & \psi(\sigma(c)_{[3]}) & = & \psi(b) & = & (b,1)(b,2)(b,3),
\end{array}$$

$$\begin{array}{lll}
\varphi((a,1)) = a, & \varphi((a,2)) = b, & \varphi((b,1)) = a, \\
\varphi((c,1)) = b, & \varphi((c,2)) = b, & \varphi((c,3)) = b.
\end{array}$$

$$\begin{array}{ll}
\mathsf{z} & = (a,1)(a,2)(b,1)(a,1)(a,2)(c,1)(c,2)(c,3)(b,1)(a,1)(a,2)(b,1)(a,1)(a,2)(b,1)\cdots. \\
\downarrow \varphi & \\
\mathsf{y} & = abaabbbbaabaaba\cdots \\
\uparrow \phi & \\
\mathsf{x} & = abacbabab\cdots.
\end{array}$$

Now we give some useful results about the lengths of the iterates of a substitution.

Lemma 20 *Let $\sigma : A \to A^*$ be a substitution. There exists a unique partition $A_1, \cdots, A_l$ of A such that for all $1 \le i \le l$ and all $a \in A_i$*

$$\lim_{n\to+\infty} \frac{|\sigma^n(a)|}{n^{d(a)}\theta(a)^n} = c(a) > 0$$

where $\theta(a)$ is the dominant eigenvalue of M restricted to A_i, $d(a)$ its Jordan order.

Proof. See Theorem II.10.2 in [SS]. □

We remark that $\Theta = \max\{\theta(a); a \in A\}$ is the dominant eigenvalue of σ. With respect to this lemma in [Pan] the author divides the family of (growing) substitutions into three sub-families The family of *quasi-uniform substitutions* : for all $a, b \in A$, $d(a) = 0$ and $\theta(a) = \theta(b)$. The family of *polynomially divergent substitutions* : for all $a, b \in A$, $\theta(a) = \theta(b)$ and $d(a) > 0$ for some $a \in A$. The family of *exponentially divergent substitutions* : there exist $a, b \in A$ such that $\theta(a) \neq \theta(b)$.
If x is a fixed point of a substitution belonging to one of these sub-families then Pansiot proved that its symbolic complexity $n \mapsto p_{\mathsf{x}}(n)$ is growing like, respectively, n, $n \log n$ or $n \log \log n$.
The substitutions we will consider in this paper are quasi-uniform substitutions. For this reason and for sake of simplicity the two following lemmas are stated only for quasi-uniform substitutions. It is clear that substitutions of constant length are quasi-uniform. It is not difficult to check that primitive substitutions are also quasi-uniform substitutions. It mainly comes from Perron's Theorem (see [Qu]).

Let $\sigma : A \to A^*$ be a quasi-uniform substitution with dominant eigenvalue Θ and fixed point x, we define

$$\begin{array}{rccl} \lambda_\sigma & : & A^* & \to \mathbb{R} \\ & & u_0 \cdots u_{n-1} & \mapsto \sum_{i=0}^{n-1} c(u_i). \end{array}$$

From Lemma 20 we deduce the following lemma.

Lemma 21 *For all $u \in A^*$ we have $\lim_{n\to+\infty} |\sigma^n(u)|/\Theta^n = \lambda_\sigma(u)$.*

We will need the following Lemma in the proof of Theorem 1.

Lemma 22 *Let $a \in A$ which has infinitely many occurrences in x. There exist a positive integer p and words $u, v, w \in A^*$ such that for all $n \in \mathbb{N}$ the word*

$$\sigma^{pn}(u)\sigma^{p(n-1)}(v)\sigma^{p(n-2)}(v)\cdots\sigma^{p}(v)vwa$$

is a prefix of x. Moreover we have

$$\lim_{n\to+\infty} \frac{|\sigma^{pn}(u)\sigma^{p(n-1)}(v)\sigma^{p(n-2)}(v)\cdots\sigma^{p}(v)vwa|}{\lambda_\sigma(u)\Theta^{pn} + \lambda_\sigma(v)\sum_{k=0}^{n-1}\Theta^{pk}} = 1.$$

Proof. Let $a \in A$ which has infinitely many occurrences in x. We set $a_0 = a$. There exists $a_1 \in A$ which has infinitely many occurrences in x and such that a_0 has an occurrence in $\sigma(a_1)$. In this way we can construct a sequence $(a_i; i \in \mathbb{N})$ of letters, where each one appears infinitely many times in x, such that $a_0 = a$ and a_i occurs in $\sigma(a_{i+1})$, for all $i \in \mathbb{N}$.
There exist $i < j$ such that $a_i = a_j = b$. It comes that a occurs in $\sigma^i(b)$ and b occurs in $\sigma^{j-i}(b)$. Hence there exist $u_1, u_2, v_1, v_2 \in A^*$ such that $\sigma^i(b) = u_1 a u_2$ and $\sigma^{j-i}(b) = v_1 b v_2$. We set $p = j - i$, $v = \sigma^i(v_1)$ and $w = u_1$. There exists u' such that $u'b$ is a prefix of x. We

remark that for all $n \in \mathbb{N}$ the word $\sigma^n(u'b)$ is a prefix of x too. We set $u = \sigma^i(u')$. We have $\sigma^p(u'b) = \sigma^p(u')v_1bv_2$. Consequently for all $n \in \mathbb{N}$

$$\sigma^{pn}(u')\sigma^{p(n-1)}(v_1)\sigma^{p(n-2)}(v_1)\cdots\sigma^p(v_1)v_1b$$

is a prefix of $\sigma^{np}(u'b)$. Then

$$\sigma^{pn}(u)\sigma^{p(n-1)}(v)\sigma^{p(n-2)}(v)\cdots\sigma^p(v)vwa$$

is a prefix of $\sigma^{np+i}(u'b)$ and consequently of x, for all $n \in \mathbb{N}$. The last part of the lemma follows from Lemma 21. □

In the rest if the section we only consider *primitive* substitutions, i.e. substitutions which matrices are primitive. In this case all the fixed points of σ are uniformly recurrent and generate the same minimal subshift, we call it the *substitution subshift generated by* σ. (For more details see [Qu] and [DHS].)

3.2 Some properties of the return words of a substitutive sequence

We say that a sequence x on a finite alphabet is *linearly recurrent (with constant $K \in \mathbb{N}$)* if it is recurrent and if, for every word u of x and all $w \in \mathcal{R}_u$ it holds

$$|w| \leq K|u|.$$

Proposition 23 *All primitive substitutive sequences are linearly recurrent.*

Proof. It suffices to prove it for fixed points of primitive substitutions. Let τ be a primitive substitution on A. Let u be a word of $L(\tau)$ and v be a return word to u. Proposition 21 implies that there exists a constant C such that for all positive integers k

$$S_k = \text{Sup}\{|\tau^k(a)|; a \in A\} \leq C\text{Inf}\{|\tau^k(a)|; a \in A\} = CI_k \ .$$

Let k be the smallest integer such that $I_k \geq |u|$. The choice of k entails that there exists a word $ab \in L(\tau)$ of length 2 such that u occurs in $\tau^k(ab)$. Let R be the largest difference between two successive occurrences of a word of length 2 of $L(\tau)$. We have

$$|v| \leq RS_k \leq RCI_k \leq RCS_1I_{k-1} \leq RCS_1|u| \ .$$

The subshift spanned by τ is linearly recurrent with constant RCS_1. □

Proposition 24 *Let x be an aperiodic linearly recurrent sequence with constant K. Then:*

1. *The number of distinct factors of length n of x is less or equal to Kn.*
2. *x is $(K+1)$-power free (i.e. $u^{K+1} \in L(\mathsf{x})$ if and only if $u = \emptyset$).*
3. *For all $u \in L(\mathsf{x})$ and for all $w \in \mathcal{R}_u$ we have $\frac{1}{K}|u| < |w|$.*
4. *For all $u \in L(\mathsf{x})$, $\text{Card}(\mathcal{R}_u) \leq K(K+1)^2$.*

Proof. We begin with a remark. Let n be a positive integer and $u \in L(\mathsf{x})$ a word of length $(K+1)n-1$. Let $v \in L(\mathsf{x})$ be a word of length n. The difference between two successive occurrences of v is less than Kn, consequently u has at least one occurrence of v. We have proved that: For each n, all words of length n occurs in each word of length $(K+1)n-1$. From this remark we deduce $\imath$).
Let $u \in L(\mathsf{x})$ be a word such that $u^{K+1} \in L(\mathsf{x})$. Each factor of x of length $|u|$ occurs in u^{K+1}. But in u^{K+1} occurs at the most $|u|$ distinct factors of length $|u|$ of x. This contradicts the aperiodicity of x.
Assume there exist $u \in L(\mathsf{x})$ and $w \in \mathcal{R}_u$ such that $|u|/K \geq |w|$. The word w is a return word to u therefore u is a prefix of wu. We deduce that w^K is a prefix of u. Hence w^{K+1} belongs to $L(\mathsf{x})$ because wu belongs to $L(\mathsf{x})$. Consequently $w = \emptyset$ and $\imath\imath\imath$) is proved.
Let u be a factor of x and $v \in L(\mathsf{x})$ be a word of length $(K+1)^2|u|$. Each word of length $(K+1)|u|$ occurs in v, hence each return word to u occurs in v. It follows from $\imath\imath\imath$) that in v will occur at the most $K(K+1)^2|u|/|u| = K(K+1)^2$ return words to u, which proves $\imath\nu$). □

3.3 Proof of Theorem 3

In the rest of the section all considered substitutions and substitutive sequences will be primitive hence we will sometimes forget to mention it. The propositions proved below will be also useful in the next section.
Let x be a uniformly recurrent sequence.

2) implies 1)

In fact it suffices to prove that 3) implies 1).
The number of u-derivative sequence of x, u being a prefix of x, is finite. Hence there exists a sequence of prefixes $(u_n; n \in \mathbb{N})$ of x such that $|u_n| < |u_{n+1}|$ and $\mathcal{D}_{u_n}(\mathsf{x}) = \mathcal{D}_{u_{n+1}}(\mathsf{x})$ for all $n \in \mathbb{N}$). Clearly this implies, for all $n \in \mathbb{N}$, that u_n is a prefix of u_{+1}. Take $u = u_1$.
The sequence x being uniformly recurrent we can choose n so large that every factor of length n of x has an occurrence of each return word to u. By Lemma 8 there exists a word some $v = u_l$ such that $|w| > n$ for all $w \in \mathcal{R}_v$. We have $R_u = R_v$, we set $R = R_u$. The set $\mathcal{R}_v$ is included in $\mathcal{R}_u^+$ (Lemma 7).
The map Θ_u is one to one (Corollary 11). Consequently we can define the morphism $\tau : R \to R$ with $\Theta_u \tau = \Theta_v$. From the choice of u and v, for every i, j in R_u the word $\Theta_u(j)u$ appears in $\Theta_v(i)$. This means that τ is a primitive substitution (for every i, j in R_u the letter j appears in the word $\tau(i)$). We have

$$\Theta_u \tau(\mathcal{D}_u(\mathsf{x})) = \Theta_v \tau(\mathcal{D}_v(\mathsf{x})) = \mathsf{x}.$$

The definition of $\mathcal{D}_u(\mathsf{x})$ gives that $\tau(\mathcal{D}_u(\mathsf{x})) = \mathcal{D}_u(\mathsf{x})$. Hence $\mathcal{D}_u(\mathsf{x})$ is a fixed point of τ. As $\mathsf{x} = \Theta_u(\mathcal{D}_u(\mathsf{x}))$, Proposition 19 achieves the proof.

1) implies 3)

Let $\sigma : A \to A$ be a primitive substitution with fixed point y and dominant eigenvalue α, and ϕ be a letter to letter morphism from A to B such that $\mathsf{x} = \phi(\mathsf{y})$. We recall that x is α-substitutive.

We easily check that if x is periodic then 1) implies 3). Hence we suppose it is not. We begin with a proposition.

Proposition 25 *Let u be a non-empty prefix of y. The derived sequence $\mathcal{D}_u(\mathsf{y})$ is the fixed point of a primitive substitution $\sigma_u : R_u \to R_u$ which satisfies*

$$\Theta_u \circ \sigma_u = \sigma \circ \Theta_u .$$

Moreover $\mathcal{D}_u(\mathsf{y})$ is α-substitutive.

Proof. Let u be a non-empty prefix of y and $i \in R_u$. The word u is a prefix of $\sigma(u)$ and $\Theta_u(i)$, hence $\sigma(\Theta_u(i))$ belongs to $\mathcal{R}_u^+$ and we can define the morphism $\sigma_u : R_u \to R_u^+$ with

$$\Theta_u \circ \sigma_u = \sigma \circ \Theta_u .$$

For all $n \in \mathbb{N}$ we have $\Theta_u \circ \sigma_u^n = \sigma^n \circ \Theta_u$ hence we easily check that σ_u is primitive and that α is the dominant eigenvalue of σ_u. Moreover

$$\Theta_u \circ \sigma_u(\mathcal{D}_u(\mathsf{y})) = \sigma \circ \Theta_u(\mathcal{D}_u(\mathsf{y})) = \sigma(\mathsf{y}) = \mathsf{y}.$$

From Corollary 11 it comes that $\sigma_u(\mathcal{D}_u(\mathsf{y})) = \mathcal{D}_u(\mathsf{y})$ and $\mathcal{D}_u(\mathsf{y})$ is α-substitutive. □

The substitution $\sigma_u : R_u \to R_u^*$ will be called *return substitution.*

Example. If $\sigma : A \to A^*$ is the substitution defined by $\sigma(a) = aba$ and $\sigma(b) = aa$ then the return substitution σ_a is defined on the alphabet $\{1, 2\}$ by $\sigma_a(1) = 1222$ and $\sigma_a(2) = 12$.

Proposition 26 *Let u be a prefix of x. Then, The sequence $\mathcal{D}_u(\mathsf{x})$ is α^k-substitutive primitive for some integer k. Moreover, there exists a morphism λ_u and a prefix v of y such that $\Theta_u \lambda_u = \phi \Theta_v$ and $\lambda_u(\mathcal{D}_v(\mathsf{y})) = \mathcal{D}_u(\mathsf{x})$.*

Proof. Let v be the unique prefix of y such that $\phi(v) = u$. If w is a return word to v then $\phi(w)$ is a concatenation of return words to u (Lemma 7). The morphism Θ_u being one to one (Proposition 11) we define the morphism $\lambda : R_v \to R_u^*$ with $\Theta_u \lambda = \phi \Theta_v$. This morphism satisfies $\lambda(\mathcal{D}_v(\mathsf{y})) = \mathcal{D}_u(\mathsf{x})$. Proposition 19 and Proposition 25 achieve the proof. □

Consequently to prove that 1) implies 3) it suffices to prove that the sets $\{\sigma_v; v \text{ prefix of } \mathsf{y}\}$ and $\{\lambda_u; u \text{ prefix of } \mathsf{x}\}$ are finite.

Proposition 27 *The sets $\{\sigma_v; v$ prefix of $\mathsf{y}\}$ and $\{\lambda_u; u$ prefix of $\mathsf{x}\}$ are finite.*

Proof. The periodic case is easy to check hence we suppose that y is non-periodic. We begin to prove that $\{\sigma_v : R_v(\mathsf{y}) \to R_v^*(\mathsf{y}); v \text{ prefix of } \mathsf{y}\}$ is finite. To do this it suffices to prove that $|R_v(\mathsf{y})|$ and $|\sigma_v(i)|$ are bounded independently of v and $i \in R_v(\mathsf{y})$.
Let v be a non-empty prefix of y, i be an element of $R_v(\mathsf{y})$ and $w = \Theta_v(i)$ be a return word to v.
The sequence y is linearly recurrent (with constant K). Thus, we have $|w| \leq K|v|$ and $|\sigma(v)| \leq S(\sigma)K|v|$, where $S(\sigma) = \text{Sup}\{|\sigma(a)|; a \in A\}$. The length of each element of $\mathcal{R}_v(\mathsf{y})$ is larger than $|v|/K$ (Proposition 24). Thus we can decompose $\sigma(w)$ in at the most

$S(\sigma)K^2$ elements of $R_v(\mathsf{y})$, so $|\sigma_v(i)| \leq S(\sigma)K^2$. Moreover we know from Proposition 24 that $|R_v(\mathsf{y})| \leq K(K+1)^2$. It ends the first part of the proof.
Let $u \in L(\mathsf{x})$ and v be the unique prefix of y such that $\phi(v) = u$. We have

$$|\lambda_u(i)|(1/K)|u| \leq |\lambda_u(i)| \mathrm{Inf}\{|w|; w \in \mathcal{R}_u(\mathsf{x})\} \leq |\Theta_{\mathsf{x},u}(\lambda_u(i))|$$

$$= |\Theta_{\mathsf{y},v}(i)| \leq \mathrm{Sup}\{|w|; w \in \mathcal{R}_v(\mathsf{y})\} \leq K|v| = M|u|.$$

Hence $|\lambda_u(i)| \leq K^2$. This completes the proof. □

3) implies 2)

We start with some notation. Let t be a word which prefix is s. By $s^{-1}t$ we mean the word r such that $t = sr$. In this way we have $ss^{-1}t = t$.
Let K be the constant given by Proposition 23.
Let u be a word of x and v such that vu is a prefix of x and u has exactly one occurrence in vu. From Proposition 24 it comes that $|vu| \leq (K+1)|u|$.
If w is a return word to vu then u is a prefix of $v^{-1}wv$ and $v^{-1}wvu$ is a word of x, hence $v^{-1}wv$ is a concatenation of return words to u. Thus, we can define $\phi_{v,u} : R_{vu} \to R_u$ with

$$\Theta_u \phi_{v,u}(i) = v^{-1}\Theta_{vu}(i)v,$$

for all $i \in R_{vu}$. We remark that $\phi_{v,u}(\mathcal{D}_{vu}(\mathsf{x})) = \mathcal{D}_u(\mathsf{x})$. The set $\{\mathcal{D}_u(\mathsf{x}); u = \mathsf{x}_{[0,n]}, n \geq 0\}$ being finite, it suffices to prove that the following set is finite to conclude

$$H = \{\phi_{v,u} : R_{vu} \to R_u; uv = \mathsf{x}_{[0,n]}, |vu| \leq (K+1)|u|, n \geq 0\}.$$

For all $i \in R_{uv}$ we have

$$|\phi_{v,u}(i)| \leq \frac{|\Theta_{vu}(i)|}{|u|/K} \leq K^2(K+1).$$

Moreover $|R_s| \leq K(K+1)^2$ for all words $s \in L(\mathsf{x})$ hence H is finite. □

3.4 Induced systems of substitutions subshifts

We obtain as a direct consequence of Theorem 3 and Proposition 14 the following theorem.

Theorem 28 *Let (X, T) be a minimal subshift. Then, it is a substitution subshift if and only if the set of its induced systems on cylinders is finite.*

This result is due to Holton and Zamboni [HZ1] but they obtain it in a different way. They remarked that if a minimal subshift is such that the set of its induced systems on cylinders is finite then it is periodic.
To study induced systems of minimal dynamical systems (X, S), where X is a Cantor set (i.e. X has a countable basis of clopen sets and no isolated point) it is often relevant to use Bratteli diagrams. For example, in the case of substitution subshift it gives a very short proof of the following result which puts in evidence the "self-similar" behavior of substitution subshifts.

Proposition 29 *Let (X,T) be a minimal subshift. If (Y,S) is some induced systems on a clopen set of (X,T) then there exists a clopen set $U \subset Y$ such that (X,T) is isomorphic to (U,S_U).*

Of course a dynamical system with positive entropy can not have this property. In fact, this property characterized minimal subshifts that are substitution subshifts. But we do not know what to say about a zero entropy dynamical system having this property. Is it a factor of a substitution subshift ? Is it measure theoretically conjugate to a substitution subshift ? Can it have nothing to do with substitutions ?

4 Cobham Theorems for primitive substitutions

In this section all substitutions and substitutive sequences will be primitive hence we will sometimes forget to mention it. We prove the following theorems, the first one is the partial answer to the conjecture announced in the introduction and from the second one we will deduce Theorem 2.

Theorem 30 *Let x be a sequence and α and β be two multiplicatively independent Perron numbers. Then, x is both α-substitutive primitive and β-substitutive primitive if and only x is periodic.*

Theorem 31 *Let x and y be two sequences and α and β be two multiplicatively independent Perron numbers. Then, x and y are respectively α-substitutive primitive and β-substitutive primitive sequences such that $L(\mathsf{x}) = L(\mathsf{y})$ if and only if x is periodic.*

4.1 A Cobham Theorem for sequences

Here we prove Theorem 30. First we establish some morphism relations between the substitutions and their return substitutions, and we find their common eigenvalues. Then we obtain some technical results about return substitutions and derivative sequences. We end with the proof.

Eigenvalues and return words. In this section τ is a primitive substitution, x one of its prefixes and u and v two prefixes of x such that $|u| < |v|$. We recall that we have

$$\Theta_u \tau_u = \tau \Theta_u \text{ and } \Theta_v \tau_v = \tau \Theta_v. \tag{1}$$

It comes that τ, τ_u and τ_v have the same dominant eigenvalue.
The word u is a prefix of v, hence a return word to v is a concatenation of return words to u. The morphism Θ_u being one to one, this allows us to define the morphism λ, from R_v to R_u^+, by $\Theta_u \lambda = \Theta_v$. Thus we obtain the relation

$$\tau_u \lambda = \lambda \tau_v.$$

Let k be an integer such that $|v| < |\tau^k(u)|$. The image by τ^k of a return word to u is a concatenation of return words to v. We define a new morphism κ, from R_u to R_v^+, by $\Theta_v\kappa = \tau^k\Theta_u$. We deduce the following morphism relations:

$$\begin{array}{rcl} \tau_v\kappa & = & \kappa\tau_u, \\ \kappa\lambda & = & \tau_v^k \quad \text{and} \\ \lambda\kappa & = & \tau_u^k. \end{array}$$

Consequently we have the following proposition:

Proposition 32 *Then there exist an integer $k \geq 1$ and two morphisms $\lambda : R_v \to R_u^+$ and $\kappa : R_u \to R_v^+$ such that*

$$\tau_v\kappa = \kappa\tau_u, \ \ \tau_u\lambda = \lambda\tau_v, \ \ \kappa\lambda = \tau_v^k \text{ and } \lambda\kappa = \tau_u^k.$$

Corollary 33 *All the return substitutions of a primitive substitution have all the same non-zero eigenvalues.*

Proof. This is a straightforward consequence of Proposition 32 and of the Perron's Theorem. The details are left to the reader. □

Working a little bit more we can obtain the following result (see [Du2]).

Proposition 34 *Let τ be a primitive substitution, x one of its fixed point and u a prefix of x. The substitutions τ and τ_u have the same eigenvalues, except perhaps 0 and roots of the unity.*

It is easy to check that if $\tau : \{0,1\} \to \{0,1\}^+$ is the Fibonacci substitution, i.e. $\tau(0) = 01$ and $\tau(1) = 0$, then we have $\tau = \tau_{01}$. Hence τ and τ_{01} have the same eigenvalues. On the other hand the set of eigenvalues of the Morse substitution, $\sigma(0) = 01$ and $\sigma(1) = 10$, is $\{0,2\}$ and the eigenvalues of σ_{011} are $0, 0, -1$ and 2.

Some technical results. For convenience in the sequel we will use alphabets $\{1, 2, \cdots, k\}$.

Proposition 35 *Let $\tau : A \to A^*$ be a primitive substitution, x be one of its fixed points and u be a prefix of x such that:*

1. *For all letters b of A, $\tau(b)$ begins by 1,*
2. *The substitutions τ and τ_u are defined on the same alphabet and are identical,*
3. *The fixed point of τ is non-periodic,*
4. *For all letters b and c of A, b has at least one occurrence in $\Theta_u(c)$.*

Let J be an infinite set of positive integers. Then there exist an infinite subset I of J, a strictly increasing sequence of positive integers $(l_p)_{p\in I}$ and a morphism $\gamma : A \to A^+$ such that for all p in I.

$$\Theta_u^{l_p}\gamma = \gamma\Theta_u^{l_p} = \tau^p.$$

Proof. Hypothesis 2 says that $A = R_u$. It is easy to check that the morphism $\Theta_u : A \to A^*$ defines a substitution. We put $\Theta = \Theta_u$. Hypothesis 4 implies that this substitution is primitive.
As the substitutions τ_u and τ are identical (hypothesis 2), they have the same fixed point x. We have seen that the fixed point of τ_u is $\mathcal{D}_u(\mathsf{x})$ (Proposition 25), hence $\mathcal{D}_u(\mathsf{x}) = \mathsf{x}$. Consequently, we have $\mathsf{x} = \Theta(\mathsf{x})$, i.e. x is a fixed point of Θ. Moreover we can remark that $\tau\Theta = \Theta\tau$.
The word u is a prefix of $\mathcal{D}_u(\mathsf{x})$, hence we can consider the sequences $(\mathcal{D}_u^n(\mathsf{x}))_{n\geq 1}$ defined by

$$\mathcal{D}_u^1(\mathsf{x}) = \mathcal{D}_u(\mathsf{x}) \text{ and } \mathcal{D}_u^{n+1}(\mathsf{x}) = \mathcal{D}_u(\mathcal{D}_u^n(\mathsf{x})) \text{ for all } n \geq 1.$$

Let us prove by induction that for all $n \geq 1$ we have:

i) $\mathcal{D}_u^n(\mathsf{x}) = \mathcal{D}_{w_n}(\mathsf{x}) = \mathsf{x}$, with $w_n = \Theta^{n-1}(u)\cdots\Theta(u)u$,

ii) $\Theta^n = \Theta_{w_n}$ and

iii) $\tau = \tau_{w_n}$.

For $n = 1$ it suffices to remark that $w_1 = u$.
Now suppose that points *i*), *ii*) and *iii*) are satisfied for some positive integer n. We have

$$\mathcal{D}_u^{n+1}(\mathsf{x}) = \mathcal{D}_u(\mathcal{D}_u^n(\mathsf{x})) = \mathcal{D}_u(\mathsf{x}) = \mathsf{x}$$

and Proposition 13 implies that:

- $\mathcal{D}_u^{n+1}(\mathsf{x}) = \mathcal{D}_u(\mathcal{D}_{w_n}(\mathsf{x})) = \mathcal{D}_w(\mathsf{x})$ and
- $\Theta_w = \Theta_{w_n}\Theta_{\mathcal{D}_{w_n}(\mathsf{x}),u} = \Theta^n\Theta_{\mathsf{x},u} = \Theta^{n+1}$

where

$$w = \Theta_{w_n}(u)w_n = \Theta^n(u)\Theta^{n-1}(u)\cdots\Theta(u)u = w_{n+1}.$$

Hence points *i*) and *ii*) are satisfied for $n + 1$.
The substitution $\tau_{w_{n+1}}$ is the return substitution to u of τ_{w_n} consequently $\Theta\tau_{w_{n+1}} = \tau_{w_n}\Theta$; that is to say $\Theta\tau_{w_{n+1}} = \tau\Theta$. But $\Theta\tau_u = \tau\Theta$ and the map Θ is one to one hence $\tau_{w_{n+1}} = \tau_u = \tau$. This completes the proof by induction of points *i*), *ii*) and *iii*).
We denote the dominant eigenvalues of M_τ and M_Θ respectively by α and β. From Lemma 21 there exists a positive number r such that for all b in A and all k in $\mathbb{N}$

$$\frac{1}{r}\alpha^k \leq |\tau^k(b)| \leq r\alpha^k \text{ and } \frac{1}{r}\beta^k \leq |\Theta^k(b)| \leq r\beta^k.$$

From this we deduce that there exists two constants c_1 and c_2 such that for all positive integers n

$$c_1\beta^n \leq |w_n| = |\Theta^{n-1}(u)\cdots\Theta(u)u| \leq c_2\beta^n.$$

From hypothesis 1 it follows that there exists an integer k_0 such that u is a prefix of all images of letters by τ^{k_0}. For every integer k, larger than k_0, we define l_k to be the greatest integer n such that w_n is a prefix of $\tau^{k-1}(1)$. For all positive integers we have

$$c_1\beta^{l_k} \leq |w_{l_k}| \leq |\tau^{k-1}(1)| \leq |w_{l_k+1}| \leq c_2\beta^{l_k+1}.$$

Thus we obtain

$$|\tau^{k-1}(1)| \leq \frac{\beta c_2}{c_1} c_1 \beta^{l_k} \leq \frac{\beta c_2}{c_1} |w_{l_k}|.$$

Let k be an integer larger than k_0. For all letters b of A the word w_{l_k} is a prefix of $\tau^k(b)$. Hence all images by τ^k of words are concatenations of return words to w_{l_k}. This remark allows us to define the morphism γ_k, from A to A^+, by $\Theta_{w_{l_k}} \gamma_k = \tau^k$. We have:

$$\Theta_{w_{l_k}} \gamma_k \Theta_{w_{l_k}} = \tau^k \Theta^{l_k} = \Theta^{l_k} \tau^k.$$

The map $\Theta_{w_{l_k}} = \Theta^{l_k}$ is one to one hence $\gamma_k \Theta_{w_{l_k}} = \tau^k$ and finally

$$\Theta_{w_{l_k}} \gamma_k = \gamma_k \Theta_{w_{l_k}}. \tag{2}$$

The substitution τ is linearly recurrent (with constant K), hence for all b in A we have

$$|\gamma_k(b)| = L_{w_{l_k}}(\tau^k(b) w_{l_k}) - 1 \leq \frac{|\tau^k(b)| + |w_{l_k}|}{K |w_{l_k}|} \leq \frac{2\beta c_2 |\tau^k(b)|}{K c_1 |\tau^{k-1}(1)|} \leq \frac{2 c_2 r^2 \alpha \beta}{K c_1},$$

where H_1 is the constant given by Theorem 23. We have proved that the length of the images by γ_k of letters are bounded independently of k. Hence the set $\{\gamma_k; k \geq k_0\}$ is finite. Thus there exists an infinite set I, included in J, such that $\gamma_p = \gamma_q$ for all elements p and q of I. Let p be an element of I, we write $\gamma = \gamma_p$. Equality (2) gives $\Theta^{l_p} \gamma = \gamma \Theta^{l_p} = \tau^p$. From this last equality it follows that the sequence $(l_p)_{p \in I}$ is strictly increasing. □

Lemma 36 *Let $\tau : A \to A^*$ and $\sigma : A \to A^*$ be two primitive substitutions having the same non-periodic fixed point x. There exist an integer l, a prefix v of x, and an arbitrarily long prefix u of $\mathcal{D}_v(\mathsf{x})$ such that the word u and the substitution τ_v^l , and u and the substitution σ_v^l, both satisfy the hypothesis of Proposition 35.*

Proof. Let $(\mathsf{x}^{(n)}; n \in \mathbb{N})$ be the sequences defined by: $\mathsf{x}^{(0)} = \mathsf{x}$ and $\mathsf{x}^{(n+1)}$ is the derived sequence of $\mathsf{x}^{(n)}$ on the word 1 (the first letter of $\mathsf{x}^{(n)}$). For all integers n we call $A^{(n)}$ the alphabet of $\mathsf{x}^{(n)}$. Let $(u^{(n)})_{n \geq 1}$ be the sequence of words defined by:

$$u^{(1)} = 1 \quad \text{and} \quad u^{(n+1)} = \Theta_{\mathsf{x}, u^{(n)}}(1) u^{(n)}.$$

According to Proposition 13, for all integers n larger than 1, the word $u^{(n)}$ is a prefix of x such that

$$\Theta_{\mathsf{x},1} \Theta_{\mathsf{x}^{(1)},1} \cdots \Theta_{\mathsf{x}^{(n-1)},1} = \Theta_{\mathsf{x}, u^{(n)}} \text{ and } \mathsf{x}^{(n)} = \mathcal{D}_{u^{(n)}}(\mathsf{x}).$$

The sets $\{\tau_u; u \text{ prefix of } \mathsf{x}\}$ and $\{\sigma_u; u \text{ prefix of } \mathsf{x}\}$ are finite by Proposition 27. Hence, there exists an infinite set I of positive integers such that for all integers p and q of I, we have $\tau_{u^{(p)}} = \tau_{u^{(q)}}$ and $\sigma_{u^{(p)}} = \sigma_{u^{(q)}}$. We remark that a fixed point of the substitutions $\tau_{u^{(p)}}$ and $\sigma_{u^{(p)}}$ is $\mathsf{x}^{(p)} = \mathcal{D}_{u^{(p)}}(\mathsf{x})$.
Let p and q be two elements of I with $p < q$. By definition of $(\mathsf{x}^{(n)}; n \in \mathbb{N})$ we have

$$\mathsf{x}^{(q)} = \underbrace{\mathcal{D}_1 \cdots \mathcal{D}_1}_{(q-p) \quad \text{times}} (\mathsf{x}^{(p)})$$

Hence (Proposition 13) there exists a prefix u of $\mathsf{x}^{(p)}$ such that

- $\mathcal{D}_u(\mathsf{x}^{(p)}) = \mathsf{x}^{(q)}$,
- $\Theta_{\mathsf{x}^{(p)},1}\Theta_{\mathsf{x}^{(p+1)},1}\cdots\Theta_{\mathsf{x}^{(q-1)},1} = \Theta_{\mathsf{x}^{(p)},u}$,
- $\Theta_{\mathsf{x}^{(p)},u}\tau_{u^{(q)}} = \tau_{u^{(p)}}\Theta_{\mathsf{x}^{(p)},u}$ and $\Theta_{\mathsf{x}^{(p)},u}\sigma_{u^{(q)}} = \sigma_{u^{(p)}}\Theta_{\mathsf{x}^{(p)},u}$.

From the last equalities it is clear that $(\tau_{u^{(p)}})_u = \tau_{u^{(q)}}$ and $(\sigma_{u^{(p)}})_u = \sigma_{u^{(q)}}$; where $(\tau_{u^{(p)}})_u$ and $(\sigma_{u^{(p)}})_u$ are respectively the return substitutions to u of $\tau_{u^{(p)}}$ and $\sigma_{u^{(p)}}$.
From the definition of $(u^{(n)})_{n\geq 1}$ we deduce that the sequence $(|u^{(n)}|)_{n\geq 1}$ is strictly increasing. Thus it follows from Lemma 8 that

$$\lim_{j\to+\infty} \min\{|\Theta_{\mathsf{x}^{(0)},u^{(j)}}(b)| = |\Theta_{\mathsf{x}^{(0)},1}\Theta_{\mathsf{x}^{(1)},1}\cdots\Theta_{\mathsf{x}^{(j)},1}(b)|; b\in A^{(j+1)}\} = +\infty.$$

and consequently that

$$\lim_{j\to+\infty} \min\{|\Theta_{\mathsf{x}^{(p)},1}\Theta_{\mathsf{x}^{(p+1)},1}\cdots\Theta_{\mathsf{x}^{(j)},1}(b)|; b\in A^{(j+1)}\} = +\infty.$$

Therefore we can suppose that q, and consequently u, is such that each letter of $A^{(p)}$ (the alphabet of $\mathsf{x}^{(p)}$) has at least one occurrence in each return word to u of $\mathsf{x}^{(p)}$. (We recall that the set of return words to u of $\mathsf{x}^{(p)}$ is $\{\Theta_{\mathsf{x}^{(p)},u}(b) = \Theta_{\mathsf{x}^{(p)},1}\Theta_{\mathsf{x}^{(p+1)},1}\cdots\Theta_{\mathsf{x}^{(q-1)},1}(b); b\in A^{(q)}\}$.)
The word $\Theta_{\mathsf{x},u^{(p)}}(1)u^{(p)}$ is a prefix of x hence we can choose an integer l such that the word $\Theta_{\mathsf{x},u^{(p)}}(1)u^{(p)}$ is a prefix of $\tau^l(1)$ and $\sigma^l(1)$. Thus the first letter of each image of $\tau^l_{u^{(p)}}$ and $\sigma^l_{u^{(p)}}$ is 1.
We set $v = u^{(p)}$ and $\gamma = \tau^l_v$. The substitution γ and the prefix u of $\mathsf{x}^{(p)}$ (which is a fixed point of γ) fulfill the hypotheses of Proposition 35. Indeed we chose the integer l to satisfy hypothesis 1. Hypothesis 2 is also satisfied because

$$\gamma_u = (\tau_{u^{(p)}})_u = \tau_{u^{(q)}} = \tau_{u^{(p)}} = \gamma,$$

where γ_u is the return substitution to u. Hypothesis 3 does not set any difficulty. Hypothesis 4 follows from the choice of q.
It is clear that $\sigma^l_{u^{(p)}}$ and u also satisfy the same hypotheses. □

Theorem 37 *If two primitive substitutions have the same non-periodic fixed point, then they have some powers which have the same eigenvalues, except perhaps* 0 *and roots of the unity.*

Proof. It follows from Proposition 35 and Lemma 36. □

Proof of Theorem 30. To begin we suppose that x is a non-periodic sequence which is both α-substitutive primitive and β-substitutive primitive. We will prove that α and β are multiplicatively independent. Let A be the alphabet of x. There exist fixed points y and z of, respectively, the substitutions $\tau : B\to B^*$ and $\sigma : C\to C^*$, a morphism ϕ, from B to A, and a morphism φ from C to A such that $\phi(\mathsf{y}) = \varphi(\mathsf{z}) = \mathsf{x}$.
Recall that by Theorem 3, if a sequence is primitive substitutive then its set of derived sequences is finite. Hence there exist three sequences, $(u^{(n)}; n\in \mathbb{N})$, $(v^{(n)}; n\in\mathbb{N})$ and $(w^{(n)}; n\in\mathbb{N})$, of prefixes of respectively y, x and z such that for all integers n we have:

- $\mathcal{D}_{u^{(n)}}(\mathsf{y}) = \mathcal{D}_{u^{(n+1)}}(\mathsf{y})$,
- $\mathcal{D}_{v^{(n)}}(\mathsf{x}) = \mathcal{D}_{v^{(n+1)}}(\mathsf{x})$,
- $\mathcal{D}_{w^{(n)}}(\mathsf{z}) = \mathcal{D}_{w^{(n+1)}}(\mathsf{z})$,
- $\phi(u^{(n)}) = \varphi(w^{(n)}) = v^{(n)}$ and $|v^{(n)}| < |v^{(n+1)}|$.

Let n be an integer. The images of words by $\phi\Theta_{u^{(n)}}$ are concatenations of return words to $v^{(n)}$. The map $\Theta_{v^{(n)}} : R^*_{v^{(n)}} \to A^*$ being one to one, this allows us to define a morphism λ_n by $\Theta_{v^{(n)}}\lambda_n = \phi\Theta_{u^{(n)}}$. In the same way we define the morphism γ_n by $\Theta_{v^{(n)}}\gamma_n = \varphi\Theta_{w^{(n)}}$. From Proposition 27 we know that the sets $\{\lambda_n; n \in \mathbb{N}\}$ and $\{\gamma_n; n \in \mathbb{N}\}$ are finite. For this reason we can suppose that for all integers n we have $\lambda_n = \lambda_{n+1}$ and $\gamma_n = \gamma_{n+1}$.
Let i be an integer. The sequence y (resp. z) is non-periodic and uniformly recurrent. Hence, according to Lemma 8, there exists an integer j larger than i such that each word $wu^{(i)}$, where w is a return word to $u^{(i)}$, has at least one occurrence in each return word to $u^{(j)}$. Consequently we can define a primitive substitution δ by $\Theta_{u^{(i)}}\delta = \Theta_{u^{(j)}}$. In the same way we define a primitive substitution ρ by $\Theta_{v^{(i)}}\rho = \Theta_{v^{(j)}}$. We have $\rho\lambda_j = \lambda_j\delta$. Indeed

$$\Theta_{v^{(i)}}\rho\lambda_j = \Theta_{v^{(j)}}\lambda_j = \phi\Theta_{u^{(j)}} = \phi\Theta_{u^{(i)}}\delta = \Theta_{v^{(i)}}\lambda_i\delta = \Theta_{v^{(i)}}\lambda_j\delta.$$

A standard application of Perron's Theorem (Theorem 15) shows that δ and ρ have the same dominant eigenvalue.
We recall that $\mathcal{D}_{u^{(i)}}(\mathsf{y}) = \mathcal{D}_{u^{(j)}}(\mathsf{y})$. Hence δ has the same fixed point as $\tau_{u^{(i)}}$, that is to say $\mathcal{D}_{u^{(i)}}(\mathsf{y})$. It follows from Theorem 37 and Proposition 34 that the dominant eigenvalues of δ and τ are multiplicatively dependent.
In the same way we prove that ρ and σ have multiplicatively dependent dominant eigenvalues. This completes the first part of the proof. The second part is given by the following proposition.

Proposition 38 *Let x be a sequence on a finite alphabet and α a Perron number. If x is periodic (resp. ultimately periodic) then it is α-substitutive primitive (resp. α-substitutive).*

Proof. Let x be a periodic sequence with period p. Hence we can suppose that $A = \{1, \cdots, p\}$ and $\mathsf{x} = (1\cdots p)^\omega$. Let M be a primitive matrix which dominant eigenvalue is α and $\sigma : B \to B^*$ a primitive substitution which matrix is M. Let y be one of its fixed points. In the sequel we construct, using σ, a new substitution τ with dominant eigenvalue α, together with a fixed point $\mathsf{z} = \tau(\mathsf{z})$, and a letter to letter morphism ϕ such that $\phi(\mathsf{z}) = \mathsf{x}$. We define the alphabet

$$D = \{(b, i) \; ; b \in B \; , \; 1 \leq i \leq p\},$$

the morphism $\psi : B \to D^*$ and the substitution $\tau : D \to D^*$ by

$$\psi(b) = (b,1)\cdots(b,p) \quad \text{and} \quad \tau((b,i)) = (\psi(\sigma(b)))_{[(i-1)|\sigma(b)|, i|\sigma(b)|-1]},$$

for all $(b,i) \in D$. The substitution τ is well defined because $|\psi(\sigma(b))| = p|\sigma(b)|$. Moreover, these morphisms are such that $\tau \circ \psi = \psi \circ \sigma$. Hence the substitution τ is primitive. Its

fixed point is $\mathsf{z} = \psi(\mathsf{y})$ and (using Perron's Theorem and the fact that $M_\tau M_\psi = M_\psi M_\sigma$) its dominant eigenvalue is α.
Let $\phi : D \to A$ be the letter to letter morphism defined by $\phi((b,i)) = i$. It is easy to see that $\phi(\mathsf{z}) = \mathsf{x}$. It follows that x is α-substitutive primitive.
Suppose now that x is ultimately periodic. Then there exists two non-empty words u and v such that $\mathsf{x} = uv^\omega$. From what precedes we know that there exists a substitution $\tau : D \to D^*$, a fixed point $\mathsf{z} = \tau(\mathsf{z})$ and a letter to letter morphism $\phi : D \to A$ such that $\phi(\mathsf{z}) = v^\omega$. Let $E' = \{a_1, a_2, \cdots, a_{|u|}\}$ be an alphabet, with $|u|$ letters, disjoint from D and consider the sequence $\mathsf{t} = a_1 a_2 \cdots a_{|u|}\mathsf{z} \in (E' \cup D)^{\mathbb{N}} = F^{\mathbb{N}}$. It suffices to prove that t is α-substitutive. We extend τ to F setting $\tau(a_i) = a_i$, $1 \le i \le |u|$. Let G be the alphabet of the words of length $|u|+1$ of t, that is to say

$$G = \{(\mathsf{t}_n \mathsf{t}_{n+1} \cdots \mathsf{t}_{n+|u|}); n \in \mathbb{N}\} \quad \text{where} \quad \mathsf{t} = \mathsf{t}_0 \mathsf{t}_1 \cdots .$$

The sequence $\bar{\mathsf{t}} = (\mathsf{t}_0 \mathsf{t}_1 \cdots \mathsf{t}_{|u|})(\mathsf{t}_1 \mathsf{t}_2 \cdots \mathsf{t}_{|u|+1}) \cdots (\mathsf{t}_n \mathsf{t}_{n+1} \cdots \mathsf{t}_{n+|u|}) \cdots$ is the fixed point of the substitution $\zeta : G \to G^*$ we define as follows. Let $(l_0 l_1 \cdots l_{|u|-1} a)$ be an element of G. Let $s_0 s_1 \cdots s_{|u|-1}$ be the suffix of length $|u|$ of $\tau(l_0 l_1 \cdots l_{|u|-1})$.
If $|\tau(a)| \le |u|$, we set $\zeta((l_0 l_1 \cdots l_{|u|-1} a)) =$

$$(s_{[0,|u|-1]}\tau(a)_0)(s_{[1,|u|-1]}\tau(a)_{[0,1]}) \cdots (s_{[|\tau(a)|-1,|u|-1]}\tau(a)_{[0,|\tau(a)|-1]}),$$

otherwise $\zeta((l_0 l_1 \cdots l_{|u|-1} a)) =$

$$(s_{[0,|u|-1]}\tau(a)_0) \cdots (s_{|u|-1}\tau(a)_{[0,|u|-1]})(\tau(a)_{[0,|u|]}) \cdots (\tau(a)_{[|\tau(a)|-|u|-1,|\tau(a)|-1]}),$$

By induction we can prove that for all $n \in \mathbb{N}$ we have

$$\zeta^n((\mathsf{t}_0 \mathsf{t}_1 \cdots \mathsf{t}_{|u|})) = (\mathsf{t}_0 \mathsf{t}_1 \cdots \mathsf{t}_{|u|})(\mathsf{t}_1 \mathsf{t}_2 \cdots \mathsf{t}_{|u|+1}) \cdots (\mathsf{t}_{|\tau^n(\mathsf{t}_{|u|})|-1} \cdots \mathsf{t}_{|\tau^n(\mathsf{t}_{|u|})|+|u|-1}).$$

Consequently $\bar{\mathsf{t}}$ is a fixed point of ζ and $\rho(\bar{\mathsf{t}}) = \mathsf{t}$ where $\rho : G \to F$ is defined by

$$\rho((r_0 r_1 \cdots r_{|u|})) = r_0.$$

Moreover we remark that for all $n \in \mathbb{N}$ we have

$$|\zeta^n((r_0 r_1 \cdots r_{|u|}))| = |\tau^n(r_{|u|})|.$$

From this and Lemma 20 it comes that for all $(r_0 r_1 \cdots r_{|u|}) \in D$ we have

$$\lim_{n \to +\infty} \frac{|\zeta^{n+1}((r_0 r_1 \cdots r_{|u|}))|}{|\zeta^n((r_0 r_1 \cdots r_{|u|}))|} = \alpha.$$

Hence α is the dominant eigenvalue of ζ and t is α-substitutive. $\square$

Could we obtain a result analogous to Theorem 37? That is to say concerning all eigenvalues. The answer is negative. Here is a counterexample: Let τ and σ be two substitutions defined respectively by

$$\begin{cases} a & \to & abab \\ b & \to & abbb \end{cases} \quad \text{and} \quad \begin{cases} a & \to & abab \\ b & \to & accc \\ c & \to & abbc \end{cases}.$$

Let x and y be their respective fixed points. Eigenvalues of the substitution τ are 1 and 4. Those of σ are 1, -2 and 4. Let $\phi : \{a,b,c\} \to \{a,b\}$ be the morphism defined by $\phi(a) = a$ and $\phi(b) = \phi(c) = b$, then $\phi(\mathsf{y}) = \mathsf{x}$. The sequence x arises from two substitutions, one has the eigenvalue -2 and the other does not.

4.2 A Cobham Theorem for languages

In this subsection we prove Theorem 31. First we recall some known results about primitive substitutions.

Let $\tau : A \to A^*$ be a primitive substitution, x one of its fixed points, M its matrix, α the dominant eigenvalue of M and $v = (v_i; i \in A) \in \mathbb{R}_+^{|A|}$ the unique eigenvector associated to α such that $\sum_{i\in A} v_i = 1$ (Theorem 15). We call it the *frequency vector of* τ (or x). We justify this name with the following theorem (Theorem V.13 and Corollary V.14 in [Qu]). Let u be a word of A^*, we define $N(u) = (n_i; i \in A)$ the vector where n_i is the number of occurrences of the letter i in the word u.

Theorem 39 *Let $\tau : A \to A^*$ be a primitive substitution, x one of its fixed points, and $v = (v_i; i \in A) \in \mathbb{R}_+^{|A|}$ its frequency vector. Then*

$$\lim_{l\to+\infty} \frac{N(\mathsf{x}_k \cdots \mathsf{x}_{k+l})}{l+1} = v$$

uniformly in k.

In other words, for all i in A the frequency of the letter i in x exists and is equal to v_i.

We keep the previous notations. Let B be an alphabet, $\varphi : A \to B^*$ be a letter to letter morphism, P its matrix and $\mathsf{y} = \varphi(\mathsf{x})$. We call *frequency vector of* y the vector Pv. We remark that the sum of its coordinates is equal to 1. The proof of the following corollary is immediate.

Corollary 40 *Let $\tau : A \to A^*$ be a primitive substitution, x one of its fixed points, $v = (v_i; i \in A) \in \mathbb{R}_+^{|A|}$ its frequency vector, B be an alphabet, $\varphi : A \to B^*$ be a letter to letter morphism, P its matrix and $\mathsf{y} = \varphi(\mathsf{x}) = (\mathsf{y}_n; n \in \mathbb{N})$. Then*

$$\lim_{l\to+\infty} \frac{N(\mathsf{y}_k \cdots \mathsf{y}_{k+l})}{l+1} = Pv$$

uniformly in k.

Proof of Theorem 31. From Theorem 3 there exists a sequence of prefixes of x, $(u_i; i \in \mathbb{N})$, satisfying for all $i \in \mathbb{N}$:

- u_i is a prefix of u_{i+1} with $u_i \neq u_{i+1}$ and
- $\mathcal{D}_{u_i}(\mathsf{x}) = \mathcal{D}_{u_{i+1}}(\mathsf{x}) = \tilde{\mathsf{x}}$.

We notice that $\tilde{\mathsf{x}}$ is α^n-substitutive primitive for some n (Lemma 26).

Let $(u_i; i \in \mathbb{N})$ be such a sequence and $(v_i; i \in \mathbb{N})$ be the unique sequence of prefixes of y verifying for all $i \in \mathbb{N}$:

- u_i is a prefix of v_i and
- v_i has exactly one occurrence of u_i ; we set $v_i = d_i u_i$.

The sequences x and y being uniformly recurrent, we can suppose (Proposition 8) that, for all $i \in \mathbb{N}$, each return word to u_i (resp. v_i) has an occurrence in each return word to u_{i+1} (resp. v_{i+1}).
From Theorem 24 there exist a number K such that for all words $u \in L(\mathsf{x}) = L(\mathsf{y})$ and all words $v \in \mathcal{R}_u$ we have $|u|/K \leq |v| \leq K|u|$. We remark that for all i in $\mathbb{N}$ the word d_i is a suffix of a return word to u_i. This implies that $|v_i| \leq (K+1)|u_i|$ for all $i \in \mathbb{N}$.
From Theorem 3, we can suppose that for all $i \in \mathbb{N}$ we have $\mathcal{D}_{v_i}(\mathsf{y}) = \mathcal{D}_{v_{i+1}}(\mathsf{y}) = \tilde{\mathsf{y}}$. The sequence $\tilde{\mathsf{y}}$ is β^m-substitutive primitive for some $m \in \mathbb{N}$ (Lemma 26).
For all $i \in \mathbb{N}$ we have $R_{u_{i+1}} = R_{u_i}$ and $R_{v_{i+1}} = R_{v_i}$. We set $R_{u_0} = C$ and $R_{v_0} = D$ and we recall that $\tilde{\mathsf{x}}$ belongs to $C^{\mathbb{N}}$ and $\tilde{\mathsf{y}}$ to $D^{\mathbb{N}}$.
Let $i \in \mathbb{N}$ and $b \in D$. The word $\Theta_{v_i}(b)v_i = \Theta_{v_i}(b)d_iu_i$ belongs to $L(\mathsf{x})$ hence $\Theta_{v_i}(b)d_i$ belongs to $L(\mathsf{x})$ too. The word v_i is a prefix of $\Theta_{v_i}(b)v_i$ hence there exists a word t such that $\Theta_{v_i}(b)v_i = d_itu_i$ and such that u_i is a prefix of tu_i. Hence t is a concatenation of return words to u_i. In other words there exists a unique word $\rho_i(b)$ in $R_{u_i}^*$ such that $t = \Theta_{u_i}\rho_i(b)$. This defined a morphism $\rho_i : D \to C^*$ verifying $\Theta_{v_i}(b)d_i = d_i\Theta_{u_i}(\rho_i(b))$. We remark that if s is an element of $L(\tilde{\mathsf{y}})$ then we have also $d_i\Theta_{u_i}(\rho_i(s)) = \Theta_{v_i}(s)d_i$.
From Theorem 24 we obtain an upper bound for $|\rho_i(b)|$.

$$|\rho_i(b)| = L_{u_i}(\Theta_{v_i}(b)d_i) \leq \frac{|\Theta_{v_i}(b)d_i|}{(1/K)|u_i|} \leq K\frac{K|v_i| + |d_i|}{|u_i|} \leq K\frac{(K+1)|v_i|}{|u_i|} \leq K(K+1)^2.$$

Consequently the set $\{\rho_i; i \in \mathbb{N}\}$ is finite. Hence we can suppose that $\rho_j = \rho_{j+1} = \rho$ for all $j \in \mathbb{N}$.
Let $i \in \mathbb{N}$. The return words to u_i are concatenations of return words to u_0 and Θ_{u_0} is one-to-one (Proposition 11), we can define a substitution $\sigma_i : C \to C^*$ with $\Theta_{u_0}\sigma_i = \Theta_{u_i}$. This substitution is primitive because each return word to u_0 has an occurrence in each return word to u_i. One of its fixed points is $\tilde{\mathsf{x}}$. In fact we have

$$\Theta_{u_0}\sigma_i(\tilde{\mathsf{x}}) = \Theta_{u_i}(\tilde{\mathsf{x}}) = \Theta_{u_i}(\mathcal{D}_{u_i}(\mathsf{x})) = x = \Theta_{u_0}(\tilde{\mathsf{x}}),$$

and Θ_{u_0} being one to one it comes $\sigma(\tilde{\mathsf{x}}) = \tilde{\mathsf{x}}$.
In the same way we define the primitive substitution $\tau_i : D \to D^*$ with $\Theta_{v_0}\tau_i = \Theta_{v_i}$. One of its fixed point is $\tilde{\mathsf{y}}$. Let α' et β' be the dominant eigenvalues of, respectively, σ_i and τ_i. It comes that $\tilde{\mathsf{x}}$ (resp. $\tilde{\mathsf{y}}$) is α^n-substitutive primitive and α'-substitutive primitive (resp. β^m-substitutive primitive and β'-substitutive primitive). Theorem 30 implies that α^n and α' (resp. β^m and β') are multiplicatively independent. For all $i \in \mathbb{N}$ there two rational numbers p_i and q_i such that α^{p_i} is the dominant eigenvalue of σ_i and β^{q_i} the one of τ_i.
Let $w \in L(\tilde{\mathsf{y}})$, we have

$$\frac{|d_i\Theta_{u_i}\rho(w)|}{|w|} = \frac{|d_i\Theta_{u_0}\sigma_i\rho(w)|}{|w|} = \frac{|d_i|}{|w|} + \frac{|\Theta_{u_0}\sigma_i\rho(w)|}{|\sigma_i\rho(w)|}\frac{|\sigma_i\rho(w)|}{|\rho(w)|}\frac{|\rho(w)|}{|w|}.$$

We call v the frequency vector of $\tilde{\mathsf{y}}$. From Theorem 39 we have

$$\lim_{w \in L(\tilde{\mathsf{y}}), |w| \to +\infty} \frac{|\rho(w)|}{|w|} = \lim_{w \in L(\tilde{\mathsf{y}}), |w| \to +\infty} ||M_\rho\left(\frac{N(w)}{|w|}\right)|| = ||M_\rho(v)|| = c_1.$$

where $||.||$ is the norm defined by $||(v_1, \cdots, v_n)|| = |v_1| + \cdots + |v_n|$. We remark that $\sigma_i \rho(w)$ belongs to $L(\tilde{x})$ and we call u the frequency vector of $\tilde{x}$. Applying once more Theorem 39 we obtain

$$\lim_{w \in L(\tilde{y}), |w| \to +\infty} \frac{|\Theta_{u_0} \sigma_i \rho(w)|}{|\sigma_i \rho(w)|} = ||M_{\Theta_{u_0}}(u)|| = c_2.$$

Then it comes

$$\lim_{w \in L(\tilde{y}), |w| \to +\infty} \frac{|d_i \Theta_{u_i} \rho(w)|}{|w|} = c_1 c_2 ||M_{\sigma_i}(u)|| = c_1 c_2 \alpha^{p_i}.$$

We note that the constants c_1 and c_2 do not depend on i. With analogous considerations we show there exists a constant c_3, which does not depend on i, such that

$$\lim_{w \in L(\tilde{y}), |w| \to +\infty} \frac{|\Theta_{v_i}(w) d_i|}{|w|} = c_3 \beta^{q_i}.$$

We obtain $c_1 c_2 \alpha^{p_i} = c_3 \beta^{q_i}$. Let j be a positive integer distinct from i. The sequences $(p_i; i \in \mathbb{N})$ and $(q_i; i \in \mathbb{N})$ tends to infinity hence we can suppose that $p_i < p_j$ and $q_i < q_j$. Then $\alpha^{p_j - p_i} = \beta^{q_j - q_i}$, i.e. α and β are multiplicatively dependent. □

4.3 A Cobham Theorem for subshifts

In this subsection we prove Theorem 2. We first recall the Curtis-Hedlund-Lyndon Theorem (Theorem 6.2.9 in [LM]).

Theorem 41 *Let (X,T) and (Y,T) be two subshifts on, respectively, the alphabets A and B, such that there exists an isomorphism $F : (X,T) \to (Y,T)$. Then there exists a map $f : A^{2r+1} \to B$ such that for all $i \in \mathbb{Z}$ and $x \in X$ we have*

$$(F(x))_i = f(x_{[i-r, i+r]}) \ .$$

The map $f : A^{2r+1} \to B$ is called a *block map*. It can naturally be extended to a map from A^n to B^{n-2r}, for all $n \geq 2r+1$, in the following way : $(f(u))_i = f(u_{[i,i+2r]})$ for all $u \in A^n$, $n \geq 2r$ and $0 \leq i < n - 2r$.
Let x be a sequence on the alphabet A. For each $n \in \mathbb{N}$ we define the sequence $x^{(n)}$ on the alphabet $A^n = \{(x_i \cdots x_{i+n-1}); i \in \mathbb{N}\}$ by $x_i^{(n)} = (x_i \cdots x_{i+n-1})$, for all $i \in \mathbb{N}$. We recall a result in [Qu].

Proposition 42 *If x is a α-substitutive primitive sequence then $x^{(n)}$ is also α-substitutive primitive for all $n \geq 1$. Moreover the subshifts generated by these sequences are isomorphic.*

Proof of Theorem 2. Let A and B be the alphabets of respectively (X,T) and (Y,T). There exists an isomorphism $F : (X,T) \to (Y,T)$ and a map $f : A^{2n+1} \to B$ such that for all $i \in \mathbb{Z}$ and $z \in X$ we have $(F(z))_i = f(z_{[i-n,i+n]})$ (Theorem 41). Hence, there exists a letter to letter morphism $\phi : A^r \to B$ such that for all $z \in X$ we have $F(z) = \phi(z^{(n)})$. The sequence $\phi(x^{(n)})$ is α-substitutive and generates (X,T), as well as y. Hence they have the same language. Theorem 31 achieves the proof. □

Let x be a α-substitutive primitive sequence and (X,T) be the subshift it generates. We set $I(X,T) = \overline{\alpha}$ where $\overline{\alpha}$ is the equivalence class of α for the equivalence relation defined in $\mathbb{R}^+$ with $\beta \equiv \gamma$ if and only if β and γ are multiplicatively dependent. Theorem 2 implies that $I(X,T)$ is a isomorphism invariant for subshifts generated by primitive substitutive sequences.

Theorem 43 *Let (X,T) and (Y,T) be two subshifts generated by primitive substitutive sequences. If (X,T) and (Y,T) are isomorphic then $I(X,T) = I(Y,T)$.*

The reciprocal is not true. The following substitutions

$$\begin{array}{llcll} \sigma(0) & = 010 & \text{et} & \tau(0) & = 001 \\ \sigma(1) & = 01 & & \tau(1) & = 10, \end{array}$$

have the same dominant eigenvalue α^2 where $\alpha = (1+\sqrt{5})/2$ but their dimension group, respectively $(\mathbb{Z}^2, \{(x,y) \in \mathbb{Z}^2; x + \alpha y > 0\}, (3,5))$ and $(\mathbb{Z}^3, \{(x,y,z) \in \mathbb{Z}^3; \alpha x + 2y + z > 0\}, (2,0,-1))$, are not isomorphic (for more details see [DHS]).

4.4 Measures of cylinders

In this subsection we just want to point out a nice result of Holton and Zamboni [HZ2] that gives a different proof of Theorem 2.

Let (X,S) be a dynamical system. An *invariant measure* for (X,S) is a probability measure μ, on the σ-algebra $\mathcal{B}(X)$ of Borel sets, with $\mu(S^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}(X)$; the measure is *ergodic* if every S-invariant Borel set has measure 0 or 1. The set of invariant measures for (X,S) is denoted by $\mathcal{M}(X,S)$. The system (X,S) is *uniquely ergodic* if $\text{Card}(\mathcal{M}(X,S)) = 1$. It is well known ([Qu]) that the subshift generated by substitutive primitive sequences are uniquely ergodic. In [HZ2] it is proved the following.

Theorem 44 *Let (X,T) be a subshift generated by a α-substitutive primitive sequence and μ be its unique ergodic measure. Then, the measures of cylinders in X lie in a finite union of geometric progressions. More precisely there exists a finite set of positive real numbers $\mathcal{F}$ such that*

$$\{\mu([u]); u \in L(X)\} \subset \bigcup_{n \in \mathbb{N}} \alpha^{-n} \mathcal{F}.$$

They proved this using what they called directed graphs. There exists a very short proof of this result using Bratteli diagrams.

Let F be an isomorphism from the subshift (X,T) to the subshift (Y,T), generated respectively by an α-substitutive primitive sequence and a β-substitutive primitive sequence. We have to prove that α and β are multiplicatively independent.

Let μ and λ be the unique ergodic measures of (X,T) and (Y,T) respectively. The measure $F\mu$ defined by $F\mu(A) = \mu(F^{-1}(A))$, for all Borel set of Y, is ergodic and consequently equal to λ. From Theorem 44 there exists a finite set $\mathcal{F}$ such that $\{\lambda([u]); u \in L(Y)\} \subset \bigcup_{n \in \mathbb{N}} \beta^{-n}\mathcal{F}$. In the proof of Theorem 20 in [Du4] is obtained the following result.

Proposition 45 *Let (X,T) be a subshift generated by a substitutive primitive sequence on the alphabet A. There exists a constant K such that for all block map $f : A^{2r+1} \to B$ we have* $\text{Card}(f^{-1}(\{u\}) \leq K$.

Consequently from Theorem 41 there exists a finite set $\mathcal{G}$ such that $\{F\mu([u]); u \in L(Y)\} \subset \bigcup_{n\in \mathbb{N}} \alpha^{-n}\mathcal{G}$. This implies that α and β have to be multiplicatively dependent.

5 Cobham Theorem for numeration systems

In this Section we prove Theorem 1. We obtain it as a corollary of the following theorem.

Theorem 46 *Let α and β be two multiplicatively independent Perron numbers. Let A be a finite alphabet. Restricted to the family of quasi-uniform sequences, a sequence $\mathsf{x} \in A^{\mathbb{N}}$ is α-substitutive and β-substitutive if and only if it is ultimately periodic.*

Where "restricted to quasi-uniform substitutions" means that the sequence x is the image by a letter to letter morphism of a fixed point of a quasi-uniform substitution. Results about recognizability by finite automaton will not be proved but all references concerning the proofs are given.

5.1 Finite Automata

Let A be a finite alphabet. An *automaton over* A, $\mathcal{A} = (Q, A, E, I, T)$, is a directed graph labelled by elements of A where Q is the set of *states*, $I \subset Q$ is the set of *initial* states, $T \subset Q$ is the set of *terminal* states and $E \subset Q \times A \times Q$ is the set of labelled edges. If $(p, a, q) \in E$, we write $p \xrightarrow{a} q$. The automaton is *finite* if Q is finite. A *path* in the automaton is a sequence $P = ((p_n, a_n, q_n); 0 \leq n \leq N)$ where $q_n = p_{n+1}$ for all $0 \leq n \leq N-1$. We say it is *admissible* if p_0 belongs to I and q_N belongs to T. We call $L(\mathcal{A})$ the set of all admissible paths. The *label* of P is the word $a_0 a_1 \cdots a_N$. The set $L(\mathcal{A})$ of labels of admissible paths is called the language of $\mathcal{A}$. A subset L of A^* is said to be *recognizable by a finite automaton* if there exists a finite automaton $\mathcal{A}$ such that $L = L(\mathcal{A})$.

5.2 Numeration systems

A *numeration system* is a strictly increasing sequence of integers $U = (U_n; n \in \mathbb{N})$ such that

1. $U_0 = 1$,
2. the set $\{\frac{U_{n+1}}{U_n}; n \in \mathbb{N}\}$ is bounded.

Let $U = (U_n; n \in \mathbb{N})$ be a numeration system and c be the upper bound of $\{\frac{U_{n+1}}{U_n}; n \in \mathbb{N}\}$. Let A_U be the alphabet $\{0, \cdots, c' - 1\}$ where c' is the upper integer part of c. Using the Euclid algorithm we can write in a unique way each integer x as follows

$$x = a_i U_i + a_{i-1} U_{i-1} + \cdots + a_0 U_0;$$

i is the unique integer such that $U_i \leq x < U_{i+1}$ and $x_i = x$, $x_j = a_j U_j + x_{j-1}$, $j \in \{1, \cdots, i\}$, where a_j is the quotient of the Euclidean division of x_j by U_j and x_{j-1} the remainder, and $a_0 = x_0$. We will say that $\rho_U(x) = a_i \cdots a_0$ is the *U-representation* of x and we set

$$L(U) = \{0^n \rho_U(x); n \in \mathbb{N}, x \in \mathbb{N}\}.$$

We say a set $E \subset \mathbb{N}$ is *U-recognizable* if the language $0^* \rho_U(E) = \{0^n \rho_U(x); n \in \mathbb{N}, x \in E\}$ is recognizable by a finite automaton. We say that U is *linear* if it is defined with a linearly recurrent relation, i.e. if there exists $k \in \mathbb{N}^*$, $d_1, \cdots, d_k \in \mathbb{Z}$, $d_k \neq 0$, such that for all $n \geq k$

$$U_n = d_1 U_{n-1} + \cdots + d_k U_{n-k}.$$

The polynomial $P(X) = X^k - d_1 X^{k-1} - \cdots - d_{k-1} X - d_k$ is called *characteristic polynomial* of U.
When $U = (p^n; n \in \mathbb{N})$ we say that E is p-recognizable and we set $\rho_U = \rho_p$ and $U = U_p$.

5.3 Some examples

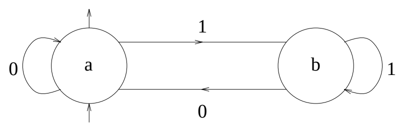
Now we can give some examples to illustrate the Second Cobham's Theorem. Let

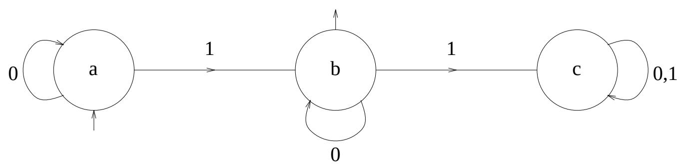
$$E_1 = \{2n; n \in \mathbb{N}\},\ E_2 = \{2^n; n \in \mathbb{N}\},\ E_3 = \left\{ n \in \mathbb{N}; \sum_{i=0}^{k} \epsilon_i \equiv 0[2], \rho_2(n) = \epsilon_k \cdots \epsilon_1 \epsilon_0 \right\}.$$

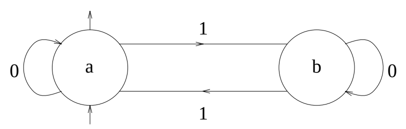
We have $0^* \rho_2(E_1) = \{w0; w \in \{0,1\}^*\}$,

$$0^* \rho_2(E_2) = \{0^n 1 0^m; n, m \in \mathbb{N}\} \text{ and } 0^* \rho_2(E_3) = \{w_0 \cdots w_n \in \{0,1\}^*; \sum_{i=0}^{k} \epsilon_i \equiv 0[2]\}.$$

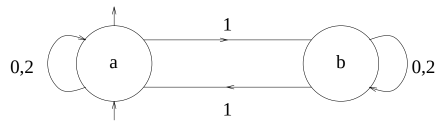
Hence these sets are 2-recognizable, respectively, by the following automata.







Where an arrow going into a state means that this state is an initial state and an arrow going outside of a state means that this state is a terminal state.
The set E_2 being an arithmetic progression, the First Cobham's Theorem asserts that E_1 is 3-recognizable. We have $0^* \rho_3(E_1) = \{w_0 \cdots w_n \in \{0,1,2\}^*; \sum_{i=0}^{k} \epsilon_i \equiv 0[2]\}$ and the automaton recognizing this set is the following.



The Second Cobham's Theorem asserts that for each of these sets we can find some substitutions of constant length that generate their characteristic sequences. We labelled the states to construct these substitutions. Let the set of states be the alphabet A of the substitution. The image of the state a is the word $w_0 \cdots w_{|A|}$ where w_i is the state you reach starting from a and passing through the arrow labelled by i. The substitutions we obtain are :

$$\begin{array}{lllllllllllllllll} \sigma_1 & : & a & \to & ab & \quad \bar{\sigma}_1 & : & a & \to & aba & \quad \sigma_2 & : & a & \to & ab & \quad \sigma_3 & : & a & \to & ab \\ & & b & \to & ab & & & b & \to & bab & & & b & \to & bc & & & b & \to & ba \\ & & & & & & & & & & & & c & \to & cc & & & & & \end{array}$$

Now call, for each substitution, x the unique fixed point starting with the letter a and identified to 1 the initial states and to 0 the other states. You obtained the characteristic sequence y of the corresponding set of integers E. For E_3 we obtain the well-known Morse sequence y :

$$\begin{array}{lcl} \mathsf{x} & = & a\ b\ b\ a\ b\ a\ a\ b\ b\ a \cdots \\ \downarrow & & \\ \mathsf{y} & = & 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 1 \cdots \\ \uparrow & & \\ E & : & 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9 \cdots . \end{array}$$

5.4 Bertrand numeration systems

A *Bertrand numeration system* U [Ber2] is a numeration system satisfying for all $n \in \mathbb{N}$:

$$w \in L(U) \text{ if and only if } w0^n \in L(U).$$

It is a natural condition because all numeration systems in base $p \geq 2$ satisfy it.
Let $\alpha > 1$ be a real number. All $x \in [0, 1]$ can be uniquely written in the following way :

$$x = \sum_{n \geq 1} a_n \alpha^{-n}, \tag{3}$$

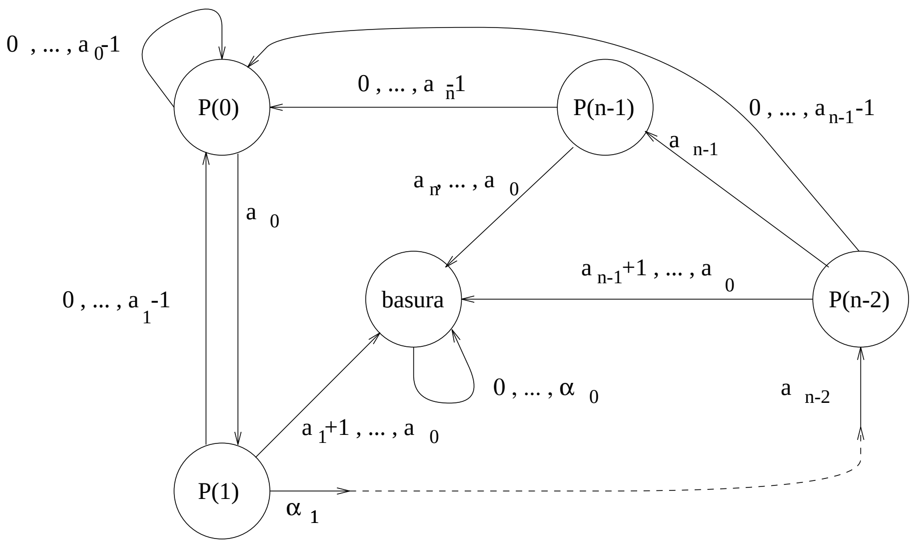
with $x_1 = x$ and for all $n \geq 1$, $a_n = [\alpha x_n]$ and $x_{n+1} = \{\alpha x_n\}$, where $[.]$ is the integer part and $\{.\}$ the fractional part. We call *α-expansion* of x the sequence $d_\alpha(x) = (a_n; n \in \mathbb{N}^*)$ and $L(\alpha)$ the set of finite words having an occurrence in some sequences $d_\alpha(x)$, $x \in [0, 1]$. If $d_\alpha(1)$ is ultimately periodic we say α is a *β-number* (for more details or informations about these numbers see [Par] or [Fr]). We remark that integers greater or equal to 2 are β-numbers. Bertrand-Mathis proved the following results :

Theorem 47 [Ber2] *Let U be a numeration system. It is a Bertrand numeration system if and only if there exists a real number $\alpha > 1$ such that $L(U) = L(\alpha)$. In this case, if U is linear then α is a root of the characteristic polynomial of U.*

Theorem 48 [Ber1] *Let $\alpha > 1$ be a real number. The language $L(\alpha)$ is recognizable with a finite automaton if and only if α is a β-number.*

5.5 ω_α-substitutive sequences

We recall that when α is an integer we have a characterization of U_α-recognizable sets of integers by means of substitutions of constant length (this is the second Theorem of Cobham), where $U = (\alpha^n; n \in \mathbb{N})$. We will see we have the same kind of characterization for Bertrand numeration systems corresponding to some non integral β-numbers.
Let U be a Bertrand numeration system such that $L(U) = L(\alpha)$ where α is a non integral β-number. It is not very difficult to be convinced that, in the case where $d_\alpha(1)$ is periodic, $\mathbb{N}$ is recognizable by the finite automaton given in the next Figure. In the ultimately periodic case the same kind of automaton could be given.



From this automaton in [Fab2] the author defines a substitution he called ω_α. The importance of these substitutions is justified by Theorem 49. They will allow us to applied results of the previous sections and to prove Theorem 1.

- If $d_\alpha(1) = a_1 \cdots a_n 0^\omega$, $a_n \neq 0$, then $(\omega_\alpha, \{1, \cdots, n\}, 1)$ is defined by

$$\begin{array}{ll} 1 & \to 1^{a_1}2; \\ \vdots & \\ n-1 & \to 1^{a_{n-1}}n; \\ n & \to 1^{a_n}; \end{array} .$$

- If $d_\alpha(1) = a_1 \cdots a_n(a_{n+1}a_{n+2} \cdots a_{n+m})^\omega$, where n and m are minimal and where $a_{n+1} + a_{n+2} + \cdots + a_{n+m} \neq 0$, then $(\omega_\alpha, \{1, \cdots, n+m\}, 1)$ is defined by

$$\begin{array}{ll} 1 & \to 1^{a_1}2; \\ \vdots & \\ n+m-1 & \to 1^{a_{n+m-1}}(n+m); \\ n+m & \to 1^{a_{n+m}}(n+1); \end{array} .$$

We remark in both cases the substitution ω_α is primitive and α is the dominant eigenvalue of M_{ω_α}. Let $\sigma : A \to A^*$ and $\tau : B \to B^*$ be two substitutions, we say that σ projects on τ if there exists a letter to letter morphism $\phi : A \to B^*$ such that $\phi \circ \sigma = \tau \circ \phi$. A substitution that projects on ω_α is called ω_α*-substitution* and we call ω_α*-substitutive* sequence (α-automatic sequence in [Fab2]) each sequence which is the image under a letter to letter morphism of a fixed point of a ω_α-substitution. In [Fab2] (Corollary 1) Fabre proved the following result :

Theorem 49 *Let U be a Bertrand numeration system such that $L(U) = L(\alpha)$ where α is a β-number. A set $E \subset \mathbb{N}$ is U-recognizable if and only if its characteristic sequence $(\mathsf{x}_n; n \in \mathbb{N})$ ($\mathsf{x}_n = 1$ if $n \in E$ and $\mathsf{x}_n = 0$ otherwise) is ω_α-substitutive.*

We remark that $\phi \circ \sigma^n = \tau^n \circ \phi$. If τ is primitive (and consequently quasi-uniform) then it comes that σ is quasi-uniform. We obtain the following corollary (see [Du3]).

Corollary 50 *Let U be a Bertrand numeration system such that $L(U) = L(\alpha)$ where α is a β-number. If the set $E \subset \mathbb{N}$ is U-recognizable then its characteristic sequence $(\mathsf{x}_n; n \in \mathbb{N})$ ($\mathsf{x}_n = 1$ if $n \in E$ and $\mathsf{x}_n = 0$ otherwise) is the image under a letter to letter morphism of a fixed point of a quasi uniform substitution which dominant eigenvalue is α.*

5.6 An application to numeration systems

To prove Theorem 46 we first prove that the letters appearing infinitely many times in x appears with bounded gaps. From this we deduce the same result for the words. And to conclude we use a particular matrix decomposition into primitive diagonal blocks and Theorem 31.

Words appear with bounded gaps. Let α and β be two multiplicatively independent Perron numbers. Let σ and τ be two uniform substitutions on the alphabets A and B, with fixed points y and z respectively. Let $\phi : A \to C$ and $\psi : B \to C$ be two letter to letter morphisms such that $\phi(\mathsf{y}) = \psi(\mathsf{z}) = \mathsf{x}$. We recall that two real numbers α and β are said to be *multiplicatively independent* if and only if for all $k, l \in \mathbb{Z}$, $\alpha^k \neq \beta^l$. The following well-known theorem will be crucial in this section (see [HW]).

Theorem 51 *Let α and β be two multiplicatively independent positive numbers. Let d and e be two non-negative integers. Then the set*

$$\left\{ \frac{\alpha^n}{\beta^m}; n, m \in \mathbb{N} \right\}$$

is dense in $\mathbb{R}^+$.

Proposition 52 *The letters of C which have infinitely many occurrences in x appear with bounded gaps in x.*

Proof. Let $c \in C$ which has infinitely many occurrences. Let $X = \{n \in \mathbb{N}; \mathsf{x}_n = c\}$. Assume that the letter c do not appear with bounded gaps. Then there exists $a \in A$ with infinitely

many occurrences in y and a strictly increasing sequence $(p_n; n \in \mathbb{N})$ of positive integers such that in $\phi(\sigma^{p_n}(a))$ do not appear the letter c. Let A'' be the set of such letters. Fix $a \in A''$. Let $u \in A^*$ such that ua is a prefix of y. Of course we can suppose that u is non-empty.
For all $n \in \mathbb{N}$ we call $\Omega_n \subset A$ the set of letters appearing in $\sigma^{p_n}(a)$. There exists two distinct integers $n_1 < n_2$ such that $\Omega_{n_1} = \Omega_{n_2}$. Let Ω be the set of letters appearing in $\sigma^{p_{n_2}-p_{n_1}}(\Omega_{n_1})$. It is easy to show that $\Omega = \Omega_{n_1} = \Omega_{n_2}$.
Consequently the set of letters appearing in $\sigma^{p_{n_2}-p_{n_1}}(\Omega)$ is equal to Ω and for all $k \in \mathbb{N}$ the set of letters appearing in $\sigma^{p_{n_1}+k(p_{n_2}-p_{n_1})}(A)$ is equal to Ω. We set $p = p_{n_1}$ and $g = p_{n_2} - p_{n_1}$. We remark that the letter c do not appear in the word $\phi(\sigma^{p+kg}(a))$ and that $[|\sigma^{p+kg}(u)|, |\sigma^{p+kg}(ua)|[\cap X = \emptyset$, for all $k \in \mathbb{N}$. Let $v = \sigma^p(u)$ and $w = \sigma^p(a)$, we have

$$[|\sigma^{kg}(v)|, |\sigma^{kg}(vw)|[\cap X = \emptyset. \tag{4}$$

We have $\lambda_\sigma(v) < \lambda_\sigma(vw)$. Consequently there exists an $\epsilon > 0$ such that

$$\lambda_\sigma(v)(1+\epsilon) < \lambda_\sigma(vw)(1-\epsilon).$$

From Lemma 21 we obtain that there exists k_0 such that for all $k \geq k_0$ we have

$$\frac{|\sigma^{kg}(v)|}{\alpha^{kg}} < \lambda_\sigma(v)(1+\epsilon) < \lambda_\sigma(vw)(1-\epsilon) < \frac{|\sigma^{kg}(vw)|}{\alpha^{kg}}. \tag{5}$$

From Lemma 22 applied to τ we have that there exists $s, t, t' \in B^*$ and $h \in \mathbb{N}^*$ such that for all $n \in \mathbb{N}$

$$\psi(\mathsf{y}_{[\tau^{hn}(s)\tau^{h(n-1)}(t)\cdots\tau^h(t)tt']}) = c.$$

From the second part of Lemma 22 it comes that there exists $\gamma \in \mathbb{R}$ such that

$$\lim_{n\to+\infty} \frac{|\tau^{hn}(s)\tau^{h(n-1)}(t)\cdots\tau^h(t)tt'|}{\beta^{hn}} = \gamma.$$

From Theorem 51 it comes that there exist two strictly increasing sequences of integers, $(m_i; i \in \mathbb{N})$ and $(n_i; i \in \mathbb{N})$, and $l \in \mathbb{R}$ such that

$$\gamma \frac{\beta^{m_i h}}{\alpha^{n_i g}} \to_{i\to+\infty} l \in]\lambda_\sigma(v)(1+\epsilon), \lambda_\sigma(vw)(1-\epsilon)[.$$

And consequently

$$\frac{|\tau^{hm_i}(s)\tau^{h(m_i-1)}(t)\cdots\tau^h(t)tt'|}{\alpha^{n_i g}} = \frac{|\tau^{hm_i}(s)\tau^{h(m_i-1)}(t)\cdots\tau^h(t)tt'|}{\gamma\beta^{m_i h}} \frac{\gamma\beta^{m_i h}}{\alpha^{n_i g}} \to_{i\to+\infty} l. \tag{6}$$

From (5) and (6) there exists $i \in \mathbb{N}$ such that

$$|\sigma^{n_i g}(v)| < |\tau^{hm_i}(s)\tau^{h(m_i-1)}(t)\cdots\tau^h(t)tt'| < |\sigma^{n_i g}(vw)|,$$

which means that $|\tau^{hm_i}(s)\tau^{h(m_i-1)}(t)\cdots\tau^h(t)tt'|$ belongs to X and is consequently in contradiction with (4). □

Corollary 53 *The words having infinitely many occurrences in* x *appear in* x *with bounded gaps.*

Proof. Let u be a word having infinitely many occurrences in x. We set $|u| = n$. To prove that u appears with bounded gaps in x it suffices to prove that the letter 1 appears with bounded gaps in the sequence $\mathsf{t} \in \{0,1\}^{\mathbb{N}}$ defined by

$$\mathsf{t}_i = 1 \text{ if } \mathsf{x}_{[i,i+n-1]} = u$$

and 0 otherwise. In the sequel we prove that t is α and β-substitutive.
The sequence $\mathsf{y}^{(n)} = ((\mathsf{y}_i \cdots \mathsf{y}_{i+n-1}); i \in \mathbb{N})$ is a fixed point of the substitution $\sigma_n : A_n \to A_n^*$ where A_n is the alphabet A^n, defined for all $(a_1 \cdots a_n)$ in A_n by

$$\sigma_n((a_1 \cdots a_n)) = (b_1 \cdots b_n)(b_2 \cdots b_{n+1}) \cdots (b_{|\sigma(a_1)|} \cdots b_{|\sigma(a_1)|+n-1})$$

where $\sigma(a_1 \cdots a_n) = b_1 \cdots b_k$ (for more details see Section V.4 in [Qu]).
Let $\rho : A_n \to A^*$ be the letter to letter morphism defined by $\rho((b_1 \cdots b_n)) = b_1$ for all $(b_1 \cdots b_n) \in A_n$. We have $\rho \circ \sigma_n = \sigma \circ \rho$, and then $M_\rho M_{\sigma_n} = M_\sigma M_\rho$. Consequently the dominant eigenvalue of σ_n is α and $\mathsf{y}^{(n)}$ is α-substitutive.
Let $f : A_n \to \{0,1\}$ be the letter to letter morphism defined by $f((b_1 \cdots b_n)) = 1$ if $b_1 \cdots b_n = u$ and 0 otherwise. It is easy to see that $f(\mathsf{y}^{(n)}) = \mathsf{t}$ hence t is α-substitutive.
In the same way we show that t is β-substitutive and Theorem 52 achieves the proof. □

Decomposition of a substitution into sub-substitutions. The following proposition is a consequence of the paragraph 4.4 and the Proposition 4.5.6 in [LM].

Proposition 54 *Let* $M = (m_{i,j})_{i,j \in A}$ *be a matrix with non-negative coefficients and no column equal to 0. There exists there positive integer* $p \neq 0$, q, l, *where* $q \leq l-1$, *and a partition* $\{A_i; 1 \leq i \leq l\}$ *of* A *such that*

$$M^p = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_q \\ A_{q+1} \\ A_{q+2} \\ \vdots \\ A_l \end{array} \begin{array}{c} \begin{array}{cccccccc} A_1 & A_2 & \cdots & A_q & A_{q+1} & A_{q+2} & \cdots & A_l \end{array} \\ \left(\begin{array}{cccccccc} M_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ M_{1,2} & M_2 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{1,q} & M_{2,q} & \cdots & M_q & 0 & 0 & \cdots & 0 \\ M_{1,q+1} & M_{2,q+1} & \cdots & M_{q,q+1} & M_{q+1} & 0 & \cdots & 0 \\ M_{1,q+2} & M_{2,q+2} & \cdots & M_{q,q+2} & 0 & M_{q+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ M_{1,l} & M_{2,l} & \cdots & M_{q,l} & 0 & 0 & \cdots & M_l \end{array} \right) \end{array}, \qquad (7)$$

where the matrices M_i, $1 \leq i \leq q$ *(resp.* $q+1 \leq i \leq l$*), are primitive or equal to zero (resp. primitive), and such that for all* $1 \leq i \leq q$ *there exists* $i+1 \leq j \leq l$ *such that the matrix* $M_{i,j}$ *is different from 0.*

In what follows we keep the notations of Proposition 54. We will say that $\{A_i; 1 \leq i \leq l\}$ is a *primitive component partition of* A *(with respect to* M*)*. If i belongs to $\{q+1, \cdots, l\}$ we will say that A_i is a *principal primitive component of* A *(with respect to* M*)*.

Let $\tau : A \to A^*$ be a substitution and $M = (m_{i,j})_{i,j\in A}$ its matrix. Let $i \in \{q+1, \cdots, l\}$. We denote τ_i the restriction $\tau^p_{/A_i} : A_i \to A^*$ of τ^p to A_i. Because $\tau_i(A_i)$ is included in A_i^* we can consider that τ_i is a morphism from A_i to A_i^* which matrix is M_i. Let $i \in \{1, \cdots, q\}$ such that M_i is not equal to 0. Let φ_i be the morphism from A to A_i^* defined by $\varphi(b) = b$ if b belongs to A_i and the empty word otherwise. Let consider the map $\tau_i : A_i \to A^*$ defined by $\tau_i(b) = \varphi_i(\tau^p(b))$ for all $a \in A_i$. We remark as previously that $\tau_i(A_i)$ is included in A_i^*, consequently τ_i defines a morphism from A_i to A_i^* which matrix is M_i.
We will say that the substitution $\tau : A \to A^*$ satisfy the condition **(C)** if:

C1. The matrix M, itself, is of the types (7) (i.e. $p = 1$) ;

C2. The matrices M_i are equal to 0 or with positive coefficients if $1 \leq i \leq q$ and with positive coefficients otherwise ;

C3. For all matrices M_i different from 0, with $i \in \{1, \cdots, l\}$, there exists $a_i \in A_i$ such that $\tau_i(a_i) = a_i u_i$ where u_i is a non-empty word of A^* if $M_i \neq [1]$ and empty otherwise.

From Proposition 54 every substitution $\tau : A \to A^*$ has a power τ^k satisfying condition **(C)**. The definition of substitutions implies that for all $q+1 \leq i \leq l$ we have $M_i \neq [1]$.
Let $\tau : A \to A^*$ be a substitution satisfying condition **(C)** (we keep the previous notations). For all $1 \leq i \leq l$ such that M_i is different from 0 and [1], the map $\tau_i : A_i \to A_i^*$ defines a substitution we will call *main sub-substitution of* τ if $i \in \{q+1, \cdots, l\}$ and *non-main sub-substitution of* τ otherwise. Moreover matrix M_i has positive coefficients which implies that the substitution τ_i is primitive. We remark that there exists at least one main sub-substitution.
In [Du3] the following results were obtained and will be used in the sequel.

Lemma 55 *Let $\sigma : A \to A^*$ et $\tau : B \to B^*$ be two substitutions satisfying condition* **(C)***, D an alphabet, $\varphi : A \to D^*$ and $\phi : B \to D^*$ two letter to letter morphisms such that $\varphi(L(\tau)) = \phi(L(\sigma))$. If $\overline{\sigma}$ is a main sub-substitution of σ then there exists a main sub-substitution $\overline{\tau}$ of τ such that $\varphi(L(\overline{\tau})) = \phi(L(\overline{\sigma}))$.*

Proof of Theorem 46. We take the notations of the first lines of the part 5.6.
Let $\overline{\sigma} : \overline{A} \to \overline{A}^*$ be a main sub-substitution of σ. From Lemma 55 there exists a main sub-substitution $\overline{\tau}$ of τ such that $\phi(L(\overline{\sigma})) = \psi(L(\overline{\tau})) = L$. From Theorem 31 it comes that L is periodic, i.e. there exists a word u such that $L = L(u^\omega)$ where $|u|$ is the least period. There exists an integer N such that all the words of length $|u|$ appear infinitely many times in $\mathsf{x}_N\mathsf{x}_{N+1}\cdots$. We set $\mathsf{t} = \mathsf{x}_N\mathsf{x}_{N+1}\cdots$ and we will prove that t is periodic and consequently x will be ultimately periodic.
The word u appear infinitely many times, consequently it appears with bounded gaps. Let $\mathcal{R}_u$ be the set of return words to u (a word w is a return word to u if $wu \in L(\mathsf{x})$, u is a prefix of wu and u has exactly two occurrences in wu). It is finite. There exists an integer N such that all the words $w \in \mathcal{R}_u \cap L(\mathsf{x}_N\mathsf{x}_{N+1}\cdots)$ appear infinitely many times in x. Hence these words appear with bounded gaps in x. We set $\mathsf{t} = \mathsf{x}_N\mathsf{x}_{N+1}\cdots$ and we will prove that t is periodic and consequently x will be ultimately periodic. We can suppose that u is a prefix of

t. Then t is a concatenation of return words to u. Let w be a return word to u. It appears with bounded gaps hence it appears in some $\phi(\overline{\sigma}^n(a))$ and there exits two words, p and q, and an integer i such that $wu = pu^i q$. As $|u|$ is the least period of L it comes that $wu = u^i$. It follows that $\mathsf{t} = u^\omega$. □

Proof of Theorem 1. It is a consequence of Theorem 46, Theorem 50 and of the Second Cobham's Theorem. □

References

[Bea] M.P. Béal, *Codage Symbolique*, Masson, Paris-Milan-Barcelone, 1993.

[Ber1] A. Bertrand-Mathis, *Développement en base θ, répartition modulo un de la suite $(x\theta^n)_{n\geq 0}$, langages codés et θ-shift*, Bull. Soc. Math. France 114 (1986), 271-323.

[Ber2] A. Bertrand-Mathis, *Comment écrire les nombres entiers dans une base qui n'est pas entière*, Acta Math. Acad. Sci. Hungar. 54 (1989), 237-241

[Bes] A. Bès, *An extension of Cobham-Semënov theorem*, preprint.

[BH1] V. Bruyère and G. Hansel, *Recognizable sets of numbers in nonstandard bases*, Lecture Notes in Comput. Sci. 911 (1995) 167-179.

[BH2] V. Bruyère and G. Hansel, *Bertrand numeration systems and recognizability*, à paraître dans Theo. Comp. Sci..

[BHMV] V. Bruyère, G. Hansel, C. Michaux and R. Villemaire, *Logic and p-recognizable sets of integers*, Bull. Belgian Math. Soc. Simon Stevin vol. 1 (1994) 191-238.

[BP] V. Bruyère and F. Point, *On the Cobham-Semënov theorem*, Theory of Computing Systems 30 (1997), 197-220.

[Bu] J. R. Büchi, *Weak second-order arithmetic and finite automata*, Z. Math. Logik Grundlag. Math. 6 (1960), 66-92.

[CKMR] G. Christol, T. Kamae, M. Mendès-France et G. Rauzy, *Suites Algébriques et Substitutions*, Bull. Soc. Math. France 10 8 (1980), 401-419.

[Co1] A. Cobham, *On the base-dependence of sets of numbers recognizable by finite automata*, Math. Syst. Theo. 3 (1969), 186-192.

[Co2] A. Cobham, *Uniform tag sequences*, Math. Syst. Theo. 6 (1972), 164-192.

[Du1] F. Durand, *A characterization of substitutive sequences using return words*, Disc. Math. 179 (1998), 89-101.

[Du2] F. Durand, *A generalization of Cobham's theorem*, Theory Comput. Syst. 31 (1998), 169-185.

[Du3] F. Durand, *Sur les ensembles d'entiers reconnaissables*, J. de Théorie des Nombres de Bordeaux 10 (1998), 65-84.

[Du4] F. Durand, *Linearly recurrent subshifts have a finite number of non-periodic subshift factors*, Ergod. Th. & Dynam. Sys. 20 (2000), 1061-1078.

[DHS] F. Durand, B. Host, C. Skau, *Substitutions, Bratteli diagrams and dimension groups*, to appear in Ergod. Th. & Dynam. Sys. 19 (1999), 953-993.

[Ei] S. Eilenberg, *Automata, Languages and Machines*, Academic Press vol. A.

[Fab1] S. Fabre, *Une généralisation du théorème de Cobham*, Acta Arithmetica LXVII.3 (1994), 197-208.

[Fab2] S. Fabre, *Substitutions et β-systèmes de numération*, Theo. Comp. Sc. 137 (1995), 219-236.

[Fag] I. Fagnot, *Sur les facteurs des mots automatiques*, Theoret. Comput. Sci. 172 (1997), 67-89.

[Fr] C. Frougny, *Number representation and finite automata*, Topics in Symbolic Dynamics and Applications, Cambridge University Press, London Mathematical Society, Lecture Notes Series 279 (2000), 207-228.

[Ha1] G. Hansel, *A propos d'un théorème de Cobham*, Actes de la fête des mots, D. Perrin Ed., GRECO de programmation, Rouen (1982).

[Ha2] G. Hansel, *Systèmes de numération indépendants et syndéticité*, Theoret. Comput. Sci. 204 (1998), 119-130.

[HW] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, 5th ed. (1979).

[HM] G. A. Hedlund and M. Morse, *Symbolic dynamics*, Am. J. Math. 60 (1938), 815-866.

[HZ1] C. Holton and L. Q. Zamboni, *Descendants of primitive substitutions*, Theory Comput. Syst. 32 (1999), 133-157.

[HZ2] C. Holton and L. Q. Zamboni, *Directed graphs and substitutions*, preprint.

[HJ] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press.

[Ho] B. Host, *Substitution subshifts and Bratteli diagrams*, Topics in Symbolic Dynamics and Applications, Cambridge University Press, London Mathematical Society, Lecture Notes Series 279 (2000), 35-55.

[LM] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press (1995).

[MV] C. Michaux and R. Villemaire, *Presburger arithmetic and recognizability of sets of natural numbers by automata: New proofs of Cobham's theorem and Semenov's theorem*, Annals of Pure and Applied Logic 77 (1996), 251-277.

[MS] F. Mignosi and P. Séébold, *If a D0L–language is k–power free then it is circular*, ICALP 1993, Lect. Notes in Comp. Sci. 700.

[Mo] B. Mossé, Puissances de mots et reconnaissabilité des points fixes d'une substitution, *Theo. Comp. Sci.* **99** (1992) 327–334.

[Pan] J.-J. Pansiot, *Complexité des facteurs des mots infinis engendrés par morphismes itérés*, Lect. Notes in Comp. Sci. 172 (1984), 380-389.

[Par] W. Parry, *On the β-expansions of real numbers*, Acta Math. Acad. Sci. Hungar 11 (1960), 401-416.

[Qu] M. Queffélec, *Substitution Dynamical Systems-Spectral Analysis*, Lecture Notes in Math. vol.1294 (1987).

[SS] A. Salomaa and M. Soittola, *Automata-theoretic aspects of formal power series*, Texts and monographs in computer science, Springer-Verlag (1978).

[Se] A. L. Semenov, *The Presburger nature of predicates that are regular in two number systems*, Siberian Math. J. 18 (1977), 289-299.

[Sh] J. Shallit, *Numeration systems, linear recurrences and regular sets*, Theo. Comp. Sci. 61 (1988), 1-16.