

# Théorème de Cobham

Titus Lupu

19 décembre 2008

## 1 Introduction

On s'intéresse aux ensembles d'entiers reconnaissables par automate. On dispose des résultats suivants (voir [1]) : Si un ensemble  $X \subseteq \mathbb{N}$  est reconnaissable en représentant un entier  $n$  par le mot de longueur  $n$  sur l'alphabet  $\{0\}$  (on dit alors que  $X$  est *rationnel*), alors pour tout entier  $k \geq 2$ ,  $X$  est reconnaissable en codant les entiers en base  $k$  (on dit alors que  $X$  est *k-rationnel*). Deux entiers  $k$  et  $l$  sont dits *multiplicativement dépendants* s'il existe  $p$  et  $q$  entiers tel que  $k^p = l^q$ . Si  $k$  et  $l$  sont multiplicativement dépendants, alors les ensembles  $k$ -rationnels et  $l$ -rationnels sont les mêmes. Le théorème de Cobham affirme que si à l'inverse  $k$  et  $l$  sont *multiplicativement indépendants*, les ensembles à la fois  $k$  et  $l$ -rationnels sont exactement les ensembles rationnels. Le but de cet exposé est de présenter une démonstration de ce résultat.

## 2 Premières propriétés

Tout au long de ce document on utilisera les notations suivantes : Soit un entier  $k \geq 2$ . Pour  $n \in \mathbb{N}$ , on note  $(n)_k$  la représentation en base  $k$  de  $n$ . Pour  $X \subseteq \mathbb{N}$ , on note  $(X)_k := \{(n)_k \mid n \in X\}$ . Si  $w$  est un mot non vide sur l'alphabet  $[[0, k-1]]$  ne commençant pas par  $0$ , on note  $[w]_k$  l'entier dont  $w$  est l'écriture en base  $k$ .

Pour un mot  $w = w_1w_2\dots w_n\dots$ , éventuellement infini, on note  $w[a, b]$  le facteur  $w_a\dots w_{b-1}$  de  $w$ .

**Proposition 1.** *Soient  $m, p \geq 0$  et  $k \geq 2$  des entiers. Si  $X \subseteq \mathbb{N}$  est  $k$ -rationnel, alors  $Y = \{m + rp \mid r \in X\}$  est  $k$ -rationnel. Si  $Y \subseteq \mathbb{N}$  est  $k$ -rationnel, alors  $X = \{r \geq 0 \mid m + rp \in Y\}$  est  $k$ -rationnel.*

*Démonstration.* Lorsqu'on dit qu'un ensemble  $X \subseteq \mathbb{N}$  est  $k$ -rationnel on suppose qu'il existe un automate qui reconnaît  $(X)_k$  en lisant les écritures en base  $k$  de gauche à droite. Mais dans ce cas, l'automate miroir reconnaît  $(X)_k$  en lisant les écritures en base  $k$  dans le sens inverse. Donc le sens de lecture n'a pas d'incidence sur la reconnaissabilité.

En fixant un entier  $m$  on peut réaliser un automate qui calcule  $(r+m)_k$  en lisant  $(r)_k$  de droite à gauche, en appliquant le même algorithme qu'on utilise

pour faire l'addition à la main. La retenue qu'on aura besoin de propager peut prendre au plus 2 valeurs. Ensuite on peut brancher à la sortie de cet automate un automate de reconnaissance par lecture de droite à gauche. De même pour la soustraction de  $m$  et la multiplication par  $p$  (la retenue peut prendre alors au plus  $k$  valeurs). Pour la division euclidienne par  $p$ , on utilise la lecture de gauche à droite. En combinant ces opérations arithmétiques, on obtient le résultat.  $\square$

**Définition 2.** Soit  $X \subseteq \mathbb{N}$ . On appelle *mot caractéristique* de  $X$  le mot infini  $w = w_1w_2\dots w_n\dots$  sur l'alphabet  $\{0,1\}$  tel que  $w_n = 1$  si et seulement si  $n \in X$ .

**Définition 3.** On dit qu'un mot infini  $w = w_1w_2\dots w_n\dots$  est *ultimement périodique* s'il existe  $N \geq 0$  et  $p \geq 1$  entiers tel que pour tout  $n \geq N$ ,  $w_n = w_{n+p}$ .  $p$  est alors appelé une période du mot  $w$ . Un ensemble  $X \subseteq \mathbb{N}$  est dit ultimement périodique si son mot caractéristique l'est.

*Exemple 4.* Si  $X$  est l'ensemble des entiers pairs, le mot caractéristique de  $X$  est 101010..., et il est ultimement périodique.

**Proposition 5.** Un ensemble  $X \subseteq \mathbb{N}$  est rationnel si et seulement s'il est ultimement périodique.

*Démonstration.* Le sens ultimement périodique implique rationnel est clair.

Réciproquement, les ensembles rationnels sont ceux qui admettent une expression rationnelle sur l'alphabet  $\{0\}$ . Montrons par induction structurale sur une expression rationnelle  $\mathcal{R}$  : l'ensemble codé par  $\mathcal{R}$  est ultimement périodique.

Si  $\mathcal{R} = \emptyset, \varepsilon$  ou 0, l'ensemble codé par  $\mathcal{R}$  est respectivement  $\emptyset, \{0\}$  ou  $\{1\}$ .

Si  $X_1, X_2 \subseteq \mathbb{N}$  sont codés respectivement par les expressions rationnelles  $\mathcal{R}_1, \mathcal{R}_2$ , soit  $X \subseteq \mathbb{N}$  codé par  $\mathcal{R}_1\mathcal{R}_2$  et  $X' \subseteq \mathbb{N}$  codé par  $\mathcal{R}_1 + \mathcal{R}_2$ . Considérons  $N_1, N_2 \geq 0$  et  $p_1, p_2 \geq 1$  tel que  $X_1$   $p_1$ -périodique à partir du rang  $N_1$  et  $X_2$   $p_2$ -périodique à partir du rang  $N_2$ . Alors  $X = \{n_1 + n_2 \mid n_1 \in X_1, n_2 \in X_2\}$  est  $\text{ppcm}(p_1, p_2)$ -périodique à partir du rang  $N_1 + N_2 + \text{ppcm}(p_1, p_2)$ .  $X' = X_1 \cup X_2$  est  $\text{ppcm}(p_1, p_2)$ -périodique à partir du rang  $\max(N_1, N_2)$ .

Si  $X \subseteq \mathbb{N}$  est codé par l'expression rationnelle  $\mathcal{R}$ , alors  $\mathcal{M}$  codé par  $\mathcal{R}^*$  est le sous-monoïde de  $\mathbb{N}$  engendré par  $X$ . Soit  $d$  le  $\text{pgcd}$  des éléments de  $X$ . Alors il existe  $N \in \mathbb{N}$  tel que  $\mathcal{M} \cap [N, +\infty[ = d\mathbb{N} \cap [N, +\infty[$ . En effet, soient  $a_1 < a_2 < \dots < a_n \in X$  tel que  $d = \text{pgcd}(a_1, a_2, \dots, a_n)$ . Soit  $\mathcal{M}'$  les sous-monoïde engendré par  $(a_1, a_2, \dots, a_n)$ . On peut écrire  $n$  relations de Bezout :

$$\begin{aligned} d &= -u_{1,1}a_1 + u_{1,2}a_2 + \dots + u_{1,n}a_n = u_{2,1}a_1 - u_{2,2}a_2 + \dots + u_{2,n}a_n = \dots \\ &= u_{n,1}a_1 + u_{n,2}a_2 + \dots - u_{n,n}a_n \end{aligned}$$

où les  $u_{i,j}$  sont des entiers positifs. En posant  $N = u_{1,1}a_1 + u_{2,2}a_2 + \dots + u_{n,n}a_n$ , on a  $\mathcal{M}' \cap [N, +\infty[ = d\mathbb{N} \cap [N, +\infty[$ . Comme  $\mathcal{M}' \subseteq \mathcal{M} \subseteq d\mathbb{N}$ , on a le résultat voulu. Donc  $\mathcal{M}$  est ultimement périodique.  $\square$

**Définition 6.** Soit  $w$  un mot infini. Un *facteur récurrent* de  $w$  est un facteur de  $w$  qui apparaît pour une infinité d'indices dans l'écriture de  $w$ .

*Exemple 7.* 010 est un facteur récurrent du mot 101010....

**Proposition 8.** Soit  $w$  un mot infini et  $\mathcal{N}(m)$  pour  $m \geq 1$  le nombre de facteurs récurrents de  $w$  de longueur  $m$ .  $w$  est ultimement périodique si et seulement s'il existe  $m \geq 1$  tel que  $\mathcal{N}(m) \leq m$ .

*Démonstration.* Dans un sens, si  $w$  est  $p$ -périodique à partir d'un certain rang,  $\mathcal{N}(p) = p$ . Réciproquement, on remarque que pour tout  $m \geq 1$ ,  $\mathcal{N}(m) \leq \mathcal{N}(m+1)$  : Soit  $u$  un facteur récurrent de  $w$  de longueur  $m$ . Comme la lettre succédant à  $u$  dans  $w$  ne peut prendre qu'un nombre fini de valeurs, une de ces valeurs apparaît une infinité de fois, et  $u$  se prolonge en un facteur récurrent de longueur  $m+1$ . Considérons alors  $m \geq 1$  minimal tel que  $\mathcal{N}(m) \leq m$ . Si  $m = 1$ , à partir d'un certain rang,  $w$  est toujours composé d'une même lettre. Sinon,  $\mathcal{N}(m-1) > m-1$ , et par croissance de  $\mathcal{N}$ ,  $\mathcal{N}(m-1) = \mathcal{N}(m) = m$  : il existe une unique manière de prolonger un facteur récurrent de longueur  $m-1$  en un facteur récurrent de longueur  $m$ . A partir de l'indice  $N$  où les seuls facteurs de longueur  $m$  qui apparaissent dans  $w$  sont les facteurs récurrents, la lettre succédant à un facteur de longueur  $m-1$  ne dépend plus de la position de celui-ci. Soit alors  $u$  un facteur récurrent de longueur  $m-1$  et  $n_2 > n_1 \geq N$  deux indices de début de  $u$ .  $n_2 - n_1$  est alors une période de  $w$ .  $\square$

### 3 Ensembles quasi-périodiques

**Définition 9.** Un ensemble  $X \subseteq \mathbb{N}$  est dit *d-quasi-périodique* si pour tout  $n \in \mathbb{N}$ ,  $[n, n+d[ \cap X \neq \emptyset$ , et *quasi-périodique* s'il est d-quasi-périodique pour un certain  $d$ .

On établit dans ce qui suit un résultat essentiel pour la démonstration du théorème de Cobham proposée.

**Théorème 10.** Si  $k$  et  $l \geq 2$  sont multiplicativement indépendants, alors un ensemble infini à la fois  $k$  et  $l$ -rationnel est quasi-périodique. (voir [3] )

*Démonstration.* On commence par donner quelques lemmes intermédiaires :

**Lemme 11.** Si  $k$  et  $l \geq 2$  sont multiplicativement indépendants, alors l'ensemble des  $\frac{k^p}{l^q}$  où  $p, q \in \mathbb{N}$  est dense dans  $\mathbb{Q}^+$ .

*Démonstration.* Cela revient à montrer que  $\{p \log(k) - q \log(l) \mid p, q \in \mathbb{N}\}$  est dense dans  $\mathbb{R}$ . Comme  $k$  et  $l$  sont multiplicativement indépendants,  $\frac{\log(k)}{\log(l)}$  est irrationnel et  $\log(k)\mathbb{Z} + \log(l)\mathbb{Z}$  est dense dans  $\mathbb{R}$ . Soit  $\eta = \min(\log(k), \log(l))$  et

$$\mathcal{E} = \{p \log(k) - q \log(l) \mid p, q \in \mathbb{N}\} \cap [-\eta, \eta].$$

On a

$$\log(k)\mathbb{Z} + \log(l)\mathbb{Z} \cap [-\eta, \eta] = \mathcal{E} \cup -\mathcal{E}.$$

Donc  $\overline{\mathcal{E} \cup -\mathcal{E}} = \overline{\mathcal{E}} \cup \overline{-\mathcal{E}} = [-\eta, \eta]$ . D'où l'adhérence de  $\{p \log(k) - q \log(l) \mid p, q \in \mathbb{N}\}$  est d'intérieur non vide. En utilisant la stabilité par dilatation de rapport entier et par translation de  $\log(k)$  et  $-\log(l)$  on conclut que c'est  $\mathbb{R}$  tout entier.  $\square$

**Lemme 12.** Soit  $\mathcal{A} = (Q, I, F, \Sigma, \delta)$  un automate fini et  $s \in Q$ . On pose

$$L_s = \{|w| \in \mathbb{N} \mid w \in \Sigma^*, \delta(s, w) \in F\}$$

$L_s$  est ultimement périodique.

*Démonstration.* D'après la Proposition 5, cela revient à montrer que  $L_s$  est rationnel. On définit l'automate  $\mathcal{A}' = (Q, \{s\}, F, \{0\}, \delta')$  avec la fonction de transition : pour  $q, q' \in Q$  et  $u \in \{0\}$ ,  $\delta' : q \xrightarrow{u} q'$  si et seulement s'il existe  $w \in \Sigma^*$  tel que  $|w| = |u|$  et  $\delta : q \xrightarrow{w} q'$ . La représentation de  $L_s$  dans  $\{0\}^*$  est le langage reconnu par  $\mathcal{A}'$ .  $\square$

**Lemme 13.** Soient  $m, n, a, b, c, d \geq 1$  entiers tel que  $n < m$  et  $k, l \geq 2$  deux entiers multiplicativement indépendants. Alors il existe des entiers  $p, q \geq 1$  tel que

$$nl^{c+dq} \leq mk^{a+bp} < (m+1)k^{a+bp} \leq (n+1)l^{c+dq}$$

*Démonstration.* Ceci revient à montrer qu'il existe des entiers  $p, q \geq 1$  tel que

$$\frac{nl^c}{mk^a} \leq \frac{(k^b)^p}{(l^d)^q} \leq \frac{(n+1)l^c}{(m+1)k^a}$$

Or  $k^b$  et  $l^d$  sont encore multiplicativement indépendants, on conclut par le Lemme 11.  $\square$

**Lemme 14.** Soit  $k \geq 2$  un entier et  $X \subseteq \mathbb{N}$  un ensemble  $k$ -rationnel infini. Alors il existe des entiers  $m, a, b \geq 1$  tel que pour tout  $p \in \mathbb{N}$ ,  $X \cap [mk^{a+bp}, (m+1)k^{a+bp}] \neq \emptyset$ . De plus,  $m$  peut-être choisi arbitrairement grand.

*Démonstration.* Soit  $\mathcal{A} = (Q, I, F, [0, k-1], \delta)$  un automate émondé reconnaissant  $(X)_k$  et  $s \in Q$  un état dans lequel on peut tomber pour une infinité d'entiers ( $s$  existe car  $X$  est infini). Alors  $L_s$  est infini. D'après le Lemme 12, il existe  $a, b \geq 1$  tel que  $\{a + bp \mid p \in \mathbb{N}\} \subseteq L_s$ . Alors pour tout entier  $m$  conduisant à  $s$  à partir d'un état initial, pour tout  $p \in \mathbb{N}$ ,  $X \cap [mk^{a+bp}, (m+1)k^{a+bp}] \neq \emptyset$ .  $\square$

**Lemme 15.** Soit  $l \geq 2$  entier et  $X$  un ensemble  $l$ -rationnel infini. Soit  $\mathcal{A} = (Q, \{i_0\}, F, [0, l-1], \delta)$  un automate déterministe émondé reconnaissant  $(X)_l$ . Supposons qu'il existe  $s \in Q$  tel que  $\mathbb{N} \setminus L_s$  infini. Alors il existe des entiers  $n, c, d \geq 1$  tel que pour tout  $q \in \mathbb{N}$ ,  $X \cap [nl^{c+dq}, (n+1)l^{c+dq}] = \emptyset$ .

*Démonstration.* Si  $s = i_0$  et 0 est le seul entier conduisant à  $i_0$ , alors au moins un autre état  $s' \neq i_0$  vérifie  $\mathbb{N} \setminus L_{s'}$  infini. On peut donc supposer qu'il existe un entier  $n \geq 1$  conduisant de  $i_0$  à  $s$ . D'après le Lemme 12,  $L_s$  est ultimement périodique, et comme  $\mathbb{N} \setminus L_s$  est infini, il existe des entiers  $c, d \geq 1$  tel que  $\{c+dq \mid q \in \mathbb{N}\} \cap L_s = \emptyset$ . Comme  $\mathcal{A}$  est déterministe, pour tout  $q \in \mathbb{N}$ ,  $X \cap [nl^{c+dq}, (n+1)l^{c+dq}] = \emptyset$ .  $\square$

**Lemme 16.** Soient  $k, l \geq 2$  entiers multiplicativement indépendants,  $X \subseteq \mathbb{N}$  ensemble infini à la fois  $k$  et  $l$ -rationnel et  $\mathcal{A} = (Q, \{i_0\}, F, [[0, l-1]], \delta)$  un automate déterministe émondé reconnaissant  $(X)_l$ . Alors pour tout  $s \in Q$ ,  $\mathbb{N} \setminus L_s$  est fini.

*Démonstration.* Supposons par l'absurde que ce n'est pas le cas.

D'après le Lemme 15, il existe des entiers  $n, c, d \geq 1$  tel que pour tout  $q \in \mathbb{N}$ ,  $X \cap [nl^{c+dq}, (n+1)l^{c+dq}] = \emptyset$ . En appliquant le Lemme 14, il existe des entiers  $m > n$ ,  $a, b \geq 1$  tel que pour tout  $p \in \mathbb{N}$ ,  $X \cap [mk^{a+bp}, (m+1)k^{a+bp}] \neq \emptyset$ . Or d'après le Lemme 13, il existe des entiers  $p, q \geq 1$  tel que

$$nl^{c+dq} \leq mk^{a+bp} < (m+1)k^{a+bp} \leq (n+1)l^{c+dq}$$

On aboutit à une contradiction.  $\square$

On est en mesure maintenant de prouver le Théorème 10 :

Soit  $\mathcal{A} = (Q, \{i_0\}, F, [[0, l-1]], \delta)$  un automate déterministe émondé reconnaissant  $(X)_l$ . D'après le Lemme 16, pour tout  $s \in Q$ ,  $\mathbb{N} \setminus L_s$  est fini. Soit alors  $N \in \mathbb{N}$  tel que pour tout  $s \in Q$ ,  $N \in L_s$ . Considérons maintenant un entier  $n > l^{N-1}$ .  $(n)_l$  est de longueur  $\geq N$ . Soit  $w$  le préfixe de  $(n)_l$  obtenu en enlevant les  $N$  dernières lettres de  $(n)_l$  et  $s = \delta(i_0, w)$ . Comme  $N \in L_s$ , on peut trouver  $u$  un mot de longueur  $N$  tel que  $\delta(s, u) \in F$ . Alors  $m = [wu]_l \in X$  et  $|m - n| \leq l^N$ . Donc  $X$  est  $l^N$ -quasi-périodique.  $\square$

Voici un résultat sur les ensembles quasi-périodiques utilisé dans la démonstration du théorème de Cobham :

**Proposition 17.** Soit  $X$  un ensemble  $d$ -quasi-périodique. Pour tous  $K, L, h \in \mathbb{N}$  et  $\eta > 0$  réel tel que  $K < L < K + \eta$ , il existe  $x \in X$  et  $y \in \mathbb{N}$  tel que

$$yL \leq xK + h \leq yL + \eta d$$

*Démonstration.* Soit  $r$  le plus petit entier positif vérifiant  $rK + h \leq rL$ . Alors pour tout  $i \geq 1$ ,  $(r-i)L < (r-i)K + h$ , que  $r-i$  soit positif ou négatif. Alors :

$$(r-i)K + h = rK + h - iK < rL - iK < rL - iL + i\eta = (r-i)L + i\eta$$

Ainsi, pour  $1 \leq i \leq d$  on a les inégalités :

$$(r-i)L < (r-i)K + h < (r-i)L + \eta d \quad (*)$$

On peut trouver  $j \in \mathbb{N}$  tel que  $jL + r - d \geq 0$  et  $jK + r - d \geq 0$ . Dans l'inégalité (\*), on ajoute à chaque terme  $jKL$  pour obtenir que pour tout  $1 \leq i \leq d$

$$(jK + r - i)L < (jL + r - i)K + h < (jK + r - i)L + \eta d$$

Comme  $X$  est  $d$ -quasi-périodique, il existe  $1 \leq i \leq d$  tel que  $x = jL + r - i \in X$ , ce qui conclut.  $\square$

## 4 Preuve du théorème de Cobham

**Théorème 18.** *Soient  $k, l \geq 2$  deux entiers multiplicativement indépendants. Alors tout ensemble  $X \subseteq \mathbb{N}$  à la fois  $k$  et  $l$ -rationnel est rationnel. (voir [4] ou [2])*

*Démonstration.* Si  $X$  est fini, le résultat est clair. Supposons donc que  $X$  est infini. On pose pour  $t, p \geq 0$  entiers l'ensemble

$$E_{t,p} = \{y \in \mathbb{N} \mid yk^p + t \in X\}$$

D'après la Proposition 1,  $E_{t,p}$  est  $l$ -rationnel. On définit sur  $\mathbb{N}$  la relation d'équivalence  $\rho_k : x \sim y \pmod{\rho_k}$  si et seulement si  $(x)_k^{-1}(X)_k = (y)_k^{-1}(X)_k$ . L'ensemble  $\mathbb{N}/\rho_k$  est fini car  $X$  est  $k$ -rationnel. Les classes d'équivalence obtenues sont  $k$ -stables dans le sens où si  $x \sim y \pmod{\rho_k}$ , alors pour tout  $p \in \mathbb{N}$ , pour tout entier  $t < k^p$ ,

$$xk^p + t \sim yk^p + t \pmod{\rho_k}.$$

Soit  $(x_1, x_2, \dots, x_n)$  un système de représentants de  $\mathbb{N}/\rho_k$ . Fixons  $1 \leq i \leq n$ . On prend pour  $j \neq i$  un mot  $z_j \in (x_i)_k^{-1}(X)_k \setminus (x_j)_k^{-1}(X)_k$  et on pose  $p_j = |z_{i,j}|$ ,  $t_j = [z_j]_k$ . La classe de  $x_i$  modulo  $\rho_k$  est

$$\bigcap_{j \neq i} E_{t_j, p_j}$$

et donc est en particulier  $l$ -rationnelle.

Pour  $1 \leq i \leq n$ , on considère un automate fini déterministe complet  $\mathcal{A}_i = (Q_i, \{q_i^0\}, F_i, [[0, l-1]], \delta_i)$  reconnaissant la classe de  $x_i$  modulo  $\rho_k$  en écriture en base  $l$ . Soit  $\mathcal{A}$  l'automate produit

$$\mathcal{A} = (Q_1 \times \dots \times Q_n, \{q^0 = (q_1^0, \dots, q_n^0)\}, F_1 \times \dots \times F_n, [[0, l-1]], \delta)$$

où  $\delta((q_1 \dots q_n), a) = (\delta_1(q_1, a), \dots, \delta_n(q_n, a))$ . On définit la relation d'équivalence  $\theta$  sur  $\mathbb{N}$  par  $x \sim y \pmod{\theta}$  si et seulement si  $\delta(q^0, (x)_l) = \delta(q^0, (y)_l)$ .  $\theta$  est un raffinement  $l$ -stable de  $\rho_k$ , c.a.d que si  $x \sim y \pmod{\theta}$  alors  $x \sim y \pmod{\rho_k}$  et pour tout  $q \in \mathbb{N}$ , pour tout entier  $t < l^q$ ,

$$xl^q + t \sim yl^q + t \pmod{\theta}.$$

On note  $c$  le cardinal de  $\mathbb{N}/\theta$ , c.a.d. le nombre d'états de  $\mathcal{A}$ .

On introduit deux mots infinis :  $u = u_1u_2\dots u_n\dots$  sur l'alphabet  $\mathbb{N}/\rho_k$  où  $u_n$  est la classe de  $n$  modulo  $\rho_k$ ; et  $v = v_1v_2\dots v_n\dots$  sur l'alphabet  $\mathbb{N}/\theta$  où  $v_n$  est la classe de  $n$  modulo  $\theta$ . Comme  $\rho_k$  est un raffinement de la relation sur  $\mathbb{N}$  :  $x \sim y$  si et seulement si  $x, y \in X$  ou  $x, y \in \mathbb{N} \setminus X$ , si  $u$  est ultimement périodique, et c'est ce que l'on va montrer, alors  $X$  l'est. Plus précisément, on va montrer qu'il existe un entier  $m$  tel que  $u$  à moins de  $m$  facteurs récurrents de longueurs  $m$ , et on conclura par la Proposition 8.

Soit  $w$  un facteur récurrent de  $u$  de longueur 2. Comme les classes d'équivalence de  $\rho_k$  sont à la fois  $k$  et  $l$ -rationnelles, l'ensemble des indices  $n \in \mathbb{N}$  tel que  $w = u_nu_{n+1}$  est à la fois  $k$  et  $l$ -rationnel (et bien sûr infini). Donc d'après le Théorème 10, cet ensemble est quasi-périodique. On prend donc un entier  $d \geq 1$  tel que pour tout  $w$  facteur récurrent de  $u$  de longueur 2, l'ensemble des indices de début de  $w$  est  $d$ -quasi-périodique. On choisit un réel  $0 < \varepsilon < 1$  tel que  $\frac{c\varepsilon}{1-\varepsilon} < \frac{1}{2}$ , et des entiers  $p, q \geq 0$  tel que  $1 < \frac{l^q}{k^p} < 1 + \frac{\varepsilon}{d}$ .

Soient  $K = k^p$ ,  $L = l^q$  et  $m = \lfloor K(1 - \varepsilon) \rfloor$ . On va montrer que pour tout facteur récurrent  $w$  de longueur  $m$  de  $u$ , il existe  $y \in \mathbb{N}$  tel que

$$u[yL, (y+1)L] = swt$$

avec  $|s| \leq \varepsilon K$ .

Tout d'abord, comme  $w$  est un facteur récurrent de  $u$  de longueur inférieure à  $K$ ,  $w$  apparaît une infinité de fois dans des facteurs de  $u$  de la forme  $u[xK, (x+2)K]$ . Or comme la relation  $\rho_k$  est  $k$ -stable, la valeur d'un facteur  $u[xK, (x+2)K]$  ne dépend que de la valeur de  $u[x, x+2]$ . Dans l'ensemble des valeurs de  $u[x, x+2]$  où  $u[xK, (x+2)K]$  est un facteur contenant  $w$ , une valeur est atteinte une infinité de fois. Il existe donc un ensemble  $d$ -quasi-périodique  $\mathcal{I}$  d'indices tel que pour tout  $x, x' \in \mathcal{I}$ ,  $x \sim x' \pmod{\rho_k}$  et  $(x+1) \sim (x'+1) \pmod{\rho_k}$ , et de plus pour tout  $x \in \mathcal{I}$ ,  $w$  est un facteur de  $u[xK, (x+2)K]$ . Les facteurs  $u[xK, (x+2)K]$  pour  $x \in \mathcal{I}$  sont égaux entre eux et s'écrivent

$$u[xK, (x+2)K] = w'ww''.$$

Posons maintenant  $h = |w'|$  et  $\eta = K\frac{\varepsilon}{d}$ . On a  $K < L < K + \eta$  par construction. En appliquant la Proposition 17. à  $\mathcal{I}$ , (qui est  $d$ -quasi-périodique) et à  $K, L, h, \eta$ , on obtient l'existence d'un entier  $y$  tel que

$$u[yL, (y+1)L] = swt$$

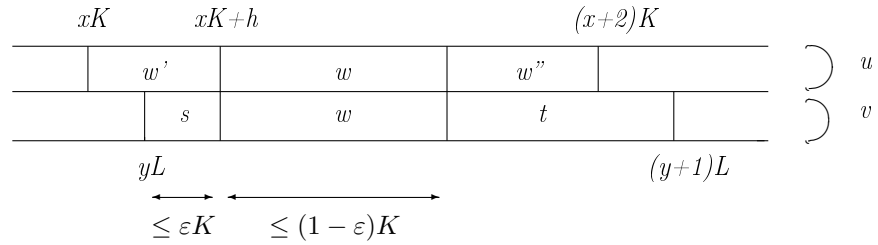
avec  $|s| \leq d\eta = \varepsilon K$ .

Or la valeur de  $u[yL, (y+1)L]$  est entièrement déterminée par la valeur de  $v_y$  car  $\theta$  est un raffinement  $l$ -stable de  $\rho_k$ . Donc  $u[yL, (y+1)L]$  peut prendre au plus  $c$  valeurs. De plus  $|s|$  peut prendre au plus  $\varepsilon K$  valeurs. Donc un facteur récurrent  $w$  de  $u$  de longueur  $m$  peut prendre au plus  $\varepsilon Kc$  valeurs. Or

$$\varepsilon Kc \leq \frac{1}{2}K(1 - \varepsilon) \leq \frac{1}{2}(m + 1) \leq m$$

Donc  $u$  est ultimement périodique. Donc  $X$  est ultimement périodique, soit rationnel.  $\square$

Position du facteur  $w$  :



## Références

- [1] O. Carton. *Langages formels, calculabilité et complexité*. Vuibert, 2008.
- [2] J.-P. Allouche et J. Shallit. *Automatic sequences*. Cambridge University Press, 2003.
- [3] M. Rigo et L. Waxweiler. A note on syndeticity, recognizable sets, and cobham's theorem. *Bull. of the EATCS*, (88) :169–173, February 2006.
- [4] D. Perrin. Finite automata. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 1, pages 1–57. Elsevier, 1990.