ACTA ARITHMETICA
LXVII.3 (1994)

Une généralisation du théorème de Cobham

par

S. Fabre (Villetaneuse)

Nous généralisons le théorème de Cobham ([2]), en démontrant qu'une partie infinie de $\mathbb{N}$ est reconnaissable en base k (k entier strictement plus grand que un) et reconnaissable dans un système de numération associé à un nombre de Pisot unitaire (ayant une propriété arithmétique supplémentaire) si et seulement si elle est ultimement périodique.

A. Rappels et notations

1. *Notations.* Nous ne travaillons qu'avec des alphabets finis $A = \{0, 1, \ldots, k\}$ avec k un entier positif, dont les éléments a sont appelés les *lettres.* L'ensemble A^* designe le monoïde libre engendré par A : un élément u de A^* est une suite finie de lettres de A appelé un *mot*, le mot vide est noté par ε. Si pour quatre mots u, v, s et t de A^* nous avons $u = vst$, v est un *préfixe* de u, t un *suffixe* de u et s un *facteur* de u.

Nous notons la longueur d'un mot u par $|u|$, en particulier $|\varepsilon| = 0$. Le mot de longueur n ne comportant que la lettre a est noté par a^n.

Nous définissons sur A^* l'ordre lexicographique (noté $\succ$) par :

- si $u = a_1 \ldots a_n$ et $v = b_1 \ldots b_n$, alors $u \succ v$ si, pour un $i \in \{1, \ldots, n\}$, $a_i > b_i$, et pour tout $j \le i$, $a_j = b_j$;
- si $u = a_1 \ldots a_m$ et $v = b_1 \ldots b_n$ avec $m < n$, alors $u \succ v$ si $u0^{n-m} \succ v$.

2. *θ-système de numération*, *ensemble U_θ-reconnaissable.* Un *nombre de Pisot* est un entier algébrique strictement plus grand que 1, dont tous les conjugués sont à l'intérieur du cercle unité. Pour θ un nombre de Pisot, soit le *θ-développement de* 1 (noté $D_\theta(1)$) la suite infinie d'entiers positifs $(\alpha_n)_{n\in\mathbb{N}}$ suivante (pour plus de précision voir [9], [1]) :

- $\alpha_0 = [\theta]$, $r_0 = \{\theta\}$ ($[x]$ désignant la partie entière de x et $\{x\}$ la partie fractionnaire de x);
- pour tout entier n, $\alpha_{n+1} = [\theta r_n]$, $r_{n+1} = \{\theta r_n\}$.

Le θ-développement de 1 d'un nombre de Pisot est soit ultimement périodique, soit fini (i.e. $D_\theta(1) = \alpha_0 \ldots \alpha_n(\alpha_{n+1} \ldots \alpha_{n+m})^\omega$ ou $D_\theta(1) = \alpha_0 \ldots \alpha_n$) ([9]). Nous pouvons alors définir à partir de $D_\theta(1)$ le polynôme $Q_\theta(x)$ de $\mathbb{Z}[X]$ suivant ([1]) :

- si $D_\theta(1) = \alpha_0 \ldots \alpha_n$, alors $Q_\theta(x) = x^{n+1} - \alpha_0 x^n - \alpha_1 x^{n-1} - \ldots - \alpha_n$;
- si $D_\theta(1) = \alpha_0 \ldots \alpha_n(\alpha_{n+1} \ldots \alpha_{n+m})^\omega$, alors

$$\begin{aligned} Q_\theta(x) = (&x^{n+m+1} - \alpha_0 x^{n+m} - \alpha_1 x^{n+m-1} - \ldots - \alpha_{n+m}) \\ &- (x^{n+1} - \alpha_0 x^n - \alpha_1 x^{n-1} - \ldots - \alpha_n). \end{aligned}$$

Si θ est toujours racine de $Q_\theta(x)$, ce polynôme peut ne pas être son polynôme minimal (noté $P_\theta(x)$). Dans toute la suite le réel θ désignera toujours un nombre de Pisot unitaire pour lequel $Q_\theta(x)$ est son polynôme minimal, et nous dirons que θ possède la *propriété* (μ).

Comme A. Bertrand ([1], mais aussi [6]), nous définissons le *θ-système de numération* $((U_n)_{n\in\mathbb{N}}; A_\theta)$ de la façon suivante :

- si $D_\theta(1) = \alpha_0 \ldots \alpha_n$, alors

$$U_0 = 1; \quad U_i = \begin{cases} \alpha_0 U_{i-1} + \alpha_1 U_{i-2} + \ldots + \alpha_{i-1} U_0 + 1 & \text{pour } 1 \le i \le n; \\ \alpha_0 U_{i-1} + \alpha_1 U_{i-2} + \ldots + \alpha_n U_{i-(n+1)} & \text{pour } i \ge n+1; \end{cases}$$

- si $D_\theta(1) = \alpha_0 \ldots \alpha_n(\alpha_{n+1} \ldots \alpha_{n+m})^\omega$, alors

$$U_0 = 1; \quad U_i = \alpha_0 U_{i-1} + \alpha_1 U_{i-2} + \ldots + \alpha_{i-1} U_0 + 1 \quad \text{pour } i \ge 1;$$

- dans les deux cas $A_\theta = \{0, 1, \ldots, [\theta]\}$.

Un entier p *admet comme représentant le mot* $a_0 \ldots a_n$ de A_θ^* si on a l'égalité $p = a_0 U_n + \ldots + a_n U_0$. Dans ce système, tout entier non nul admet un représentant unique ne comportant aucun facteur plus grand ou égal à $D_\theta(1)$ pour l'ordre lexicographique ([1] et [6]), et ne débutant pas par la lettre 0 : c'est le *représentant normalisé* de l'entier, et nous notons par $[\theta]^*$ l'ensemble des représentants normalisés.

Pour la suite, $\Pi_\theta(x)$ (respectivement $\Pi_k(x)$) désigne la valeur numérique du mot x dans le θ-système de numération (respectivement en base k), et $\nu_\theta(n)$ (resp. $\nu_k(n)$) le représentant normalisé de n dans le θ-système de numération (resp. le représentant de n en base k).

Exemple 1. Nous notons $\phi = (1+\sqrt{5})/2$ le "nombre d'or". Nous avons $D_\phi(1) = 11$, ce qui implique que la base du système de numération associé à ϕ est la suite de Fibonacci (soit 1, 2, 3, 5 etc.), et l'alphabet $A_\phi = \{0, 1\}$. Nous trouvons le système de numération de Fibonacci ([12]) dans lequel le nombre 8 admet trois représentants ne débutant pas par un 0 : les mots 10000, 1011 et 1100, le premier étant le représentant normalisé de 8.

Pour $\phi^2 = (3+\sqrt{5})/2$, $D_{\phi^2}(1) = 21^\omega$. Ici la base du système est la suite des entiers apparaîssant à des rangs impairs dans la suite de Fibonacci (soit 1, 3, 8 etc.), l'alphabet étant $A_{\phi^2} = \{0, 1, 2\}$. Le nombre 8 admet deux représentants : les mots 100 et 22, le premier est le représentant normalisé de 8.

Dans ces deux cas $Q_\theta(x)$ est le polynôme minimal de θ.

Soit ν le nombre de Pisot unitaire dont le polynôme minimal est $P_\nu(x) = x^3 - x - 1$. Par calcul, nous trouvons $D_\nu(1) = 10001$, ce qui implique que $Q_\nu(x) = x^5 - x^4 - 1 = P_\nu(x)(x^2 - x + 1)$. Nous voyons que $Q_\nu(1) \neq P_\nu(1)$, et ν sort du cadre de notre étude.

Soit S une partie de $\mathbb{N}$; nous définissons le langage $L_\theta(S)$ comme l'ensemble des représentants normalisés des éléments de S dans le θ-système de numération. Par analogie avec la notion de k-reconnaissabilité, nous disons que la partie S de $\mathbb{N}$ est *U_θ-reconnaissable* si $L_\theta(S)$ est reconnaissable par un automate ayant un nombre fini d'états (pour la théorie des automates et la k-reconnaissabilité nous renvoyons à [4]).

Un *θ-automate* est un automate qui n'accepte que les mots ne comportant aucun facteur supérieur ou égal à $D_\theta(1)$. Par définition les θ-automates ne sont pas complets dans l'alphabet sur lequel ils sont définis. Cependant, il est clair que si une partie S de $\mathbb{N}$ est U_θ-reconnaissable, $L_\theta(S)$ pourra être reconnu par un θ-automate. De plus, cet automate pourra être choisi déterministe et minimal.

3. *Substitution de longueur θ, suite θ-automatique.* Pour nous, une *substitution* est un triplet $\omega = (\omega, A, a_0)$ où : A est un alphabet fini; ω une application de A dans A^* (prolongeable en un morphisme de A^* dans A^*); a_0 une lettre de A telle que $\omega(a_0) = a_0 u$ (avec u un mot de A^*, différent du mot vide). On peut alors définir dans $A^{\mathbb{N}}$ le point fixe de la substitution, $X_\omega = \lim_{n\to\infty}(\omega(a_0))^n$.

Pour k entier fixé, si toutes les lettres ont pour image par ω un mot de longueur k, ω est une *substitution de longueur constante k*.

Si le mot X_ω est périodique (resp. ultimement périodique) la substitution ω est dite *périodique* (resp. *ultimement périodique*).

Nous rappelons brièvement quelques définitions et des résultats donnés dans [5].

• *Morphisme.* Soient $\omega = (\omega, A, a_0)$ et $\tau = (\tau, B, b_0)$ deux substitutions; nous disons que h est un *morphisme* de ω dans τ si h est une application de A dans B telle que

$$h(a_0) = b_0; \qquad h(\omega(a)) = \tau(h(a)) \quad \forall a \in A.$$

• *Conjugaison.* Les substitutions $\omega = (\omega, A, a_0)$ et $\tau = (\tau, B, b_0)$ sont *conjuguées* s'il existe une substitution $\sigma = (\sigma, C, c_0)$ et deux morphismes h_1 et h_2 de σ dans ω et de σ dans τ, respectivement.

Propriété. ω et τ sont conjuguées si et seulement si pour tout entier n, $|\omega(a_n)| = |\tau(b_n)|$ (a_n et b_n désignent respectivement les n-ièmes lettres de X_ω et X_τ).

• *Substitution ω_θ.* A partir de $D_\theta(1)$, nous définissons une substitution ω_θ de la façon suivante :

• si $D_\theta(1) = (\alpha_0 \ldots \alpha_n)$, alors $\omega_\theta = (\omega_\theta, \{0, 1, \ldots, n\}, 0)$:

$$0 \to 0^{\alpha_0}1, \quad 1 \to 0^{\alpha_1}2, \ \ldots, \ n-1 \to 0^{\alpha_{n-1}}n, \quad n \to 0^{\alpha_n};$$

• si $D_\theta(1) = (\alpha_0\alpha_1 \ldots \alpha_n)(\alpha_{n+1}\alpha_{n+2} \ldots \alpha_{n+m})^\omega$, alors $\omega_\theta = (\omega_\theta, \{0, 1, \ldots, n+m\}, 0)$:

$$0 \to 0^{\alpha_0}1, \ \ldots, \ n+m-1 \to 0^{\alpha_{n+m-1}}(n+m), \quad n+m \to 0^{\alpha_{n+m}}(n+1).$$

Propriétés. Le polynôme caractéristique de la matrice d'occurrences de ω_θ est égal à $Q_\theta(x)$. La suite $(|\omega_\theta^n(0)|)_{n\in\mathbb{N}}$ est égale à la base du système de numération associé à θ.

• *Substitution de longueur θ.* Nous appelons *substitution de longueur θ* toute substitution conjuguée avec ω_θ (définie comme précédemment).

Par extension de la définition des suites k-automatiques (voir par exemple [4]), nous disons que la suite $(b_n)_{n\in\mathbb{N}}$ appartenant à $B^{\mathbb{N}}$ (B un alphabet fini) est *θ-automatique* si elle est obtenue par codage littéral du point fixe d'une θ-substitution.

Soit $\Delta = \{Q, q_0, F, \{0, \ldots, [\theta]\}, \delta\}$ un θ-automate; la *suite des états* de Δ est la suite infinie $[\delta(q_0, \nu_\theta(n))]_{n\in\mathbb{N}}$. Nous obtenons alors un théorème et son corollaire :

Théorème 0. *La suite infinie des états d'un θ-automate est point fixe d'une θ-substitution $\sigma = (\sigma, Q, q_0)$. Inversement, le point fixe d'une θ-substitution est la suite des états d'un θ-automate.*

Corollaire 0. *Une partie infinie de $\mathbb{N}$ est une partie U_θ-reconnaissable si et seulement si sa suite caractéristique* (*appartenant à* $\{0,1\}^{\mathbb{N}}$) *est θ-automatique.*

Remarque. Le théorème 0 n'est que la version pour les θ-systèmes de numération d'un théorème dû à Cobham pour le cas entier ([3]).

Nous généralisons ici le théorème de Cobham ([2], voir aussi [7] pour une autre démonstration) : "Soient k et j deux entiers strictement plus grands que 1, multiplicativement indépendants (i.e. $\log k/\log j \notin \mathbb{Q}$). Une partie infinie de $\mathbb{N}$ est une partie j-reconnaissable et k-reconnaissable si et seulement si elle est ultimement périodique," par le théorème : "Soient θ un nombre de Pisot unitaire ayant la propriété (μ) et k un entier strictement plus grand que 1. Une partie infinie de $\mathbb{N}$ est k-reconnaissable et U_θ-reconnaissable si et seulement si elle est ultimement périodique."

Signalons d'ores et déjà que C. Frougny a montré qu'une partie ultimement périodique est U_θ-reconnaissable ([6]).

B. Indépendance entre les k-substitutions et les θ-substitutions

1. *Deux propositions*

PROPOSITION 1. *Soient θ un nombre de Pisot unitaire ayant la propriété (μ), k un entier strictement plus grand que 1, et $\omega = (\omega, A, a_0)$ une k-substitution non ultimement périodique. S'il existe une lettre a de A telle que $\omega(a)$ soit le mot a^k, alors X_ω n'est pas θ-automatique.*

Preuve. Soit $\omega = (\omega, A, a_0)$ une k-substitution vérifiant les hypothèses de la proposition, et supposons que X_ω soit θ-automatique : cela signifie qu'il existe une substitution $\sigma = (\sigma, B, b_0)$ de longueur θ, et une application Φ de B dans A, telles que $\Phi(X_\sigma) = X_\omega$.

Nous pouvons considérer le mot X_ω comme la suite infinie des états d'un k-automate $\Delta_k = \{A, a_0, \{0, \ldots, k-1\}, \delta_k\}$ ([3]), et le mot X_σ comme celle d'un θ-automate $\Delta_\theta = \{B, b_0, \{0, \ldots, [\theta]\}, \delta_\theta\}$ (théorème 0).

Nous avons dans Δ_k :

$$\exists u \in \{0, \ldots, k-1\}^* \quad \text{tel que} \quad \delta_k(a_0, u) = a;$$
$$\forall v \in \{0, \ldots, k-1\}^*, \qquad \delta_k(a, v) = a \qquad (\text{car } \omega(a) = a^k).$$

Donc

$$\forall v \in \{0, \ldots, k-1\}^*, \qquad \delta_k(a_0, uv) = a.$$

On obtient

$$\forall \alpha \in \mathbb{N}\ \forall n \in [\Pi_k(u0^\alpha), \Pi_k(u0^\alpha) + k^\alpha[\,, \qquad \delta_k(a_0, \nu_k(n)) = a.$$

Comme X_ω n'est pas ultimement périodique, il existe une lettre récurrente b de B (i.e. b apparaît une infinité de fois dans X_σ) telle que $\Phi(b)$ soit différente de a. Cela implique que l'état b est récurrent dans Δ_θ : il existe $(w, s, t) \in \{0, \ldots, [\theta]\}^*$ et w différent du mot vide tel que pour tout $\beta \in \mathbb{N}$, $\delta_\theta(b_0, sw^\beta t) = b$.

Nous allons obtenir une contradiction en montrant qu'il existe un couple d'entiers (α, β) tel que

$$\text{(a)} \qquad \Pi_k(u0^\alpha) \le \Pi_\theta(sw^\beta t) < \Pi_k(u0^\alpha) + k^\alpha.$$

Comme θ est un nombre de Pisot on peut voir que $\Pi_\theta(sw^\beta t) = M\theta^{\beta|w|} + L + o(\varrho^{\beta|w|})$, où L et M sont deux constantes réelles dépendant de (w, s, t), $|\varrho| < 1$, et $\lim_{\beta\to\infty} o(\varrho^{\beta|w|}) = 0$. Les inégalités (a) sont alors équivalentes à

$$\alpha \log k + \log(\Pi_k(u)) \le \log M + \beta \log \theta^{|w|} + \log\left(1 + \frac{L + o(\varrho^{\beta|w|})}{M\theta^{\beta|w|}}\right)$$
$$< \alpha \log k + \log(\Pi_k(u) + 1).$$

Nous supposons β grand pour pouvoir négliger le terme $\log(1+\ldots)$. Nous obtenons donc les inégalités suivantes :

$$\alpha+\frac{\log(\Pi_k(u))}{\log k}\le\beta\frac{\log\theta^{|w|}}{\log k}+\frac{\log M}{\log k}<\alpha+\frac{\log(\Pi_k(u)+1)}{\log k}.$$

L'entier β recherché doit vérifier

$$\left\{\frac{\log(\Pi_k(u))-\log M}{\log k}\right\}\le\left\{\beta\frac{\log\theta^{|w|}}{\log k}\right\}<\left\{\frac{\log(\Pi_k(u)+1)-\log M}{\log k}\right\}.$$

Mais $\log\theta^{|w|}/\log k$ est irrationnel, ce qui implique que la suite $(n\log\theta^{|w|}/\log k)_{n\in\mathbb{N}}$ est équirépartie modulo 1. Il existe donc une infinité d'entiers $(\beta_j)_{j\in\mathbb{N}}$ vérifiant l'inégalité précédente. Pour chaque β_j on pose

$$\alpha_j=\left[\beta_j\frac{\log\theta^{|w|}}{\log k}\right]-\left[\frac{\log(\Pi_k(u))-\log M}{\log k}\right].$$

Le couple (α_j,β_j) vérifie alors les inégalités (a) dès que β_j est assez "grand". Nous obtenons ainsi la contradiction souhaitée, ce qui démontre la proposition 1.

DÉFINITION. Soient $\omega=(\omega,A,a_0)$ une substitution et $[0\ldots N]$ le mot de longueur N préfixe de X_ω. Soit E_t l'ensemble des entiers défini par $E_t=\{0\}\cup\{|\omega^t([0\ldots N])|,\ \forall N\in\mathbb{N}\}$. La substitution ω est *reconnaissable à droite et à gauche* à l'ordre t s'il existe un entier L_t (dépendant de t) tel que si n appartient à E_t, et si, pour m entier, les mots $[n-L_t\ldots n+L_t]$ et $[m-L_t\ldots m+L_t]$ de X_ω centrés en n et m respectivement sont identiques, alors m appartient à E_t.

Dans le cas où ω est une k-substitution, cela implique que m est congru à n modulo k^t.

Si quel que soit l'entier t, ω est reconnaissable à droite et à gauche à l'ordre t, nous disons que ω est *reconnaissable à tous les ordres*.

Remarque. Cette notion a été étudiée par B. Mossé ([8]), dont nous utiliserons certains résultats. Nous renvoyons aussi à [10] pour la notion de reconnaissabilité.

PROPOSITION 2. *Soient θ un nombre de Pisot unitaire ayant la propriété (μ), k un entier strictement plus grand que 1 et $\omega=(\omega,A,a_0)$ une k-substitution reconnaissable à tous les ordres. Alors X_ω n'est pas θ-automatique.*

Avant de donner la preuve de la proposition, nous énonçons deux lemmes :

LEMME 1. *Soit une θ-substitution $\sigma=(\sigma,B,b_0)$, et b une lettre de B. Alors*, $\lim_{n\to\infty}|\sigma^n(b)|=\infty$.

Preuve. Evident en utilisant la définition de la conjugaison.

LEMME 2. *Soit* $(U_n)_{n\in\mathbb{N}}$ *une suite strictement croissante d'entiers positifs définie à partir d'une relation de récurrence de la forme*

$$U_{n+m+1} = a_0U_{n+m} + \ldots + a_{m-1}U_{n+1} + a_mU_n$$

avec $a_0, \ldots, a_{m-1}$ *des entiers relatifs*, a_m *égal à* ± 1 *et* $U_0, \ldots, U_m$ *des entiers naturels. Soit* k *un entier positif strictement plus grand que* 1. *Considérons la suite infinie d'entiers* $(U_{n,t})_{n\in\mathbb{N}}$ *appartenant à* $\{0, \ldots, k^t - 1\}$ (t *entier* > 1), *définie par la suite* $(U_n)_{n\in\mathbb{N}}$ *modulo* k^t. *Alors*, *la suite* $(U_{n,t})_{n\in\mathbb{N}}$ *est périodique de période* T_t *et* $\lim_{t\to\infty} T_t = \infty$.

Preuve. Nous savons que $(U_{n,t})_{n\in\mathbb{N}}$ est ultimement périodique car engendrée par une relation de récurrence linéaire ([11]). Comme $(U_n)_{n\in\mathbb{N}}$ est une suite strictement croissante et que a_m est égal à ± 1, il est facile de montrer que la suite $(U_{n,t})_{n\in\mathbb{N}}$ est périodique pour t entier > 1.

Pour t entier > 1, on peut toujours trouver s entier tel que $U_{T_t} < k^s$, et comme $(U_{n,s})_{n\in\mathbb{N}}$ est une suite périodique, nous avons $T_t < T_s$. Nous en déduisons que $\lim_{t\to\infty} T_t = \infty$.

Preuve de la proposition 2. Soit $\omega = (\omega, A, a_0)$ une substitution satisfaisant à toutes les hypothèses de la proposition. Nous allons supposer que X_ω est θ-automatique : il existe donc une θ-substitution $\sigma = (\sigma, B, b_0)$ et une application Φ de B dans A telles que $\Phi(X_\sigma) = X_\omega$.

Soit $Y \in B^{\mathbb{N}}$ avec $X_\sigma = b_0Y$. Soit l'application $f : \mathbb{N} \to B$, $n \mapsto f(n) = b$ si b est la première lettre de $\sigma^n(Y)$. L'application f est ultimement périodique :

$$\exists (p, r) \in \mathbb{N}^2\ \forall n \geq r, \quad f(n+p) = f(n).$$

Ceci implique que pour tout $n \geq r$ (avec $f(n) = b$), et pour tous $(i, j) \in \mathbb{N}^2$ tel que $j \geq i$, le mot $\sigma^{ip}(b)$ est un préfixe du mot $\sigma^{n+jp}(Y)$.

Considérons la suite $(U_n)_{n\in\mathbb{N}}$ base du θ-système de numération. Grâce aux propriétés de la conjugaison nous avons $|\sigma^n(b_0)| = U_n$ pour tout $n \in \mathbb{N}$.

Fixons t entier tel que la période (notée $T = T_t$) de la suite $(U_{n,t})_{n\in\mathbb{N}}$ soit supérieure strictement à p (notations du lemme 2) : cela est possible car la suite $(U_n)_{n\in\mathbb{N}}$ vérifie les hypothèses du lemme 2. On peut donc trouver n (avec $f(n) = b$) supérieur à r tel que U_n ne soit pas congru modulo k^t à U_{n+p}. Mais alors, pour tout entier i, U_{n+ipT} n'est pas congru modulo k^t à $U_{n+p+ipT}$. Cependant d'après une remarque ci-dessus, pour tout entier i, le mot $\sigma^{ipT}(b)$ est préfixe du mot $\sigma^{n+ipT}(Y)$ et du mot $\sigma^{n+p+ipT}(Y)$.

Dans X_σ, nous avons donc trouvé une suite de mots $(\sigma^{ipT}(b))_{i\in\mathbb{N}}$ tels que $\lim_{i\to\infty} |\sigma^{ipT}(b)| = \infty$ (lemme 1), et pour tout entier i, le mot $\sigma^{ipT}(b)$ apparaît dans X_σ à des rangs non congrus modulo k^t. Comme par hypôthèse $\Phi(X_\sigma) = X_\omega$, en considérant la suite de mots $(\Phi[\sigma^{ipT}(b)])_{i\in\mathbb{N}}$, nous obtenons une contradiction avec la reconnaissabilité à l'ordre t de la substitution ω. Ceci achève la preuve de la proposition 2.

2. *Une classification des substitutions*. Soient $\omega = (\omega, A, a_0)$ une substitution et a une lettre de A; considérons le sous-ensemble $A(a)$ de A suivant :

$$A(a) = \{a' \in A : \exists (i,j) \in \mathbb{N}^* \text{ avec } a' \in \omega^i(a) \text{ et } a \in \omega^j(a')\}.$$

Propriétés immédiates :

1. $A(a_0) \neq \emptyset$ (car $a_0 \in A(a_0)$).
2. $A(a)$ peut être vide même si a est une lettre récurrente de A.
3. Si $a' \in A(a)$ alors $A(a') = A(a)$.
4. Si $X_\omega \neq (a_0)^\omega$, il existe $a \neq a_0$ tel que $A(a) \neq \emptyset$.
5. Soit $a \neq a_0$ et $a \in A(a')$; alors a est une lettre récurrente de A.

Exemple 2. Soit $\omega = (\omega, \{0, 1, \ldots, 5\}, 0)$ la 2-substitution suivante :

$$\omega : \quad \begin{array}{ll} 0 \to 01 & 3 \to 13 \\ 1 \to 23 & 4 \to 54 \\ 2 \to 45 & 5 \to 45. \end{array}$$

Sur cet exemple nous avons : $A(0) = \{0\}$, $A(1) = A(3) = \{1, 3\}$, $A(4) = A(5) = \{4, 5\}$. La lettre 2 est récurrente bien que l'ensemble $A(2)$ soit vide.

Nous pouvons classer les substitutions $\omega = (\omega, A, a_0)$ en trois catégories :

(a) Toutes les lettres de A appartiennent à $A(a_0)$ (toutes les lettres sont alors récurrentes).

(b) Il existe au moins deux ensembles différents $A(a)$ et $A(a')$ non vides ne contenant que des lettres récurrentes.

(c) La lettre a_0 n'est pas récurrente, et toutes les lettres récurrentes appartiennent au même ensemble $A(a)$ (ce qui implique en particulier que si $a' \neq a_0$ et $A(a') \neq \emptyset$, alors $a' \in A(a)$).

Soit $\omega = (\omega, A, a_0)$ une substitution rentrant dans la catégorie (b) ci-dessus, et notons :

- $A_1, \ldots, A_n$ tous les sous-ensembles deux à deux distincts de la forme $A(a)$ de A;
- $a_1, \ldots, a_m$ les lettres récurrentes de A n'appartenant à aucun A_i, $i \in \{1, \ldots, n\}$;
- $\Omega = \{A_1, \ldots, A_n\} \cup \{a_1, \ldots, a_m\}$.

Nous formons des chaînes $(x \to y \to \ldots \to z)$ d'éléments de Ω de la façon suivante :

- $A_i \to A_j$, pour i différent de j, s'il existe $a \in A_i$ et $a' \in A_j$ telles que $a' \in \omega(a)$;
- $A_i \to a_j$ s'il existe $a \in A_i$ telle que $a_j \in \omega(a)$;
- $a_j \to A_i$ s'il existe $a \in A_i$ telle que $a \in \omega(a_j)$;
- $a_i \to a_j$ si $a_j \in \omega(a_i)$.

Il est alors impossible d'avoir simultanément $x \to y$ et $y \to x$ (pour x et y deux éléments différents de Ω). De plus, nous interdisons $A_i \to A_i$. Nous obtenons :

PROPOSITION 3. *Il existe dans Ω des chaînes de longueur maximale* (*construites comme ci-dessus*). *L'extrémité finale de ces chaînes est un élément de $(A_i)_{i\in\{1,\dots,n\}}$* (*noté A_{f}*) *qui vérifie* :

$$\forall a \in A_{\mathrm{f}}, \quad \omega(a) \text{ est un mot formé de lettres de } A_{\mathrm{f}}.$$

Preuve. Comme un élément de Ω ne peut pas apparaître deux fois dans la même chaîne (car sinon nous obtenons une contradiction avec les définitions), et que le cardinal de Ω est fini, toutes les chaînes que nous pouvons construire sont de longueur finie. Considérons une chaîne de longueur maximale: $x \to y \to \ldots \to z$. Comme pour toute lettre a de z le mot $\omega(a)$ est défini, z ne peut pas être un élément de $(a_j)_{j\in\{1,\dots,m\}}$, et doit contenir toutes les lettres de $\omega(a)$.

EXEMPLE 3. Reprenons la substitution $\omega = (\omega, \{0,1,\dots,5\}, 0)$ de l'exemple 2. Alors $\Omega = \{A(0), A(1), A(4), 2\}$. Il y a une chaîne de longueur quatre: $A(0) \to A(1) \to 2 \to A(4)$. L'ensemble $A(4)$ vérifie bien que pour tout $a \in A(4)$, $\omega(a)$ est un mot formé de lettres de $A(4)$.

3. *Un théorème*

THÉORÈME 1. *Soient θ un nombre de Pisot unitaire ayant la propriété* (μ), *k un entier strictement plus grand que* 1 *et $\omega = (\omega, A, a_0)$ une k-substitution non ultimement périodique*; *alors X_ω n'est pas une suite θ-automatique.*

Preuve. Nous allons démontrer le théorème selon l'appartenance de la substitution ω à l'une des trois catégories (a)–(c) définies au paragraphe B.2 :

(a) La substitution ω est primitive, elle est alors reconnaissable à tous les ordres ([8]). Le mot X_ω n'est pas une suite θ-automatique (proposition 2).

(b) Avec les notations de la proposition 3, soit A_{f} l'extrémité finale d'une chaîne de longueur maximale d'éléments de Ω. Nous construisons la k-substitution non ultimement périodique $\tau = (\tau, (A \setminus A_{\mathrm{f}}) \cup \{b\}, a_0)$ (où b n'est pas une lettre de A) suivante :

- pour $a \in A \setminus A_{\mathrm{f}}$, $\tau(a)$ est le mot de $((A \setminus A_{\mathrm{f}}) \cup \{b\})^*$ obtenu en remplaçant dans $\omega(a)$ toutes les lettres de A_{f} par b;
- $\tau(b)$ est le mot b^k.

Il existe un morphisme h (§A.2) de τ dans ω défini par

$$h : A \to (A \setminus A_{\mathrm{f}}) \cup \{b\}, \qquad a \mapsto h(a) = \begin{cases} a & \text{si } a \in A \setminus A_{\mathrm{f}}, \\ b & \text{si } a \in A_{\mathrm{f}}. \end{cases}$$

Comme le mot X_τ ne peut pas être une suite θ-automatique (proposition 1), X_ω ne l'est donc pas.

(c) Si la matrice d'occurrences de la substitution ω est apériodique, alors ω est reconnaissable à tous les ordres ([8]) et donc X_ω n'est pas une suite θ-automatique (proposition 2).

Si la matrice d'occurrences de la substitution ω est périodique, de période d, alors ω^d est une substitution de la catégorie (b) ci-dessus, et comme $X_\omega = X_{\omega^d}$, X_ω n'est pas une suite θ-automatique.

Toutes les k-substitutions non ultimement périodiques appartiennent à l'une des trois catégories ci-dessus, le théorème en découle.

C. Une généralisation du théorème de Cobham

1. *Une propriété des ensembles U_θ-reconnaissables*

PROPOSITION 4. *Soient θ un nombre de Pisot unitaire ayant la propriété (μ), et S une partie infinie U_θ-reconnaissable de $\mathbb{N}$. Pour deux entiers quelconques a et b, nous définissons l'ensemble $S(a,b) = \{n \in \mathbb{N} : an + b \in S\}$. Alors $S(a,b)$ est U_θ-reconnaissable.*

Preuve. Nous ne donnons que la démarche utilisée pour démontrer le résultat. La proposition est équivalente à dire que pour toute suite $(x_n)_{n\in\mathbb{N}}$ θ-automatique, les suites extraites $(x_{an+b})_{n\in\mathbb{N}}$ (a, b entiers) sont θ-automatiques. Ceci se montre en deux étapes :

- on montre que les suites extraites $(x_{n+b})_{n\in\mathbb{N}}$ sont θ-automatiques,
- on montre que les suites extraites $(x_{an})_{n\in\mathbb{N}}$ sont θ-automatiques.

La deuxième étape est la partie technique de la démonstration.

2. *Une généralisation du théorème de Cobham*

PROPOSITION 5. *Soient θ un nombre de Pisot unitaire ayant la propriété (μ), k un entier strictement plus grand que 1 et S une partie infinie de $\mathbb{N}$ k-reconnaissable et U_θ-reconnaissable. Si $\Delta_k = \{Q, q_0, F, \{0,\ldots,k-1\}, \delta\}$ est le k-automate minimal complet reconnaissant S, alors, quel que soit l'état q de Δ_k, l'ensemble $S(q) = \{n \in \mathbb{N} : \delta(q_0, \nu_k(n)) = q\}$ est k-reconnaissable et U_θ-reconnaissable.*

Preuve. En utilisant le résultat de la proposition 4, la preuve est identique à celle du lemme 1 de l'article de Cobham ([2]).

COROLLAIRE 1. *Avec les hypothèses et les notations de la proposition 5, la suite infinie des états $(q_n)_{n\in\mathbb{N}}$ $(= [\delta(q_0, \nu_k(n))]_{n\in\mathbb{N}})$ de Δ_k est θ-automatique.*

Preuve. Supposons que Δ_k se compose de $i+1$ états $q_0, q_1, \ldots, q_i$. Pour chaque état q_j, nous considérons le θ-automate $\Delta_\theta(j)$ reconnaissant l'ensemble $S(q_j)$ (proposition 5). Nous construisons alors le θ-automate Δ_θ, produit des θ-automates $\Delta_\theta(j)$. Chaque état de Δ_θ est un $(i+1)$-uplé $(p_0, p_1, \ldots, p_i)$ où p_j est un état de $\Delta_\theta(j)$. Ces $(i+1)$-uplés comportent un seul état p_j final dans $\Delta_\theta(j)$ (sinon contradiction avec la définition des automates $\Delta_\theta(j)$).

Nous considérons l'application Φ définie par $\Phi((p_0, p_1, \ldots, p_i)) = q_j$ si p_j est final dans $\Delta_\theta(j)$. Si $((p_0, p_1, \ldots, p_i)_n)_{n\in\mathbb{N}}$ désigne la suite des états de Δ_θ, nous avons alors $\Phi[((p_0, p_1, \ldots, p_i)_n)_{n\in\mathbb{N}}] = (q_n)_{n\in\mathbb{N}}$. La suite infinie des états de Δ_k est donc θ-automatique.

THÉORÈME 2. *Soient θ un nombre de Pisot unitaire ayant la propriété (μ) et k un entier strictement plus grand que* 1. *Une partie infinie de* $\mathbb{N}$ *est k-reconnaissable et U_θ-reconnaissable si et seulement si elle est ultimement périodique.*

Preuve. Nous savons que les sous-ensembles ultimement périodiques de $\mathbb{N}$ sont U_θ-reconnaissables ([6]), et k-reconnaissables (propriété ancienne, voir [4] par exemple).

Si S est une partie infinie, k-reconnaissable et U_θ-reconnaissable de $\mathbb{N}$, alors la suite infinie des états du k-automate minimal, complet, reconnaissant S est θ-automatique (corollaire 1). Cependant, cette suite peut être considérée comme le point fixe d'une k-substitution ([3]); elle est donc ultimement périodique (théorème 1), et S est une partie ultimement périodique de $\mathbb{N}$.

Remarques générales. Une démarche analogue peut être utilisée pour démontrer le théorème original de Cobham avec deux entiers k et j premiers entre eux (voir §A).

On peut aussi appliquer la même démarche, et obtenir un théorème analogue, en considérant un nombre de Pisot θ (pas forcément unitaire) ayant la propriété (μ), et un entier k (>1) premier avec le coefficient constant du polynôme minimal de θ.

Remerciements. Je remercie les professeurs G. Rauzy et B. Host pour leurs conseils et leurs encouragements qui m'ont permis de mener à bien ce travail.

Bibliographie

[1] A. Bertrand, *Répartition modulo* 1 *des suites exponentielles et systèmes dynamiques symboliques*, Thèse d'état, Université de Bordeaux 1, 1986.

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

[3] —, *Uniform tag sequence*, ibid. 6 (1972), 164–192.

[4] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, 1974.

[5] S. Fabre, *Substitutions et β-systèmes de numération*, preprint.
[6] C. Frougny, *Systèmes de numération linéaires et automates finis*, Thèse d'état, Université Paris-VII, 1989.
[7] G. Hansel, *A propos d'un théorème de Cobham*, Acte de la fête des mots, D. Perrin (ed.), Rouen, 1982, Greco de programmation, C.N.R.S.
[8] B. Mossé, *Puissances de mots et reconnaissabilité des points fixes d'une substitution*, Theor. Comput. Sci. 99 (1992), 327–334.
[9] W. Parry, *On the β-expansions of real numbers*, Acta Math. Acad. Sci. Hungar. 11 (1960), 401–416.
[10] M. Queffélec, *Substitution Dynamical Systems. Spectral Analysis*, Lecture Notes in Math. 1294, Springer, Berlin, 1987.
[11] G. Rauzy, *Relations de récurrence modulo m*, Séminaire Delange–Pisot, 5ième année (1963–1964), (2-01)–(2-10).
[12] E. Zeckendorf, *Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas*, Bull. Soc. Roy. Sci. Lièges 3–4 (1960), 179–182.

UNIVERSITÉ PARIS XIII
DÉPARTEMENT DE MATHÉMATIQUES
AVENUE J.-B. CLÉMENT
93430 VILLETANEUSE, FRANCE

Reçu le 25.11.1992
et révisé le 16.11.1993 (2345)