Journal de Théorie des Nombres
de Bordeaux **10** (1998), 65–84

Sur les ensembles d'entiers reconnaissables

par Fabien DURAND

RÉSUMÉ. Soient U et V deux systèmes de numération de Bertrand, α et β deux β-nombres multiplicativement indépendants tels que $L(U) = L(\alpha)$ et $L(V) = L(\beta)$, et E un sous-ensemble de $\mathbb{N}$. Si E est U-reconnaissable et V-reconnaissable alors E est une réunion finie de progressions arithmétiques.

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

1. Introduction

Etant donné un sous-ensemble de $\mathbb{N} = \{0, 1, \cdots\}$ peut-on trouver un "algorithme simple" (i.e un automate) acceptant les éléments de E et rejetant ceux qui n'y appartiennent pas. En 1969 Cobham a montré que l'existence d'un tel algorithme dépend de la base dans laquelle les éléments de E sont écrits. Le Théorème de Cobham [Co1] s'énonce ainsi : *Soient p et q deux entiers positifs multiplicativement indépendants et $E \subset \mathbb{N}$. L'ensemble E est p-reconnaissable et q-reconnaissable si et seulement si E est une réunion finie de progressions arithmétiques.* Rappelons qu'un ensemble $E \subset \mathbb{N}$ est p-reconnaissable si le langage formé des écritures en base $p \in \mathbb{N}$ des éléments de E est reconnaissable par automate (pour plus de détails voir [Ei]).
Durant les années 70, et jusqu'à la fin des années 80, peu d'articles complétèrent ou poursuivirent les travaux de Cobham [CKMR, Ei, Ha1, Se].
Les ensembles d'entiers p-reconnaissables peuvent se définir de façon équivalente en termes arithmétiques, algébriques, de la théorie des langages ou de la logique du premier ordre (voir [BHMV]). Ces définitions sont à l'origine de généralisations récentes et variées du Théorème de Cobham [Bes, BP, Du2, Fa1, Fag1, Fag2, MV]. Dans cet article nous étendons le Théorème de Cobham à des systèmes de numération "non-standard" :

Manuscrit reçu le 18 octobre 1997.

Théorème 1. *Soient U et V deux systèmes de numération de Bertrand, α et β deux β-nombres multiplicativement indépendants tels que $L(U) = L(\alpha)$ et $L(V) = L(\beta)$, et E un ensemble d'entiers positifs. Si E est U-reconnaissable et V-reconnaissable alors E est une réunion finie de progressions arithmétiques.*

Ce théorème étend les résultats de Bès [Bes] et de Fagnot [Fag2] qui considérent le cas où α et β sont des nombres de Pisot dont les polynômes minimaux sont les polynômes caractéristiques des relations de récurrence définissant U et V. Nous ne faisons d'hypothèse de ce type et les méthodes employées sont totalement différentes.
Gardons les notations du Théorème 1. Dans [Fa2] Fabre montre que E est U-reconnaissable si et seulement si sa suite caractéristique est ω_α-substitutive. Ce résultat est rappelé dans la section 5, Théorème 22. Cela induit une formulation équivalente du Théorème 1 en termes de substitutions. Pour cette raison les sections 2, 3 et 4 sont consacrées aux substitutions.
La Section 2 est consacrée aux définitions liées aux substitutions. Nous y rappelons la notion de *mot de retour*, ainsi que quelques unes de leurs propriétés obtenues dans [Du1, Du2, DHS]. Nous y ferons souvent référence. Dans la section 3 nous étendons aux langages substitutifs primitifs un résultat obtenu dans [Du2]

Théorème 2. *Soient α et β deux réels. Soient $\mathbf{x}$ et y deux suites non-périodiques respectivement α-substitutive primitive et β-substitutive primitive telles que $L(\mathbf{x}) = L(\mathsf{y})$. Alors α et β sont multiplicativement dépendants.*

Dans [Fag1] le même résultat a été obtenu pour les langages engendrés par des substitutions (non-nécessairement primitives) de longueur constante. Ensuite nous faisons quelques remarques sur les systèmes dynamiques engendrés par des substitutions. Dans la Section 4 nous étendons le Théorème 2 aux langages substitutifs (non-nécessairement primitifs) engendré par des substitutions se projetant sur des substitutions primitives et vérifiant une condition leur évitant d'être "trop lacunaires" (Proposition 18). La Section 5 a pour but de rappeler les résultats classiques concernant les systèmes de numération et la notion de reconnaissabilité d'ensembles d'entiers. Ces rappels étant faits nous utilisons les résultats des sections précédentes et un résultat de Hansel [Ha2] concernant la syndéticité d'ensembles d'entiers afin de prouver le Théorème 1.

2. Définitions, terminologie et rappels

2.1. **Mots et suites.** Un *alphabet* est un ensemble fini, ses éléments sont des *lettres*. Soit A un alphabet, un *mot* sur A est un élément du monoïde libre, noté A^*, engrendré par A. Soit $x = x_0x_1 \cdots x_n$ (avec $x_i \in A$, $0 \leq i \leq$

$n-1$) un mot sur A, sa *longueur* est n et se note $|x|$. Le *mot vide* se note ϵ, $|\epsilon| = 0$. L'ensemble des mots non-vides de A^* se note A^+. L'ensemble $A^{\mathbb{N}}$ est composé de *suites*. Si $\mathbf{x} = \mathbf{x}_0\mathbf{x}_1\cdots$ est une suite sur A (avec $\mathbf{x}_i \in A$, $i \in \mathbb{N}$) et $I = [k,l]$ un intervalle de $\mathbb{N}$, nous posons $\mathbf{x}_I = \mathbf{x}_k\mathbf{x}_{k+1}\cdots\mathbf{x}_l$; $\mathbf{x}_I$ est un *facteur* de $\mathbf{x}$. Lorsque $k=0$ le mot $\mathbf{x}_I$ est appelé *préfixe* de $\mathbf{x}$. L'ensemble des facteurs de longueur n de $\mathbf{x}$ se note $L_n(\mathbf{x})$ et l'ensemble des facteurs de $\mathbf{x}$, c'est à dire le *langage de* $\mathbf{x}$, se note $L(\mathbf{x})$. Les *occurrences* d'un mot u de $L(\mathbf{x})$ sont les entiers i tels que $\mathbf{x}_{[i,i+|u|-1]} = u$. Lorsque $\mathbf{x}$ est un mot nous utilisons la même terminologie et des définitions analogues. Soient u et v deux mots, nous notons $L_u(v)$ le nombre d'occurences de u dans v. Le mot u est un *suffixe* de v lorsqu'il existe un mot $x \in A^*$ tel que $v = xu$.

La suite $\mathbf{x}$ est *ultimement périodique* s'il existe un mot u et un mot non-vide v tels que $\mathbf{x} = uv^\omega$, où v^ω est la concaténation infinie du mot v. Quand u sera le mot vide nous dirons que $\mathbf{x}$ est *périodique*. Une suite est *non-périodique* si elle n'est pas ultimement périodique.

La suite $\mathbf{x}$ est *uniformément récurrente* si pour chaque facteur u de $\mathbf{x}$ il existe une constante K telle que pour toutes occurrences successives i et j de u dans $\mathbf{x}$ nous avons $|i-j| \leq K$. On dit également que tout facteur apparaît à *lacunes bornées* dans $\mathbf{x}$.

2.2. **Morphismes et matrices.** Soient A, B et C trois alphabets. Un *morphisme* τ est une application de A dans B^*. Un morphisme définit par concaténation une application de A^* dans B^*. Si $\tau(A)$ est inclus dans B^+, cela définit une application de $A^{\mathbb{N}}$ dans $B^{\mathbb{N}}$. Par abus nous notons toutes ces applications τ.

A un morphisme τ, de A dans B^*, est associé la matrice $M_\tau = (m_{i,j})_{i\in B, j\in A}$ où $m_{i,j}$ est le nombre d'occurrences de la lettre i dans le mot $\tau(j)$. A la composition de morphismes correspond la multiplication de matrices. Par exemple, soient $\tau_1 : B \to C^*$, $\tau_2 : A \to B^*$ et $\tau_3 : A \to C^*$ trois morphismes tels que $\tau_1\tau_2 = \tau_3$. Nous avons l'égalité suivante : $M_{\tau_1}M_{\tau_2} = M_{\tau_3}$. En particulier si τ est un morphisme de A dans A^* nous avons $M_{\tau^n} = M_\tau^n$.

Toute matrice carrée M à coefficients positifs a une valeur propre réelle positive r, ayant un vecteur propre à coefficients positifs, telle que le module de toute autre valeur propre de M a un module est inférieur ou égal à r. Nous l'appelons la *valeur propre dominante de* M (voir [LM]). Une matrice carrée est *primitive* si elle a une puissance dont les coefficients sont strictement positifs. Un morphisme de A dans A^* est *primitif* si sa matrice l'est. Dans ce cas la valeur propre dominante est une racine simple du polynôme caractéristique, elle est strictement supérieure au module de toute autre valeur propre et elle a un vecteur propre à coordonnées strictement positives ; c'est le Théorème de Perron-Frobenius (voir [LM]).

2.3. Substitutions et suites substitutives. Une *substitution* est un triplet (τ, A, a), où A est un alphabet, a est une lettre de A telle que la première lettre du mot $\tau(a)$ est a et τ est un morphisme de A dans A^* tel que $\lim_{n\to+\infty} |\tau^n(b)| = +\infty$ pour toute lettre b de A. Pour simplifier l'écriture nous écrirons souvent τ au lieu de (τ, A, a).
Soit (τ, A, a) une substitution. Il existe une unique suite $\mathbf{x} = (\mathbf{x}_n; n \in \mathbb{N})$ de $A^{\mathbb{N}}$ telle que $\mathbf{x}_0 = a$ et $\tau(\mathbf{x}) = \mathbf{x}$. La suite $\mathbf{x}$ est le *point fixe* de τ, nous le notons $\mathbf{x}_\tau$. Nous posons $L(\tau) = L(\mathbf{x}_\tau)$. Quitte à prendre un sous-alphabet de A on peut toujours supposer que A est égal à l'ensemble des lettres ayant une occurrence dans $\mathbf{x}_\tau$.
Si (τ, A, a) est une substitution primitive alors la suite $\mathbf{x}_\tau$ est uniformément récurrente (pour plus de détails voir [Qu]).
Soient A et B deux alphabets, le morphisme ϕ de A dans B^* est un *morphisme lettre à lettre* si $\phi(A) = B$. Une suite y est *substitutive* (resp. *substitutive primitive*) s'il existe une substitution (resp. substitution primitive) τ et un morphisme lettre à lettre ϕ tels que $\mathsf{y} = \phi(\mathbf{x}_\tau)$. Nous dirons également que y *est engendrée par* τ. Cette suite est α*-substitutive* si α est la valeur propre dominante de τ (i.e. de M_τ). Remarquons que les suites substitutives primitives sont uniformément récurrentes.
Une légère réorganisation de la preuve de la Proposition 3.1 dans [Du1] permet d'établir la proposition suivante.

Proposition 3. *Soient* $\mathbf{x}$ *une suite* α*-substitutive sur* A*,* B *un alphabet et* φ *un morphisme de* A *dans* B^+*. Alors il existe* $n \in \mathbb{N}$ *tel que la suite* $\varphi(\mathbf{x})$ *est* α^n*-substitutive. De plus si* $\mathbf{x}$ *est* α*-substitutive primitive alors il existe* $n \in \mathbb{N}$ *tel que* $\varphi(\mathbf{x})$ *est* α^n*-substitutive primitive.*

Preuve. Il existe une substitution (ζ, C, e) de point fixe y et un morphisme lettre à lettre ρ de C dans A^* tels que $\mathbf{x} = \rho(\mathsf{y})$. Posons $\phi = \varphi\rho$, nous avons $\varphi(\mathbf{x}) = \phi(\mathsf{y})$.
Soient $D = \{(c, k); c \in C \,\text{et}\, 0 \le k \le |\phi(c)| - 1\}$ et $\psi : C \to D^*$ le morphisme défini par:

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

Il existe un entier n tel que $|\zeta^n(c)| \ge |\phi(c)|$ pour tout c dans C.
Soit τ le morphisme de D dans D^* défini par:

$$\begin{array}{rrcll} & \tau((c,k)) & = & \psi(\zeta^n(c)_{[k,k]}) & \text{si } 0 \le k < |\phi(c)| - 1, \\ \text{et} & \tau((c, |\phi(c)| - 1)) & = & \psi(\zeta^n(c)_{[|\phi(c)|-1, |\zeta^n(c)|-1]}) & \text{sinon.} \end{array}$$

Pour tout c dans C nous avons

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

par conséquent $\tau(\psi(\mathsf{y})) = \psi(\zeta^n(\mathsf{y})) = \psi(\mathsf{y})$. Ainsi $\psi(\mathsf{y})$ est le point fixe de la substitution $(\tau, D, (e,0))$ et $\tau\psi = \psi\zeta^n$. De plus remarquons que la valeur propre dominante de τ est α^n.

Soit χ le morphisme lettre à lettre de D dans B défini par $\chi((c,k)) = \phi(c)_{[k,k]}$ pour tout (c,k) dans D. Pour tout c dans C on obtient

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

puis $\chi(\psi(\mathsf{y})) = \phi(\mathsf{y})$. Par conséquent $\varphi(\mathbf{x})$ est α^n-substitutive.

La relation $\tau\psi = \psi\zeta^n$ implique que pour tout $k \in \mathbb{N}$ nous avons $\tau^k\psi = \psi\zeta^{kn}$. Par conséquent si ζ est primitive alors τ l'est également et si $\mathbf{x}$ est α-substitutive primitive alors $\varphi(\mathbf{x})$ est α^n-substitutive primitive. $\square$

Le résultat suivant est le résultat principal de l'article [Du2]. Il peut être vu comme une généralisation du Théorème de Cobham pour les substitutions primitives.

Théorème 4. *Soient α et β deux réels et $\mathbf{x}$ une suite non-périodique qui est α-substitutive primitive et β-substitutive primitive. Alors α et β sont multiplicativement dépendants.*

Rappelons un résultat essentiel sur la structure des suites substitutives primitives établi dans [Du1] (Théorème 4.4) (mais obtenu auparavant dans [Mo] pour les points fixes de substitutions primitives) qui sera utile dans la preuve du Théorème 1.

Théorème 5. *Soit $\mathbf{x}$ une suite substitutive primitive non-périodique. Alors il existe un entier N tel que u^N appartient à $L(\mathbf{x})$ si et seulement si u est le mot vide.*

2.4. **Mots de retour.** Dans toute cette sous-section $\mathbf{x}$ sera une suite uniformément récurrente sur l'alphabet A et u un mot de $L(\mathbf{x})$. Un mot w est un *mot de retour sur u de $\mathbf{x}$* s'il existe deux occurrences successives, i et j, de u dans $\mathbf{x}$ telles que $w = \mathbf{x}_{[i,j-1]}$.

La suite $\mathbf{x}$ étant uniformément récurrente l'ensemble des mots de retour sur u, que nous notons $\mathcal{R}_u$, est fini. On peut vérifier sans peine qu'un mot w est un mot de retour sur u si et seulement s'il satisfait les trois conditions suivantes :

1. $wu \in L(\mathbf{x})$;
2. u est un préfixe de wu ;
3. le mot wu contient exactement deux occurrences de u.

Remarquons qu'un mot de $L(\mathbf{x})$ vérifiant 1. et 2. est une concaténation de mots de retour sur u.

Proposition 6. (Lemme 3.2 dans [Du1]) *Soit* $\mathbf{x}$ *une suite uniformément récurrente non-périodique, alors*

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

Munissons $\mathcal{R}_u$ de l'ordre total suivant : $w \leq v$, $w, v \in \mathcal{R}_u$, si et seulement si la première occurrence de wu dans $\mathbf{x}$ est inférieure ou égale à la première occurrence de vu dans $\mathbf{x}$. Cet ordre définit une bijection $\Theta_u : R_u \to \mathcal{R}_u \subset A^*$, où $R_u = \{1, \cdots, \mathrm{Card}(\mathcal{R}_u)\}$, de la façon suivante : $\Theta_u(k)$ est le k^{eme} mot de retour pour cet ordre.

Proposition 7. ([Du1, DHS])
Les applications $\Theta_u : R_u^* \to A^*$ *et* $\Theta_u : R_u^{\mathbb{N}} \to A^{\mathbb{N}}$ *sont injectives.*

Cette proposition nous permet de définir la notion de suite dérivée introduite dans [Du1]. Pour tout préfixe v de $\mathbf{x}$ la *suite dérivée de* $\mathbf{x}$ *par rapport à* v est l'unique suite $\mathcal{D}_v(\mathbf{x})$ sur R_v telle que $\Theta_v(\mathcal{D}_v(\mathbf{x})) = \mathbf{x}$. Dans [Du1] (Théorème 2.5) a été prouvée la caractérisation suivante des suites substitutives primitives.

Théorème 8. *Soient* A *un alphabet et* $\mathbf{x}$ *une suite sur* A. *La suite* $\mathbf{x}$ *est substitutive primitive si et seulement si l'ensemble* $\{\mathcal{D}_u(\mathbf{x}); u \text{ préfixe de } \mathbf{x}\}$ *est fini.*

Le théorème suivant a été prouvé dans [DHS] (Théorème 24 et Proposition 25).

Théorème 9. *Soit* y *une suite substitutive primitive non-périodique. Il existe une constante* K *telle que : Pour tout mot* u *de* $L(\mathsf{y})$

1. *Pour tout mot* w *appartenant à* $\mathcal{R}_u$, $|u|/K \leq |w| \leq K|u|$;
2. $\mathrm{Card}(\mathcal{R}_u) \leq K(K+1)^2$.

L'inégalité $|u|/K \leq |w|$ du théorème précèdent implique que le nombre d'occurrences du mot u dans un mot de longueur n de $L(\mathsf{y})$ est inférieur ou égal à $Kn/|u|$.

3. Un Théorème de Cobham pour les langages substitutifs primitifs

Dans cette section nous allons prouver le Théorème 2. Avant d'en donner une preuve rappelons quelques résultats classiques sur les substitutions. Soient (τ, A, a) une substitution primitive, $\mathbf{x}$ son point fixe, M sa matrice, α sa valeur propre dominante et $v = (v_i; i \in A)$ un vecteur propre associé à α tel que $\sum_{i\in A} v_i = 1$. D'après le Théorème de Perron-Frobenius ce vecteur est unique et à coefficients strictement positifs. Nous l'appellerons le *vecteur fréquence de* τ (ou de $\mathbf{x}$). Ce nom se justifie par le prochain théorème. Soit u un mot de A^*, nous notons $N(u) = (n_i; i \in A)$ le vecteur où n_i est le nombre d'occurrences de la lettre i dans le mot u. Dans [Qu] (Théorème V.13 et Corollaire V.14) est prouvé le résultat suivant.

Théorème 10. *Soient* (τ, A, a) *une substitution primitive,* $\mathbf{x} = (\mathbf{x}_n; n \in \mathbb{N})$ *son point fixe et* $v = (v_i; i \in A)$ *le vecteur fréquence de* τ*. Alors*

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

uniformément en k*.*

Autrement dit, pour tout i dans A la fréquence de la lettre i dans $\mathbf{x}$ existe et vaut v_i. Gardons les notations qui précèdent le Théorème 10. Soient B un alphabet, $\varphi : A \to B^*$ un morphisme lettre à lettre, P sa matrice et $\mathsf{y} = \varphi(\mathbf{x})$. Nous appelons *vecteur fréquence de* Y le vecteur Pv. Remarquons que la somme des coordonnées de ce vecteur vaut 1. La preuve du corollaire suivant est immédiate.

Corollaire 11. *Soient* (τ, A, a) *une substitution primitive,* $\mathbf{x}$ *son point fixe,* v *son vecteur fréquence,* B *un alphabet,* $\varphi : A \to B^*$ *un morphisme lettre à lettre,* P *sa matrice et* $\mathsf{y} = \varphi(\mathbf{x}) = (\mathsf{y}_n; n \in \mathbb{N})$*. Alors*

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

uniformément en k*.*

Rappelons un résultat établi dans [Du1] (Proposition 5.1).

Proposition 12. *Soient* $\mathbf{x}$ *le point fixe de la substitution primitive* (τ, A, a) *et* v *l'un de ses préfixes non-vides. Alors la suite dérivée* $\mathcal{D}_v(\mathbf{x})$ *est le point fixe de la substitution primitive* $(\tau_v, R_v, 1)$ *définie par* $\Theta_v \tau_v = \tau \Theta_v$*.*

Remarquons dans la proposition précèdente que les substitutions (τ, A, a) et $(\tau_v, R_v, 1)$ ont la même valeur propre dominante.

Lemme 13. *Soient* y *une suite* α*-substitutive primitive et* u *un préfixe de* y*. La suite* $\mathcal{D}_u(\mathsf{y})$ *est* α^k*-substitutive primitive pour un certain entier* k*.*

Preuve. Soient φ un morphisme lettre à lettre , τ une substitution primitive (dont la valeur propre dominante est α) et $\mathbf{x}$ son point fixe tels que $\varphi(\mathbf{x}) = \mathsf{y}$. Soit v l'unique préfixe de $\mathbf{x}$ tel que $\varphi(v) = u$. Si w est un mot de retour sur v alors $\varphi(w)$ est une concaténation de mots de retour sur u. L'application Θ_u étant injective (Proposition 7) cela permet de définir le morphisme $\lambda : R_v \to R_u^*$ par $\Theta_u \lambda = \varphi \Theta_v$. Ce morphisme vérifie $\lambda(\mathcal{D}_v(\mathbf{x})) = \mathcal{D}_u(\mathsf{y})$. Les propositions 3 et 12 terminent la preuve. $\square$

Preuve du Théorème 2. D'après le Théorème 8 il existe une suite de préfixes $(u_i; i \in \mathbb{N})$ de $\mathbf{x}$ vérifiant pour tout $i \in \mathbb{N}$:

- u_i est un préfixe strict de u_{i+1} (i.e. $u_i \neq u_{i+1}$) et
- $\mathcal{D}_{u_i}(\mathbf{x}) = \mathcal{D}_{u_{i+1}}(\mathbf{x}) = \tilde{\mathbf{x}}$.

Notons que $\tilde{\mathbf{x}}$ est α^n-substitutive primitive pour un certain n (Lemme 13). Soient $(u_i; i \in \mathbb{N})$ une telle suite et $(v_i; i \in \mathbb{N})$ l'unique suite de préfixes de y vérifiant pour tout $i \in \mathbb{N}$:

- u_i est un suffixe de v_i et
- v_i a exactement une occurrence de u_i ; nous posons $v_i = d_i u_i$.

Les suites $\mathbf{x}$ et y étant uniformément récurrentes nous pouvons supposer (Proposition 6) que, pour tout $i \in \mathbb{N}$, chaque mot de retour sur u_i (resp. v_i) a une occurrence dans chaque mot de retour sur u_{i+1} (resp. v_{i+1}).
Rappelons (Théorème 9) qu'il existe un réel K tel que pour tout mot $u \in L(\mathbf{x}) = L(\mathsf{y})$ et tout mot $v \in \mathcal{R}_u$ nous avons $|u|/K \leq |v| \leq K|u|$, puis remarquons que pour tout i dans $\mathbb{N}$ le mot d_i est un suffixe d'un mot de retour sur u_i. Ceci implique que $|v_i| \leq (K+1)|u_i|$ pour tout $i \in \mathbb{N}$.
D'après le Théorème 8, quitte à prendre une suite extraite de $(v_i; i \in \mathbb{N})$, nous pouvons supposer que pour tout $i \in \mathbb{N}$ nous avons $\mathcal{D}_{v_i}(\mathsf{y}) = \mathcal{D}_{v_{i+1}}(\mathsf{y}) = \tilde{\mathsf{y}}$. Donc $\tilde{\mathsf{y}}$ est β^m-substitutive primitive pour un certain $m \in \mathbb{N}$ (Lemme 13).
Rappelons que si u est un préfixe de $\mathbf{x}$ alors R_u désigne l'alphabet de la suite $\mathcal{D}_u(\mathbf{x})$. Pour tout $i \in \mathbb{N}$ nous avons $R_{u_{i+1}} = R_{u_i}$ et $R_{v_{i+1}} = R_{v_i}$. Posons $R_{u_0} = C$ et $R_{v_0} = D$ et rappelons que $\tilde{\mathbf{x}}$ appartient à $C^{\mathbb{N}}$ et $\tilde{\mathsf{y}}$ à $D^{\mathbb{N}}$. Soient $i \in \mathbb{N}$ et $b \in D$. Le mot $\Theta_{v_i}(b)v_i = \Theta_{v_i}(b)d_i u_i$ appartient à $L(\mathbf{x})$ donc $\Theta_{v_i}(b)d_i$ y appartient également. Puisque v_i est un préfixe de $\Theta_{v_i}(b)v_i$, il existe un mot t tel que $\Theta_{v_i}(b)v_i = d_i t u_i$ et tel que u_i est un préfixe de tu_i. Donc t est une concaténation de mots de retour sur u_i, autrement dit il existe un unique mot $\rho_i(b)$ dans $R_{u_i}^*$ tel que $t = \Theta_{u_i}\rho_i(b)$. Ceci définit un morphisme $\rho_i : D \to C^*$ vérifiant $\Theta_{v_i}(b)d_i = d_i\Theta_{u_i}(\rho_i(b))$. Remarquons que si s est un élément de $L(\tilde{\mathsf{y}})$ alors nous avons également $d_i\Theta_{u_i}(\rho_i(s)) = \Theta_{v_i}(s)d_i$.
L'utilisation du Théorème 9 permet de donner une majoration de la longueur de $\rho_i(b)$.

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

Par conséquent l'ensemble $\{\rho_i; i \in \mathbb{N}\}$ est fini. Quitte à prendre une suite extraite nous pouvons supposer que $\rho_j = \rho_{j+1} = \rho$ pour tous les entiers positifs j.
Soit $i \in \mathbb{N}$. Puisque les mots de retour sur u_i sont des concaténations de mots de retour sur u_0 et que Θ_{u_0} est injective (Proposition 7), nous pouvons

définir une substitution $\sigma_i : C \to C^*$ par $\Theta_{u_0}\sigma_i = \Theta_{u_i}$. Cette substitution est primitive car chaque mot de retour sur u_0 a une occurrence dans chaque mot de retour sur u_i. Elle a pour point fixe la suite $\tilde{\mathbf{x}}$. En effet nous avons

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

or Θ_{u_0} est injective (Proposition 7) d'où $\sigma(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}$.
De même nous définissons la substitution primitive $\tau_i : D \to D^*$ par $\Theta_{v_0}\tau_i = \Theta_{v_i}$. Elle a pour point fixe la suite $\tilde{\mathbf{y}}$.
Soient α' et β' les valeurs propres dominantes respectives de σ_i et τ_i. Il vient que $\tilde{\mathbf{x}}$ (resp. $\tilde{\mathbf{y}}$) est α^n-substitutive primitive et α'-substitutive primitive (resp. β^m-substitutive primitive et β'-substitutive primitive). Le Théorème 4 implique que α^n et α' (resp. β^m et β') sont multiplicativement indépendants, i.e. il existe deux rationnels positifs p_i et q_i tels que α^{p_i} soit la valeur propre dominante de σ_i et β^{q_i} celle de τ_i.
Soit $w \in L(\tilde{\mathbf{y}})$, nous avons

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

Appelons v le vecteur fréquence de $\tilde{\mathbf{y}}$. D'après le Théorème 10 nous avons

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

où $||.||$ est la norme définie par $||(v_1,\cdots,v_n)|| = |v_1|+\cdots+|v_n|$. Remarquons que $\sigma_i\rho(w)$ appartient à $L(\tilde{\mathbf{x}})$ et appelons u le vecteur fréquence de $\tilde{\mathbf{x}}$. En appliquant de nouveau le Théorème 10 nous obtenons

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

Ainsi nous obtenons

$$\lim_{w\in L(\tilde{\mathbf{y}}),|w|\to+\infty} \frac{|d_i\Theta_{u_i}\rho(w)|}{|w|} = c_1c_2||M_{\sigma_i}(u)|| = c_1c_2\alpha^{p_i}.$$

Notons que les constantes c_1 et c_2 ne dépendent pas de i. Avec des considérations analogues nous montrons qu'il existe une constante c_3 ne dépendant pas de i telle que

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

Nous obtenons $c_1c_2\alpha^{p_i} = c_3\beta^{q_i}$. Soit j un entier positif distinct de i. Les suites $(p_i; i \in \mathbb{N})$ et $(q_i; i \in \mathbb{N})$ tendant vers l'infini nous pouvons supposer $p_i < p_j$ et $q_i < q_j$. Ainsi $\alpha^{p_j-p_i}$ est égal à $\beta^{q_j-q_i}$, i.e. α et β sont multiplicativement dépendants. $\square$

Une remarque sur les systèmes dynamiques substitutifs. Un *système dynamique* est un couple (X,T) où X est un espace métrique compact et T un homéomorphisme de X sur X. Deux systèmes dynamiques (X,T) et (Y,S) sont *isomorphes* s'il existe une bijection continue $f : X \to Y$ telle que $f \circ T = S \circ f$.
On dit que (X,T) est un *système dynamique symbolique* sur l'alphabet A (ou *subshift* sur A) lorsque X est un fermé de $A^{\mathbb{Z}}$ (pour la topologie produit infini des topologies discrètes) tel que $T(X) = X$ où $T : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ est défini par $T((\mathbf{x}_n; n \in \mathbb{Z})) = (\mathbf{x}_{n+1}; n \in \mathbb{Z})$ pour tout $(\mathbf{x}_n; n \in \mathbb{Z}) \in A^{\mathbb{Z}}$. Soit $\mathbf{x} = (\mathbf{x}_n; n \in \mathbb{Z}) \in A^{\mathbb{Z}}$. Posons $\Omega(\mathbf{x}) = \{\mathsf{y} \in A^{\mathbb{Z}}; L(\mathbf{x}) = L(\mathsf{y})\}$. On vérifie aisément que $(\Omega(\mathbf{x}), T)$ est un système dynamique, nous dirons que c'est le système dynamique engendré par $\mathbf{x}$.
Soient $\mathbf{x}$ une suite α-substitutive primitive et (X,T) le système dynamique engendré par $\mathbf{x}$ (i.e. $X = \Omega(\mathbf{x})$). Posons $I(X,T) = \overline{\alpha}$ où $\overline{\alpha}$ est la classe d'équivalence de α pour la relation d'équivalence définie sur $\mathbb{R}^+$ par $\beta \equiv \gamma$ si et seulement si β et γ sont multiplicativement dépendants. Le Théorème 2 implique que $I(X,T)$ est un invariant d'isomorphisme pour les systèmes dynamiques engendrés par des suites substitutives primitives.

Théorème 14. *Soient (X,T) et (Y,T) deux systèmes dynamiques engendrés par des substitutions primitives. Si (X,T) et (Y,T) sont isomorphes alors $I(X,T) = I(Y,T)$.*

La réciproque n'est pas vraie car les substitutions σ et τ, définies respectivement par

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

ont la même valeur propre dominante α^2 où $\alpha = (1+\sqrt{5})/2$ mais leur groupe de dimension, respectivement$(\mathbb{Z}^2, \{(x,y) \in \mathbb{Z}^2; x + \alpha y > 0\}, (3,5))$ et $(\mathbb{Z}^3, \{(x,y,z) \in \mathbb{Z}^3; \alpha x + 2y + z > 0\}, (2,0,-1))$, ne sont pas isomorphes (pour plus de détails voir [DHS]).

4. Le cas des langages substitutifs qui ne sont pas primitifs

4.1. **Décomposition d'une substitution en sous-substitutions.** La proposition suivante est une conséquence du paragraphe 4.4 et de la Proposition 4.5.6 de [LM].

Proposition 15. *Soit $M = (m_{i,j})_{i,j \in A}$ une matrice à coefficients positifs ou nuls dont aucune des colonnes n'est nulle. Il existe trois entiers positifs*

$p \neq 0$, q, l, où $q \leq l-1$, et une partition $\{A_i; 1 \leq i \leq l\}$ de A tels que

(2)

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

où les matrices M_i, $1 \leq i \leq q$ (resp. $q+1 \leq i \leq l$), sont primitives ou nulles (resp. primitives), et tels que pour tout $1 \leq i \leq q$ il existe $i+1 \leq j \leq l$ tel que la matrice $M_{i,j}$ soit non-nulle.

Dans ce qui suit nous gardons les notations de la Proposition 15. Nous dirons que $\{A_i; 1 \leq i \leq l\}$ est la *partition en composantes primitives de A (par rapport à M)*. Si i appartient à $\{q+1, \cdots, l\}$ nous dirons que A_i est une *composante primitive principale de A (par rapport à M)*.
Soient (τ, A, a) une substitution et $M = (m_{i,j})_{i,j \in A}$ sa matrice. Soit $i \in \{q+1, \cdots, l\}$. Nous noterons τ_i la restriction $\tau^p_{/A_i} : A_i \to A^*$ de τ^p à A_i. Puisque $\tau_i(A_i) \subset A_i^*$ nous pouvons considérer que τ_i est un morphisme de A_i dans A_i^* dont la matrice est M_i. Soient $i \in \{1, \cdots, q\}$ tel que M_i soit non-nulle. Définissons φ_i le morphisme de A dans A_i^* qui à b associe b si b appartient à A_i et le mot vide sinon. Considérons l'application $\tau_i : A_i \to A^*$ définie par $\tau_i(b) = \varphi_i(\tau^p(b))$ pour tout $a \in A_i$. Remarquons comme précédemment que $\tau_i(A_i) \subset A_i^*$, par conséquent τ_i définit un morphisme de A_i dans A_i^* dont la matrice est M_i.
Nous dirons que la substitution (τ, A, a) vérifie la condition **(C)** si:

- C1. La matrice M, elle-même, se met sous la forme (2) (i.e. $p = 1$) ;
- C2. Les matrices M_i sont nulles ou à coefficients strictement positifs si $1 \leq i \leq q$ et à coefficients strictement positifs sinon ;
- C3. Pour toute matrice M_i non-nulle, $i \in \{1, \cdots, l\}$, il existe $a_i \in A_i$ tel que $\tau_i(a_i) = a_i u_i$ où u_i est un mot non-vide de A^* si $M_i \neq [1]$ et vide sinon.

D'après la Proposition 15 toute substitution (τ, A, a) a une puissance (τ^k, A, a) vérifiant la condition **(C)**. La définition des substitutions implique que pour tout $q+1 \leq i \leq l$ on a $M_i \neq [1]$.

Soient (τ, A, a) une substitution vérifiant la condition **(C)** (nous gardons les mêmes notations que précédemment). Pour tout $1 \leq i \leq l$ tel que M_i soit non-nulle et différente de la matrice [1], l'application $\tau_i : A_i \to A_i^*$ définit une substitution (τ_i, A_i, a_i) que nous appellerons *sous-substitution principale de* τ si $i \in \{q+1, \cdots, l\}$ et *sous-substitution non-principale de* τ sinon. De plus la matrice M_i étant à coefficients strictement positifs cela implique que la substitution (τ_i, A_i, a_i) est primitive. Remarquons qu'il existe au moins une sous-substitution principale.

Lemme 16. *Soient* (σ, A, a) *et* (τ, B, b) *deux substitutions vérifiant la condition* **(C)**, D *un alphabet,* $\varphi : A \to D^*$ *et* $\phi : B \to D^*$ *deux morphismes lettre à lettre tels que* $\varphi(L(\tau)) = \phi(L(\sigma))$. *Si* $\overline{\sigma}$ *est une sous-substitution principale de* σ *alors il existe une sous-substitution principale* $\overline{\tau}$ *de* τ *telle que* $\varphi(L(\overline{\tau})) = \phi(L(\overline{\sigma}))$.

Preuve. Soit $(\overline{\sigma}, \overline{A}, \overline{a})$ une sous-substitution principale de σ. Pour tout $n \in \mathbb{N}$ définissons k_n le plus grand entier k pour lequel il existe une lettre $c \in B$ telle que $\varphi(\tau^k(c))$ ait une occurrence dans $\phi(\overline{\sigma}^n(\overline{a}))$. L'alphabet B étant fini il existe une lettre $c \in B$ et deux suites d'entiers strictement croissantes, $(i_n; n \in \mathbb{N})$ (extraite de $(k_n; n \in \mathbb{N})$) et $(j_n; n \in \mathbb{N})$, telles que $\varphi(\tau^{i_n}(c))$ ait une occurrence dans $\phi(\overline{\sigma}^{j_n}(\overline{a}))$ pour tout $n \in \mathbb{N}$.
Dans ce qui suit nous utilisons les notations de la Proposition 15 (pour la matrice B). Montrons que le mot $\tau^{2l}(c)$ a une occurrence d'une lettre $\overline{b}$ appartenant à une composante primitive principale $\overline{B}$ de B.
Soit $d \in B_i$ avec $1 \leq i \leq q$. Si $M_i = 0$ alors il existe $j \geq i+1$ et $d' \in B_j$ tels que d' ait une occurrence dans $\tau(d)$ (car aucune des colonnes correspondant à B_i n'est nulle).
D'autre part si $M_i \neq 0$ (i.e. M_i est à coefficients strictement positifs) alors il existe $j \geq i+1$ et $d' \in B_j$ tels que d' ait une occurrence dans $\tau^2(d)$ (car il existe $i+1 \leq k \leq l$ tel que $M_{i,k}$ est non-nulle).
Donc, par induction finie, pour tout e dans B le mot $\tau^{2l}(e)$ a une occurrence d'une lettre appartenant à une composante primitive de B. A fortiori $\tau^{2l}(c)$ a une occurrence d'une lettre $\overline{b}$ appartenant à une composante primitive principale $\overline{B}$ de B.
Par conséquent la sous-substitution principale $\overline{\tau}$ associée à $\overline{B}$ est telle que $\{\varphi(\overline{\tau}^{i_n - 2l}(\overline{b})); i_n - 2l \geq 0, n \in \mathbb{N}\}$ est contenu dans $\phi(L(\overline{\sigma}))$. Les substitutions $\overline{\tau}$ et $\overline{\sigma}$ étant primitives leurs points fixes sont uniformément récurrents, nous en déduisons $\varphi(L(\overline{\tau})) = \phi(L(\overline{\sigma}))$. $\square$

4.2. **Le cas des substitutions se projetant sur des substitutions primitives.** La définition suivante a été introduite dans [Fa2] (voir également [BH1, BH2]) afin d'étendre le Théorème de Cobham à des systèmes de numération non-standard.

Définition. Soient (σ, A, a) et (τ, B, b) deux substitutions. Nous dirons que (σ, A, a) *se projette sur* (τ, B, b) s'il existe un morphisme lettre à lettre $\varphi : A \to B^*$ tel que $\varphi(a) = b$ et que pour toute lettre $c \in A$

$$\varphi\sigma(c) = \tau\varphi(c).$$

Remarques. Soient $\mathbf{x}$ et y les points fixes respectifs de σ et τ. Notons que $\varphi(\mathbf{x}) = \mathsf{y}$ et que $\varphi\sigma^n = \tau^n\varphi$ pour tout $n \in \mathbb{N}$, i.e. (σ^n, A, a) se projette sur (τ^n, B, b) pour tout $n \in \mathbb{N}$.
Soit u (resp. v) un vecteur propre, à coefficients positifs ou nuls, de la valeur propre dominante α (resp. β) de M_σ (resp. de M_τ^T, la transposée de la matrice de τ). Puisque φ est un morphisme lettre à lettre $M_\varphi u$ est non-nulle. Donc α est une valeur propre de M_τ (car $\alpha(M_\varphi u) = M_\tau(M_\varphi u)$) et $\alpha \leq \beta$. Montrons que β est inférieur à α. Si $(v^T)M_\varphi$ est nulle alors $\varphi(A) \neq B$. Cela impliquerait l'existence d'au moins une lettre de B n'apparaissant pas dans y, ce qui n'est pas possible. Par conséquent les valeurs propres dominantes des matrices de σ et de τ sont identiques.

Exemple. Il est possible qu'une substitution non-primitive se projette sur une substitution primitive. Soient $(\sigma, \{a, b, c\}, a)$ et $(\tau, \{0, 1\}, 0)$ deux substitutions et $\varphi : \{a, b, c\} \to \{0, 1\}^*$ le morphisme définis par

$$\begin{array}{lllll} \sigma(a) = ab, & \tau(0) = 01 & \text{et} & \varphi(a) = 0 \\ \sigma(b) = c & \tau(1) = 0 & & \varphi(b) = 1 \\ \sigma(c) = cb & & & \varphi(c) = 0 \end{array}$$

La substitution τ est primitive sans que σ le soit.

Lemme 17. *Soit (σ, A, a) une substitution vérifiant la condition* **(C)** *et se projetant sur une substitution primitive (τ, B, b). Soit σ' une sous-substitution de σ. Si σ' est principale alors σ' et τ ont la même valeur propre dominante, sinon la valeur propre dominante de σ' est strictement inférieure à celle de τ.*

Preuve. Soit $\psi : A \to B^*$ un morphisme lettre à lettre tel que $\psi\sigma = \tau\psi$. Soient $A_1, \cdots, A_l$ les composantes primitives de A (par rapport à M_σ) et $q \leq l - 1$ l'entier tel que A_i est principale si et seulement si $q + 1 \leq i \leq l$. Soient $\sigma_1, \cdots, \sigma_l$ les sous-substitutions de σ associées respectivement à $A_1, \cdots, A_l$.
Soit $q + 1 \leq i \leq l$, nous notons ψ_i la restriction de ψ à A_i. Nous avons $\psi_i\sigma_i(e) = \tau\psi_i(e)$ pour tout $e \in A_i$. Une remarque faite précédemment indique que les valeurs propres dominantes de τ et σ_i sont identiques. La première partie du lemme est prouvée.
Soit $0 \leq i \leq q$. Par définition des sous-substitutions la matrice de σ_i, M_i, est à coefficients strictement positifs et différente de la matrice $[1]$. Soient α et β les valeurs propres dominantes respectives des substitutions

primitives τ et σ_i. Un résultat classique sur les substitutions primitives ([Qu], Proposition V.7) donne l'existence de deux constantes positives non-nulles K_1 et K_2 telles que

$$K_1\alpha^n \leq |\tau^n(d)| \leq K_2\alpha^n \text{ et } K_1\beta^n \leq |\sigma_i^n(c)| \leq K_2\beta^n$$

pour tout $n \in \mathbb{N}$, tout $d \in B$ et tout $c \in A_i$. Soit $c \in A_i$, nous avons pour tout $n \in \mathbb{N}$

$$K_2\alpha^n \geq |\tau^n\psi(c)| = |\psi\sigma^n(c)| = |\sigma^n(c)| \geq |\sigma_i^n(c)| \geq K_1\beta^n.$$

Par conséquent $\beta \leq \alpha$. Nous allons montrer que l'inégalité est stricte ; i.e. $\alpha \neq \beta$.

Soit $c \in A_i$. Comme dans la preuve du Lemme 16 il existe une lettre f, appartenant à une composante primitive principale A_j, ayant une occurrence dans $\sigma^{2l}(c)$. Posons $s = 2l$. Puisque f appartient à A_j nous avons $\sigma^n(f) = \sigma_j^n(f)$, $n \in \mathbb{N}$. La valeur propre dominante de la substitution primitive $\sigma_j : A_j \to A_j^*$ est α (car $q+1 \leq j \leq l$) par conséquent pour tout $n \in \mathbb{N}$ nous avons

$$K_1\alpha^n \leq |\sigma^n(f)| \leq K_2\alpha^n.$$

D'après la Proposition V.7 et la Proposition V.9 de [Qu] il existe une constante $L > 0$ telle que pour tout $n \in \mathbb{N}$

$$L\beta^n \leq L_c(\sigma_i^n(c)).$$

Ainsi pour tout $n \in \mathbb{N}$ et tout $k \in \mathbb{N}^*$ nous avons

$$|\sigma^{n+l}(\sigma_i^k(c))| = |\sigma^n(\sigma^s\sigma_i^k(c))| \geq |\sigma^n\sigma_i^{k+s}(c)| + |\sigma^n(f)|L_c(\sigma_i^k(c)).$$

Par conséquent

$$(3)\quad \begin{aligned} |\sigma^{(n+1)s}(c)| &\geq |\sigma^{ns}\sigma_i^s(c)| + |\sigma^{ns}(f)| \geq |\sigma^{(n-1)s}\sigma_i^{2s}(c)| \\ &\qquad + |\sigma^{(n-1)s}(f)|L_c(\sigma_i^s(c)) + |\sigma^{ns}(f)| \\ &\geq \cdots \geq |\sigma_i^{(n+1)s}(c)| + \sum_{k=0}^{n}|\sigma^{(n-k)s}(f)|L_c(\sigma_i^{ks}(c)) \geq \sum_{k=0}^{n} LK_1\alpha^{(n-k)s}\beta^{ks}. \end{aligned}$$

Or $|\sigma^{(n+1)s}(c)| \leq K_2\alpha^{(n+1)s}$ pour tout $n \in \mathbb{N}$, par conséquent $\alpha \neq \beta$. □

Soient A un alphabet et $\mathbf{x}$ une suite sur A. Nous dirons que $\mathbf{x}$ est à *puissances de mots bornées* s'il existe un entier positif k tel que : $u^k \in L(\mathbf{x})$ si et seulement si u est le mot vide. Nous dirons qu'un langage $L \subset A^*$ vérifie la condition $(*)$ s'il est *factoriel* (i.e. si tous les mots ayant une occurrence dans un mot de L appartiennent à L) et s'il existe une suite $\mathsf{y} \in A^{\mathbb{N}}$ à puissances de mots bornées telle que $L(\mathsf{y}) \subset L$. Autrement dit, un langage L vérifie la condition $(*)$ si et seulement s'il existe un entier positif k et une infinité de mots appartenant à L n'ayant pas d'occurrence de puissance k-ième de mot non-vide. L'une des implications est directe, montrons la

réciproque. Supposons que le langage L est tel qu'il existe un entier k et une infinité de mots appartenant à L n'ayant pas d'occurrence de puissance k-ième de mot non-vide. Appelons $\mathcal{L}$ l'ensemble des mots de L n'ayant pas d'occurrence de puissance k-ième de mot non-vide. L'ensemble $\mathcal{L}$ étant infini et l'alphabet étant fini il existe une suite $(u_i\,;\, i \in \mathbb{N})$ d'éléments de $\mathcal{L}$ telle que u_i est un préfixe de u_{i+1} et $|u_i| < |u_{i+1}|$ pour tout $i \in \mathbb{N}$. Soit y l'unique suite de $A^{\mathbb{N}}$ telle que pour tout $i \in \mathbb{N}$ le mot u_i est un préfixe de y. Puisque $L(\mathsf{y})$ est contenu dans $\mathcal{L}$, y est à puissances de mots bornées.
Nous dirons que $\mathbf{x} \in A^{\mathbb{N}}$ vérifie la condition $(*)$ si $L(\mathbf{x})$ la vérifie.
La proposition suivante est centrale dans la preuve du résultat principal de la section suivante.

Proposition 18. *Soient (σ, A, a) et (τ, B, b) deux substitutions, se projetant respectivement sur les substitutions primitives (σ', A', a') et (τ', B', b'), telles qu'il existe deux morphismes lettre à lettre $\varphi : A \to C^*$ et $\psi : B \to C^*$ vérifiant $\varphi(L(\sigma)) = \psi(L(\tau)) = L$. Si L vérifie la condition $(*)$ alors les valeurs propres dominantes de σ et de τ sont multiplicativement dépendantes.*

Preuve. Quitte à prendre (σ^k, A, a) et (τ^k, B, b), pour un certain k, nous pouvons supposer que (σ, A, a) et (τ, B, b) vérifient la condition **(C)**.
Pour commencer montrons qu'il existe nécessairement une sous-substitution principale $\overline{\sigma}$ de σ de point fixe z tel que $\varphi(\mathsf{z})$ n'est pas périodique.
Supposons que le langage L vérifie la condition $(*)$: il existe une suite $\mathsf{y} \in C^{\mathbb{N}}$ telle que y soit à puissances de mots bornées. En utilisant des arguments analogues à ceux de la preuve du Lemme 16 nous montrons qu'il existe une sous-substitution principale $(\overline{\sigma}, \overline{A}, \overline{a})$ de σ, une lettre $c \in \overline{A}$ et une suite d'entiers strictement croissante $(i_n; n \in \mathbb{N})$ telles que $\varphi(\{\overline{\sigma}^{i_n}(c); n \in \mathbb{N}\}) \subset L(\mathsf{y})$. Soit z le point fixe de $\overline{\sigma}$. Puisque y n'est pas périodique, $\varphi(\mathsf{z})$ ne l'est pas non plus.
D'après le Lemme 16 il existe une sous-substitution principale, que nous notons $(\overline{\tau}, \overline{B}, \overline{b})$, de (τ, B, b) telle que $\varphi(L(\overline{\sigma})) = \psi(L(\overline{\tau}))$. D'après le Théorème 2 $\overline{\sigma}$ et $\overline{\tau}$ ont des valeurs propres dominantes multiplicativement dépendantes. Le Lemme 17 permet de conclure : les valeurs propres dominantes de σ et τ sont multiplicativement dépendantes. □

5. Systèmes de numération et ensembles U-reconnaissables

5.1. **Définitions et rappels.** Un *système de numération* est une suite $U = (U_n; n \in \mathbb{N})$ strictement croissante d'entiers telle que

1. $U_0 = 1$,
2. l'ensemble $\{\frac{U_{n+1}}{U_n}; n \in \mathbb{N}\}$ est borné supérieurement.

Soient $U = (U_n; n \in \mathbb{N})$ un système de numération et c la borne supérieure de $\{\frac{U_{n+1}}{U_n}; n \in \mathbb{N}\}$. Notons A_U l'alphabet $\{0, \cdots, c' - 1\}$ où c' est la partie entière supérieure de c. En utilisant l'algorithme d'Euclide nous pouvons écrire de façon unique chaque entier positif x sous la forme

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

i.e. i est l'unique entier tel que $U_i \leq x < U_{i+1}$ et $x_i = x$, $x_j = a_j U_j + x_{j-1}$, $j \in \{1, \cdots, i\}$, où a_j est le quotient de la division euclidienne de x_j par U_j et x_{j-1} le reste, et $a_0 = x_0$. Nous dirons que $\rho_U(x) = a_i \cdots a_0$ est la *U-représentation* de x et nous posons

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

Nous dirons qu'un ensemble $E \subset \mathbb{N}$ est *U-reconnaissable* si le langage $\{0^n \rho_U(x); n \in \mathbb{N}, x \in E\}$ est reconnaissable par un automate. Pour la notion de langage reconnaissable par automate nous dirigeons le lecteur vers [Ei]. Nous dirons que U est *linéaire* s'il est défini par une relation de récurrence linéaire, i.e. s'il existe $k \in \mathbb{N}^*$, $d_1, \cdots, d_k \in \mathbb{Z}$, $d_k \neq 0$, tels que pour tout $n \geq k$

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

Le polynôme $P(X) = X^k - d_1 X^{k-1} - \cdots - d_{k-1} X - d_k$ est appelé *polynôme caractéristique* de U. Dans [Sh] Shallit a montré :

Théorème 19. *Soit U un système de numération. Si $\mathbb{N}$ est U-reconnaissable alors U est linéaire.*

5.2. **Systèmes de numération de Bertrand.** Un *système de numération* U est *de Bertrand* [Ber3] si : $w \in L(U)$ si et seulement si $w0^n \in L(U)$ pour tout $n \in \mathbb{N}$. C'est une condition naturelle puisque tous les systèmes de numération en base $p \in \mathbb{N}$ la vérifient.

Soit $\alpha > 1$ un réel positif. Tout $x \in [0, 1]$ s'écrit de façon unique sous la forme

$$x = \sum_{n \geq 1} a_n \alpha^{-n}, \tag{4}$$

avec $x_1 = x$ et pour tout $n \geq 1$, $a_n = [\alpha x_n]$ et $x_{n+1} = \{\alpha x_n\}$, où $[.]$ désigne la partie entière inférieure et $\{.\}$ la partie fractionnaire (pour plus de détails voir [Par]). Nous appelons *α-développement* de x la suite $d_\alpha(x) = (a_n; n \in \mathbb{N}^*)$. Nous notons $L(\alpha)$ l'ensemble des mots finis ayant une occurrence dans l'une des suites $d_\alpha(x)$, $x \in [0, 1]$. Si $d_\alpha(1)$ est ultimement périodique nous dirons que α est un *β-nombre* (pour plus de détails sur ces nombres voir [Par]). Bertrand-Mathis a montré les résultats suivants:

Théorème 20. [Ber3] *Soit U un système de numération. C'est un système de Bertrand si et seulement s'il existe un réel $\alpha > 1$ tel que $L(U) = L(\alpha)$. Dans ce cas si U est linéaire alors α est une racine du polynôme caractéristique de U.*

Théorème 21. [Ber1] *Soit $\alpha > 1$ un réel. Le langage $L(\alpha)$ est reconnaissable par automate si et seulement si α est un β-nombre.*

5.3. **Les suites ω_α-substitutives.** Pour tout β-nombre α définissons la substitution ω_α de la façon suivante [Fa2] :

- Si $d_\alpha(1) = a_1 \cdots a_n 0^\omega$, $a_n \neq 0$, alors $(\omega_\alpha, \{1, \cdots, n\}, 1)$ est définie par
$$\begin{array}{ll} 1 & \to 1^{a_1}2; \\ \vdots & \\ n-1 & \to 1^{a_{n-1}}n; \\ n & \to 1^{a_n}; \end{array} .$$
- Si $d_\alpha(1) = a_1 \cdots a_n (a_{n+1} a_{n+2} \cdots a_{n+m})^\omega$, où n et m sont minimaux et où $a_{n+1} + a_{n+2} + \cdots + a_{n+m} \neq 0$, alors $(\omega_\alpha, \{1, \cdots, n+m\}, 1)$ est définie par
$$\begin{array}{ll} 1 & \to 1^{a_1}2; \\ \vdots & \\ n+m-1 & \to 1^{a_{n+m-1}}(n+m); \\ n+m & \to 1^{a_{n+m}}(n+1); \end{array} .$$

Remarquons que dans les deux cas la substitution ω_α est primitive et que α est la valeur propre dominante de M_{ω_α}. Pour montrer cette dernière propriété il suffit de calculer le polynôme caractéristique de ω_α et d'exhiber un vecteur propre de α ayant des coordonnées strictement positives. Nous appellerons *ω_α-substitution* toute substitution qui se projette sur la substitution ω_α et nous appellerons suite *ω_α-substitutive* (α-automatique dans [Fa2]) toute suite qui est l'image par un morphisme lettre à lettre du point fixe d'une ω_α-substitution. Dans [Fa2] (Corollaire 1) Fabre a montré le résultat suivant :

Théorème 22. *Soit U un système de Bertrand tel que $L(U) = L(\alpha)$ où α est un β-nombre. Une partie E de $\mathbb{N}$ est U-reconnaissable si et seulement si sa suite caractéristique $(\mathbf{x}_n; n \in \mathbb{N})$ (i.e. $\mathbf{x}_n = 1$ si $n \in E$ et $\mathbf{x}_n = 0$ sinon) est ω_α-substitutive.*

5.4. **Une application aux systèmes de numération.** Nous dirons que $E \subset \mathbb{N}$ vérifie la condition $(*)$ si sa suite caractéristique la vérifie.

Proposition 23. *Soient U et V deux systèmes de numération de Bertrand tels que $L(U) = L(\alpha)$ et $L(V) = L(\beta)$ où α et β sont deux β-nombres, et E un ensemble d'entiers positifs U-reconnaissable et V-reconnaissable. Si E vérifie la condition $(*)$ alors α et β sont multiplicativement dépendants.*

Preuve. Soit $\mathbf{x}$ la suite caractéristique de E. D'après le Théorème 22 il existe deux substitutions (σ, A, a) et (τ, B, b) se projetant respectivement sur ω_α et ω_β, et deux morphismes lettre à lettre $\varphi : A \to \{0,1\}^*$ et $\psi : B \to \{0,1\}^*$ tels que $\mathbf{x} = \varphi(\mathbf{x}_\sigma) = \psi(\mathbf{x}_\tau)$. Puisque les substitutions ω_α

et ω_β sont primitives et que $\mathbf{x}$ vérifie la condition $(*)$, la Proposition 18 permet de conclure. $\square$

L'énoncé de la Proposition 23 peut être légèrement amélioré (sans modification de la preuve) :

Proposition 24. *Soient U et V deux systèmes de numération de Bertrand, α et β deux β-nombres tels que $L(U) = L(\alpha)$ et $L(V) = L(\beta)$, et E (resp. E') un ensemble d'entiers positifs U-reconnaissable (resp. V-reconnaissable) dont la suite caractéristique est $\mathbf{x}$ (resp. $\mathbf{x}'$). Si $\mathbf{x}$ vérifie la condition $(*)$ et $L(\mathbf{x}) = L(\mathbf{x}')$ alors α et β sont multiplicativement dépendants.*

Un ensemble $E \subset \mathbb{N}$ est *syndétique* s'il existe $p \in \mathbb{N}$ tel que pour tout $n \in \mathbb{N}$ on ait $E \cap [n, n+p] \neq \emptyset$; i.e la lettre 1 apparaît à lacunes bornées dans la suite caractéristique de E. Récemment Hansel [Ha2] a prouvé un résultat très général sur la syndéticité des ensembles d'entiers reconnaissables. Dans le cas qui nous intéresse son résultat s'énonce ainsi :

Théorème 25. *Soient U et V deux systèmes de numération de Bertrand tels que $L(U) = L(\alpha)$ et $L(V) = L(\beta)$ où α et β sont deux β-nombres multiplicativement indépendants. Si E est un ensemble infini d'entiers positifs U-reconnaissable et V-reconnaissable alors E est syndétique.*

Preuve du Théorème 1. Supposons que E est U-reconnaissable et V-reconnaissable. Soit $\mathbf{x} = (\mathbf{x}_n; n \in \mathbb{N})$ la suite caractéristique de E. Soient A et B les alphabets respectifs de ω_α et ω_β. Il existe deux substitutions (σ, A', a) et (τ, B', b) et quatre morphismes lettre à lettre $\varphi : A' \to A^*$, $\varphi' : B' \to B^*$, $\psi : A' \to \{0,1\}^*$ et $\psi' : B' \to \{0,1\}^*$ tels que $\varphi\sigma = \omega_\alpha\varphi$, $\varphi'\tau = \omega_\beta\varphi'$, $\psi(\mathsf{y}) = \psi'(\mathsf{z}) = \mathbf{x}$, où $\mathsf{y} = (\mathsf{y}_n; n \in \mathbb{N})$ et $\mathsf{z} = (\mathsf{z}_n; n \in \mathbb{N})$ sont les points fixes respectifs des substitutions (σ, A', a) et (τ, B', b).
Soit y' le point fixe d'une composante primitive principale $\overline{\sigma}$ de σ. Puisque α et β sont multiplicativement indépendants, E ne vérifie pas la condition $(*)$ (c'est la Proposition 23).
Par conséquent $\psi(\mathsf{y}')$ est périodique (c'est le Théorème 5). Posons $\psi(\mathsf{y}') = u^\omega$ où $|u|$ est la plus petite période de $\psi(\mathsf{y}')$.
Montrons que u apparaît à lacunes bornées dans $\mathbf{x}$. Posons $n = |u|$. La suite $\mathsf{y}^{(n)} = ((\mathsf{y}_i \cdots \mathsf{y}_{i+n-1}); i \in \mathbb{N})$ est le point fixe de la substitution $(\sigma_n, A_n, (\mathsf{y}_0\mathsf{y}_1 \cdots \mathsf{y}_{n-1}))$, où $A_n = A^n$, définie pour tout $(a_1 \cdots a_n)$ dans A_n par

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

où $\sigma(a_1 \cdots a_n) = b_1 \cdots b_k$ (pour plus de détails voir la section V.4 de [Qu]).
Soit $\rho : A_n \to A^*$ le morphisme lettre à lettre défini par $\rho((b_1 \cdots b_n)) =$

b_1 pour tout $(b_1 \cdots b_n) \in A_n$. Nous avons $\rho\sigma_n = \sigma\rho$, donc $\mathsf{y}^{(n)}$ est ω_α-substitutive car

$$\varphi\rho\sigma_n = \varphi\sigma\rho = \omega_\alpha\varphi\rho.$$

De la même façon nous montrons que la suite $\mathsf{z}^{(n)}$ est ω_β-substitutive.
Soit $F = \{i \in \mathbb{N}; \mathbf{x}_{[i,i+n-1]} = u\}$. On remarque sans difficulté que la suite caractéristique de F est une projection lettre à lettre de $\mathsf{y}^{(n)}$ mais également de $\mathsf{z}^{(n)}$. Par conséquent F est U-reconnaissable et V-reconnaissable, c'est le Théorème 22. De plus $\overline{\sigma}$ étant une composante primitive principale de σ, u apparaît une infinité de fois dans $\mathbf{x}$ (i.e F est infini). Par conséquent F est syndétique (c'est le Théorème 25) ; i.e u apparaît à lacunes bornées dans $\mathbf{x}$.
Bien que $\mathbf{x}$ n'est pas nécessairement uniformément récurrente, la notion de mot de retour sur u a un sens. L'ensemble $\mathcal{R}_u$ des mots de retour sur u est fini (car u apparaît à lacunes bornées dans $\mathbf{x}$), donc il existe $i \in \mathbb{N}$ tel que tout $w \in \mathcal{R}_u \cap L((\mathbf{x}_n; n \geq i))$ apparaît une infinité de fois dans $\mathbf{x}$. De plus le même raisonnement que précédemment montre que les mots $w \in \mathcal{R}_u \cap L((\mathbf{x}_n; n \geq i))$ apparaissent à lacunes bornées dans $\mathbf{x}$.
Soient $w \in \mathcal{R}_u \cap L((\mathbf{x}_n; n \geq i))$ et $K = \max\{i_{n+1} - i_n; n \in \mathbb{N}\}$ où $(i_n; n \in \mathbb{N})$ est la suite strictement croissante des occurrences de wu dans $\mathbf{x}$. Dans tout mot de longueur $K + |w| + |u|$ apparaît le mot wu.
Nous remarquons sans peine que pour tout $n \in \mathbb{N}$ le mot u^n appartient à $L(\mathbf{x})$ car $L(\mathsf{y}')$ est contenu dans $L(\mathbf{x})$. Soit $n \in \mathbb{N}$ tel que $|u^n| > K + |w| + |u|$. Alors le mot wu a une occurrence dans u^n et par conséquent il existe $j \in \mathbb{N}$ tel que $w = u^j$. Cela implique que $\mathbf{x}$ est ultimement périodique. $\square$

Remerciements. Ce travail a été partiellement supporté par le programme "FONDAP-matematicas aplicadas, programa en modelamiento estocastico".

Bibliographie

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

[Ber3] A. Bertrand-Mathis, *Comment écrire les nombres entiers dans une base qui n'est pas entière*, Acta Math. Acad. Sci. Hungar. 54 (1989), 237-241.

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

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

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

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

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

[CKMR] G. Christol, T. Kamae, M. Mendès-France et G. Rauzy, *Suites Algébriques et Substitutions*, Bull. Soc. Math. France 108 (1980), 401-419.

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

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

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

[Du2] F. Durand, *A generalization of Cobham's theorem*, à paraître dans Theory of Computing Systems.

[DHS] F. Durand, B. Host and C. Skau, *Substitutions, Bratteli diagrams and dimension groups*, à paraître dans Ergod. Th. & Dynam. Sys..

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

[Fa1] S. Fabre, *Une généralisation du théorème de Cobham*, Acta Arithmetica LXVII.3 (1994) 197-208.

[Fa2] S. Fabre, *Substitutions et β-systèmes de numération*, Theo. Comp. Sc. 137 (1995), 219-236.

[Fag1] I. Fagnot, *Sur les facteurs des mots automatiques*, à paraître dans Theo. Comp. Sci..

[Fag2] I. Fagnot, *Cobham's theorem and automaticity in non-standard bases*, preprint.

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

[Ha2] G. Hansel, *Systèmes de numération indépendants et syndéticité*, preprint.

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

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

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

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

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

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

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

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

Fabien Durand
Institut de Mathématiques de Luminy
CNRS-UPR 9016
Case 907, 163, avenue de Luminy
13288 Marseille Cedex 9, France
E-mail : durand@iml.univ-mrs.fr