

# *Astérisque*

ANNE BERTRAND

## **Nombres de Perron et problèmes de rationalité**

*Astérisque*, tome 198-199-200 (1991), p. 67-76

[http://www.numdam.org/item?id=AST\\_1991\\_\\_198-199-200\\_\\_67\\_0](http://www.numdam.org/item?id=AST_1991__198-199-200__67_0)

© Société mathématique de France, 1991, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (<http://smf4.emath.fr/Publications/Asterisque/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# NOMBRES DE PERRON ET PROBLEMES DE RATIONALITE

par

Anne BERTRAND

## 1. Le Théorème de Perron.

On dit qu'une matrice carrée  $B$  est primitive s'il existe un entier  $k$  tel que  $B^k$  ait tous ses coefficients strictement positifs. Le Théorème de PERRON, qui date du siècle dernier, affirme que :

**THÉORÈME DE PERRON :** *Toute matrice primitive  $B$  à coefficients positifs ou nuls admet une valeur propre  $\lambda$  strictement positive telle que pour toute autre valeur propre  $\mu$  de  $B$  :*

$$|\mu| < \lambda.$$

*Le nombre  $\lambda$  est dit valeur propre strictement dominante de  $B$ .*

Ceci est vrai en particulier pour les matrices à coefficients strictement positifs, qui sont forcément primitives.

On dit qu'une matrice carrée d'ordre  $B = (b_{ij})$  est réductible s'il existe une partition de  $\{1, \dots, n\}$  en deux ensembles non vides  $I$  et  $J$  tels que

$$\forall (i, j) \in I \times J \quad b_{ij} = 0.$$

Si cela n'est pas vrai  $B$  est dite irréductible. FROBENIUS a complété les résultats de PERRON en établissant que :

**THÉORÈME DE FROBENIUS :** *Soit  $B$  une matrice carrée irréductible à coefficients positifs ou nuls ; alors  $B$  admet une valeur propre réelle positive  $\lambda$  telle que toute autre valeur propre  $\mu$  vérifie*

$$|\mu| \leq \lambda$$

*et si  $|\mu| = \lambda$ , alors il existe une racine de l'unité  $e^{2i\pi/h}$  telle que  $\mu = e^{2i\pi/h} \lambda$  ; le spectre de  $B$  est invariant par rotation d'angle  $\frac{2\pi}{h}$ .*

S.M.F.

Astérisque 198-199-200 (1991)

*Exemples* : la matrice  $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$  est primitive : son carré est  $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  et ses valeurs propres sont  $\frac{1 + \sqrt{5}}{2}$  ( $\sim 1,6$ ) et  $\frac{1 - \sqrt{5}}{2}$  ( $\sim -0,6$ ).

La matrice  $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$  est irréductible, la matrice  $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$  ne l'est pas ( $b_{ij} = 0$  si  $(i, j) \in \{1, 2\} \times \{3\}$ ).

Dans le cas où  $B$  est à coefficients dans  $\mathbb{N}$  (c'est à ce cas précis que nous nous intéresserons), les valeurs propres de  $B$  sont des entiers algébriques racines du polynôme caractéristique de  $B$  ; si  $\lambda$  est valeur propre strictement dominante d'une matrice primitive  $B$  sur  $\mathbb{N}$ , les conjugués de  $\lambda$  sont aussi valeurs propres de  $B$  et donc tous les conjugués de  $\lambda$  distincts de  $\lambda$  sont en module strictement inférieur à  $\lambda$ . Douglas LIND a baptisé *Nombres de Perron* les entiers algébriques qui sont strictement supérieurs aux modules de leurs conjugués. On peut se poser la question suivante : quels sont exactement les nombres entiers algébriques qui sont valeur propre strictement dominante d'au moins une matrice primitive à coefficients dans  $\mathbb{N}$  ? Et bien, ce sont exactement les nombres de PERRON comme le montre le théorème suivant :

**THÉORÈME (LIND - HANDLEMAN, 1980-1984)** : *Soit  $\lambda$  un nombre de Perron ; alors il existe une matrice primitive  $B$  à coefficients dans  $\mathbb{N}$  dont  $\lambda$  est la valeur propre strictement dominante.*

*Exemple* :  $\frac{1 + \sqrt{5}}{2}$  est un Perron, mais par  $\sqrt{2}$  dont le conjugué est  $-\sqrt{2}$ . Les démonstrations de LIND et HANDLEMAN sont de nature géométrique ; nous voulons présenter ici la trame d'une preuve algébrique basée sur des problèmes de rationalité (de langages et de séries).

## 2. Langages.

Soit  $A$  un alphabet fini, c'est à dire un ensemble fini de symboles ; soit  $A^*$  l'ensemble des mots sur  $A$ , c'est à dire des suites finies sur  $A$  (y compris le mot vide) ; on munit  $A^*$  du produit de concaténation (le concaténé de  $u_1 \cdots u_k$  et  $v_1 \cdots v_h$  est  $u_1 \cdots u_k v_1 \cdots v_h$ ) ; on appelle langage une partie de  $A^*$ .

Etant donné un langage  $L$  sur un alphabet  $A$ , on définit une relation d'équivalence sur les mots de  $A^*$  :  $u \sim v$  si et seulement si

$$\forall a, b \in A^* \quad aub \in L \iff avb \in L$$

( $u$  et  $v$  ont les mêmes contextes dans  $L$ ).

Le langage  $L$  est dit rationnel si le nombre de classes de  $A^*$  modulo cette relation d'équivalence est fini. Par exemple, si  $A = \{0, 1\}$ , si  $L_1$  est l'ensemble des mots sur  $A$  dans lesquels n'apparaissent jamais deux 1 consécutifs ( $L_1 = \{0, 1, 00, 01, 10, 000, 001, 010, 100, 101, \dots\}$ ) alors  $L_1$  est rationnel car les classes modulo  $L_1$  ont pour système de représentants dans  $A^*$  :

- 0  $(a0b \in L_1 \text{ si } a, b \in L_1)$
- 1  $(a1b \in L_1 \text{ si } a, b \in L_1, a \text{ finit et } b \text{ débute par } 0)$
- 01  $(a01b \in L_1 \text{ si } a, b \in L_1, b \text{ commence par } 0)$
- 10  $(a10b \in L_1 \text{ si } a, b \in L_1 \text{ et } a \text{ finit par } 0)$
- 11  $(a11b \text{ n'est jamais dans } L_1).$

L'ensemble des langages rationnels est aussi la plus petite classe de langages contenant les langages finis et stables pour la réunion, le "produit"  $L_1L_2 = \{uv; u \in L_1, v \in L_2\}$  et par l'opération "étoile" :  $L^* = L \cup L^2 \cup L^3 \cup \dots$  (ceci constitue le Théorème de Kleene).

*Exemple* : si  $C = \{0, 10\}$ , alors  $C^*$  est l'ensemble des mots de  $L_1$  finissant par 0 ; on remarque que

$$L_1 = C^*U(C^*1).$$

### 3. Matrices sur $\mathbb{N}$ et systèmes dynamiques.

Un langage  $L$  est dit factoriel si, dès que  $uvw$  est dans  $L$ ,  $v$  y est aussi ; il est dit prolongeable si pour tout  $u$  appartenant à  $L$  on peut trouver des mots de  $A^*$ ,  $a$  et  $b$ , tels que  $aub$  soit dans  $L$  ; il est dit transitif si, dès que  $u$  et  $v$  sont dans  $L$ , on peut trouver un mot  $w$  de  $A^*$  tel que  $uvw$  soit dans  $L$ .

Le langage  $A^*$ , le langage  $L_1$  du §.2 possèdent ces trois propriétés.

**Systèmes dynamiques symboliques.** Considérons l'ensemble  $A^{\mathbb{N}} = \{a_1a_2a_3 \dots; a_i \in A\}$  des suites sur  $A$ , muni de la transformation  $T$  dite shift :  $T(a_1a_2a_3 \dots) = (a_2a_3a_4 \dots)$ . Etant donné un langage  $L$  factoriel et prolongeable on appelle système dynamique symbolique  $S$  associé à  $L$  l'ensemble  $S = \{a_1a_2a_3 \dots; \forall n, p, a_{n+1} \dots a_{n+p} \in L\}$  dont on vérifie qu'il est invariant par  $T$ . Lorsque  $L$  est rationnel on dit que  $L$  est un système sofique ; le système est dit transitif si le langage  $L$  est transitif et il est dit mélangeant s'il existe un entier  $k$  tel que pour tout  $h \geq k$  et pour tout couple  $u, v$  de mots de  $L$  (apparaissant donc dans les suites de  $S$ ) on peut trouver un mot  $w$  de longueur  $h$  (la longueur d'un mot est le nombre de lettres qui le composent) tel que  $uvw$  soit encore dans  $L$ .

Si, par exemple,  $L = L_1$ ,  $S_1$  est l'ensemble des suites de 0 et 1 dans lesquelles on n'observe jamais deux 1 consécutifs.

Les systèmes de MARKOV sont un cas particulier des systèmes sofiques : soit  $P$  un ensemble fini de mots, et  $L$  le langage formé par les mots ne contenant aucun mot de  $P$  ; il est facile de voir qu'il existe un entier  $k$  et un ensemble fini  $M$  de mots de longueur  $k + 1$  tel que le système dynamique associé à  $L$  soit  $S = \{a_1 a_2 \dots; \forall n, a_n \dots a_{n+k} \in M\}$  : la connaissance des  $k$  lettres  $a_n \dots a_{n+k-1}$  détermine les valeurs possibles de  $a_{n+k}$  ; de tels langages sont bien sûr rationnels. Par exemple le système  $S_1$  est un système de MARKOV mélangeant d'ordre 1 pour lequel  $M = \{00, 01, 10\}$ . On peut associer à ce système dynamique la matrice suivante :

$$\begin{pmatrix} b_{00} & b_{10} \\ b_{01} & b_{11} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} : \begin{pmatrix} 1 \text{ car } 00 \in M & 1 \text{ car } 10 \in M \\ 1 \text{ car } 01 \in M & 0 \text{ car } 11 \in M \end{pmatrix}$$

*Autre exemple* : considérons le système d'ordre 2 dans lequel 111 n'apparaît jamais ; on peut lui associer la matrice d'ordre 4 sur  $\mathbb{N}$  :

$$\begin{pmatrix} b_{00,00} & b_{00,01} & b_{00,10} & b_{00,11} \\ b_{01,10} & b_{01,01} & b_{01,10} & b_{01,11} \\ b_{10,00} & b_{10,01} & b_{10,10} & b_{10,11} \\ b_{11,00} & b_{11,01} & b_{11,10} & b_{11,11} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

dans laquelle  $b_{c_1 c_2, d_1 d_2} = 1$  s'il existe un mot  $u_1 u_2 u_3$  de  $L \cap A^3$  tel que  $u_1 u_2 = c_1 c_2$  et  $u_2 u_3 = d_1 d_2$  ;  $b_{c_1 c_2, d_1 d_2} = 0$  sinon. De même, à tout système de MARKOV d'ordre  $k$  sur un alphabet de  $p$  lettres on associe la matrice carrée d'ordre  $p^k$   $b = (b_{c_1 \dots c_k, d_1 \dots d_k})$  où  $b_{c_1 \dots c_k, d_1 \dots d_k}$  vaut 1 si  $d_1 \dots d_{k-1} = c_2 \dots c_k$  et  $d_1 \dots d_k \in M$  et 0 sinon. Lorsque le système est mélangeant la matrice  $B$  est primitive ; d'une façon générale et par un procédé plus sophistiqué (sic), à *tout système sofique mélangeant on peut associer une matrice primitive  $B$  sur  $\mathbb{N}$* .

Que représente pour le système la valeur propre  $\lambda$  strictement dominante de  $B$  ? Remarquons que le nombre de mots de longueur  $k + 1$  apparaissant dans une chaîne de MARKOV est égal à la somme des coefficients de la matrice  $B$  ; de même, le nombre de mots de longueur  $k + n$  apparaissant dans une chaîne de MARKOV est égal à la somme des coefficients de la matrice  $B^n$  ; par exemple, pour le système  $L_1$  avec la matrice  $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $B^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ ,  $B^3 = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \dots$  ; la somme des coefficients de  $B$  vaut 3, celle de  $B^2$ , 5, celle de  $B^3$ , 8 ; or il y a trois mots de longueur deux (00,01,10), cinq mots de longueur 3 (000,001,010,100,101) et huit mots de longueur 4 (0000,0001,0010,0100,0101,1000,1001,1010). Les 2-mots sont ainsi formés à partir des mots de longueur 1 :

$$0 \xrightarrow{b_{00}=1} 00$$

ou bien

$$0 \xrightarrow{b_{01}=1} 01$$

et

$$1 \xrightarrow{b_{10}=1} 10$$

Les 3-mots sont obtenus comme suit :

$$0 \xrightarrow{b_{00}=1} 00 \xrightarrow{b_{00}=1} 000$$

ou bien

$$0 \xrightarrow{b_{01}=1} 01 \xrightarrow{b_{10}=1} 010$$

et

$$1 \xrightarrow{b_{10}=1} 10 \xrightarrow{b_{00}=1} 100$$

ou bien

$$1 \xrightarrow{b_{10}=1} 10 \xrightarrow{b_{01}=1} 101$$

et

$$\begin{aligned} B^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1.1 + 1.1 & 1.1 \\ 1.1 & 1.1 \end{pmatrix} \\ &= \begin{pmatrix} b_{0,0} & b_{0,0} + b_{0,1} & b_{1,0} & b_{0,0} & b_{0,1} + b_{0,1} & b_{1,1} \\ b_{1,0} & b_{0,0} + b_{1,1} & b_{1,0} & b_{1,0} & b_{0,1} + b_{1,1} & b_{1,1} \end{pmatrix} \end{aligned}$$

qui correspondent aux mots :

$$\begin{pmatrix} 000 & \text{et} & 010 & 001 \\ 100 & & & 101 \end{pmatrix}$$

Or la somme des coefficients de  $B^n$  est de l'ordre de  $\lambda^n$  ; il s'ensuit que si  $\omega_n$  désigne le nombre de mots de longueur  $n$  du langage associé à  $S$  alors

$$\lim_{n \rightarrow \infty} \sqrt[n]{\omega_n} = \lambda .$$

Dans notre exemple ou  $L = L_1$ ,  $\lambda = \frac{1 + \sqrt{5}}{2}$  et  $(\omega_1, \omega_2, \omega_3 \dots) = (2, 3, 5, 8 \dots)$  où l'on retrouve la suite de FIBONACCI :  $\omega_n$  peut se mettre sous la forme

$$\alpha \left( \frac{1 + \sqrt{5}}{2} \right)^n + B \left( \frac{1 - \sqrt{5}}{2} \right)^n \sim \alpha \left( \frac{1 + \sqrt{5}}{2} \right)^n .$$

Le nombre  $\text{Log } \lambda$  est dit *entropie* du système dynamique ; on définit de même l'entropie de tout système dynamique symbolique.

Pour terminer ce paragraphe, donnons l'exemple d'un système sofique qui n'est pas un système de Markov : soit

$$S_2 = \{a_1 a_2 \dots ; a_i \in \{0, 1, 2\} ; \forall i, j : a_i \dots a_{i+j} \neq 211 \dots 12\}$$

sur l'alphabet  $\{0, 1, 2\}$ . Ce n'est pas un système de Markov car le nombre de mots exclus (les 211...12) ne peut être fini et si l'on voit arriver 11...1, il faut aller regarder vers la gauche, à une distance qui peut être aussi grande qu'on veut, le premier terme différent de 1 pour savoir si après on peut mettre 2 où si l'on doit se limiter à 0 ou 1. Par contre il est sofique (les classes modulo  $L_2$  sont représentées par (0,1,21,20,121,02,12)). L'entropie du système est  $\text{Log} \left( \frac{3 + \sqrt{5}}{2} \right)$ .

Bien entendu, si  $\text{Log } \lambda$  est l'entropie d'un tel système, alors  $\lambda$  est toujours un nombre de Perron et *pour prouver le théorème de Lind, il suffit de montrer qu'à tout nombre de Perron on peut associer un système sofique  $S$  dont l'entropie est  $\text{Log } \lambda$  :  $\lambda$  sera alors valeur propre strictement dominante de la matrice primitive  $B$  associée à  $S$ .*

#### 4. Séries $\mathbb{N}$ -rationnelles.

Considérons l'ensemble  $E$  des séries rationnelles  $\sum_{n \geq 0} a_n X^n$  à coefficients dans  $\mathbb{C}$  ; on dit qu'une série est propre si le coefficient  $a_0$  est nul et on définit alors l'"étoile"  $S^*$  de la série propre  $S = \sum_{n \geq 1} a_n X^n$  comme :

$$S^* = \frac{1}{1 - S} = 1 + S + S^2 + \dots + S^n + \dots$$

Définissons maintenant l'ensemble des séries  $\mathbb{N}$ -rationnelles : on l'obtient à partir des polynômes (séries "finies") sur  $\mathbb{N}$  comme l'ensemble des langages

rationnels à partir des langages finis : l'ensemble des séries  $\mathbb{N}$ -rationnelles est par définition le plus petit sous-ensemble de  $E$  contenant les polynômes à coefficients dans  $\mathbb{N}$ , stable par addition, par multiplication et par l'opération "étoile" appliquée aux séries propres.

*Remarque* : si dans la définition ci-dessus on remplace  $\mathbb{N}$  par  $\mathbb{C}$  on obtient l'ensemble des fractions rationnelles habituelles sans pôle en 0 ; si on remplace  $\mathbb{N}$  par  $\mathbb{Z}$  on obtient l'ensemble des fractions rationnelles s'écrivant  $\frac{P(X)}{Q(X)}$  avec  $P$  et  $Q$  dans  $\mathbb{Z}[X]$  et  $Q(0) = 1$ . Une série entière à coefficients dans  $\mathbb{Z}$  est rationnelle si et seulement si ses coefficients vérifient une relation de récurrence linéaire à coefficients dans  $\mathbb{Z}$ .

Une série  $\mathbb{N}$ -rationnelle est forcément  $\mathbb{Z}$ -rationnelle et à coefficient dans  $\mathbb{N}$  ; mais toutes les séries  $\mathbb{Z}$ -rationnelles à coefficients dans  $\mathbb{N}$  ne sont pas  $\mathbb{N}$ -rationnelles :

**THÉORÈME (SOITTOLA)** : *Une série  $\mathbb{Z}$ -rationnelle à coefficients dans  $\mathbb{N}$  est  $\mathbb{N}$ -rationnelle si et seulement si elle est l'emboîtement de séries rationnelles possédant une racine strictement dominante*

*Quelques explications* : on dit que la série  $V = v_0 + v_1X + v_2X^2 + \dots$  est l'emboîtement des séries  $S_0, \dots, S_{p-1}$  si pour tout  $m$ ,  $v_{mp+i}$  est le  $m^{\text{ème}}$  coefficient de  $S_i$  : par exemple l'emboîtement de  $1 + 3X + 9X^2 + \dots$  et  $14 + 2X + 2X^2 + \dots$  est  $1 + 14X + 3X^2 + 2X^3 + 9X^4 + 2X^5 + \dots$

Mettons une série  $\mathbb{Z}$ -rationnelle  $V$  sous la forme  $\frac{P}{Q}$  où  $P$  et  $Q$  sont des polynômes de  $\mathbb{Z}[X]$  premiers entre eux avec  $Q(0) = 1$  ; soit  $\lambda$  le rayon de convergence de la série : alors  $\frac{1}{\lambda}$  est racine de  $Q$  et est inférieur ou égal aux modules des autres racines ; si  $\frac{1}{\lambda}$  est strictement inférieur aux modules des autres racines de  $Q$  on dit que la série admet  $\lambda$  comme série strictement dominante ;  $\lambda$  est alors un nombre de Perron !

En plagiant la preuve de ce théorème on peut montrer que :

**LEMME** : *Soit  $\lambda$  un nombre de Perron ; alors il existe une série  $\mathbb{N}$ -rationnelle  $\sum_{n \geq 1} a_n X^n$  telle que  $1 = \frac{a_1}{\lambda} + \frac{a_2}{\lambda^2} + \dots + \frac{a_n}{\lambda^n} + \dots$  ; de plus on peut supposer que le PGCD des entiers  $n$  tels que  $a_n \neq 0$  vaut 1.*

Comme les  $a_i$  sont  $\geq 0$  et le PGCD égal à 1,  $\lambda$  est racine strictement



dominante de la série

$$\frac{1}{1 - (a_1X + a_2X^2 + \dots)}$$

*Exemples de séries  $\mathbf{N}$ -rationnelles* : si  $L$  est un langage rationnel, si  $a_n$  désigne le nombre de mots de longueur  $n$  de  $L$  (la longueur d'un mot est par définition le nombre de mots qui le composent) alors la série  $\sum_{n \geq 1} a_n X^n$  est  $\mathbf{N}$ -rationnelle.

### 5. Codes et systèmes dynamiques.

On appelle code sur un alphabet  $A$  un ensemble  $C$  de mots de  $A^*$  tels que si  $m, p, q, r$  sont des mots de  $C$  alors  $mp = qr$  implique  $m = q$  et  $p = r$  ( $C$  engendre par concaténation un monoïde libre). Par exemple  $\{0, 10\}$  est un code ;  $C_2 = \{0, 1, 20, 210, 2110, \dots, 21111 \dots 10, \dots\}$  en est un aussi ;  $\{a, ab, ba\}$  n'en est pas car  $aba = (ab)a = a(ba)$ . Soit  $C^* = \{m, m_2 \dots m_k; k = 1, 2 \dots; m_i \in C\}$ . Soit  $L(C)$  l'ensemble des mots facteurs de mots de  $C^*$  (un mot  $q$  est facteur d'un mot  $r$  s'il existe deux mots  $u$  et  $v$  sur  $A$  tels que  $r = uqv$ ) ;  $L(C)$  est un langage factoriel prolongeable et transitif ; à ce langage on peut associer comme au §3 un système dynamique symbolique  $S(C)$  ; par exemple si  $C = \{0, 12\}$  ,  $C^*$  est l'ensemble des mots commençant et finissant par 0 et ne comportant jamais deux 1 consécutifs ;  $L(C)$  est le langage  $L_1$  du §2. On peut montrer que si  $C$  est un langage rationnel le langage  $L(C)$  est rationnel et  $S(C)$  est sofique ; s'il est fini c'est un système de Markov (conditions suffisantes mais non nécessaires) ; soit  $c_n$  le nombre de mots de longueur  $n$  du code  $C$  : si le PGCD des entiers  $n$  avec  $c_n \neq 0$  est 1, le système  $S(C)$  est mélangeant (c'est le cas pour  $C = \{0, 10\}$  ; si  $C = \{00, 11\}$  le nombre de 0 entre 10 et 01 est toujours pair dans  $L(C)$  : il n'y a pas mélange). La série  $c_n X^n$  est  $\mathbf{N}$ -rationnelle si le code est rationnel. L'entropie du système est égale à  $\text{Log } \lambda$  où  $\lambda$  est le seul réel  $> 0$  tel que

$$1 = \frac{c_1}{\lambda} + \frac{c_2}{\lambda^2} + \dots$$

car le nombre de mots de longueur  $n$  de  $L(C)$  est comparable au nombre de mots  $d_n$  de longueur  $n$  de  $C^*$  et

$$\sum_{n \geq 0} d_n X^n = \frac{1}{1 - (c_1 X + c_2 X^2 + c_3 X^3 + \dots)}$$

Si le système est mélangeant,  $\lambda$  est racine strictement dominante de  $\sum d_n X^n$ . Selon le §3 il existe une matrice primitive  $B$  dont  $\lambda$  est la valeur propre strictement dominante.

Pour prouver le Théorème de Lind, nous allons donc associer à tout nombre de Perron  $\lambda$  un code  $C$  de PGCD 1 tel que si  $c_n$  est le nombre de mots de longueur  $n$  de  $C$  alors  $1 = \frac{c_1}{\lambda} + \frac{c_2}{\lambda^2} + \dots$ . Pour cela il suffit de prouver le lemme :

LEMME : A toute série propre  $\mathbb{N}$ -rationnelle  $\sum_{n \geq 1} c_n X^n$  on peut associer un code rationnel  $C$  comportant  $c_n$  mot de longueur  $n$ .

Ceci se montre en prouvant que l'ensemble des séries "génératrices" des codes  $\sum_{n \geq 1} c_n X^n$  contient les polynômes sur  $\mathbb{N}$  (par exemple pour  $3X + 2X^2 + X^3$  prendre un alphabet de 10 lettres  $q_1 \dots q_{10}$  et le code  $\{a_1, a_2, a_3, a_4 a_5, a_6 a_7, a_8 a_9 a_{10}\}$  est stable pour l'addition (prendre la réunion de deux codes sur deux alphabets différents), par multiplication (concatener deux codes sur deux alphabets distincts) et par l'opération étoile (au coefficient constant près !).

Or la plus petite classe possédant ces propriétés est celle des séries  $\mathbb{N}$ -rationnelles propres.

## 6. Trame de la preuve du théorème.

Prenons un nombre de Perron  $\lambda$  ; trouvons une série  $\mathbb{N}$ -rationnelle  $\sum_{n \geq 1} c_n X^n$  telle que  $1 = \frac{c_1}{\lambda} + \dots + \frac{c_n}{\lambda^n} + \dots$  et telle que le PGCD des entiers  $n$  avec  $c_n \neq 0$  soit 1 ; à cette série associons un code  $C$  rationnel, à ce code associons un système dynamique sofique mélangeant  $S$  ; à ce système associons une matrice primitive  $B$  sur  $\mathbb{N}$  ;  $\lambda$  est alors valeur propre strictement dominante de  $B$ . ■

N.B. Remarquons qu'on peut choisir  $B$  à coefficient 0 ou 1 ; par ailleurs les conjugués de  $\lambda$  sont valeurs propres de  $B$  mais en général  $B$  en a d'autres : ces "parasites" sont-ils toujours les mêmes ? y-a-t-il des "conjugués" sur  $\mathbb{N}$  ? Quelle est la taille minimale de  $B$  ? autant de questions sans réponse.

## REFERENCES

- [1] LIND D., *The entropies of topological Markov shifts and a related class of algebraic integers. Ergodic Th. and Dynam. Systems* (1984) 283-300.
- [2] BERSTEL B.J. et PERRIN D., *Theory of codes*, Orlando Academic Press, (1985).
- [3] WEISS B., *Subshifts of finite type and sofic systems. Monats. Math. 77* (1973) 462-474.

- [4] WILLIAMS R.F, *Classification of subshifts of finite type*. Ann. of Math. 98 (1973), 120-153 and 99 (1974), 380-381.
- [5] BLANCHARD F. et HANSEL G., *Systèmes codés*. Theor. Comp. Sci. North Holland 44 (1986) 17-44.
- [6] BERSTEL B.J. et REUTENAUER C., *Les séries rationnelles et leurs langages*. Masson, Paris (1984).

Anne BERTRAND  
Département de Mathématiques  
Université de Poitiers  
40, avenue du Recteur Pineau  
86022 POITIERS

*(Reçu le 23 Novembre 1989)*