COBHAM'S THEOREM FOR ABSTRACT NUMERATION SYSTEMS

ÉMILIE CHARLIER, JULIEN LEROY, AND MICHEL RIGO

Abstract. Abstract numeration systems generalize numeration systems whose representation map is increasing and such that the language of all the representations is regular. We show that if a set of integers is recognized by some finite automaton within two independent abstract numeration systems, then this set is a finite union of arithmetic progressions.

1. Introduction

Finite automata make up the simplest model of computation in terms of the Chomsky hierarchy. With such a computational perspective, one can consider sets of integers whose base k expansions form a language accepted by some finite automaton. These sets are said to be k-*recognizable*. In 1969, Cobham obtained a fundamental result initiating the systematic study of k-recognizable sets and showing that k-recognizability depends on the chosen base [Cob69]. Let $k, \ell \geq 2$ be two multiplicatively independent integers, *i.e.*, $\log k / \log \ell$ is irrational. If a subset X of $\mathbb{N}$ is simultaneously k-recognizable and ℓ-recognizable, then it is ultimately periodic, *i.e.*, there exists N, p such that for all $n \geq N$, $n \in X$ if and only if $n + p \in X$. See for instance [BHMV94] for a survey on k-recognizable sets.

This result is also motivating the introduction of non-standard numeration systems: It is meaningful to consider other numeration systems to handle new sets of integers recognizable by finite automata. The bibliography in [Dur11] provides many pointers to various extensions of Cobham's theorem. One can for instance replace the sequence $(k^n)_{n\geq 0}$ with an increasing sequence of integers satisfying a linear recurrence relation. As an example, one can consider the Zeckendorf numeration system and greedy representations of the integers written as sums of non-consecutive Fibonacci numbers [**?**]. Recognizability is extended to this context of non-standard numeration systems [Lot02, Chap. 7]. For instance, numeration systems based on linear recurrences whose characteristic polynomial is the minimal polynomial of a Pisot number carry the main properties of the integer base systems [BH97].

When dealing with non-standard systems, a minimal natural requirement is to impose that the language of representations of all the integers is recognized by some finite automaton. Given a word, one can decide in linear time whether or not it is a valid representation. This requirement is formalized as follows. An *abstract numeration system* S is defined by an infinite regular language L over a totally ordered alphabet. In this setting, the integer $n \geq 0$ is uniquely represented by the word of rank n in the radix ordered language L with respect to the ordering of the alphabet. A set X of integers is said to be S-*recognizable*, if the representations within this system S of the elements belonging to X constitute a language recognized by some finite automaton. See [BR10, Chap. 3]. In particular, linear systems built upon a Pisot number have a regular language of numeration and are therefore special cases of abstract numeration systems.

Let S be an abstract numeration system. It is well known that any ultimately periodic set is S-recognizable: the periodic decimation of a radix ordered language yields a regular language

[**?**]. Taking into account this positive result, in the context of the theorem of Cobham, it is therefore natural to formulate the following questions.

- What is the analogue of multiplicatively independent bases: how can we say that two abstract numeration systems S and T are *independent*?
- Let S and T be two independent numeration systems. If a subset of $\mathbb{N}$ is simultaneously S-recognizable and T-recognizable, is this set necessarily ultimately periodic?

In this paper, we answer these questions concluding a research initiated 15 years ago. See for instance [**?**].

Back to integer base systems, k-recognizable sets can be characterized in terms of morphic sequences. In 1972, Cobham proved that a set $X \subset \mathbb{N}$ is k-recognizable if and only if its characteristic sequence $\mathbf{1}_X$ is k-automatic, *i.e.*, $\mathbf{1}_X$ is the image under a coding of the fixed point of a morphism of constant length k [**?**]. Hence the first theorem of Cobham can be restated as follows. Let $k, \ell \geq 2$ be two multiplicatively independent integers. If an infinite sequence $\mathbf{w}$ is both k-automatic and ℓ-automatic, then this sequence is ultimately periodic: $\mathbf{w} = uv^\omega$ for some finite words u, v.

Therefore another point of view to generalize numeration systems, is to consider a larger class of infinite sequences: an infinite sequence is *morphic* if it is the image under a coding of the fixed point of a morphism. In that setting, a series of papers has led Durand to a generalization of the theorem of Cobham [**?**, **?**, **?**, Dur11]. Roughly, if an infinite sequence can be generated by two "independent" morphisms (and some extra codings), then this sequence must be ultimately periodic. The precise statement relies on the Perron eigenvalue of matrices associated with morphisms and is recalled in Section 3.3.

In this paper, we make use of the fact that morphic sequences are linked to abstract numeration systems. Namely, the theorem of Cobham from 1972 can be extended to S-recognizable sets [RM02] as follows. Let X be a set of integers. There exists an abstract numeration system S such that X is S-recognizable if and only if $\mathbf{1}_X$ is a morphic sequence. The proof of this result introduces erasing morphisms. If X is S-recognizable, one can show that $\mathbf{1}_X$ is the image under a (possibly erasing) morphism of the fixed point of a (possibly erasing) morphism.

In this paper, we gather all the necessary tools to deal with these erasing morphisms and keep track of the corresponding matrices and their spectral properties. Indeed, many authors have considering the problem of getting rid of erasing morphisms when dealing with morphic sequences [Cob68, Pan83, AS03, **?**]. It is well known that any morphic sequence of the kind $g(f^\omega(a))$ can be obtained as $\tau(\sigma^\omega(b))$ for some non-erasing morphism σ and coding, or letter-to-letter morphism, τ. Our task is to relate carefully the properties of the matrix associated with f to the one associated with σ before taking advantage of the results of Durand. Moreover, most of Durand's results apply for morphisms with an exponential growth rate. The case of morphisms with dominating eigenvalue equal to 1 is partially covered in [DR09].

This paper is organized as follows. In Section 2, we make use of the theorem of Perron–Frobenius and we discuss properties of non-negative matrices. In particular, we introduce the notion of dilated matrix and show that a non-negative matrix and any of its dilated version have the same dominating eigenvalue. Section 3 contains the main discussion about erasing morphisms. We explain how to get rid off these morphisms and relate the spectrum of the new non-erasing morphism with the former erasing ones. In Section 4, we review the notion of abstract numeration system. In particular, we make precise the link between abstract numeration systems and morphic sequences. Finally, in the last section, we are able to state an analogue of Cobham's theorem for independent abstract numeration systems.

2. Operations on matrices

When dealing with morphisms and automata, it is natural to associate a matrix with such a morphism and automaton. In this section, we introduce dilatation of matrices and we recall some results about non-negative matrices and their dominating eigenvalues. This notion of dilatation provides information on the transformations we apply to the morphisms and automata. However it is not crucial for the results we obtain. This section thus could be skipped at a first reading.

Roughly speaking, when dilating a matrix M, each element $M_{i,j}$ is replaced in a convenient way by a matrix of size $k_i \times k_j$ whose lines all sum up to $M_{i,j}$.

Definition 1. Let M be a real square matrix of size m. A real square matrix D of size $n \geq m$ is called a *dilated matrix of* M if there exist positive integers $k_1, \ldots, k_m$ such that

(1) $\sum_{i=1}^{m} k_i = n$;
(2) rows and columns are both indexed by pairs (i,k) for $1 \leq i \leq m$ and $1 \leq k \leq k_i$;
(3) D satisfies the following property:

$$\forall i,j \in \{1,\ldots,m\},\ \forall k \in \{1,\ldots,k_i\}, \qquad \sum_{\ell=1}^{k_j} D_{(i,k),(j,\ell)} = M_{i,j}. \tag{1}$$

The vector $(k_1, k_2, \ldots, k_m)$ is called the *dilatation vector of* D. We let $\mathrm{Dil}(M)$ denote the set of dilated matrices of M.

In other words, given a square matrix M of size m, a dilated matrix with dilatation vector $(k_1, \ldots, k_m)$ of M is a block matrix

$$D = \begin{pmatrix} B_{1,1} & \cdots & B_{1,m} \\ \vdots & \ddots & \vdots \\ B_{m,1} & \cdots & B_{m,m} \end{pmatrix}$$

where each block $B_{i,j}$ has k_i rows and k_j columns and such that for all $k \in \{1,\ldots,k_i\}$, one has

$$\sum_{\ell=1}^{k_j} (B_{i,j})_{k,\ell} = M_{i,j}.$$

Definition 1 can be adapted to column vectors instead of matrices. The idea is to repeat several times a given entry to be compatible with the multiplication of a matrix with a column vector.

Definition 2. Let $\vec{x} \in \mathbb{R}^m$ be a vector. A vector $\vec{d} \in \mathbb{R}^n$, $n \geq m$, is a *dilated vector of* $\vec{x}$ if there exist positive integers $k_1, \ldots, k_m$ such that $\sum_{i=1}^{m} k_i = n$. Elements of $\vec{d}$ are indexed by pairs (i,k) for $1 \leq i \leq m$ and $1 \leq k \leq k_i$. For all $i \in \{1,\ldots,m\}$ and all $k \in \{1,\ldots,k_i\}$, we set $d_{(i,k)} = x_i$.

Example 3. Consider the following matrix M and vector $\vec{x}$

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \vec{x} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

The matrix D and the vector $\vec{d}$ below are respectively dilated matrix of M and dilated vector of $\vec{x}$ with dilatation vector $(1,2,2)$.

$$D = \left(\begin{array}{c|cc|cc} 1 & 1 & 0 & 0 & 1 \\ \hline 2 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1/2 & 1/2 \\ \hline 1 & \sqrt{2} & 1-\sqrt{2} & 1 & -1 \\ 1 & 0 & 1 & 0 & 0 \end{array}\right) \quad \text{and} \quad \vec{y} = \left(\begin{array}{c} 1 \\ \hline 0 \\ 0 \\ \hline 2 \\ 2 \end{array}\right).$$

Definition 4. Let M be a square matrix. The *spectrum of* M is the multiset of its eigenvalues (repeated with respect to their algebraic multiplicities). It is denoted by $\mathrm{Spec}(M)$. The *spectral radius of* M is the real number $\rho(M) = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(M)\}$.

Lemma 5. *Let M be a real square matrix of size m and let D be a dilated matrix of M. Each eigenvalue of M is also an eigenvalue of D.*

Proof. Assume that D is a dilated matrix of M with dilatation vector $(k_1,\ldots,k_m)$. Let λ be an eigenvalue of M and let $\vec{x}$ be an eigenvector of M such that $M\vec{x} = \lambda\vec{x}$. Let $\vec{y}$ be a dilated vector of $\vec{x}$ with dilatation vector $(k_1,\ldots,k_m)$. The vector $\vec{y}$ is non-zero and for all i in $\{1,\ldots,m\}$ and all $k \in \{1,\ldots,k_i\}$, we have

$$(D\vec{y})_{(i,k)} = \sum_{j=1}^{m}\sum_{\ell=1}^{k_j} D_{(i,k),(j,\ell)}\, y_{(j,\ell)} = \sum_{j=1}^{m}\left(\sum_{\ell=1}^{k_j} D_{(i,k),(j,\ell)}\right) x_j = \sum_{j=1}^{m} M_{i,j}x_j = \lambda\, x_i = \lambda\, y_{(i,k)}.$$

Hence, λ is an eigenvalue of D. □

Definition 6. A non-negative square matrix M of size m is said to be *irreducible* if for all $i,j \in \{1,\ldots,m\}$, there exists a positive integer $k = k(i,j)$ such that $(M^k)_{i,j} > 0$. Otherwise, it is said to be *reducible.*

Theorem 7 (Perron–Frobenius Theorem [**?**])**.** *Let M be a non-negative square matrix. If M is irreducible, then the real number $\rho(M)$ is an eigenvalue of M which is algebraically simple.*

Remark 8. Let M be a non-negative square matrix. If M is reducible, then there exists a permutation matrix P such that $P^{-1}MP$ is a lower block triangular matrix where each diagonal block is either irreducible or zero [LM95, p. 119]. Furthermore, by taking a convenient power of M, $P^{-1}M^nP$ is a lower block triangular matrix where each diagonal block is either primitive or zero. See [LM95, Section 4.5] for details about the cyclic structure of irreducible matrices.

As a consequence of Perron–Frobenius Theorem, we get the following theorem.

Theorem 9. *If M is a non-negative square matrix, then $\rho(M)$ is an eigenvalue of M.*

Let M be a non-negative square matrix. Note that if $\alpha \in \mathbb{C}$ is an eigenvalue of M, then $|\alpha| \le \rho(M)$. This is the reason why the eigenvalue $\rho(M)$ is also called the *dominating eigenvalue of M.*

Proposition 10. *Let M be a non-negative square matrix. For any non-negative matrix D in* $\mathrm{Dil}(M)$*, M and D have the same dominating eigenvalue.*

Proof. We follow the lines of the proof of [NR07, Proposition 7]. Due to Lemma 5 and Theorem 9, we have $\rho(D) \ge \rho(M)$. Let us prove that we also have $\rho(D) \le \rho(M)$.

The Collatz–Wielandt formula (see for instance [**?**, Chap. 8]) states that, for any irreducible matrix N of size m,

$$\rho(N) = \max_{\substack{\vec{y}\in\mathbb{R}^m\\ \vec{y}\geq 0}} \min_{\substack{1\leq i\leq m\\ y_i\neq 0}} \frac{(N\vec{y})_i}{y_i}.$$

Let m (resp. n) be the size of M (resp. D). Let us first suppose that M and D are irreducible. Let us prove that for all non-negative vectors $\vec{y}\in\mathbb{R}^n$ there is a non-negative vector $\vec{x}\in\mathbb{R}^m$ such that

$$\min_{\substack{1\leq i\leq n\\ y_i\neq 0}} \frac{(D\vec{y})_i}{y_i} \leq \min_{\substack{1\leq i\leq m\\ x_i\neq 0}} \frac{(M\vec{x})_i}{x_i}.$$

Let $\vec{y}$ be a non-negative vector in $\mathbb{R}^n$ and let $(k_1,\ldots,k_m)$ be the dilatation vector of D. With the convention taken in Definition 2, we index the components of $\vec{y}$ by the ordered pairs (i,k) for $1\leq i\leq m$ and $1\leq k\leq k_i$. Let us define the non-negative vector $\vec{x}\in\mathbb{R}^m$ by

$$x_i = \max_{1\leq k\leq k_i} y_{(i,k)}.$$

We have

$$\begin{aligned}
\min_{\substack{1\leq i\leq m\\ 1\leq k\leq k_i\\ y_{(i,k)}\neq 0}} \frac{(D\vec{y})_{(i,k)}}{y_{(i,k)}} &= \min_{\substack{1\leq i\leq m\\ 1\leq k\leq k_i\\ y_{(i,k)}\neq 0}} \frac{1}{y_{(i,k)}} \sum_{j=1}^{m}\sum_{\ell=1}^{k_j} D_{(i,k),(j,\ell)} y_{(j,\ell)}\\
&\leq \min_{\substack{1\leq i\leq m\\ 1\leq k\leq k_i\\ y_{(i,k)}\neq 0}} \frac{1}{y_{(i,k)}} \sum_{j=1}^{m}\left(\sum_{\ell=1}^{k_j} D_{(i,k),(j,\ell)}\right) x_j\\
&\leq \min_{\substack{1\leq i\leq m\\ 1\leq k\leq k_i\\ y_{(i,k)}\neq 0}} \frac{1}{y_{(i,k)}} \sum_{j=1}^{m} M_{i,j} x_j\\
&= \min_{\substack{1\leq i\leq m\\ x_i\neq 0}} \frac{1}{x_i} \sum_{j=1}^{m} M_{i,j} x_j.
\end{aligned}$$

This concludes the case of irreducible matrices.

Now suppose that M or D is reducible. Let J be the $n\times n$ matrix whose entries are all equal to 1. Let C be the $m\times m$ matrix defined by $C_{i,j}=k_j$ for all i,j. We can consider sequences of matrices $(M_s)_{s\geq 1}$ and $(D_s)_{s\geq 1}$ where $M_s = M+\frac{1}{s}C$ (resp. $D_s = D+\frac{1}{s}J$). Note that M_s and D_s are positive matrices, hence irreducible. Moreover, J is a dilated matrix of C with dilatation vector $(k_1,\ldots,k_m)$. Hence the same holds for D_s and M_s. We can therefore apply the same reasoning as in the first part of the proof and obtain $\rho(D_s)\leq\rho(M_s)$ for all $s\geq 1$. Since $\lim_{s\to+\infty}\rho(M_s)=\rho(M)$ and $\lim_{s\to+\infty}\rho(D_s)=\rho(D)$, we conclude that $\rho(D)\leq\rho(M)$. □

In what follows we are mainly concerned with matrices in $\mathbb{N}^{m\times m}$. The next two lemmas are classical results of linear algebra.

Lemma 11. *Let M be a square matrix in $\mathbb{N}^{m\times m}$. There is a positive integer k such that for all $i,j\in\{1,\ldots,m\}$, there exist $\lambda(i,j)\in\operatorname{Spec}(M)$ and $d(i,j)\in\mathbb{N}$ such that $(M^{kn})_{i,j} =$*

$\Theta((\lambda(i,j))^n n^{d(i,j)})$. In particular, if M is irreducible, then $\lambda(i,j) = \rho(M)$ for all $i,j \in \{1,\dots,m\}$.

Lemma 12. *Let M be a lower block triangular square matrix in $\mathbb{N}^{m\times m}$*

$$M = \begin{pmatrix} D_1 & 0 & \cdots & 0 \\ \star & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \star & \cdots & \star & D_k \end{pmatrix}$$

where for all $h \in \{1,\dots,k\}$ the diagonal block D_h has size $m_h \geq 1$ and is either irreducible or zero. Let k be a positive integer such that for all $i,j \in \{1,\dots,m\}$ and all $n \in \mathbb{N}_{\geq 1}$, $M_{i,j}^{kn} \neq 0$ if and only if $M_{i,j}^{k(n+1)} \neq 0$. Then, with the notation of the previous lemma, for all $h,h' \in \{1,\dots,k\}$, $h' \leq h$, all $i \in \{1+\sum_{g<h} m_g, \sum_{g\leq h} m_g\}$ and all $j \in \{1+\sum_{g'<h'} m_{g'}, \sum_{g'\leq h'} m_{g'}\}$, we have

$$\begin{aligned}\lambda(i,j) = \max\Bigg\{\rho(D_f) \mid (1 \leq f \leq k) &\wedge \left(\exists i' \in \{1+\sum_{e<f} m_e, \sum_{e\leq f} m_e\}, j' \in \{1+\sum_{g'<h'} m_{g'}, \sum_{g'\leq h'} m_{g'}\} : M_{i',j'}^k \geq 1\right) \\ &\wedge \left(\exists i'' \in \{1+\sum_{g<h} m_g, \sum_{g\leq h} m_g\}, j'' \in \{1+\sum_{e<f} m_e, \sum_{e\leq f} m_e\} : M_{i'',j''}^k \geq 1\right)\Bigg\}\end{aligned}$$

3. Morphic sequences

We recall classical definitions on sequences that can be obtained as the image under a morphism g of the infinite sequence generated by applying iteratively another morphism f on an initial letter a. It is well known that such a sequence $g(f^\omega(a))$ can also be obtained with a coding τ and a non-erasing morphism σ, *i.e.*, $g(f^\omega(a)) = \tau(\sigma^\omega(b))$. We discuss this result to relate precisely the eigenvalues associated with f and σ. Finally, we present the analogue of the theorem of Cobham from 1969 adapted to such sequences.

3.1. **Basic definitions.** Let A be an alphabet. The set of finite words over A is denoted by A^*. Endowed with the concatenation product, this set is a monoid whose neutral element is the empty word ε. The length of a word $w \in A^*$ is denoted by $|w|$ and the number of occurrences of the letter a in w is denoted by $|w|_a$. We set $A^+ = A^* \setminus \{\varepsilon\}$. A morphism $f : A^* \to B^*$ is a *coding*, if for all $a \in A$, $|f(a)| = 1$. It is said to be *non-erasing*, if for all $a \in A$, $|f(a)| \geq 1$. Moreover, morphisms defined on A^* can naturally be extended to morphisms defined over $A^{\mathbb{N}}$.

Definition 13. Let A be an alphabet and $f : A^* \to A^*$ be a morphism. We call a letter $a \in A$ *mortal (w.r.t. f)* if there is a positive integer n such that $f^n(a) = \varepsilon$. A non-mortal letter is called *immortal (w.r.t. f)*. We let $A_{\mathcal{M},f}$ (or simply $A_{\mathcal{M}}$) denote the set of mortal letters and $A_{\mathcal{I},f}$ (or simply $A_{\mathcal{I}}$) the set of immortal letters. Let B be a subset of the alphabet A. We let $\kappa_{A,B} : A^* \to A^*$ denote the morphism defined by $\kappa_{A,B}(a) = \varepsilon$ if $a \in B$ and $\kappa_{\mathcal{M},f}(a) = a$ otherwise.

Definition 14. Let $f : A^* \to A^*$ be a morphism and let $B \subset A$ be a sub-alphabet. If $f(B) \subset B^*$, we say that $f_B := f|_{B^*} : B^* \to B^*$ is a *sub-morphism* of f.

Definition 15. Let $f : A^* \to A^*$ be a morphism. The *incidence matrix* of f is the matrix M_f defined, for all $a, b \in A$, by

$$(M_f)_{a,b} = |\sigma(b)|_a.$$

For all sub-alphabets $B \subset A$, we let

$$(M_f)_B$$

denote the sub-matrix of M_f obtained from M_f by taking rows and columns corresponding to letters in B. A morphism $f : A^* \to A^*$ is said to be *irreducible* if its incidence matrix M_f is irreducible. The eigenvalues and the spectrum of M_f are called respectively the eigenvalues and the spectrum of f. In particular, since M_f is non-negative, thanks to Theorem 9 we can also talk about the dominating eigenvalue of f.

The next result is a direct consequence of Lemma 11 and of the fact that for any morphism f and for all $n \in \mathbb{N}$, we have $M_{f^n} = M_f^n$.

Dans le lemme 11, on prend une puissance de la matrice. Ici, j'ai l'impression que ce n'est pas necessaire, si?

Lemma 16. *Let $f : A^* \to A^*$ be a morphism. For all $a \in A$, there exists $d(a) \in \mathbb{N}$ and $\lambda(a) \in \mathrm{Spec} f$ such that $|f^n(a)| \in \Theta(n^{d(a)}\lambda(a)^n)$.*

Definition 17. A morphism $f : A^* \to A^*$ is *prolongable* on a letter $a \in A$ if $f(a) = au$ for some $u \in A^*$ and $\lim_{n\to+\infty} |f^n(a)| = +\infty$. Convergence of a sequence of finite words to an infinite sequence is classical, see for instance [BR10]. A sequence $\mathbf{w}$ over A is said to be *pure morphic* if there is a morphism $f : A^* \to A^*$ prolongable on a such that $\mathbf{w} = f^\omega(a) := \lim_{n\to+\infty} f^n(a)$. Moreover, if all letters of A occur in $\mathbf{w}$ and λ is the dominating eigenvalue of f, then, with the notation of Lemma 16, we deduce from Lemma 12 that $\lambda(a) = \lambda$. Furthermore, as f is prolongable, we have either $\lambda > 1$, or $\lambda = 1$ and $d := d(a) \geq 1$. The pair (λ, d) is called the *growth type* of f (w.r.t. a). In this case we say that $\mathbf{w}$ is (λ, d)*-pure morphic*. A sequence is *morphic* (resp. (λ, d)*-morphic*) if it is a morphic image of a pure (resp. (λ, d)-pure) morphic sequence. When $\lambda > 1$, we simply talk about λ-(pure) morphic sequences. A morphism whose dominating eigenvalue λ is greater than 1 is said to be *exponential*. Otherwise, it is *polynomial (of degree d)* where $(1, d)$ is the growth type of f.

Remark 18. Observe that if $\mathbf{w} = f^\omega(a)$ is a pure morphic sequence, then $\mathbf{w}$ is a fixed point of f (*i.e.*, $f(\mathbf{w}) = \mathbf{w}$) and the letter a is not mortal.

Remark 19. As in [Dur11], we impose in the definition of pure morphic sequence that all letters of the alphabet of the morphism occur in $\mathbf{w}$. This is required to have well-defined (λ, d)-pure morphic sequences. Indeed, consider the morphism $f : \{0, 1, 2\}^* \to \{0, 1, 2\}^*$ defined by $f(0) = 0001$, $f(1) = 12$ and $f(2) = 21$. The dominating eigenvalue of f is 3, but we do not want to say that $f^\omega(1)$ is 3-pure morphic. With the definition we consider, it is 2-pure morphic.

Lemma 20. *Let $f : B^* \to B^*$ and $g : B^* \to A^*$ be morphisms. There is a positive constant C such that for all $b \in B$ and all $n \in \mathbb{N}$, $|g(f^n(b))| \leq C|f^n(b)|$. Furthermore, if g is non-erasing then for all $b \in B$, $|g(f^n(b))| = \Theta(|f^n(b)|)$.*

3.2. **Avoiding erasing morphisms.**

Theorem 21. [Cob68] *Let* $\mathbf{w} = g(f^\omega(a))$ *be a morphic sequence over an alphabet* A*. There exists a non-erasing morphism* $\sigma : B^* \to B^*$*, a letter* $b \in B$ *and a coding* $\tau : B \to A$ *such that* $\mathbf{w} = \tau(\sigma^\omega(b))$*.*

Proofs of this result can be found in [Pan83, AS03, **?**]. Cassaigne and Nicolas [CN03] gave a constructive proof of this result in which two steps are needed. First, one shows that the morphisms f and g can be chosen to be non-erasing. This step is omitted in [CN03]. The second step builds the morphisms σ and τ. Since our aim is to compare the growth type of σ with the growth type of f, we will recall the algorithm.

Lemma 22. *Let* $\mathbf{w} = f^\omega(a)$ *be a* (λ, d)*-pure morphic sequence over an alphabet* A*. Let* k *be the number of mortal letters of* f*. If* k *is not zero, then the morphism* $f_\mathcal{I} := (\kappa_{A,A_\mathcal{M}} \circ f)|_{A_\mathcal{I}^*} : A_\mathcal{I}^* \to A_\mathcal{I}^*$ *is non-erasing and such that* $\mathbf{w} = f^k(f_\mathcal{I}^\omega(a))$*. Furthermore,* $f_\mathcal{I}$ *also has growth type* (λ, d) *(w.r.t.* a*),* $|f^k(f_\mathcal{I}^n(a))| = \Theta(\lambda^n n^d)$ *and we have* $M_{f_\mathcal{I}} = (M_f)_{A_\mathcal{I}}$ *and* $\mathrm{Spec}(f_\mathcal{I}) = \mathrm{Spec}(f) \setminus \{\underbrace{0, 0, \ldots, 0}_{k \text{ times}}\}$*.*

Proof. Since k is the number of mortal letters, it follows that $f^k(b) = \varepsilon$ for all $b \in A_\mathcal{M}$. Indeed, proceed by contradiction and suppose that $f^k(b) \neq \varepsilon$ for some $b \in A_\mathcal{M}$. Then $b, f(b), \ldots, f^k(b)$ are words over $A_\mathcal{M}$ and for each i, $f^{i+1}(b)$ must contain a letter not occurring in $b, \ldots, f^i(b)$. Hence the number of mortal letters would be greater than k.

We set $\kappa_\mathcal{M} = \kappa_{A,A_\mathcal{M}}$. Observe that $\mathbf{w} = f^k(\mathbf{w})$. Then we also have $\mathbf{w} = f^k \circ \kappa_\mathcal{M}(\mathbf{w})$.

It remains to prove that $\kappa_\mathcal{M}(\mathbf{w}) = f_\mathcal{I}^\omega(a)$. First, we show by induction on $\ell \geq 1$ that

$$(\kappa_\mathcal{M} \circ f)^\ell = \kappa_\mathcal{M} \circ f^\ell. \tag{2}$$

The result is obvious for $\ell = 1$. We get

$$(\kappa_\mathcal{M} \circ f)^{\ell+1} = \kappa_\mathcal{M} \circ f \circ (\kappa_\mathcal{M} \circ f)^\ell = \kappa_\mathcal{M} \circ f \circ \kappa_\mathcal{M} \circ f^\ell$$

where we used the induction hypothesis for the last equality. To conclude with the induction step, observe that $\kappa_\mathcal{M} \circ f \circ \kappa_\mathcal{M} = \kappa_\mathcal{M} \circ f$. It is a consequence of the fact that, for all $b \in A_\mathcal{M}$, $f(b) \in A_\mathcal{M}^*$.

On the one hand, $\kappa_\mathcal{M} \circ f^\ell(a)$ tends to $\kappa_\mathcal{M}(\mathbf{w})$ as $\ell \to +\infty$. On the other hand, thanks to (2), for all $\ell \geq 1$, $\kappa_\mathcal{M} \circ f^\ell(a) = (\kappa_\mathcal{M} \circ f)^\ell(a) = f_\mathcal{I}^\ell(a)$ which tends to $f_\mathcal{I}^\omega(a)$ as $\ell \to +\infty$. By uniqueness of the limit, it follows that $\kappa_\mathcal{M}(\mathbf{w}) = f_\mathcal{I}^\omega(a)$.

Note that, for all $b \in A_\mathcal{I}$, $f(b)$ contains at least a symbol in $A_\mathcal{I}$. Hence the morphism $f_\mathcal{I}$ is non-erasing. In particular, for all $n \geq k$, we get $|f^n(a)| = |f^k \circ f^{n-k}(a)| = |\kappa_\mathcal{M} \circ f^k \circ f^{n-k}(a)| = |\kappa_\mathcal{M} \circ f^k \circ \kappa_\mathcal{M} \circ f^{n-k}(a)| = |(\kappa_\mathcal{M} \circ f)^k \circ (\kappa_\mathcal{M} \circ f)^{n-k}(a)| = |f_\mathcal{I}^n(a)|$ so $f_\mathcal{I}$ has the same growth type as f (w.r.t. a).

To conclude with the proof, up to a permutation (corresponding to a reordering of the alphabet where appear first all the immortal letters), the matrix M_f can be written as

$$\begin{pmatrix} (M_f)_{A_\mathcal{I}} & 0 \\ \star & (M_f)_{A_\mathcal{M}} \end{pmatrix}.$$

Hence $M_{f_\mathcal{I}} = (M_f)_{A_\mathcal{I}}$. From the above discussion, $((M_f)_{A_\mathcal{M}})^k = 0$ then $\mathrm{Spec}(f_{A_\mathcal{M}}) = \mathrm{Spec}((M_f)_{A_\mathcal{M}}) = \{\underbrace{0, 0, \ldots, 0}_{k \text{ times}}\}$. Similarly, we have $\mathrm{Spec}(f_\mathcal{I}) = \mathrm{Spec}(f) \setminus \mathrm{Spec}(f_{A_\mathcal{M}})$. $\square$

The idea of the next statement is to remove the largest sub-morphism of f whose alphabet is erased by g.

Lemma 23. *Let $\mathbf{w} = g(f^\omega(a))$ be a morphic sequence with $g : B^* \to A^*$ and $f : B^* \to B^*$ a non-erasing morphism. Let C be a sub-alphabet of $\{b \in B \mid g(b) = \varepsilon\}$ such that f_C is a sub-morphism of f. The morphisms $f_\varepsilon := (\kappa_{B,C} \circ f)|_{(B\setminus C)^*}$ and $g_\varepsilon := g_{|(B\setminus C)^*}$ are such that $\mathbf{w} = g_\varepsilon(f_\varepsilon^\omega(a))$. In particular, for all $n \geq 1$, we have $g_\varepsilon \circ f_\varepsilon^n = (g \circ f^n)|_{(B\setminus C)^*}$. Furthermore, we have $M_{f_\varepsilon} = (M_f)_{B\setminus C}$ and $\operatorname{Spec}(f_\varepsilon) = \operatorname{Spec}(f) \setminus \operatorname{Spec}(f_C)$. Finally, if f has growth type (λ, d) (w.r.t. a) and if $|g(f^n(a))| = \Theta(\lambda^n n^{d'})$ for $d' \leq d$, then f_ε has growth type (λ, e) (w.r.t. a) for some e such that $d' \leq e \leq d$.*

Proof. It is easily seen that $\operatorname{Spec}(f_\varepsilon) = \operatorname{Spec}((M_f)_{B\setminus C}) = \operatorname{Spec}(f) \setminus \operatorname{Spec}(f_C)$. Indeed, from the construction of f_ε, we have (up to a permutation matrix)

$$M_f = \begin{pmatrix} M_{f_\varepsilon} & 0 \\ \star & M_{f_C} \end{pmatrix}.$$

Let us prove that $\mathbf{w} = g_\varepsilon(f_\varepsilon^\omega(a))$. We have $g = g_\varepsilon \circ \kappa_{B,C}$. Since, for all $b \in C$, $f(b) \in C^*$, we can use exactly the same reasoning as in (2) and get, for all $n \geq 1$,

$$\kappa_{B,C} \circ f^n = \kappa_{B,C} \circ f^n \circ \kappa_{B,C} = (\kappa_{B,C} \circ f)^n.$$

Hence, we get

$$\begin{aligned} \mathbf{w} &= g(f^\omega(a)) \\ &= g_\varepsilon \circ \kappa_{B,C} \circ f^n(f^\omega(a)) \\ &= g_\varepsilon \circ \kappa_{B,C} \circ f^n \circ \kappa_{B,C}(f^\omega(a)) \\ &= g_\varepsilon \circ (\kappa_{B,C} \circ f)^n \circ \kappa_{B,C}(f^\omega(a)) \\ &= g_\varepsilon \circ f_\varepsilon^n \circ \kappa_{B,C}(f^\omega(a)) \end{aligned}$$

We have $a \notin C$ and $f_\varepsilon(a) \in a(B \setminus C)^+$, otherwise $\mathbf{w}$ would be finite. Thus, f_ε is prolongable on a and

$$\kappa_{B,C}(f^\omega(a)) = f_\varepsilon^\omega(a).$$

In particular for all $n \geq 1$and all $b \in B \setminus C$, we have

$$g \circ f^n(b) = g_\varepsilon \circ \kappa_{B,C} \circ f^n(b) = g_\varepsilon \circ (\kappa_{B,C} \circ f)^n(b) = g_\varepsilon \circ f_\varepsilon(b).$$

To conclude with the proof, we get by Lemma 16 that $|f_\varepsilon^n(a)| = \Theta(\lambda(a) n^{d(a)})$ for some $\lambda(a) \in \operatorname{Spec}(f_\varepsilon)$ and $d(a) \in \mathbb{N}$. Then, using Lemma 20, we observe that there are some positive constants K_1, K_2, K_3, K_4 such that

$$K_1 \lambda^n n^{d'} \leq |g \circ f^n(a)| = |g_\varepsilon \circ f_\varepsilon^n(a)| \leq K_2 |f_\varepsilon^n(a)| \leq K_3 |f^n(a)| \leq K_4 n^d \lambda^n.$$

meaning that $\lambda(a) = \lambda$ and that $d' \leq d(a) \leq d$. □

Lemma 24. *Let $\mathbf{w} = g(f^\omega(u_0))$ be a morphic sequence with $f : B^* \to B^*$ and $g : B^* \to A^*$. Let B' be the set of letters b in B such that $g(f^n(b)) \neq \varepsilon$ for infinitely many integers n. There are some non-erasing morphisms $f' : B'^* \to B'^*$, $g' : B'^* \to A^*$ such that $\mathbf{w} = g'(f'^\omega(u_0))$ and there exists some n such that $M_{f'} = (M_{f^n})_{B'}$. Furthermore, if f has growth type (λ, d) (w.r.t. u_0) and $|g(f^n(u_0))| = \Theta(\lambda^n n^{d'})$ for some $d' \leq d$, then f' has growth type (λ^n, e) (w.r.t. u_0) for some e such that $d' \leq e \leq d$.*

Proof. First, using Remark 8, we replace f with a convenient power f^n is such a way that M_{f^n} is equal (up to a permutation) to a lower block triangular matrix whose diagonal blocks are either primitive or zero.

We just need to iterate the operations of the previous two lemmas to get the morphisms f' and g'. The growth type of f' under the last assumption directly follows from the growth type (λ^n, d) of f^n and by those of the morphisms obtained at each applied lemma.

First, Lemma 22 provides[1] a morphism g_1 and a non-erasing morphism f_1 defined over $B_1 \subset B$ such that $\mathbf{w} = g_1(f_1^\omega(u_0))$ and we have $M_{f_1} = (M_{f^n})_{B_1}$.

We apply Lemma 23 to f_1, g_1 and the largest sub-alphabet C of $B_1 \cap g_1^{-1}(\varepsilon)$ such that $(f_1)_C$ is a sub-morphism of f_1. We obtain new morphisms g_2 and f_2 defined over $B_2 \subset B_1$ such that $\mathbf{w} = g_2(f_2^\omega(u_0))$. We have $M_{f_2} = (M_{f_1})_{B_2} = (M_{f^n})_{B_2}$.

Observe that the new morphism f_2 might be erasing: This is the case when a letter $a \in B_1$ is not erased by g_1, but is such that $g_1(f_1(a)) = \varepsilon$ (such a letter is called *moribund* in [AS03, Definition 7.7.2]). In that case, we iterate the process: Lemma 22 followed by Lemma 23 (applied to the largest possible sub-alphabet) provide new morphisms $f_3 : B_3^* \to B_3^*$ and $g_3 : B_3^* \to A^*$ such that $\mathbf{w} = g_3(f_3^\omega(u_0))$ and $M_{f_3} = (M_{f^n})_{B_3}$. We iterate this process until f_ℓ is non-erasing for some ℓ (this always happens since the two applied lemmas remove letters from a finite alphabet). As a consequence of Lemma 23, for all letters $b \in B_\ell$

$$g_\ell(f_\ell^j(b)) \neq \varepsilon \quad \text{for infinitely many } j. \tag{3}$$

Moreover, we have

$$M_{f_\ell} = (M_{f^n})_{B_\ell}.$$

By the choice of n and using the fact that f_ℓ is non-erasing, the diagonal blocks of M_{f_ℓ} are all primitive (*i.e.*, each letter corresponding to such a block appears in $f_\ell^t(b)$ for all letters b corresponding to that block and all large enough t). Hence, we can strengthen (3): for all letters $b \in B_\ell$, there is a positive integer N_b such that $g_\ell(f_\ell^j(b)) \neq \varepsilon$ for all $j \geq N_b$. Let $N = \max\{N_b \mid b \in B_\ell\}$. The morphism $g' := g_\ell \circ f_\ell^N$ is non-erasing and such that $\mathbf{w} = g'(f_\ell^\omega(u_0))$. To conclude the proof, take $B' = B_\ell$ and $f' = f_\ell$. □

The proof of the following result can be found in [CN03].

Lemma 25. [CN03, Lemme 4] *Let $\mathbf{w} = g(f^\omega(a))$ be a morphic sequence for some morphisms $f : B^* \to B^*$ and $g : B^* \to A^*$. There exists some positive integers p and q such that*

$$|(g \circ f^p)(f^q(a))| > |(g \circ f^p)(a)| \quad \text{and} \quad |(g \circ f^p)(f^q(b))| \geq |(g \circ f^p)(b)|, \ \forall b \in B.$$

We recall the algorithm of Cassaigne and Nicolas to get the morphisms σ and τ given in Theorem 21 (the correctness of this algorithm is provided by Proposition 26). Let $\mathbf{w} = g(f^\omega(a))$ be a morphic sequence with $f : B^* \to B^*$ and $g : B^* \to A^*$. Thanks to Lemma 24, f and g can be taken non-erasing. Next applying Lemma 25 and replacing f with f^q and g with $g \circ f^p$, we can suppose that f and g are non-erasing and satisfy

$$|g(f(a))| > |g(a)| \quad \text{and} \quad |g(f(b))| \geq |g(b)|, \ \forall b \in B. \tag{4}$$

Note that this is the second time that we replace f with one of its power (the first time was in Lemma 24).

[1] With the notation of Lemma 22, $g_1 = g \circ f^k$, $f_1 = f_{\mathcal{I}}$, $B_1 = A_{\mathcal{I}}$.

Algorithm 1. The input is two non-erasing morphisms $f : B^* \to B^*$ and $g : B^* \to A^*$ satisfying (4). The output is two new morphisms σ and τ defined on a new alphabet Π.

Since g is non-erasing, we define the alphabet

$$\Pi = \{(b, i) \mid b \in B, 0 \leq i < |g(b)|\}$$

and the morphism

$$\alpha : B^* \to \Pi^*,\ b \mapsto (b, 0)(b, 1) \cdots (b, |g(b)| - 1).$$

We also define the coding

$$\tau : \Pi^* \to A^*, (b, i) \mapsto (g(b))_i$$

where $(g(b))_i$ denotes the ith letter occurring in $g(b)$, $0 \leq i < |g(b)|$. It is clear that $\tau \circ \alpha = g$.

Since $|\alpha(f(b))| = |g(f(b))| \geq |g(b)|$, $\alpha(f(b))$ can be factorized (not necessarily in a unique way) into $|g(b)|$ non-empty words. Pick such a factorization

$$\alpha(f(b)) = w_{b,0}\, w_{b,1} \,\cdots\, w_{b,|g(b)|-1} \tag{5}$$

with $w_{b,i} \in \Pi^+$ for all i. We moreover impose $|w_{a,0}| \geq 2$ (recall that $|g(f(a))| > |g(a)|$). Now, define

$$\sigma : \Pi^* \to \Pi^*, (b, i) \mapsto w_{b,i}.$$

Proposition 26. *Let $\mathbf{w} = g(f^\omega(a))$ be a (λ, d)-morphic sequence such that $f : B^* \to B^*$ and $g : B^* \to A^*$ are two non-erasing morphisms satisfying* (4)*. The morphisms τ and σ built in Algorithm 1 are such that $\mathbf{w} = \tau(\sigma^\omega((a, 0)))$, σ is non-erasing, τ is a coding. Moreover, M_σ is a dilated matrix of M_f and σ has growth type* (λ, d) *(w.r.t.* $(a, 0)$*).*

Proof. It is clear that τ is a coding and that σ is non-erasing and prolongable on the first letter of $w_{a,0}$ which is $(a, 0)$. Recall that $\tau \circ \alpha = g$. Let $\mathbf{u} = f^\omega(a)$. We have $\tau(\alpha(\mathbf{u})) = \mathbf{w}$. Let us show that $\sigma^\omega((a, 0)) = \alpha(\mathbf{u})$. From (5) and since g is non-erasing, we observe that

$$\alpha \circ f = \sigma \circ \alpha$$

which implies that $\alpha(\mathbf{u})$ is a fixed point of σ: $\sigma(\alpha(\mathbf{u})) = \alpha(f(\mathbf{u})) = \alpha(\mathbf{u})$.

Let us prove that $M_\sigma \in \mathrm{Dil}(M_f)$ with dilatation vector $(|g(b)|)_{b\in B}$. For all $b_i, b_j \in B$ and for all $k \in \{0, 1, \ldots, |g(b_i)| - 1\}$, we have

$$\begin{aligned}
\sum_{\ell=0}^{|g(b_j)|-1} (M_\sigma)_{(b_i,k),(b_j,\ell)} &= \sum_{\ell=0}^{|g(b_j)|-1} |\sigma((b_j, \ell))|_{(b_i,k)} \\
&= \sum_{\ell=0}^{|g(b_j)|-1} |w_{b_j,\ell}|_{(b_i,k)} \\
&= |w_{b_j,0}\, w_{b_j,1} \,\cdots\, w_{b_j,|g(b_j)|-1}|_{(b_i,k)} \\
&= |\alpha(f(b_j))|_{(b_i,k)} \\
&= (M_f)_{b_i,b_j}.
\end{aligned}$$

Indeed, if $(M_f)_{b_i,b_j} = x$ for some $b_i, b_j \in B$, then the word $(b_i, 0)(b_i, 1) \cdots (b_i, |g(b_i)|-1)$ occurs x times in $\alpha(f(b_j))$. Therefore, we also have $|\alpha(f(b_j))|_{(b_i,k)} = x$ for all $k \in \{0, 1, \ldots, |g(b_i)|-1\}$.

Finally, let us prove that σ has the same growth type as f. As g is non-erasing, by Lemma 20 we have $|g(f^n(a))| = \Theta(\lambda^n n^d)$. Then, since $|g(a)| = |\alpha(a)|$ for all $a \in A$, and $\alpha \circ f = \sigma \circ \alpha$, we get that $|\sigma^n \circ \alpha(a)| = \Theta(\lambda^n n^d)$. Finally, as $\sigma^n((a, 0))$ converges to $\alpha(\mathbf{u}) = \sigma(\alpha(\mathbf{u}))$ when n increases, there are some integers $k_1, k_2 \in \mathbb{N}$ such that $\sigma \circ \alpha(a)$ is a prefix of $\sigma^{k_1}((a, 0))$, itself a prefix of $\sigma^{k_2} \circ \alpha(a)$. Thus, for all n we have $|\sigma^n(\alpha(a))| \le |\sigma^{n+k_1}((a,0))| \le |\sigma^{n+k_2}(\alpha(a))|$, meaning that σ has growth type (λ, d) (w.r.t. $(a,0)$). □

3.3. **Cobham's theorem for morphic sequences.** Two real numbers $\lambda, \lambda' > 1$ are *multiplicatively independent* if the only non-negative integers k, ℓ such that $\lambda^k = \lambda'^\ell$ are $k = \ell = 0$. Otherwise, λ, λ' are *multiplicatively dependent*. It is clear that multiplicative dependence is an equivalence relation over $(1, +\infty)$ and, for all $n \ge 1$, λ and λ^n are multiplicatively dependent.

F. Durand proved the following result [Dur11].

Theorem 27. [Dur11] *Let $\lambda, \lambda' > 1$ be two multiplicatively independent real numbers. Let* $\mathbf{u}$ *(resp.* $\mathbf{v}$*) be a λ-pure morphic (resp. λ'-pure morphic) sequence. Let ϕ and ψ be two non-erasing morphisms. If $\mathbf{w} = \phi(\mathbf{u}) = \psi(\mathbf{v})$, then* $\mathbf{w}$ *is ultimately periodic.*

If a sequence $\mathbf{w}$ is both λ- and λ'-morphic with $\lambda, \lambda' > 1$ being multiplicatively independent, then $\mathbf{w}$ needs not be ultimately periodic. As an example, consider the morphism $f : a \mapsto abc, b \mapsto ba, c \mapsto ccccc$ and the morphism $g : a \mapsto a, b \mapsto b, c \mapsto \varepsilon$. Therefore, the sequence $f^\omega(a)$ is 5-pure morphic and with our definition, $g(f^\omega(a))$ is thus 5-morphic. Nevertheless, observe that $g(f^\omega(a))$ is the well-known Thue–Morse sequence which is 2-pure morphic and not ultimately periodic.

In the latter example, observe that the letter c provides the dominating eigenvalue of f which is 5 but c is erased by g. This is the reason why we introduce the next definition.

Definition 28. Let $f : B^* \to B^*$ and $g : B^* \to A^*$ be two morphisms. Let λ be the dominating eigenvalue of f. Let B' be the set of letters b in B such that $g(f^n(b)) \ne \varepsilon$ for infinitely many integers n. We say that (f, g) satisfies the *dominating eigenvalue property* (DEV-property) if $\lambda \in \mathrm{Spec}((M_f)_{B'})$.

Proposition 29. *Let $\lambda, \lambda' > 1$ be two multiplicatively independent real numbers. Suppose that $\mathbf{w} = g(f^\omega(a)) = g'(f'^\omega(a'))$ with f of growth type (λ, d) (w.r.t. a) and f' of growth type (λ', d') (w.r.t. a'). If both (f, g) and (f', g') satisfy the DEV-property, then* $\mathbf{w}$ *is ultimately periodic.*

Proof. Applying Lemma 24 and Lemma 25 to $f : B^* \to B^*$ and $g : B^* \to A^*$ provides non-erasing morphisms $F : B'^* \to B'^*$ and $G : B'^* \to A^*$ such that $\mathbf{w} = G(F^\omega(a))$ where B' is the set of letters b in B such that $g(f^j(b)) \ne \varepsilon$ for infinitely many integers j. Moreover, we have $M_F = (M_{f^n})_{B'}$ for some n. Since (f, g) satisfies the DEV-property, the dominating eigenvalue of F is λ^n. Thus, $\mathbf{w}$ is the image under the non-erasing morphism G of the λ^n-pure morphic word $F^\omega(a)$.

The same holds for g' and f'. We get that $\mathbf{w}$ is the image under a non-erasing morphism G' of a $(\lambda')^m$-pure morphic word $(F')^\omega(a')$. Note that λ^n and $(\lambda')^m$ are multiplicatively independent. Hence we can apply Theorem 27. □

4. Abstract numeration systems

4.1. **Basic definitions.**

Definition 30. A *deterministic finite automaton* (DFA for short) over A is a labeled directed graph $\mathcal{A} = (Q, q_0, A, \delta, T)$ where Q is the set of *states*, $\delta : Q \times A \to A$ is the *transition function*, $q_0 \in Q$ is the *initial state* and T is the set of *terminal states*. Generally, δ is a partial function, *i.e.*, its domain is a subset of $Q \times A$. A DFA is *complete* if δ is a total function, *i.e.*, its domain is $Q \times A$. The function δ can be extended to $Q \times A^*$ by $\delta(q, \varepsilon) = q$ and $\delta(q, ub) = \delta(\delta(q, u), b)$ for all $q \in Q$, $u \in A^*$, $b \in A$.

A state q of $\mathcal{A}$ is *accessible* if it can be reached from the initial state: there exists $u \in A^*$ such that $\delta(q_0, u) = q$. A state q of $\mathcal{A}$ is *co-accessible* if there exists $u \in A^*$ such that $\delta(q, u) \in T$. A DFA is *trim* if all its states are accessible and co-accessible.

The reason why we consider trim automata is similar to assuming that all letters of A occur in a pure morphic sequence $\mathbf{w}$ generated by a morphism $f : A^* \to A^*$ (see Definition 17 and Remark 19). States that are not accessible or co-accessible could induce irrelevant eigenvalues associated with the automaton.

Definition 31. Let $\mathcal{A} = (Q, q_0, A, \delta, T)$ be a DFA. A word $u \in A^*$ is said to be *accepted* (or *recognized*) by $\mathcal{A}$ if (q_0, u) belongs to the domain of δ and $\delta(q_0, u) \in T$. The *language accepted by* $\mathcal{A}$ is the set of words accepted by $\mathcal{A}$. It is denoted by $L(\mathcal{A})$. A language is *regular* if it is the language accepted by some DFA. Two DFAs are *equivalent* if they accept the same language. Removing states that are not accessible or not co-accessible, any DFA is equivalent to a trim DFA.

Let L be a regular language. Up to a renaming of the states, there exists a unique DFA with a minimal number of states accepting L. This automaton is called the *minimal automaton* of L. For a general reference, see for instance [**?**]. We let $\mathcal{M}_L$ denote the trim minimal automaton of L.

Definition 32. A *DFA with output* (DFAO for short) $\mathcal{O} = (Q, q_0, A, \delta, B, \alpha)$ is a complete DFA (Q, q_0, A, δ, T) where the set of terminal states T is replaced with an *output function* $\alpha : Q \to B$, B being an alphabet.

As mentioned in the introduction, it is desirable to consider numeration systems where the valid representations of the integers make up a regular language: the language of the numeration is accepted by some DFA. To fulfil this requirement, abstract numeration systems provide a unified framework [BR10, Chap. 3].

Definition 33. [**?**] An *abstract numeration system* is a triple $S = (L, A, <)$ where $(A, <)$ is a totally ordered alphabet and L is an infinite regular language over A. The words in L are totally ordered by the radix order over A^* induced by $<$. Words are ordered by increasing length and two words of the same length are ordered thanks to the lexicographic order. This order is an increasing one-to-one correspondence $\mathrm{rep}_S : \mathbb{N} \to L$. The word $\mathrm{rep}_S(n)$ is the $(n+1)$th word in L.

Definition 34. Let $S = (L, A, <)$ be an abstract numeration system. A set $X \subset \mathbb{N}$ is *S-recognizable* if $\mathrm{rep}_S(X)$ is a regular language.

The following definition is a natural generalization of the concept of k-automatic sequence where base k expansions are the inputs of some DFAO.

Definition 35. Let $S = (L, A, <)$ be an *abstract numeration system.* A sequence $\mathbf{w} = w_0 w_1 w_2 \cdots$, $w_i \in B$ for all i, is *S-automatic* if there exists a DFAO $\mathcal{O} = (Q, q_0, A, \delta, B, \alpha)$ such that

$$w_i = \alpha(\delta(q_0, \mathrm{rep}_S(i))), \quad \forall i \geq 0. \tag{6}$$

If X is a subset of $\mathbb{N}$, the *characteristic sequence* of X is $\mathbf{1}_X = x_0 x_1 x_2 \cdots$, where $x_i = 1$ if $i \in X$ and $x_i = 0$ otherwise.

Remark 36. Let S be an abstract numeration system. A subset X of $\mathbb{N}$ is S-recognizable if and only if its characteristic sequence $\mathbf{1}_X$ is S-automatic.

4.2. **The S-automatic sequences are exactly the morphic sequences.** The next result is an extension of the theorem of Cobham from 1972 characterizing morphic sequences generated by a constant length morphism.

Theorem 37. [Rig00, RM02] *A sequence is S-automatic for some abstract numeration system S if and only if it is morphic.*

Given an abstract numeration system S and a DFAO $\mathcal{O}$, we let $\mathbf{w}_{S,\mathcal{O}}$ denote the corresponding S-automatic sequence defined by (6). The proof of Theorem 37 is constructive. The algorithm below provides morphisms showing that an S-automatic sequence is morphic.

Algorithm 2. The input is an abstract numeration system $S = (L, A, <)$ and a DFAO $\mathcal{O} = (R, r_0, A, \delta_{\mathcal{O}}, B, \alpha)$ generating a sequence $\mathbf{w}_{S,\mathcal{O}}$ over B defined by (6). Let $\mathcal{M}_L = (Q, q_0, A, \delta_L, T)$ be the trim minimal automaton of L. For convenience, we can assume that $A = \{1 < \cdots < n\}$. An intermediate output is an auxiliary product automaton. The final output is two morphisms $f_{S,\mathcal{O}}$ and $g_{S,\mathcal{O}}$ such that

$$\mathbf{w}_{S,\mathcal{O}} = g_{S,\mathcal{O}}(f^{\omega}_{S,\mathcal{O}}(z))$$

for some letter z.

Consider the Cartesian product automaton $\mathcal{P} = \mathcal{M}_L \times \mathcal{O}$ defined as follows. The set of states of $\mathcal{P}$ is a subset P of $Q \times R$. The initial state is (q_0, r_0) and the alphabet is A. The transition function $\Delta : (Q \times R) \times A^* \to Q \times R$ of $\mathcal{P}$ is defined as follows. First $((q, r), w) \in (Q \times R) \times A^*$ belongs to the domain of Δ if and only if (q, w) belongs to the domain of δ_L. In that case,

$$\Delta((q, r), w) = (\delta_L(q, w), \delta_{\mathcal{O}}(r, w)).$$

We only consider the accessible part of $\mathcal{P}$: its set of states is given by

$$P = \{(q, r) \in Q \times R \mid \exists w \text{ such that } ((q_0, r_0), w) \in \operatorname{dom} \Delta \text{ and } \Delta((q_0, r_0), w) = (q, r)\}.$$

Define a morphism $f_{S,\mathcal{O}} : (P \cup \{z\})^* \to (P \cup \{z\})^*$ prolongable on a letter z which does not belong to P as follows: $f_{S,\mathcal{O}}(z) = z f_{S,\mathcal{O}}(q_0, r_0)$ and $f_{S,\mathcal{O}}(q, r)$ is obtained by concatenating the states reached when reading the different letters in A:

$$f_{S,\mathcal{O}}(q, r) = \Delta((q, r), 1) \cdots \Delta((q, r), n). \tag{7}$$

If $((q, r), i)$ does not belong to the domain of Δ, we set $\Delta((q, r), i) = \varepsilon$ in the above formula. Note that $f_{S,\mathcal{O}}$ can therefore be erasing. This situation happens exactly when $\mathcal{M}_L$ has some terminal state without outgoing transition.

Define $g_{S,\mathcal{O}} : (P \cup \{z\})^* \to B^*$ by

$$g_{S,\mathcal{O}}(z) = g_{S,\mathcal{O}}(q_0, r_0) \quad \text{and} \quad g_{S,\mathcal{O}}(q, r) = \begin{cases} \alpha(r), & \text{if } q \in T, \\ \varepsilon, & \text{otherwise.} \end{cases}$$

By definition, $g_{S,\mathcal{O}}$ can be erasing.

Example 38. Consider the language L over $\{1, 2\}$ of the words avoiding the factor 22 and not having a suffix 2. In Figure 1, the minimal automaton of L is depicted on top, a DFAO $\mathcal{O}$ with three states is depicted on the left. The output alphabet is $\{a, b, c\}$ and the output associated with each state is written inside the state (we identify these states with their outputs). The resulting Cartesian product $\mathcal{P} = \mathcal{M}_L \times \mathcal{O}$ occupies the central position (where the non-accessible state $(2, a)$ has been removed). If the initial state of $\mathcal{M}_L$ is 1 and the other

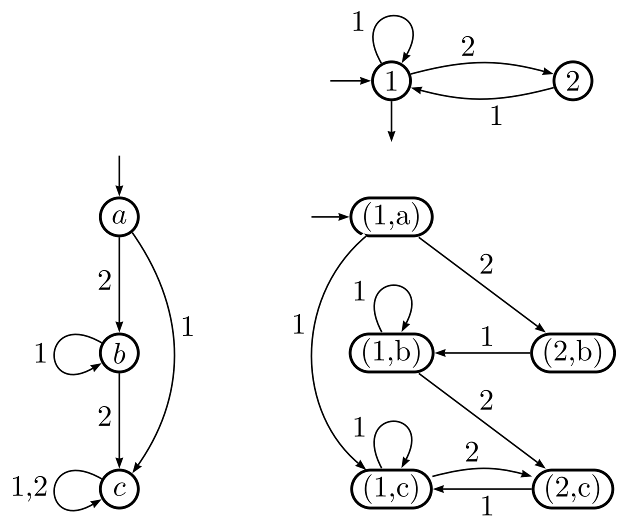

FIGURE 1. A trim minimal DFA, a DFAO and the corresponding Cartesian product.

state of $\mathcal{M}_L$ is 2, the morphism $f_{S,\mathcal{O}}$ is given by $z \mapsto z(1, c)(2, b)$ and

$$(1, a) \mapsto (1, c)(2, b),\ (1, b) \mapsto (1, b)(2, c),\ (1, c) \mapsto (1, c)(2, c),\ (2, b) \mapsto (1, b),\ (2, c) \mapsto (1, c)$$

and, for $x \in \{a, b, c\}$,

$$g_{S,\mathcal{O}} : (1, x) \mapsto x,\ (2, b), (2, c) \mapsto \varepsilon.$$

We get

$$f^{\omega}_{S,\mathcal{O}}(z) = z(1, c)\,(2, b)\,(1, c)\,(2, c)\,(1, b)\,(1, c)\,(2, c)\,(1, c)\,(1, b)\,(2, c)\,(1, c)\,(2, c)\,(1, c)\,(1, c)\cdots$$

and

$$g_{S,\mathcal{O}}(f^{\omega}_{S,\mathcal{O}}(z)) = accbccbcccbccccccbcccccccccccbcccccccc\cdots$$

Definition 39. Let $\mathcal{A} = (Q, q_0, A, \delta, T)$ be a DFA. The *incidence matrix* of $\mathcal{A}$ is the matrix $M_{\mathcal{A}} \in \mathbb{N}^{Q\times Q}$ defined by $(M_{\mathcal{A}})_{p,q} = \mathrm{Card}\,(\{a \in A \mid \delta(p, a) = q\})$. We can again say that the eigenvalues and the spectrum of $M_{\mathcal{A}}$ are respectively the eigenvalues and the spectrum of $\mathcal{A}$. Let L be regular language. In particular, since the matrix $M_{\mathcal{M}_L}$ is non-negative, we can talk of the *dominating eigenvalue* of L.

Example 40. Continuing the above example, the incidence matrix of $\mathcal{M}_L$ and $\mathcal{P}$ are respectively

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \left(\begin{array}{ccc|cc} 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

Note that the second matrix is a dilated matrix of the first one with a dilatation vector $(3, 2)$.

Proposition 41. *Let $S = (L, A, <)$ be an abstract numeration system and $\mathcal{O}$ be a DFAO. Let $\mathcal{P}$ be the Cartesian product automaton $\mathcal{M}_L \times \mathcal{O}$ given in Algorithm 2. Then $M_{\mathcal{P}} \in \mathrm{Dil}\,(M_{\mathcal{M}_L})$.*

Proof. Let $\mathcal{M}_L = (Q, q_0, A, \delta_L, T)$ and $\mathcal{O} = (R, r_0, A, \delta_{\mathcal{O}}, B, \alpha)$. Rows and columns of $M_{\mathcal{P}}$ are both indexed by pairs $(q, r) \in Q \times R$. Recall that we only consider the pairs in P corresponding to accessible states in $\mathcal{P}$. We fix orderings of Q and R and consider the corresponding lexicographic order of $Q \times R$: $(x, y) < (x', y')$ if $x < x'$ or, $x = x'$ and $y < y'$.

Let q, s be states in $\mathcal{M}_L$. Since the DFAO $\mathcal{O}$ is complete, for all $a \in A$, we have $\delta_L(q, a) = s$ in $\mathcal{M}_L$ if and only if, for all $r \in R$, $\Delta((q, r), a) = (s, \delta_{\mathcal{O}}(r, a))$. Moreover, for all $r \in R$ and all $a \in A$, there is exactly one state $t \in R$ such that $\delta_{\mathcal{O}}(r, a) = t$. Hence, for all r with $(q, r) \in P$,

$$\sum_{\substack{t \in R \text{ with} \\ (s,t) \in P}} (M_{\mathcal{P}})_{(q,r),(s,t)} = (M_{\mathcal{M}_L})_{q,s}.$$

□

As a consequence of Proposition 10, we immediately obtain the following result.

Corollary 42. *The matrices $M_{\mathcal{P}}$ and $M_{\mathcal{M}_L}$ have the same dominating eigenvalue.*

The transpose of a matrix M is denoted by $\widetilde{M}$.

Proposition 43. *Let $S = (L, A, <)$ be an abstract numeration system and $\mathcal{O}$ be a DFAO. Let $f_{S,\mathcal{O}}$ be the morphism obtained from Algorithm 2. The matrix $M_{f_{S,\mathcal{O}}}$ is (up to a permutation) a lower block triangular matrix*

$$\begin{pmatrix} 1 & 0 \\ v & \widetilde{M_{\mathcal{P}}} \end{pmatrix}$$

where v is the column vector of $\tilde{M}_{\mathcal{P}}$ corresponding to (q_0, r_0), i.e., for all $(q, r) \in Q \times R$, $v_{(q,r)} = (M_{\mathcal{P}})_{(q_0,r_0),(q,r)}$. In particular, $M_{\mathcal{P}}$ and $f_{S,\mathcal{O}}$ have the same dominating eigenvalue.

Proof. It follows directly from (7). To get the matrix given in the statement, we simply assume that the special symbol z corresponds to the first entry of the matrix and we consider the same ordering as the one used to build $M_{\mathcal{P}}$ in the proof of Proposition 41. □

5. Cobham's Theorem for abstract numeration systems

Definition 44. Let $S = (L, A, <)$ be an abstract numeration system. For all $n \in \mathbb{N}$, we let $v_L(n)$ denote the number of words of length at most n in L.

The next result is a direct consequence of Lemma 11 and Lemma 12.

Lemma 45. *Let $S = (L, A, <)$ be an abstract numeration system. Let λ_S be the dominating eigenvalue of L. There exists a non-negative integer $d_S \in \mathbb{N}$ such that $v_L(n) = \Theta(n^{d_S} \lambda_S^n)$.*

Definition 46. Let $S = (L, A, <)$ be an abstract numeration system. The couple (λ_S, d_S) of the previous lemma is called the *growth type* of S. An abstract numeration system is said to be *polynomial* (resp. *exponential*) if $\lambda_S = 1$ (resp. $\lambda_S > 1$).

Definition 47. Let $S = (L, A, <)$ and $S' = (L', A', <')$ be abstract numeration systems of growth type (λ_S, d_S) and $(\lambda_{S'}, d_{S'})$ respectively. Then S and S' are said to be *independent* if one of the following conditions holds true.

(1) λ_S and $\lambda_{S'}$ are greater than 1 and are multiplicatively independent;
(2) $\max\{\lambda_S, \lambda_{S'}\} > 1$ and $\min\{\lambda_S, \lambda_{S'}\} = 1$;
(3) $\lambda_S = \lambda_{S'} = 1$ and $\gcd(d_S, d_{S'}) = 1$.

Recall the following result.

Theorem 48 (Lecomte and Rigo [**?**])**.** *Any finite union of arithmetic progressions is S-recognizable for all abstract numeration systems S*

The main result of this section is the following.

Theorem 49. *Let $S = (L, A, <)$ and $S' = (L', A', <')$ be two independent abstract numeration systems. A set X is both S- and S'-recognizable if and only if it is a finite union of arithmetic progressions.*

We divide the proof into three cases: the *exponential case* of two exponential ANSs, the *polynomial case* of two polynomial ANSs and the *mixed case* of one polynomial ANS and one exponential ANS. One of the key points of the proof is the following result.

Lemma 50 ([CR11])**.** *Let $S = (L, A, <)$ be an abstract numeration system and $\mathcal{O}$ be a DFAO. Let $f_{S,\mathcal{O}}$ and $g_{S,\mathcal{O}}$ be the morphisms obtained from Algorithm 2 with $f_{S,\mathcal{O}}$ prolongable on (q_0, r_0). For all n, we have $|g_{S,\mathcal{O}}(f^n_{S,\mathcal{O}}((q_0, r_0)))| = v_L(n)$.*

5.1. **Proof of the mixed case.** This part is known and due to Durand and Rigo.

Theorem 51. [DR09] *Let S be a polynomial abstract numeration system and let S' be an exponential abstract numeration system. If $X \subset \mathbb{N}$ is both S- and S'-recognizable, then it is a finite union of arithmetic progressions.*

5.2. **Proof of the exponential case.** This part almost directly follows from Theorem 27.

Theorem 52. *Let S and S' be two exponential abstract numeration system with multiplicatively independent dominating eigenvalue λ_S and $\lambda_{S'}$ respectively. If $X \subset \mathbb{N}$ is both S- and S'-recognizable, then it is a finite union of arithmetic progressions.*

Proof. The set X being both S- and S'-recognizable, its characteristic sequence $\mathbf{1}_X$ is both S- and S'-automatic. Thanks to Algorithm 2, we can write

$$\mathbf{1}_X = g(f^\omega(a)) = g'(f'^\omega(a'))$$

for some morphisms f, g, f', g'.

Since $\mathcal{M}_L$ and $\mathcal{M}_{L'}$ are trim, the pairs (f, g) and (f', g') satisfy the DEV-property. Furthermore, thanks to Proposition **??** and Corollary 42, the dominating eigenvalue of f and f' are λ_S and $\lambda_{S'}$ respectively. Applying Proposition 29, we obtain that $\mathbf{1}_X$ is ultimately periodic, that is, X is a finite union of arithmetic progressions. □

5.3. **Proof of the polynomial case.**

Theorem 53. *Let S and S' be two polynomial abstract numeration system of co-prime degrees d_S and $d_{S'}$ respectively. If $X \subset \mathbb{N}$ is both S- and S'-recognizable, then it is a finite union of arithmetic progressions.*

The way we prove the result is standard to prove Cobham-like theorems. Using a result of Charlier and Rampersad, we first prove that X is syndetic. Then, we prove that the finite words occurring infinitely often in $\mathbf{1}_X$ occur with bounded gaps. Then, we suppose that $\mathbf{1}_X$ is not ultimately periodic and we reach a contradiction.

Let us recall the following result. For all n, we let $t_X(n)$ denote the $(n+1)$th element of X. Charlier and Rampersad proved the following result [CR11].

Theorem 54 (Charlier and Rampersad [CR11])**.** *Let S be a polynomial abstract numeration system of degree d. If X is an infinite S-recognizable set, then there is a positive integer $c \leq d$ such that $v_{\mathrm{rep}_S(X)}(n) = \Theta(n^c)$ and $t_X(n) = \Theta(n^{\frac{d}{c}})$.*

The next easy lemma allows us to prove that X is syndetic.

Lemma 55. *Two positive integers d and d' are co-prime if and only if $\{\frac{d}{c} \mid 1 \leq c < d\} \cap \{\frac{d'}{c'} \mid 1 \leq c' < d'\} = \emptyset$.*

Proof. Suppose that d and d' are co-prime and that there exists $c \in \{1, \ldots, d-1\}$ and $c' \in \{1, \ldots, d'-1\}$ such that $\frac{d}{c} = \frac{d'}{c'}$. Then, $c = \frac{c'd}{d'}$, meaning that d' divides c' which is a contradiction. For the other part, if d and d' have a common divisor $x \in \mathbb{N}_{>1}$, then it belongs to $\{\frac{d}{c} \mid 1 \leq c < d\} \cap \{\frac{d'}{c'} \mid 1 \leq c' < d'\}$. Indeed, $x = \frac{d}{\frac{d}{x}} = \frac{d'}{\frac{d'}{x}}$ with $\frac{d}{x} \in \{1, \ldots, d-1\}$ and $\frac{d'}{x} \in \{1, \ldots, d'-1\}$. □

Proposition 56. *Let S and S' be two polynomial abstract numeration system of co-prime degrees d_S and $d_{S'}$ respectively. If $X \subset \mathbb{N}$ is an infinite S- and S'-recognizable set, then X is syndetic.*

Proof. By Theorem 54, there are some integers $c \in \{1, \ldots, d\}$ and $c' \in \{1, \ldots, d'\}$ such that $t_X(n) \in \Theta(n^{\frac{d}{c}}) \cap \Theta(n^{\frac{d'}{c'}})$, which, by the previous lemma, is possible if and only if $c = d$ and $c' = d'$. □

The next result allows us to prove that not only X is syndetic, but also the finite words occurring infinitely often in 1_X occur in it with bounded gaps. Let $\mathbf{1}_X = x_0x_1x_2\cdots$ be the characteristic sequence of X. For all $u \in \{0,1\}^*$, we let $X_u \subset \mathbb{N}$ denote the set

$$X_u = \{i \in \mathbb{N} \mid x_ix_{i+1}\cdots x_{i+|u|-1} = u\}.$$

Proposition 57. *Let S be an abstract numeration system. If X is a S-recognizable set, then for any word $u \in \{0,1\}^*$, the set X_u is also S-recognizable.*

Proof. The case where X or X_u is finite is trivial. Thus, let us suppose that X and X_u are infinite. Let k be the smallest integer such that $x_k\cdots x_{k+|u|-1} = u$. For all $i \in \{0, 1, \ldots, |u|-1\}$, we let X_i (resp. Y_i) denote the set of non-negative integers whose characteristic sequence is $0^{k+i}u^\omega$ (resp. $0^{k+i}(10^{|u|-1})^\omega$). For all i, the sets X_i and Y_i are ultimately periodic, hence S-recognizable.

Suppose that the first letter of u is 1. By construction, $X \cap X_i \cap Y_i$ is the set of non-negative integers j such that $x_j\cdots x_{j+|u|-1} = u$ and $j \equiv k+i \mod |u|$. The set X_u is thus $\bigcup_{0\leq i<|u|} X \cap X_i \cap Y_i$ and this set is S-recognizable because S-recognizability is closed under finite unions and intersections. The case where the first letter of u is 0 is similarly obtained by replacing X by $\mathbb{N} \setminus X$ and X_i by $\mathbb{N} \setminus X_i$. □

The next result is a direct consequence of Proposition 56 and Proposition 57.

Corollary 58. *Let S and S' be two polynomial abstract numeration system of co-prime degrees d_S and $d_{S'}$ respectively. Suppose that $X \subset \mathbb{N}$ is both S- and S'-recognizable. For all words $u \in \{0,1\}^*$, if X_u is infinite, then it is syndetic.*

We will need Lemma 59 below concerning $(1,d)$-pure morphic words. It directly follows from Pansiot's word about pure morphic words [Pan83, Proof of Théorème 4.1]] and from the fact that arbitrary large words only composed of non-growing letters[2] occur in $(1,d)$-pure morphic words.

[2]With the notation of Lemma 16, a letter b is *growing* if $\lambda(b) > 1$ or $d(b) \geq 1$. Otherwise it is *non-growing*.

Lemma 59. *Let $f : A^* \to A^*$ be a non-erasing polynomial morphism prolongable on $a \in A$ and denote $\mathbf{w} = f^\omega(a)$. There exists a word $u \in A^+$ such that for all n, u^n occurs infinitely often in $\mathbf{w}$.*

Now we can end the proof of the polynomial case.

Proof of Theorem 53. Let f, g and f', g' be the morphisms given by Algorithm 2 and let a and a' be such that $g(f^\omega(a)) = \mathbf{1}_X = g'((f')^\omega(a'))$. We deduce from Proposition 43 that f and f' are polynomial of degree d_S and $d_{S'}$ respectively. Furthermore, using Lemma 50, Lemma 24 and Algorithm 1, we can suppose that f and f' are non-erasing and that g and g' are codings.

By the previous lemma, there exists a word u such that for all n, u^n occurs infinitely often in $f^\omega(a)$. Similarly, there exists a word v such that for all m, v^m occurs infinitely often in $(f')^\omega(a')$.

Suppose that $|u| \leq |v|$; the other case is similar. Let $k \geq 1$ be an integer such that $|u^k| \geq 3|v|$. Since u^k occurs infinitely often in $f^\omega(a)$, $g(u^k)$ occurs infinitely often in $\mathbf{1}_X$. Thus, by Corollary 58, $g(u^k)$ occurs with bounded gaps in $\mathbf{1}_X$.

Proceed by contradiction and suppose that $\mathbf{1}_X$ is not ultimately periodic. Thus there exists a word $w \in \{0,1\}^*$ such that $g(u^k)wg(u^k)$ occurs in $\mathbf{1}_X$ and is not $|u|$-periodic[3]. Let W be the set of such words w. Since $g(u^k)$ occurs with bounded gaps, W is finite. Thus, there exists $w' \in W$ such that $g(u^k)w'g(u^k)$ occurs infinitely often in $\mathbf{1}_X$, hence with bounded gaps (again by Corollary 58). Let $m \in \mathbb{N}$ be such that $g'(v^m)$ contains an occurrence of $g(u^k)w'g(u^k)$. Since $|u^k| \geq 3|v|$ and $g(u^k)$ is $|u|$-periodic, $g'(v^m)$ is also $|u|$-periodic. Thus $g(u^k)w'g(u^k)$ is $|u|$-periodic which is a contradiction. $\square$

Remark 60. Even if we do not have any proof of it, we believe that Theorem 49 still holds true if we replace Condition 3 of Definition 47 by $\lambda_S = \lambda_{S'} = 1$ and $d_S \neq d_{S'}$. We intend to attack this problem is a future paper.

6. Acknowledgement

This research is supported by the internal research project F1R-MTH-PUL-12RDO2 of the University of Luxembourg.

References

[AS03] J.-P. Allouche and J. Shallit. *Automatic sequences*. Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations.

[BH97] V. Bruyère and G. Hansel. Bertrand numeration systems and recognizability. *Theor. Comput. Sci.*, 181(1):17–43, 1997.

[BHMV94] V. Bruyère, G. Hansel, C. Michaux, and R. Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc*, 1:191–238, 1994.

[BR10] V. Berthé and M. Rigo, editors. *Combinatorics, automata, and number theory.*, volume 135 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2010. 636 pages.

[CN03] J. Cassaigne and F. Nicolas. Quelques propriétés des mots substitutifs. *Bull. Belg. Math. Soc. Simon Stevin*, 10(suppl.):661–676, 2003.

[Cob68] A. Cobham. On the hartmanis-stearns problem for a class of tag machines. In *SWAT (FOCS)*, pages 51–60, 1968.

[Cob69] Alan Cobham. On the base-dependence of sets of numbers recognizable by finite automata. *Mathematical Systems Theory*, pages 186–192, 1969.

[3] A word $w = w_1, \cdots w_n$ is k-periodic if for all $i \in \{1, \ldots n-k\}$, $w_i = w_{i+k}$.

[CR11] É. Charlier and N. Rampersad. The growth function of S-recognizable sets. *Theoret. Comput. Sci.*, 412(39):5400–5408, 2011.

[DR09] F. Durand and M. Rigo. Syndeticity and independent substitutions. *Adv. in Appl. Math.*, 42(1):1–22, 2009.

[Dur11] F. Durand. Cobham's theorem for substitutions. *J. Eur. Math. Soc. (JEMS)*, 13(6):1799–1814, 2011.

[LM95] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.

[Lot02] *Algebraic Combinatorics on Words*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2002.

[NR07] S. Nicolay and M. Rigo. About frequencies of letters in generalized automatic sequences. *Theoret. Comput. Sci.*, 374(1-3):25–40, 2007.

[Pan83] J.-J. Pansiot. Hiérarchie et fermeture de certaines classes de tag-systèmes. *Acta Inform.*, 20(2):179–196, 1983.

[Rig00] M. Rigo. Generalization of automatic sequences for numeration systems on a regular language. *Theoret. Comput. Sci.*, 244(1-2):271–281, 2000.

[RM02] M. Rigo and A. Maes. More on generalized automatic sequences. *J. Autom. Lang. Comb.*, 7(3):351–376, 2002.

Université de Liège, Institut de mathématique, Grande traverse 12 (B37), 4000 Liège, Belgium
echarlier@ulg.ac.be, M.Rigo@ulg.ac.be

Mathematics Research Unit, FSTC, University of Luxembourg, 6, rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg
julien.leroy@uni.lu