SYNDETICITY AND INDEPENDENT SUBSTITUTIONS

FABIEN DURAND AND MICHEL RIGO

Abstract. We associate in a canonical way a substitution to any abstract numeration system built on a regular language. In relationship with the growth order of the letters, we define the notion of two independent substitutions. Our main result is the following. If a sequence x is generated by two independent substitutions, at least one being of exponential growth, then the factors of x appearing infinitely often in x appear with bounded gaps. As an application, we derive an analogue of Cobham's theorem for two independent substitutions (or abstract numeration systems) one with polynomial growth, the other being exponential.

1. Introduction

A set $E \subset \mathbb{N}$ is *p-recognizable* for some $p \in \mathbb{N} \setminus \{0, 1\}$, if the language consisting of the p-ary expansions of the elements in E is recognizable by a finite automaton [Ei]. In 1969, A. Cobham obtained the following result [Co1]. *Let $p, q \geq 2$ be two multiplicatively independent integers (i.e., $p^k \neq q^\ell$ for all integers $k, \ell > 0$). A set $E \subset \mathbb{N}$ is both p-recognizable and q-recognizable if and only if E is a finite union of arithmetic progressions.*

A key part in all known proofs of this seminal theorem (and this remark stands also for generalizations to non-standard positional numeration systems) is to show that E is *syndetic* (i.e., the difference between two consecutive elements of E is bounded), see [Ha, Du1, Du2].

In this paper we study this syndeticity problem for a larger class of numeration systems namely, for numeration systems built on infinite regular languages, the so-called *abstract numeration systems* [LR]. In particular, these systems contain classical numeration systems like the k-ary system or the Fibonacci system, but also more "exotic" systems for which the language of the numeration contains a number of words of length n bounded by a polynomial in n (which is contrasting with the usual exponential paradigm).

In 1972, A. Cobham characterized p-recognizable sets of integers in terms of constant length substitutions. It turns out to be mainly the same for abstract numeration systems (this is the purpose of Section 3). Hence we will often say that a set of integers recognizable with respect to some abstract numeration system is *generated by a substitution*. This will enable us to solve the syndeticity problem for abstract numeration systems in terms of substitutions. Let us also observe that with the formalism of substitutions and in connection with the constructions of Section 3, Cobham's theorem obtained in [Du1] can be directly translated for a large class of abstract numeration systems (namely, those giving rise to substitutions of exponential growth satisfying the assumptions of [Du1]).

This work has been supported by a grant "Accord de coopération CNRS/CGRI-FNRS, Commissariat Général des Relations Internationales de la Communauté Française".

In [Co1, Co2, Du1, Du2], the involved substitutions σ are (exponentially) growing, meaning that the length of $\sigma^n(a)$ goes to infinity with n, for all letters a. (This implies in particular that one of the letter is of exponential growth and that none of them has polynomial growth.) The substitutions corresponding to abstract numeration systems do not have this latter property: they can be non-growing (in the polynomial case) and even worse, erasing. We take care of this extra difficulty in Section 4.
The notion of multiplicatively independent integers can be generalized to these substitutions by considering the maximal growth rate of the letters and we are thus able to define "independent" substitutions. Our main result (Theorem 17) can be roughly stated as follows:
If a set of integers E is generated by two independent substitutions (one having exponential growth), then E is syndetic. We are not able to give a complete proof in the case of two independent substitutions both having polynomial growth.
To conclude this paper, we obtain easily from the syndeticity an analogue of Cobham's theorem for two substitutions (or equivalently for two abstract numeration systems): one of exponential growth and the other one of polynomial growth. Combined with the main result of [Du1], an extended version of Cobham's theorem follows.

2. Words, morphisms, substitutions and numeration systems

The aim of this section is just to recall classical definitions and notation.

2.1. **Words and sequences.** An *alphabet* A is a finite set of elements called *letters*. A *word* over A is an element of the free monoid generated by A, denoted by A^*. Let $x = x_0x_1 \cdots x_{n-1}$ (with $x_i \in A$, $0 \le i \le n-1$) be a word, its *length* is n and is denoted by $|x|$. The number of occurrences of a letter $a \in A$ in the word w is denoted $|w|_a$ and if E is a subset of A, then $|w|_E$ is a shorthand for $\sum_{e\in E} |w|_e$. The *empty word* is denoted by ϵ, $|\epsilon| = 0$. The set of non-empty words over A is denoted by A^+. The elements of $A^{\mathbb{N}}$ are called *sequences*. If $x = x_0x_1 \cdots$ is a sequence (with $x_i \in A$, $i \in \mathbb{N}$) and $I = [k, l]$ an interval of $\mathbb{N}$ we set $x_I = x_kx_{k+1} \cdots x_l$ and we say that x_I is a *factor* of x. If $k = 0$, we say that x_I is a *prefix* of x. The set of factors of length n of x is written $L_n(x)$ and the set of factors of x, or the *language* of x, is noted $L(x)$. The *occurrences* in x of a word u are the integers i such that $x_{[i,i+|u|-1]} = u$. When x is a word, we use the same terminology with similar definitions.
The sequence x is *ultimately periodic* if there exist a word u and a non-empty word v such that $x = uv^\omega$, where $v^\omega = vvv\cdots$. Otherwise we say that x is *non-periodic*. It is *periodic* if u is the empty word. A sequence x is *uniformly recurrent* if every factor of x appears infinitely often in x and for each factor u the greatest difference of two successive occurrences of u is bounded.

2.2. **Morphisms and matrices.** Let A and B be two alphabets. A *morphism* τ is a map from A to B^*. Such a map induces by concatenation a morphism from A^* to B^*. If $\tau(A)$ is included in B^+, it induces a map from $A^{\mathbb{N}}$ to $B^{\mathbb{N}}$. These two maps are also called τ. With the morphism τ 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 $\tau(j)$. Let M be a square matrix, we call *dominant eigenvalue of* M an eigenvalue r such that the modulus of all the other eigenvalues do not exceed the modulus of r. A square matrix is called *primitive* if it has a power with positive coefficients. In this case the dominant eigenvalue is unique, positive and it is a simple root of the characteristic polynomial. This is Perron-Frobenius Theorem (see for instance [LM]).

2.3. **Substitutions and substitutive sequences.** A *substitution* is a morphism $\tau : A \to A^*$. In all this paper, and without exception, a substitution τ is assumed to fulfill the following hypothesis : There exists a letter $a \in A$ with

(1) $\lim_{n\to+\infty} |\tau^n(a)| = +\infty$ and
(2) $\tau(a) = au$ for some $u \in A^*$.

Whenever the matrix associated to τ is primitive we say that τ is a *primitive substitution.* We say τ is a *growing* substitution if $\lim_{n\to+\infty} |\tau^n(b)| = +\infty$ for all $b \in A$. We say τ is *erasing* if there exists $b \in A$ such that $\tau(b)$ is the empty word. A fixed point of τ is a sequence $x = (x_n; n \in \mathbb{N})$ such that $\tau(x) = x$. We say it is a *proper fixed point* if all letters of A have an occurrence in x. We observe that all proper fixed points of τ have the same language. Notice that each substitution has at least one proper fixed point. Let x be a proper fixed point of τ. We define

$$L(\tau) = \left\{ x_{[i,j]}; i, j \in \mathbb{N}, i \leq j \right\}.$$

Example 1. The substitution τ defined by $\tau(a) = aaab$, $\tau(b) = bc$ and $\tau(c) = b$ has two fixed points, one is starting with the letter a and is proper and the other one is starting with the letter b and is not proper.

Let B be another alphabet and $y \in B^{\mathbb{N}}$. Let $\mathcal{S}$ be a set of substitutions. We say that y is *substitutive in* $\mathcal{S}$ if $y = \phi(x)$ where $x \in A^{\mathbb{N}}$ is a proper fixed point of $\tau \in \mathcal{S}$ and $\phi : A \to B^*$ is a *letter-to-letter morphism*, i.e., $\phi(A)$ is a subset of B.

2.4. **Automata.** We assume that the reader has some basic knowledge in automata theory, see for instance [Ei]. A *deterministic finite automaton* over A or simply a DFA is a 5-tuple $\mathcal{M} = (Q, q_0, F, A, \delta)$ where Q is the finite set of states, $q_0 \in Q$ is the initial state, $F \subseteq Q$ is the set of final states and $\delta : Q \times A \to A$ is the (partial) transition function. A DFA is *complete* if δ is a total function. As usual, δ can be naturally extended to $Q \times A^*$. With the DFA $\mathcal{M}$ is associated the matrix $M_{\mathcal{M}} = (m_{i,j})_{i,j\in Q}$ where $m_{i,j} = \#\{a \in A; \delta(j, q) = i\}$.

If L is a regular language then the trim minimal automaton of L is said to be the *canonical automaton* of L. Recall that an automaton is *trim* (or reduced) if it accessible and coaccessible, i.e., every state is reachable from q_0 and every state reaches a final state. Let $\mathcal{M} = (Q, q_0, F, A, \delta)$ be a DFA and $L \subseteq A^*$ be a regular language with $\mathcal{A} = (Q', q_0', F', A, \delta')$ as canonical automaton. Then $\mathcal{M}$ is said to be an L-*automaton* if there exists an onto mapping $\Phi : Q \to Q'$ such that

(1) $\Phi(q_0) = q_0'$,
(2) $\Phi(F) \subseteq F'$,
(3) $\forall q \in Q$, $\forall a \in A$: $\Phi(\delta(q, a)) = \delta'(\Phi(q), a)$.

In the latter condition, if $\delta(q, a)$ is not defined then $\delta'(\Phi(q), a)$ is not defined, and conversely. Notice that this kind of definition can also be found in [BH] where linear numeration systems related to a Pisot number are investigated.

Remark 2. Changing the set of final states in an L-automaton allows this automaton to recognize exactly the language L. With the same notation as before, it suffices to take $\Phi^{-1}(F')$ as set of final states for the L-automaton.

2.5. **Abstract numeration systems.** If the alphabet A is totally ordered then we can enumerate the words of A^* by the *genealogical ordering* defined as follows. Let x, y be two words over A, we say $x < y$ if $|x| < |y|$ or if $|x| = |y|$ and there exist $a, b \in A$, $u, x', y' \in A^*$ such that $a < b$, $x = uax'$ and $y = uby'$. Enumerating the words of an infinite regular language L over a totally ordered alphabet $(A, <)$ by increasing genealogical order gives a one-to-one correspondence between $\mathbb{N}$ and L (see [LR]). We say that the $(n+1)$th word w in the genealogically ordered language L is the representation of n in the *abstract numeration system* $S = (L, A, <)$ and we write $\mathrm{rep}_S(n) = w$. In particular, if E is a subset of $\mathbb{N}$ then $\mathrm{rep}_S(E)$ is a subset of L. We say that E is *S-recognizable* if $\mathrm{rep}_S(E)$ is a regular language. The *characteristic sequence* of E is the sequence $\chi_E = x_0 x_1 \cdots \in \{0,1\}^{\mathbb{N}}$ such that $x_i = 1$ if and only if i belongs to E.

Example 3. Let $A = \{0, \ldots, k-1\}$ for some $k \geq 2$. The language

$$L = \{\epsilon\} \cup \{1, \ldots, k-1\}\{0, \ldots, k-1\}^*$$

genealogically ordered with the usual ordering of the digits gives the classical k-ary system. Let $B = \{0, 1\}$. Enumerating the words of $M = \{\epsilon\} \cup 1\{0, 01\}^*$ gives exactly the Fibonacci system. These two examples are special cases of linear numeration systems whose characteristic polynomial is the minimal polynomial of a Pisot number (in this setting, it is well known that the language of the numeration is regular [BH]). All the systems of this kind are therefore special cases of abstract numeration systems.

Example 4. Let us consider an abstract numeration system which is no more positional (i.e., not built on a strictly increasing sequence of integers). Let $A = \{a, b\}$ with $a < b$. The first words of $L = a^*b^*$ enumerated by genealogical order are

$$\epsilon, a, b, aa, ab, bb, aaa, aab, abb, bbb, aaaa \ldots$$

For instance, $\mathrm{rep}_S(5) = bb$ and $\mathrm{rep}_S^{-1}(a^*) = \{0, 1, 3, 6, 10, \ldots\} = E_a$ is an S-recognizable subset of $\mathbb{N}$ (formed of triangular numbers). For such a system, $\mathrm{rep}_S^{-1}(a^p b^q) = \frac{1}{2}(p+q)(p+q+1) + q$ and we cannot mimic positional systems where one can define "weight" to the "digits" a and b. Moreover we can already notice that $\#(L \cap A^n) = n+1$ has a polynomial behavior (contrasting with systems built on Pisot numbers which always have an exponential behavior).

3. The link between substitutions and numeration systems.

In this section, we associate a substitution σ to any S-recognizable set E of integers for a given abstract numeration system S. One of the fixed point z of σ is such that $f(z) = \chi_E$ for some (possibly erasing) morphism f.

Lemma 5. *Let $S = (L, A, <)$ be a numeration system. A set $E \subset \mathbb{N}$ is S-recognizable if and only if $\mathrm{rep}_S(E)$ is accepted by an L-automaton.*

Proof. Assume that E is S-recognizable. So there exists a complete and accessible DFA $\mathcal{M} = (Q, q_0, F, A, \delta)$ accepting exactly $\mathrm{rep}_S(E)$. We denote by $\mathcal{A} = (Q', q_0', F', A, \delta')$ the canonical automaton of L. Consider the "product" automaton

$$\mathcal{P} = (Q' \times Q, (q_0', q_0), F' \times F, A, \mu)$$

where the transition function μ is defined, for all $(q,r) \in Q' \times Q$ and all $a \in A$ such that $\delta'(q,a)$ exists, by

$$\mu((q,r),a) = (\delta'(q,a), \delta(r,a)).$$

Clearly, $\mathcal{P}$ is an L-automaton accepting $\mathrm{rep}_S(E)$. It suffices to consider the application $\Phi : Q' \times Q \to Q'$ mapping (q,r) onto q. $\square$

Definition 6. Let $A = \{a_1 < \cdots < a_k\}$ be a totally ordered alphabet. To any DFA $\mathcal{M} = (Q, q_0, F, A, \delta)$, if $s \notin Q$ then one can associate a substitution $\sigma_{\mathcal{M}} : Q \cup \{s\} \to (Q \cup \{s\})^*$ defined by

$$\sigma_{\mathcal{M}} : \begin{cases} s & \mapsto s\, q_0 \\ q & \mapsto \delta(q,a_1) \cdots \delta(q,a_k),\ \forall q \in Q \end{cases}$$

where in the last expression, if $\delta(q,a)$ is not defined for some a, then it is replaced by ϵ. Observe that $\sigma_{\mathcal{M}}$ can be erasing. This kind of substitution was introduced for instance in [RM]. The substitution associated to the canonical automaton of L is said to be the *canonical substitution* of L and is denoted σ_L.

Let L be a regular language and $\sigma_L : B \to B^*$ be its canonical substitution and let $\tau : A \to A^*$ be a substitution. If there exists an onto mapping $\Phi : A \to B$ such that for all $a \in A$,

$$\Phi(\tau(a)) = \sigma_L(\Phi(a))$$

then τ is said to be an *L-substitution*. Clearly, if $\mathcal{M}$ is an L-automaton then $\sigma_{\mathcal{M}}$ is an L-substitution.

Proposition 7. *Let $S = (L, A, <)$ be a numeration system and $E \subset \mathbb{N}$ be an S-recognizable set. Then there exists an L-substitution $\sigma : B \to B^*$ having $x \in B^\omega$ as fixed point and a morphism $f : B \to \{0,1\} \cup \{\epsilon\}$ such that*

$$f(x) = \chi_E.$$

Proof. We denote by $\mathcal{A} = (Q', q_0', F', A, \delta')$ the canonical automaton of L. By Lemma 5, $\mathrm{rep}_S(E)$ is accepted by some L-automaton $\mathcal{M} = (Q, q_0, F, A, \delta)$. Let $\Phi : Q \to Q'$ be the mapping related to the L-automaton and let F'' be the set of states of $\mathcal{M}$ given by

$$F'' = \Phi^{-1}(F').$$

Observe that $F \subseteq F'' \subseteq Q$. Consider the alphabet $B = Q \cup \{s\}$ ($s \notin Q$), the L-substitution $\sigma_{\mathcal{M}} : B \to B^*$ and the mapping $f : B \to \{0,1\} \cup \{\epsilon\}$ defined by

$$f : \begin{cases} s & \mapsto \epsilon; \\ q & \mapsto 1,\ \text{if } q \in F; \\ q & \mapsto 0,\ \text{if } q \in F'' \setminus F; \\ q & \mapsto \epsilon,\ \text{if } q \in Q \setminus F''. \end{cases}$$

It is easy to show that $\lim_{n\to\infty} f(\sigma_{\mathcal{M}}^n(s)) = \chi_E$. $\square$

Example 8. We continue Example 4. The canonical automaton of $L = a^*b^*$ has two states A and B such that $\delta(A,a) = A$, $\delta(A,b) = B$ and $\delta(B,b) = b$. Proceeding as in Definition 6, we get the substitution

$$\sigma_{\mathcal{M}} : s \mapsto sA,\ A \mapsto AB,\ B \mapsto B$$

having

$$w = sAABABBABBBABBBBABBBBB \cdots$$

as fixed point. Applying the morphism $f : s \mapsto \epsilon, A \mapsto 1, B \mapsto 0$ to this word w, we get the characteristic sequence of the S-recognizable set $E_a = \{0, 1, 3, 6, 10, \ldots\}$.

4. Growth type and erasures

In this section, we first consider the growth order of the length of the iterates of a substitution for any letter. From this we define the notion of growth order of a letter. Then we give arguments that allow us to get rid of erasing substitutions. In the third part of this section, we exhibit sub-alphabets which are invariant for the substitution. All of these results will play an important rôle in the proof of our main result.
Finally, we consider the relationship of the growth order of the substitution with abstract numeration systems. This will lead to an easy adaptation of Cobham's theorem given in terms of substitutions to these abstract numeration systems.

4.1. **Growth type.**

In this subsection we recall some lemmata and definitions appearing in [Du1].
Notice that in the following lemma, the substitutions σ and τ can be erasing. As we will see in the detailed proof of the result, the technical procedure of replacing τ with one of its power, allows us to get rid of irreducible components to the benefit of irreducible ones.

Lemma 9. *Let $\tau : A \to A^*$ be a substitution on the finite alphabet A. There exists p such that for $\sigma = \tau^p$ and for all $a \in A$, one of the following two situations occurs, either*

$$\exists N \in \mathbb{N} : \forall n > N,\ |\sigma^n(a)| = 0,$$

or there exist $d(a) \in \mathbb{N}$ and algebraic numbers $c(a), \alpha(a)$ such that

$$\lim_{n\to+\infty} \frac{|\sigma^n(a)|}{c(a)\, n^{d(a)}\, \alpha(a)^n} = 1.$$

Moreover, if the latter situation occurs then for all $i \in \{0, \ldots, d(a)\}$ there exists a letter $b \in A$ appearing in $\sigma^j(a)$ for some $j \in \mathbb{N}$ and such that

$$\lim_{n\to+\infty} \frac{|\sigma^n(b)|}{c(b)\, n^{i}\, \alpha(a)^n} = 1.$$

Proof. With σ we associate an automaton $\mathcal{A}_\sigma$ in the classical way: the set of states of $\mathcal{A}_\sigma$ is A, the alphabet is $\{1, \ldots, \max_{a\in A} |\sigma(a)|\}$ and the transition function δ is defined as follows. If b appears in $\sigma(a)$ at position $i \geq 1$ then $\delta(a, i) = b$. Notice that $\delta(a, k)$ is not defined if $k > |\sigma(a)|$. So $\mathcal{A}_\sigma$ is possibly not a complete automaton. From the definition of $\mathcal{A}_\sigma$, it follows that $|\sigma^n(a)|$ is exactly the number of paths of length n in $\mathcal{A}_\sigma$ starting from a.
We write $a \to b$ if there exists a path in $\mathcal{A}_\sigma$ from a to b. We define an equivalence relation $\sim$ over A as follows. We define for all $a, b \in A$,

$$a \sim b \Leftrightarrow (a = b) \text{ or } (a \to b \text{ and } b \to a).$$

As usual, an equivalence class for $\sim$ is said to be a *communicating class*. Proceeding as in [LM, p. 119], the communicating classes and the corresponding states of

$\mathcal{A}_\sigma$ can be ordered in such a way that the matrix associated with σ has a block triangular form

$$M_\sigma = \begin{pmatrix} M_1 & 0 & 0 & \dots & 0 \\ * & M_2 & 0 & \dots & 0 \\ * & * & M_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \dots & M_k \end{pmatrix}. \tag{4.1}$$

We denote by C_j the communicating class related to M_j. Each $M_j \neq 0$ is irreducible. Let p_j be the corresponding period (i.e., the smallest integer t such that $(M_j)^t$ has positive entries on the main diagonal). Let $p = \mathrm{lcm}_{j=1,\dots,k}\, p_j$. Replacing σ with σ^p does not affect its fixed point. The communicating classes C_j related to primitive blocks M_j are the same in $\mathcal{A}_\sigma$ and $\mathcal{A}_{\sigma^p}$ but each communicating class in $\mathcal{A}_\sigma$ related to a nonzero block which is not primitive is split into several communicating classes in $\mathcal{A}_{\sigma^p}$ related to primitive blocks (see for instance [LM, Section 4.5]). Assuming that σ has been replaced by σ^p (this has no consequence for the rest of this paper because we are mainly interested in the fixed points of σ, so we may assume that the substitutions we consider have such a property), we may assume in what follows that each M_j's appearing in (4.1) is either primitive or zero. Let α_j be the Perron-Frobenius eigenvalue associated with $M_j \neq 0$. If $M_j = 0$, we set $\alpha_j = 0$. One can already notice that α_j is algebraic since M_j has only integer entries. Notice also that $\alpha_j = 1 \Leftrightarrow M_j = (1)$. The number of words of length n starting from and ending to a state related to M_j is of the form $\sim c_j\, \alpha_j^n$. Since c_j can be computed from left and right Perron eigenvectors of M_j (see [LM, Thm 4.5.12]), it is clear that c_j is an algebraic number (computations take place in $\mathbb{Q}(\alpha_j)$).

We now estimate the number $|\sigma^n(a)|$ of paths of length n in $\mathcal{A}_\sigma$ starting from a given state a belonging to C_k. In the graph of the communicating classes (we use once again the terminology of [LM, p. 119]), consider the set $\mathcal{P}_k$ of all paths starting in C_k and ending in a leaf. Let $C_{k,0} = C_k, C_{k,1}, \dots, C_{k,\ell}$ be such a path $\mathfrak{p}$ (we will only consider classes such that $M_{k,i} \neq 0$, if no such a class exists then the corresponding number of words of length n is zero for n large enough). The contribution of $\mathfrak{p}$ to $|\sigma^n(a)|$ is

$$\sim c_{k,0} \dots c_{k,\ell} \sum_{n_0+\dots+n_\ell=n} \alpha_{k,0}^{n_0} \cdots \alpha_{k,\ell}^{n_\ell}$$

Let

$$\beta = \max_{i=0,\dots,\ell} \alpha_{k,i}$$

and $C_{k,j_1}, \dots, C_{k,j_t}$ be the communicating classes having β as Perron-Frobenius eigenvalue, $t \geq 1$. Therefore the contribution of $\mathfrak{p}$ to $|\sigma^n(a)|$ is

$$\sim c_{k,0} \dots c_{k,\ell}\, n^{t-1}\, \beta^n.$$

In particular, it follows that the Jordan-decomposition of the incidence matrix of $\mathcal{A}_\sigma$ restricted to the states occurring in $\mathfrak{p}$ contains a Jordan block of size t for the eigenvalue β. To conclude the first part of the proof, we just have to sum expressions like the one obtained above for all paths in $\mathcal{P}_k$.

The particular case is immediate, with the same notation as above, if b belongs to C_{k,j_m}, $m \in \{2, \dots, t\}$, then the contribution of $\mathfrak{p}$ to $|\sigma^n(b)|$ is proportional to

$n^{t-m}\beta^n$. Moreover, since a belongs to $C_{k,0}$, it is clear that $a \to b$, i.e., there exists j such that b appears in $\sigma^j(a)$. $\square$

Notice that the following definition is mainly relevant for non-erasing substitutions (and the next subsection allows us to only consider such substitutions).

Definition 10. Let σ be a non-erasing substitution possibly replaced by a convenient power as in the proof of the previous lemma. For all $a \in A$ we will call *growth type* of a the couple

$$(d(a), \alpha(a))$$

as introduced in the previous lemma. If (d, α) and (e, β) are two growth types we say that (d, α) is *less than* (e, β) (or $(d, \alpha) < (e, \beta)$) whenever $\alpha < \beta$ or, $\alpha = \beta$ and $d < e$.

Consequently if the growth type of $a \in A$ is less than the growth type of $b \in A$ then $\lim_{n \to +\infty} |\sigma^n(a)|/|\sigma^n(b)| = 0$. We say that $a \in A$ is a *growing letter* if

$$(d(a), \theta(a)) > (0, 1)$$

or equivalently, if $\lim_{n \to +\infty} |\sigma^n(a)| = +\infty$.
We set

$$\Theta := \max\{\theta(a) \mid a \in A\}, \quad D := \max\{d(a) \mid \theta(a) = \Theta \mid a \in A\}$$

and $A_{max} := \{a \in A \mid \theta(a) = \Theta, d(a) = D\}$. The dominant eigenvalue of M is Θ. We will say that the letters of A_{max} are *of maximal growth* and that (D, Θ) is the *growth type* of σ. Consequently, we say that a substitutive sequence y is (D, Θ)-*substitutive* if the underlying substitution is of growth type (D, Θ).
Observe that if $\Theta = 1$, then in view of the last part of Lemma 9, there exists at least one non-growing letter of growth type $(0, 1)$. Otherwise stated, if a letter has a polynomial growth, then there exists at least one non-growing letter. Consequently σ is growing (i.e., all its letters are growing) if and only if $\theta(a) > 1$ for all $a \in A$.
We define

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

where $c : A \to \mathbb{R}_+$ is defined in Lemma 9 and $\mathbf{1}_A$ is the usual characteristic function of the set A. From Lemma 9 we deduce the following lemma.

Lemma 11. *For all $u \in A^*$ we have $\lim_{n \to +\infty} |\sigma^n(u)|/n^D\Theta^n = \lambda_\sigma(u)$.*

We say that the word $u \in A^*$ is of *maximal growth* if $\lambda_\sigma(u) \neq 0$.

Corollary 12. *For all $k \geq 1$, the growth type of σ^k is (D, Θ^k).*

4.2. **Erasing morphisms.** In view of Proposition 7, we will have to deal with erasing substitutions and also with erasing morphisms. The following two propositions show how to get rid of the erasing behavior.

Proposition 13. *Let x be a proper fixed point of a substitution $\sigma : A \to A^*$ with growth type (D, Θ). Then, there exists a non-erasing substitution $\tau : C \to C^*$ with a proper fixed point y, a letter-to-letter morphism $\psi : C \to A$ and a morphism $\phi : A \to C^*$ verifying*

1. *$x = \psi(y)$;*
2. *There exists $l \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have $\tau^n \circ \phi = \phi \circ \sigma^{ln}$;*
3. *Each line and each column of the matrix of ϕ has a non-zero coefficient;*

(4) *The growth type of τ is (D, Θ^l).*

Proof. The statement (1), (2) and (3) can be found in [AS, Theorem 7.5.1, p. 227] and (4) is a consequence of (2) and (3). □

Proposition 14. *Let x be a proper fixed point of a substitution $\sigma : A \to A^*$ with growth type (D, Θ), $\Gamma \subset A$ and $\zeta : A \to A \setminus \Gamma$ defined by $\zeta(a) = \epsilon$ if $a \in \Gamma$ and a otherwise. Then, there exists a non-erasing substitution $\tau : C \to C^*$ with a proper fixed point y, a letter-to-letter morphism $\psi : C \to A$ and a morphism $\phi : A \to C^*$ verifying*

(1) *$\zeta(x) = \psi(y)$;*
(2) *There exists $l \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have $\psi \circ \tau^n \circ \phi = \zeta \circ \sigma^{l(n+1)}$;*
(3) *Each line and each column of the matrix of ϕ has a non-zero coefficient;*
(4) *The growth type of τ is (D, Θ^l).*

Proof. The statement (1), (2) and (3) can be found in [AS, pp. 232–236] and (4) is a consequence of (2) and (3). □

4.3. **Invariant alphabets.** Let $\Delta(w) \subseteq A$ be the set of letters having an occurrence in the word $w \in A^*$.

Lemma 15. *Let $\sigma : A \to A^*$ be a non-erasing substitution. There exists $N \geq 1$ such that for all $a \in A$ and all $n \geq 1$,*

$$\Delta((\sigma^N)^n(a)) = \Delta(\sigma^N(a)).$$

Proof. We set $A = \{a_1, \ldots, a_{|A|}\}$. The alphabet A being finite, the sequence of sub-alphabets $(\Delta(\sigma^n(a_1)))_{n\in\mathbb{N}}$ is ultimately periodic, i.e., there exist p and q such that

$$\Delta(\sigma^{q+np+i}(a_1)) = \Delta(\sigma^{q+mp+i}(a_1))$$

for all $m, n, i \in \mathbb{N}$. Hence, for k such that $kp \geq q$, we have for all $n \geq 1$

$$\Delta((\sigma^{kp})^n(a_1)) = \Delta(\sigma^{kp}(a_1)).$$

Now take a_2 and consider σ^{kp}. Proceeding as before we find r such that

$$\Delta((\sigma^r)^n(a_1)) = \Delta(\sigma^r(a_1)) \text{ and } \Delta((\sigma^r)^n(a_2)) = \Delta(\sigma^r(a_2)).$$

We conclude continuing like this with $a_3, \ldots, a_{|A|}$. □

The following corollary is just a reformulation of the previous lemma.

Corollary 16. *Let $\sigma : A \to A^*$ be a non-erasing substitution. There exists $N \geq 1$ such that for all $a, b \in A$ and $n \geq 1$*

$$a \in A \text{ appears in } (\sigma^N)^n(b) \text{ if and only if } a \text{ appears in } (\sigma^N)^{n+1}(b).$$

Replacing σ by one of its power σ^N does not alter its fixed points (we will use this argument repeatedly). Therefore we will often require that σ has the following property:

(4.2)
$$\forall a, b \in A,\ \forall n \geq 1, \quad a \in A \text{ appears in } \sigma^n(b) \text{ if and only if } a \text{ appears in } \sigma^{n+1}(b).$$

4.4. **Linking the growth order with numeration systems.** Let L be a regular language having $\mathcal{A} = (Q', q'_0, F', A, \delta')$ as canonical automaton and $M_{\mathcal{A}}$ as associated matrix.

(1) As in section 4.1, for all states $q' \in Q'$ we can define the growth type (d, α) of q' (corresponding to the number of words of length n accepted in $\mathcal{A}$ from q') and consequently, we can define the growth type of $\mathcal{A}$ as the largest growth type of the states in Q'.

(2) If $\mathcal{M} = (Q, q_0, F, A, \delta)$ is an L-automaton then $\mathcal{M}$ and $\mathcal{A}$ have the same growth type. Indeed, for any $q' \in Q'$, we denote by $p_{\mathcal{A},q'}(n)$ the number of paths of length n in $\mathcal{A}$ starting in q'. If $\Phi : Q \to Q'$ is the mapping defining the L-automaton, then for any $q \in \Phi^{-1}(q')$,
$$p_{\mathcal{A},q'}(n) \geq p_{\mathcal{M},q}(n)$$
and also
$$p_{\mathcal{A},q'}(n) \leq \sum_{q \in \Phi^{-1}(q')} p_{\mathcal{M},q}(n).$$
This means that q' and at least one of the states $q \in \Phi^{-1}(q')$ are of the same growth type and that none of the states $q \in \Phi^{-1}(q')$ is of a larger growth type than q'.

(3) If $\mathcal{M}$ is of growth type (D, Θ), $\Theta > 1$, then $\sigma_{\mathcal{M}}$ is of the same growth type. But notice that if $\mathcal{M}$ is of growth type $(D, 1)$ then $\sigma_{\mathcal{M}}$ is of growth type $(D + 1, 1)$.

As a consequence of theses observations, if L is a regular language having a canonical automaton of growth type (D, Θ), $\Theta > 1$, (resp. $(D, 1)$, $D \geq 1$) and if $E \subset \mathbb{N}$ is S-recognizable for the numeration system $S = (L, A, <)$ then from Propositions 7, 13 and 14 the sequence χ_E is (D, Θ^l)-substitutive for some l (resp. $(D + 1, 1)$-substitutive). This obersevation will be helpful in the last section of this paper (Corollary 27 and Remark 28).

5. The words appear with bounded gaps

This section is devoted to the proof of the main result of this paper:

Theorem 17. *Let $d, e \in \mathbb{N} \setminus \{0\}$ and $\alpha, \beta \in [1, +\infty[$ such that $(d, \alpha) \neq (e, \beta)$ and satisfying one of the following three conditions:*

(1) *α and β are multiplicatively independent;*
(2) *$\alpha, \beta > 1$ and $d \neq e$;*
(3) *$(\alpha, \beta) \neq (1, 1)$ and, $\beta = 1$ and $e \neq 0$, or, $\alpha = 1$ and $d \neq 0$;*

Let C be a finite alphabet. If $x \in C^{\mathbb{N}}$ is both (d, α)-substitutive and (e, β)-substitutive then the letters of C which have infinitely many occurrences in x appear in x with bounded gaps.

For the proof of this result we will proceed into three parts. The first part consists of arithmetical lemmata about density in $\mathbb{R}$. In the second part we give bounds for gaps created by some letters. In subsections 5.3 and 5.4 we exhibit an important sequence of integers and we fix some useful constants. Finally from subsection 5.5 to 5.8 we proceed to a case study depending on the growth order of the considered substitutions. Let us first fix the context we will be dealing with.

Let σ and τ be two substitutions on the alphabets A and B, with fixed points y and z and with growth types (d,α) and (e,β) respectively. Taking powers of σ and τ does not alter the fixed points y and z and does not change the multiplicative dependence. Thus, in the proof we will sometimes replace the substitution by some convenient power of itself (and this also allows us to assume that condition (4.2) is satisfied). In particular, when α and β are multiplicatively dependent we may suppose that $\alpha = \beta$.
Let $\phi : A \to C$ and $\psi : B \to C$ be two letter-to-letter morphisms such that $\phi(y) = \psi(z) = x$. Lemma 13 allows us to suppose that σ and τ are non-erasing. We call A_+ the set of growing letters of A with respect to σ.

5.1. **Some density lemmata.** Recall that $\alpha, \beta \in [1, +\infty[$ are *multiplicatively independent* whenever $\alpha^k = \beta^\ell$, $\ell, k \in \mathbb{N}$, implies $k = 0$ or $\ell = 0$. In [Du1, Corollary 11] the following result is proved. Observe that this result is well known when $d = e = 0$ (and is sometimes stated as a Kronecker's theorem). Moreover it does not take into account the case $\alpha = 1$ or $\beta = 1$.

Theorem 18. *Let α and β be multiplicatively independent elements of $]1, +\infty[$. Let d and e be non-negative integers. Then the set*

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

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

Lemma 19. *Let $d, e \in \mathbb{N}$ and $\alpha \in]1, +\infty[$. Then,*

(1) *$d, e \geq 1$ if and only if the set $\left\{ \frac{n^d}{m^e} ; n, m \in \mathbb{N} \right\}$ is dense in $\mathbb{R}_+$;*
(2) *$e \neq 0$ if and only if $\left\{ \frac{\alpha^n n^d}{m^e} ; n, m \in \mathbb{N} \right\}$ is dense in $\mathbb{R}_+$;*
(3) *$d \neq e$ if and only if $\left\{ \frac{\alpha^n n^d}{\alpha^m m^e} ; n, m \in \mathbb{N} \right\}$ is dense in $\mathbb{R}_+$.*

Proof. (1) Suppose $d, e \geq 1$. Let $l \in \mathbb{R}_+ \setminus \{0\}$ and $\epsilon > 0$. It suffices to find $n, m \in \mathbb{N}$ such that $|l - n^d/m^e| < \epsilon$.
Let $m \in \mathbb{N}$ be such that $\max(d, 2^d l/\epsilon) < (lm^e)^{1/d} - 1$ and $1/m^e < l$. There exists $n \in \mathbb{N}$ such that $n^d/m^e < l \leq (n+1)^d/m^e$. We observe that this implies that $n > d$ and $2^d l/n < \epsilon$. Consequently, we get

$$0 < l - \frac{n^d}{m^e} \leq \frac{(n+1)^d - n^d}{m^e} \leq \frac{2^d n^{d-1}}{m^e} = \frac{2^d}{n} \frac{n^d}{m^e} < \frac{2^d l}{n} < \epsilon.$$

Hence the set $\{n^d/m^e; n, m \in \mathbb{N}\}$ is dense in $\mathbb{R}_+$.

(2) Suppose $e \neq 0$. Let $l \in \mathbb{R}_+ \setminus \{0\}$ and $\epsilon > 0$. It suffices to find $n, m \in \mathbb{N}$ such that $|l - \alpha^n n^d/m^e| < \epsilon$.
Let $m_0 \in \mathbb{N}$ be such that $e \ln(1 + 1/m_0) < \ln(1 + \epsilon/l)$. Let n be such that $e \ln(m_0) < d \ln(n) + n \ln(\alpha) - \ln(l)$ and $m \geq m_0$ be such that $e \ln(m) \leq d \ln(n) + n \ln(\alpha) - \ln(l) \leq e \ln(m+1)$. Then we have

$$0 \leq d \ln(n) + n \ln(\alpha) - \ln(l) - e \ln(m) \leq e \ln\left(1 + \frac{1}{m}\right) < \ln\left(1 + \frac{\epsilon}{l}\right).$$

Hence the set $\{\alpha^n n^d/m^e; n, m \in \mathbb{N}\}$ is dense in $\mathbb{R}_+$.

(3) Suppose $d \neq e$. Let $l \in \mathbb{R}_+ \setminus \{0\}$ and $\epsilon > 0$. It suffices to find $n, m \in \mathbb{N}$ such that $|l - n^d\alpha^n / m^e\alpha^m| < \epsilon$.
We can suppose $d > e$ because $\{\alpha^n n^d/\alpha^m m^e; n, m \in \mathbb{N}\}$ is dense in $\mathbb{R}_+$ if and only if $\{\alpha^m m^e/\alpha^n n^d; n, m \in \mathbb{N}\}$ is dense in $\mathbb{R}_+$.
Let n_0 be such that

$$\frac{d-e}{2\ln\alpha}\ln n_0 \leq \frac{d-e}{\ln\alpha}\ln n_0 - \frac{\ln(l+\epsilon)}{\ln\alpha} \leq \frac{d-e}{\ln\alpha}\ln n_0 - \frac{\ln(l)}{\ln\alpha} \leq \frac{3(d-e)}{2\ln\alpha}\ln n_0.$$

Choose b_0 with $(\epsilon\alpha^{b_0})^{\frac{1}{d-e}} \geq 1$. Then for all $b \geq b_0$ there exists n_b such that

$$l\alpha^b \leq n_b^{d-e} \leq (l+\epsilon)\alpha^b.$$

The sequence (n_b) goes to infinity, consequently we can choose b and n such that $n = n_b \geq n_0$ and $1 - \epsilon/l \leq (n/n+b)^e$. Then we have

$$\frac{d-e}{2\ln\alpha}\ln n \leq b \leq \frac{3(d-e)}{2\ln\alpha}\ln n$$

Now consider $m = n + b$. This gives

$$l - \epsilon \leq l\left(\frac{n}{b+n}\right)^e \leq \frac{n^d\alpha^n}{m^e\alpha^m} \leq (l+\epsilon)\left(\frac{n}{b+n}\right)^e \leq l + \epsilon.$$

Suppose $d = e$. If $n \leq m$ then $\alpha^n n^d/\alpha^m m^e \leq 1$ and if $n > m$ then $\alpha^n n^d/\alpha^m m^e \geq \alpha$. This concludes the proof. □

Corollary 20. *Let $d, e \in \mathbb{N}$ and $\alpha, \beta \in [1, +\infty[$. We set*

$$\Omega = \left\{\frac{\alpha^n n^d}{\beta^m m^e}; n, m \in \mathbb{N}\right\}.$$

Then Ω is dense in $\mathbb{R}_+$ if and only if one of the following two conditions holds:

1. *α and β are multiplicatively independent.*
2. *$\alpha, \beta > 1$ and $d \neq e$.*
3. *$\beta = 1$ and $e \neq 0$, or, $\alpha = 1$ and $d \neq 0$;*

Proof. It follows from Theorem 18 and Lemma 19. □

We will say that two substitutions are *independent* whenever their respective growth type (d, α) and (e, β) are different and satisfy Hypothesis (1), (2) or (3) in the previous corollary. Notice that in Theorem 17 the assumptions mean that the substitutions are independent and are not both of polynomial growth (this corresponds to the hypothesis $(\alpha, \beta) \neq (1, 1)$).

5.2. **Growth type of gaps.** In this subsection we give two results on the gaps created by the letters of some sub-alphabet in prefixes of fixed points and in iterates of letters. They will be key arguments in the proof of Theorem 17.
Let $E \subset A$. For all $N \geq 1$, we set

$$M(N, x, E) := \max\{k \in \mathbb{N} : \exists i \in [0, N-k+1], |x_{[i,i+k]}|_E = k\}.$$

In what follows, if x and E are clear from the context, we simply write $M(N)$.

Proposition 21. *Let $x = (x_n)_{n\geq 0}$ be a proper fixed point of the non-erasing substitution σ of growth type (d, α) on the finite alphabet A. Assume σ is such that each letter of A has an occurrence in $\sigma(x_0)$ and σ satisfies (4.2).*
Let $E \subset A$. Suppose there exists a letter $e \in A$ such that $\sigma(e) \in E^*$ and call E' the set of all such letters. Let (d', α') be the greatest growth order among the elements of E'. Then, in each of the following situations, there exist two constants $C_1, C_2 > 0$ such that

(1) *If $(\alpha', d') = (\alpha, d)$ then, for all N,*
$$C_1 N \leq M(N) = M(N, x, E) \leq C_2 N.$$
(2) *If $\alpha = \alpha' > 1$ and $d' < d$ then, for all N,*
$$C_1 N(\log N)^{d'-d} \leq M(N) \leq C_2 N(\log N)^{d'-d}.$$
(3) *If $\alpha > \alpha' > 1$ then, for all N,*
$$C_1(\log N)^{d'-d\frac{\log\alpha'}{\log\alpha}} N^{\frac{\log\alpha'}{\log\alpha}} \leq M(N) \leq C_2(\log N)^{d'-d\frac{\log\alpha'}{\log\alpha}} N^{\frac{\log\alpha'}{\log\alpha}}.$$
(4) *If $\alpha > \alpha' = 1$ then for all N,*
$$C_1 \left(\frac{\log N}{\log\alpha}\right)^{d'} \leq M(N) \leq C_2 \left(\frac{\log N}{\log\alpha}\right)^{d'+1}.$$
(5) *If $\alpha = \alpha' = 1$ and $d' < d$ then for all N*
$$C_1 N^{d'/d} \leq M(N) \leq C_2 N^{(d'+1)/d}.$$

Remark 22. Notice that as usual the assumptions on σ made in the statement of Proposition 21 are easily satisfied by taking a convenient power of σ if needed.

Proof. Let $N \in \mathbb{N}$. There exists $n \in \mathbb{N}$ such that

$$|\sigma^{n-1}(x_0)| \leq N \leq |\sigma^n(x_0)|. \tag{5.1}$$

We start proving (1). As there exists a letter $e \in E'$ of maximal growth having an occurrence in $\sigma(x_0)$ (and since (4.2) is satisfied, $\sigma^k(e) \in E^*$, for all $k \geq 1$) we obtain

$$|\sigma^{n-2}(e)| \leq M(N) \leq |\sigma(x_0)| \max_{l\in A} |\sigma^{n-1}(l)|$$

and from Lemma 9 we deduce that there exist two constants C_1 and C_2 not depending on n such that

$$C_1 \alpha^n n^d \leq M(N) \leq C_2 \alpha^n n^d.$$

Let us prove (2). The assertion (3) can be proved following the same arguments. We start proving the left inequality. Proceeding as before we obtain a constant C_1' depending neither on n nor N such that

$$C_1' \alpha^n n^{d'} \leq M(N).$$

Moreover, from (5.1), we deduce there exist two constants C_1'', C_2'' depending neither on n nor N such that

$$C_1'' \log(N) \leq n \leq C_2'' \log(N) \text{ if } \alpha > 1 \text{ and} \tag{5.2}$$

$$C_1'' N^{1/d} \leq n \leq C_2'' N^{1/d} \text{ if } \alpha = 1. \tag{5.3}$$

This together with Lemma 9 gives the left inequality.
Let us prove the right inequality. Let i be such that $|x_{[i,i+M(N)]}|_E = M(N)$. We set $u = x_{[i,i+M(N)]}$. There exist $u_1 \in E'^*$ having an occurrence in $\sigma^{n-1}(x_0)$ and $p_1, s_1 \in E^*$ such that $u = s_1\sigma(u_1)p_1$ and $|s_1|$, $|p_1|$ less than $m = \max\{|\sigma(a)|; a \in E\}$. In the same way there exist $u_2 \in A^*$ having an occurrence in $\sigma^{n-2}(x_0)$ and $p_2, s_2 \in E'^*$ such that $u_1 = s_2\sigma(u_2)p_2$ and $|s_2|$, $|p_2|$ less than m. We remark that $\sigma^2(u_2)$ belongs to E^*. From Hypothesis (4.2), we conclude that u_2 belongs to E'^*. Hence there exist $u_1, \ldots, u_{n-1} \in E'^*$, $p_1, s_1 \in E^*$, $p_2, \ldots, p_{n-1}, s_2, \ldots, s_{n-1} \in E'^*$ such that

$$u = s_1\sigma(s_2)\cdots\sigma^{n-2}(s_{n-1})\sigma^{n-1}(u_{n-1})\sigma^{n-2}(p_{n-1})\cdots\sigma(p_2)p_1, \tag{5.4}$$

$|p_i|$ and $|s_i|$ are less than m. From this expression and Lemma 9 we deduce that there exists a constant C_2''' such that

$$|u| \leq C_2''' \alpha^n n^{d'}. \tag{5.5}$$

We conclude using (5.1) (together with Lemma 9) and (5.2).

We now prove (4). For the left inequality it works as before. For the right inequality it also works as before except that once we obtain the decomposition (5.4) we find some constant C such that $|u| \leq C\sum_{j=1}^{n-1} j^{d'}$. Consequently for some other constant $|u| \leq Cn^{d'+1}$. We conclude using Lemma 9 and (5.2)

For (5) we proceed as in the previous case except we use (5.3). □

We suppose there exists a letter $c \in C$ with infinitely many occurrences in x and that does not appear with bounded gaps in x. Projecting to $\{0,1\}$ we can suppose $C = \{0,1\}$ and $c = 1$. W.l.o.g. we may assume that σ and τ both satisfy (4.2) (as usual taking a power of the substitution does not alter its fixed points). There exist $a \in A$ with infinitely many occurrences in y and a strictly increasing sequence $(p_n)_{n\in\mathbb{N}}$ of positive integers such that the letter c does not appear in $\phi(\sigma^{p_n}(a))$. Let $A(c)$ be the set of such letters. We define $B(c)$ and B_+ as $A(c)$ and A_+ but with respect to τ and B.
The sets $A(c)$ and $B(c)$ are non-empty. Then, there exist a letter $a \in A(c) \cap A_+$ and a letter $b \in B(c) \cap B_+$ having infinitely many occurrences in y and z, with growth type $(d', \alpha') \leq (d, \alpha)$ and $(e', \beta') \leq (e, \beta)$, respectively, where (d', α') and (e', β') are maximal with respect to $A(c)$ and $B(c)$.
Because $M(N, y, \phi^{-1}(\{0\})) = M(N, z, \psi^{-1}(\{0\}))$, from Proposition 21 we deduce that we have necessarily one of the following five situations:

$$\begin{cases} (\alpha', d') = (\alpha, d) \text{ and } (\beta', e') = (\beta, e); \\ \alpha = \alpha' > 1,\ \beta' = \beta \text{ and } d - d' = e - e'; \\ \alpha > \alpha' > 1 \text{ and } \beta > \beta' > 1; \\ \alpha > \alpha' = 1 \text{ and } \beta > \beta' = 1; \\ \alpha = \alpha' = 1,\ d' < d \text{ and } \beta = \beta' = 1,\ e' < e. \end{cases} \tag{5.6}$$

We will consider these cases separately. Before we establish some general facts that will be used in the treatment of these cases.
Let $w = w_0 \cdots w_n$ be a word belonging to $L(y)$ (resp. $L(z)$), we call $\text{gap}(w)$ the largest integer k such that there exists $i \in [0, n-k+1]$ for which the letter c does not appear in $\phi(w_i \cdots w_{i+k-1})$ (resp. in $\psi(w_i \cdots w_{i+k-1})$).
The next lemma is stated for σ but of course it also holds for τ. Moreover we can assume the constant K' is the same for the two substitutions.

Lemma 23. *With notation introduced before, there exists a constant $K^{'}$ such that for all $a'' \in A$ we have:*

$$\text{gap}(\sigma^n(a'')) \leq K^{'} n^{d'} {\alpha'}^n \text{ if } \alpha' > 1 \text{ and}$$
$$\text{gap}(\sigma^n(a'')) \leq K^{'} n^{d'+1} \text{ if } \alpha' = 1$$

for all $n \in \mathbb{N}$.

Proof. It suffices to proceed as we did before to obtain (5.4) and then (5.5). □

From Lemma 9, the following limits exist and are finite and they deserve specific notation

$$\lim_{n\to+\infty} \frac{|\sigma^n(a)|}{n^{d'}{\alpha'}^n} =: \mu(a) \quad \text{and} \quad \lim_{n\to+\infty} \frac{|\tau^n(b)|}{n^{e'}{\beta'}^n} =: \mu(b).$$

5.3. **Some choices when $\alpha, \beta > 1$.** Here we suppose that the set Ω of Corollary 20 is dense in $\mathbb{R}_+$. There exist infinitely many prefixes of y (resp. z) of the type $u_1 a u_2 a'$ (resp. $v_1 b v_2 b'$) fulfilling the conditions $\imath$) and $\imath\imath$) below:

$\imath$) The growth type of $u_1 \in A^*$ and $a' \in A$ (resp. $v_1 \in B^*$ and $b' \in B$) is maximal (Lemma 9 allows such a configuration).
$\imath\imath$) The word u_2 (resp. v_2) does not contain a letter of maximal growth.

We notice this is not the case when the growth type is $(d, 1)$ because in this case there is exactly one letter of growth type $(d, 1)$ and it appears exactly once in the fixed point: this is the first letter of the fixed point.
Let $u_1 a u_2 a'$ be a prefix of y and $v_1 b v_2 b'$ be a prefix of z fulfilling the conditions $\imath$) and $\imath\imath$).
From Corollary 20 there exist four strictly increasing sequences of integers $(m_i)_{\in\mathbb{N}}$, $(n_i)_{\in\mathbb{N}}$, $(p_i)_{\in\mathbb{N}}$ and $(q_i)_{\in\mathbb{N}}$ such that

$$\lim_{i\to+\infty} \frac{n_i^d \alpha^{n_i}}{m_i^e \beta^{m_i}} = \frac{2\lambda_\tau(v_1)}{2\lambda_\sigma(u_1) + 2\lambda_\sigma(a) + \lambda_\sigma(a')} =: \gamma_1 \quad \text{and} \tag{5.7}$$

$$\lim_{i\to+\infty} \frac{p_i^e \beta^{p_i}}{q_i^d \alpha^{q_i}} = \frac{2\lambda_\sigma(u_1)}{2\lambda_\tau(v_1) + 2\lambda_\tau(b) + \lambda_\tau(b')} =: \gamma_2. \tag{5.8}$$

As a consequence of (5.7) and (5.8), we have

$$\lim_{i\to+\infty} \frac{n_i}{m_i} = \frac{\log\beta}{\log\alpha} \quad \text{and} \quad \lim_{i\to+\infty} \frac{p_i}{q_i} = \frac{\log\alpha}{\log\beta}. \tag{5.9}$$

The sequence z has infinitely many occurrences of letters of maximal growth. Hence, in this case, we can take v_1 so long that

$$\frac{2K'(2\gamma_2)^{\frac{\log\alpha'}{\log\alpha}}}{\mu(a)}\cdot\left(\frac{\log\alpha}{\log\beta}\right)^{e'-e\frac{\log\alpha'}{\log\alpha}}<1. \tag{5.10}$$

Using Lemma 11 there exists i_0 such that for all $i \geq i_0$ we have

$$\frac{|\sigma^{n_i}(u_1au_2)|}{|\tau^{m_i}(v_1)|}\leq 1\leq\frac{|\sigma^{n_i}(u_1au_2a')|}{|\tau^{m_i}(v_1bv_2)|}\quad\text{and} \tag{5.11}$$

$$\frac{|\tau^{p_i}(v_1bv_2)|}{|\sigma^{q_i}(u_1)|}\leq 1\leq\frac{|\tau^{p_i}(v_1bv_2b')|}{|\sigma^{q_i}(u_1au_2)|}. \tag{5.12}$$

It comes that the word $\psi(\tau^{m_i}(bv_2))$ (resp. $\phi(\sigma^{q_i}(au_2))$) has an occurrence in $\phi(\sigma^{n_i}(a'))$ (resp. $\psi(\tau^{p_i}(b'))$). To obtain a contradiction it suffices to have some $j \geq i_0$ such that $\mathrm{gap}(\sigma^{n_j}(a'))/\mathrm{gap}(\tau^{m_j}(b)) < 1$ or $\mathrm{gap}(\tau^{p_j}(b'))/\mathrm{gap}(\sigma^{q_j}(a)) < 1$. We observe that $\mathrm{gap}(\tau^{m_j}(b)) = |\tau^{m_j}(b)|$ and $\mathrm{gap}(\sigma^{q_j}(a)) = |\sigma^{q_j}(a)|$. We set $S_j = \mathrm{gap}(\sigma^{n_j}(a'))/|\tau^{m_j}(b)|$ and $T_j = \mathrm{gap}(\tau^{p_j}(b'))/|\sigma^{q_j}(a)|$. Then,

$$\text{it suffices to find some } j \text{ with } S_j<1 \text{ or } T_j<1. \tag{5.13}$$

We have

$$S_i \leq\frac{K'n_i^{d'}{\alpha'}^{n_i}}{\mu(b)m_i^{e'}{\beta'}^{m_i}}\cdot\frac{\mu(b)m_i^{e'}{\beta'}^{m_i}}{|\tau^{m_i}(b)|}\leq\frac{2K'}{\mu(b)}\cdot\frac{n_i^{d'}(\alpha^{n_i})^{\frac{\log\alpha'}{\log\alpha}}}{m_i^{e'}{\beta'}^{m_i}} \tag{5.14}$$

$$\leq\frac{2K'}{\mu(b)}\cdot\frac{n_i^{d'}}{m_i^{e'}{\beta'}^{m_i}}\cdot\left(2\gamma_1\frac{m_i^e\beta^{m_i}}{n_i^d}\right)^{\frac{\log\alpha'}{\log\alpha}} \tag{5.15}$$

$$\leq\frac{2K'(2\gamma_1)^{\frac{\log\alpha'}{\log\alpha}}}{\mu(b)}\cdot\frac{n_i^{d'-d\frac{\log\alpha'}{\log\alpha}}}{m_i^{e'-e\frac{\log\alpha'}{\log\alpha}}}\cdot\exp\left(m_i\left(\frac{\log\alpha'}{\log\alpha}\log\beta-\log\beta'\right)\right) \tag{5.16}$$

and, with the same kind of computations

$$T_i \leq\frac{K'p_i^{e'}\beta'^{p_i}}{\mu(a)q_i^{d'}{\alpha'}^{q_i}}\cdot\frac{\mu(a)q_i^{d'}{\alpha'}^{q_i}}{|\sigma^{q_i}(a)|}\leq\frac{2K'}{\mu(a)}\cdot\frac{p_i^{e'}\beta'^{p_i}}{q_i^{d'}(\alpha^{q_i})^{\frac{\log\alpha'}{\log\alpha}}} \tag{5.17}$$

$$=\frac{2K'(2\gamma_2)^{\frac{\log\alpha'}{\log\alpha}}}{\mu(a)}\cdot\frac{p_i^{e'-e\frac{\log\alpha'}{\log\alpha}}}{q_i^{d'-d\frac{\log\alpha'}{\log\alpha}}}\cdot\exp\left(p_i\left(\log\beta'-\frac{\log\alpha'}{\log\alpha}\log\beta\right)\right). \tag{5.18}$$

5.4. **Remarks when α and β are multiplicatively independent.** In this case we necessarily have $\alpha > 1$ and $\beta > 1$. There exists $K \geq 2$ and j_0 such that for all $i \geq j_0$ we have

$$\frac{1}{K}\leq\frac{n_i}{m_i}\leq K,\quad \frac{1}{K}\leq\frac{p_i}{q_i}\leq K,$$
$$\frac{n_i^d\alpha^{n_i}}{m_i^e\beta^{m_i}}\leq 2\gamma_1,\quad \frac{p_i^e\beta^{p_i}}{q_i^d\alpha^{q_i}}\leq 2\gamma_2,$$
$$\frac{\mu(a)q_i^{d'}\alpha'^{q_i}}{|\sigma^{q_i}(a)|}\leq 2,\quad \frac{\mu(b)m_i^{e'}\beta'^{m_i}}{|\tau^{m_i}(b)|}\leq 2.$$

In the sequel we intensively use the previous inequalities and Lemma 23. We can now proceed to a case study. In view of the hypothesis of Theorem 17, we will not consider the last case occurring in (5.6). Subsections 5.5 to 5.8 correspond to these first four cases.

5.5. $\boxed{(\alpha', d') = (\alpha, d) \textbf{ and } (\beta', e') = (\beta, e)}$**.** We necessarily have $\alpha, \beta > 1$. From (5.18) we get

$$T_i \leq \frac{4\gamma_2 K'}{\mu(a)}$$

and we conclude using (5.10) and the argument (5.13).

5.6. $\boxed{\alpha' = \alpha > 1, d' < d, \textbf{ and } \beta' = \beta > 1\textbf{,}\ e' < e}$**.** From Proposition 21, it comes that $d - d' = e - e'$.

5.6.1. *α and β are multiplicatively independent.* From (5.17) and (5.9) we have

$$T_i \leq \frac{4\gamma_2 K'}{\mu(a)} \frac{p_i^{e'-e}}{q_i^{d'-d}} = \frac{4\gamma_2 K'}{\mu(a)} \left(\frac{p_i}{q_i}\right)^{e'-e} \longrightarrow \frac{4\gamma_2 K'}{\mu(a)} \left(\frac{\log \alpha}{\log \beta}\right)^{e'-e} < 1.$$

Using (5.10) we obtain T_i is strictly smaller than 1 for some large enough i. We conclude with the argument (5.13).

5.6.2. *α and β are multiplicatively dependent.* We can suppose $\alpha = \beta$. From the hypothesis, we necessarily have $d \neq e$. From (5.17) and for $i \geq j_0$ we have:

$$T_i \leq \frac{4\gamma_2 K'}{\mu(a)} \left(\frac{p_i}{q_i}\right)^{e'-e}.$$

From (5.9) we observe that $\lim_{i\to\infty} p_i/q_i = 1$. We conclude using (5.10).

5.7. $\boxed{\alpha > \alpha' > 1 \textbf{ and } \beta > \beta' > 1}$**.** From Proposition 21, we necessarily have

$$\frac{\log \alpha'}{\log \alpha} = \frac{\log \beta'}{\log \beta} \quad \text{and} \quad e' - e\frac{\log \alpha'}{\log \alpha} = d' - d\frac{\log \alpha'}{\log \alpha}.$$

5.7.1. *α and β multiplicatively independent.* From (5.18) and (5.9) we have:

$$T_i \leq \frac{2K'(2\gamma_2)^{\frac{\log \alpha'}{\log \alpha}}}{\mu(a)} \cdot \left(\frac{p_i}{q_i}\right)^{e'-e\frac{\log \alpha'}{\log \alpha}} \longrightarrow \frac{2K'(2\gamma_2)^{\frac{\log \alpha'}{\log \alpha}}}{\mu(a)} \cdot \left(\frac{\log \alpha}{\log \beta}\right)^{e'-e\frac{\log \alpha'}{\log \alpha}}$$

which is, from (5.10), strictly smaller than 1 for large enough i.

5.7.2. *α and β are multiplicatively dependent.* We can suppose $\alpha = \beta$. We necessarily have $d \neq e$. It suffices to proceed as in the paragraph 5.6.2.

5.8. $\boxed{\alpha > \alpha' = 1 \textbf{ and } \beta > \beta' = 1}$**.** From Proposition 21 we obtain that $e' - d' \leq 1$ and $d' - e' \leq 1$, hence $|d' - e'| \leq 1$.

5.8.1. *α and β are multiplicatively independent.* From (5.14) and (5.17) and for $i \geq j_0$ we have:

$$S_i \leq \frac{K'}{\mu(b)} \frac{n_i^{d'}}{m_i^{e'}} \text{ and } T_i \leq \frac{K'}{\mu(a)} \frac{p_i^{e'}}{q_i^{d'}}.$$

a) Suppose $|e' - d'| = 1$. From (5.9) we deduce that either $(T_i)_{i\in\mathbb{N}}$ or $(S_i)_{i\in\mathbb{N}}$ tends to 0 for i tending to infinity.
b) Suppose $e' = d'$. In this case for i sufficiently large we have

$$T_i \leq \frac{K'}{\mu(a)} \left(\frac{p_i}{q_i}\right)^{e'} \leq \frac{2K'}{\mu(a)} \left(\frac{\log \alpha}{\log \beta}\right)^{e'}.$$

We conclude using (5.10).

5.8.2. *α and β multiplicatively dependent.* W.l.o.g. we suppose $\alpha = \beta$. We necessarily have $\alpha = \beta > 1$ and $d \neq e$. From Proposition 21, we obtain $|d' - e'| \leq 1$.

a) Suppose $e' = d'$. From (5.17) and for $i \geq j_0$ we have:

$$T_i \leq \frac{2K'}{\mu(a)} \left(\frac{p_i}{q_i}\right)^{e'}.$$

But, from (5.9) we know $(p_i/q_i)_i$ tends to 1. We conclude using (5.10).

b) $d' = e' + 1$. From (5.17) and for $i \geq j_0$ we have:

$$T_i \leq \frac{2K'}{\mu(a)} \frac{p_i^{e'}}{q_i^{e'+1}}.$$

Using (5.9) (T_i) clearly goes to 0.

c) $e' = d' + 1$. It can be treated as the case b).

5.9. **Consequence for the words and application to abstract numeration systems.** In the previous section we proved under the assumptions of Theorem 17 that the letters having infinitely many occurrences in x appear in x with bounded gaps. In this section we deduce that the same result holds not only for letters but also for words.
Consequently, we obtain an analogue of Cobham's theorem for one substitution of polynomial growth (the other being exponential). Theorem 26 combined with the main theorem of [Du1] leads therefore to an extended version of Cobham's theorem. This latter result expressed in terms of subsitutions can easily be translated into the formalism of abstract numeration systems (see Corollary 27 and Remark 28).

Corollary 24. *Under the assumptions of Theorem 17, the words having infinitely many occurrences in x appear in x with bounded gaps.*

Proof. The proof is essentially the same as in [Du1]. 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 $t \in \{0,1\}^{\mathbb{N}}$ defined by

$$t_i = 1, \quad \text{if } x_{[i,i+n-1]} = u;$$

and 0 otherwise.

The sequence $y^{(n)} = ((y_i \cdots y_{i+n-1}); i \in \mathbb{N})$ 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] for example). 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 $\rho \circ \sigma_n^k = \sigma^k \circ \rho$. σ is of growth type (α, d) then $y^{(n)}$ is (α, d)-substitutive.
Let $f : A_n \to \{0, 1\}$ be the letter-to-letter morphism defined by

$$f((b_1 \cdots b_n)) = 1 \text{ if } b_1 \cdots b_n = u \text{ and } 0 \text{ otherwise.}$$

It is easy to see that $f(y^{(n)}) = t$ hence t is (α, d)-substitutive. We proceed in the same way with τ and Theorem 17 concludes the proof. □

Lemma 25. [Pa, Théorème 4.1] *Let x be a proper fixed point of a substitution $\sigma : A \to A^*$. Let B be the set of non-growing letters of A. If in x occur arbitrarily long words belonging to B^*, then there exists a growing letter $a \in A$ and $i \in \mathbb{N}$ such that $\sigma^i(a) = vau$ (or uav) with $u \in B \setminus \{\epsilon\}$.*

Theorem 26. *Let $x \in C^{\mathbb{N}}$ being both (d, α)-substitutive and (e, β)-substitutive for two substitutions satisfying the point (3) of the hypothesis of Theorem 17. Then x is ultimately periodic.*

Proof. From Theorem 24 we know that the words appearing infinitely many times in x occur with bounded gaps in x. Suppose $\beta = 1$ and let z be the fixed point of τ that projects onto x. The substitution τ being polynomial one can prove that there exists a word u for which u^n occurs in z for all n (for the sake of completeness, we recall Lemma 25). We assume there is no shorter word having this property. Then using the arguments of Theorem 18 in [Du1] we achieve the proof. □

Corollary 27. *Let $S = (L, \Sigma, <)$ (resp. $T = (M, \Gamma, \prec)$) be an abstract numeration system where L is a polynomial regular language (resp. M is an exponential regular language). If a set X of integers is both S-recognizable and T-recognizable, then X is a finite union of arithmetic progressions.*

Proof. This is a direct consequence of Theorem 26 and the discussion made in subsection 4.4. □

Remark 28. If $S = (L, \Sigma, <)$ and $T = (M, \Gamma, \prec)$ are abstract numeration systems built on two exponential languages then a Cobham's theorem holds with the same assumptions as the ones considered in [Du1].

We address the following conjecture for which partial answers are given here and in [Du1].

Conjecture 29. *Let σ and τ be two independent substitutions having proper fixed points mapped on the sequence x by letter-to-letter morphisms. Then x is ultimately periodic.*

References

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

[BH] V. Bruyère, G. Hansel, *Bertrand numeration systems and recognizability*, Latin American Theoretical INformatics (Valparaíso, 1995), Theoret. Comput. Sci. **181** (1997), 17–43.

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

[Co2] A. Cobham, *Uniform tag sequences*, Math. Systems Theory **6** (1972), 164–192.

[Du1] F. Durand, *A theorem of Cobham for non-primitive substitutions*, Acta Arith. **104** (2002), 225–241.

[Du2] F. Durand, *Sur les ensembles d'entiers reconnaissables*, J. Théo. Nombres Bordeaux **10** (1998), 65–84.

[Ei] S. Eilenberg, *Automata, languages, and machines*, Vol. A. Pure and Applied Mathematics, Vol. 58. Academic Press, New-York, 1974.

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

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

[LR] P. B. A. Lecomte, M. Rigo, *Numeration systems on a regular language*, Theory Comput. Syst. **34** (2001), 27–44.

[Pa] J.-J. Pansiot, *Complexité des facteurs des mots infinis engendrés par morphismes itérés*, Lect. Notes in Comp. Sci. **172**, 380–389, Springer, Berlin, 1984.

[Qu] M. Queffélec, *Substitution Dynamical Systems-Spectral Analysis*, Lect. Notes in Math. **1294**, Springer-Verlag, Berlin, 1987.

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

(F.D.)
Université de Picardie Jules Verne
Laboratoire Amiénois de Mathématiques Fondamentales et Appliquées
CNRS-UMR 6140
33 rue Saint Leu
80039 Amiens Cedex
France.
E-mail address: `fabien.durand@u-picardie.fr`

(M.R.)
Université de Liège
Institut de Mathématique
Grande traverse 12 (B 37)
B-4000 Liège
Belgium.
E-mail address: `M.Rigo@ulg.ac.be`