

# Points génériques de Champernowne sur certains systèmes codes; application aux $\theta$ -shifts

ANNE BERTRAND-MATHIS

L.A. au C.N.R.S. no 226, U.E.R. de Math. et d'Informatique, Université de Bordeaux I, 351, cours de la Libération, F-33405 Talence Cedex, France

(Received 29 October 1985 and revised 3 March 1986)

*Abstract.* We show that if messages of length one, two, three . . . of a code  $X$  which verifies  $\rho_{X^*} < \rho_X$  are concatenated, we obtain a point generic for an invariant measure on the dynamical system associated to the code; this measure is induced by the maximal measure on the tower studied by Hansel and Blanchard.

*Résumé.* Nous montrons que si nous alignons les messages de longueur 1, puis 2, et ainsi de suite, d'un code  $X$  vérifiant  $\rho_{X^*} < \rho_X$ , nous obtenons un point générique pour une mesure invariante sur le système dynamique associé au code en question; nous montrons que cette mesure se trouve être la mesure induite par la mesure d'entropie maximale sur la tour associée au code introduite par Hansel et Blanchard. Nous en déduisons des points génériques pour les systèmes sofiques et les  $\theta$ -shifts munis de leur mesure maximale.

## I. Introduction

Champernowne [4] a montré que le nombre  $x$  dont le développement décimal est  $0,1234567891011121314\dots$  est normal en base 10, c'est à dire que tout bloc de  $k$  chiffres compris entre 0 et 9 y apparait avec la fréquence  $10^{-k}$ , ou de façon équivalente que la suite  $(x10^n)_{n \geq 0}$  est équirépartie modulo un; il avait commencé par montrer ce résultat pour le nombre  $y = 0,12\dots900010203\dots$  obtenu en écrivant tous les blocs de longueur 1, puis ceux de longueur 2, puis 3 et ainsi de suite. Cette construction a été réutilisée par nombre d'auteurs: R. Adler et P. Shields ont montré qu'en concaténant tous les blocs admissibles de longueur 1, 2, . . . etc. d'une chaîne de Markov 'd'entropie maximale' on obtient un nombre (ou point) normal et Smorodinsky a construit des nombres normaux pour les sous-shifts de Markov en combinant la construction de Champernowne avec un procédé de sélection des blocs de longueur donnée; il fait la remarque que ceci s'applique aux  $\theta$ -shifts et en effet S. Ito et I. Shiokawa ont montré qu'en concaténant les blocs admissibles de longueur croissante du  $\theta$ -shift on obtient un point  $v_1 v_2 v_3 \dots$  générique pour la mesure maximale du  $\theta$ -shift; il est prouvé dans [2] que si  $\theta$  est un nombre de Pisot et si

$$x = \frac{v_1}{\theta} + \frac{v_2}{\theta^2} + \frac{v_3}{\theta^3} + \dots$$

alors la suite  $(x\theta^n)_{n \geq 0}$  est équirépartie modulo un.

Nous montrons ici que l'on peut obtenir par le procédé de Champernowne des points 'normaux' des systèmes dynamiques codés définis par G. Hansel et F. Blanchard, pour peu que le code utilisé vérifie une condition assez largement répandue. Les systèmes de Markov et les  $\theta$ -shifts rentrent dans le cadre de ces systèmes codés.

II. *Résultats et définitions*

Soit  $A$  un ensemble fini, et soit  $A^*$  le monoïde libre engendré par  $A$ , c'est-à-dire l'ensemble des mots ou suites finies sur  $A$  muni du produit de concaténation: si  $u = u_1 \dots u_k$  et  $v = v_1 \dots v_k$  alors  $uv = u_1 \dots u_k v_1 \dots v_k$ . Nous noterons  $1$  le mot vide et  $A^+$  l'ensemble  $A^*$  privé du mot vide. Etant donné un mot  $u_1 \dots u_k$  nous noterons  $|u|$  sa longueur, c'est-à-dire l'entier  $k$ .

Nous appellerons langage une partie de  $A^*$ . Nous dirons qu'un mot  $v$  est facteur d'un mot  $u$  s'il existe deux mots  $a$  et  $b$  de  $A^*$  tels que

$$u = avb.$$

Nous dirons qu'une partie  $X$  de  $A^+$  est un code sur l'alphabet  $A$  si elle engendre un sous-monoïde libre  $X^*$  de  $A^*$  et donc si tout mot de  $X^+$  admet une décomposition unique en mots de  $X$ . Nous dirons que le code  $X$  est préfixe si aucun mot  $u$  de  $X$  ne s'écrit  $u = u_1v$  où  $u_1$  appartient à  $X$  et  $v$  à  $A^+$ .

Nous appellerons langage associé à  $X$  l'ensemble des facteurs des mots de  $X^*$ . Nous le noterons  $F(X^*)$ .

Soit  $b_n$  le nombre de mots de  $X$  de longueur  $n$  et soit  $c_n$  le nombre de mots de  $X^*$  de longueur  $n$ , mots que nous appellerons messages de longueur  $n$ , et dont nous noterons  $C_n$  l'ensemble; nous conviendrons que  $b_0 = c_0 = 1$ .

Nous appellerons rayon de convergence du code  $X$  le rayon de convergence  $\rho_X$  de la série  $\sum_{n>0} b_n z^n$  et rayon de convergence de  $X^*$  le rayon  $\rho_{X^*}$  de la série  $\sum_{n>0} c_n z^n$ .

*Systèmes dynamiques symboliques associés à un code*

Considérons l'alphabet  $A$  muni de la topologie discrète,  $A^{\mathbb{N}}$  muni de la topologie produit et du shift  $T$ :

$$T((e_n)_{n \geq 0}) = (e_{n+1})_{n \geq 0}.$$

Nous appellerons cylindre un ensemble de la forme:

$$[u]^n = [u_1 \dots u_k]^n = \{(e_n)_{n \geq 0} \in A^{\mathbb{N}}; e_{n+1} \dots e_{n+k} = u_1 \dots u_k\}.$$

Les cylindres forment une base pour la topologie de  $A^{\mathbb{N}}$ .

Nous appellerons système dynamique symbolique sur  $A$  un fermé  $T$ -invariant de  $A^{\mathbb{N}}$ .

Etant donné un code préfixe  $X_1$  nous appellerons sous-système  $S$  associé au code  $X$  le système formé par les suites  $(e_n)_{n \geq 0}$  de  $A^{\mathbb{N}}$  telles que pour tout  $m$  et tout  $k$  le mot  $e_m, \dots, e_{m+k}$  appartienne au langage  $F(X^*)$  associé à  $X$ . Nous entendrons par mesure sur  $X$  une mesure borélienne positive de masse 1 (probabiliste).

Etant donné une mesure probabiliste  $T$ -invariante  $\mu$  sur un système dynamique symbolique et un élément  $(e_n)_{n \geq 0}$  du système, nous dirons que la mesure  $\mu$  est associée à  $(e_n)_{n \geq 0}$  s'il existe un sous-ensemble infini  $F$  de  $\mathbb{N}$  tel que pour tout mot

$u_1, \dots, u_k$  de  $A^*$

$$\lim_{\substack{N \rightarrow \infty \\ N \in F}} \frac{1}{N} \sum 1 = \mu([u_1, \dots, u_k]^0)$$

où la somme est prise sur l'ensemble des  $p$  tels que

$$e_{p+1}, \dots, e_{p+k} = u_1, \dots, u_k.$$

Nous dirons que le point  $(e_n)_{n \geq 0}$  est générique pour la mesure  $\mu$  si l'on peut prendre  $F = \mathbb{N}$ :  $\mu$  est alors la seule mesure associée à  $(e_n)_{n \geq 0}$ .

Nous appellerons entropie d'une mesure  $T$ -invariante  $\mu$ :

$$h_T(\mu) = \lim_{N \rightarrow \infty} -\frac{1}{N} \sum_{a_1, \dots, a_n \in A^n} \mu([a_1, \dots, a_n]^0) \text{Log } \mu([a_1, \dots, a_n]^0).$$

Nous dirons que deux partitions  $P = P_1, \dots, P_k$  et  $Q = Q_1, \dots, Q_h$  de  $A^{\mathbb{N}}$  sont  $\varepsilon$ -indépendantes (relativement à la mesure  $\mu$ ) si:

$$\sum_{i,j} |\mu(P_i \cap Q_j) - \mu(P_i)\mu(Q_j)| < \varepsilon.$$

Etant donné une partition  $P = P_1, \dots, P_k$ , soit  $T^{-h}(P) = T^{-h}P_1, \dots, T^{-h}P_k$ , et soit pour  $m$  et  $n$  appartenant à  $\mathbb{N}$ :

$$P_{m,n} = \bigvee_{i=m}^n T^{-i}(P)$$

la partition dont les atomes sont les

$$T^{-m}P_{i_m} \cap T^{-m-1}P_{i_{m-1}} \cap \dots \cap T^{-n}P_{i_n}$$

où  $i_k \in \{1, \dots, k\}$ .

Nous dirons que  $P$  est un générateur faible de Bernoulli si la réunion des partitions  $P_{m,n}$  ainsi définies engendre la tribu des boréliens de  $A^{\mathbb{N}}$  et si pour tout  $\varepsilon > 0$  il existe  $q = q(\varepsilon)$  tel que pour tout  $n$  les partitions

$$\bigvee_{i=0}^n T^{-i}P \quad \text{et} \quad \bigvee_{i=q+2n}^{q+2n} T^{-i}P \quad \text{soient } \varepsilon \text{ indépendantes.}$$

L'extension naturelle à un système bilatéral d'un sous-système de  $A^{\mathbb{N}}$  qui possède un générateur faible de Bernoulli est isomorphe au sens de la mesure à un schéma de Bernoulli [10].

### Mesures de Champernowne

PROPOSITION I. Soit  $X$  un code préfixe (vérifiant  $\rho_{X^*} < \rho_X$ ) sur un alphabet  $A$ .

Supposons le P.G.C.D. des longueurs des mots de  $X$  égal à 1. Si nous alignons les messages de longueur 1, puis ceux de longueur 2, puis ceux de longueur 3 et ainsi de suite (dans un ordre quelconque à l'intérieur de chaque groupe de messages de même longueur) pour former une suite  $(e_n)_{n \geq 0}$ , alors la suite  $(e_n)_{n \geq 0}$  appartient au système dynamique associé à  $X$  et elle est générique pour une mesure  $\mu_X$  que nous qualifierons de mesure de Champernowne induite par  $X$ .

La mesure  $\mu_X$  est fortement mélangeante, charge tous les mots, son entropie vaut  $-\text{Log } \rho_{X^*}$  et la partition  $([a]^0, a \in A)$  est un générateur faible de Bernoulli.

Supposons maintenant le P.G.C.D. des longueurs des mots de  $X$  égal à  $q$ . Si nous alignons les messages de longueur  $q$ , puis ceux de longueur  $2q$ , et ainsi de suite pour

former une suite  $(e_n)_{n \geq 0}$ , alors la suite  $(e_n)_{n \geq 0}$  est un élément du système dynamique associé à  $X$ , générique pour une mesure  $\mu_X$  que nous appellerons encore mesure de Champernowne.

La mesure  $\mu_X$  est ergodique, change tous les mots et admet  $-\text{Log } \rho_{X^*}$  pour entropie.

Lorsqu'elle est fortement mélangente (ce qui n'est pas toujours le cas) la partition  $\{[a]^0; a \in A\}$  est un générateur faible de Bernoulli.

Si par exemple  $A = \{0, 1\}$  et  $X = \{0, 10\}$  l'une des suites  $(e_n)_{n \geq 0}$  commence par

0 00 10 000 010 100 0000 0010 0100 1000 1010 ...

Si  $A = (a, b)$  et  $X = \{aa, bbbb\}$  l'une des suites  $(e_n)_{n \geq 0}$  commence par:

aa aaaa bbbb aaaaaa aabbbb bbbbaa aaaaaaaaa ...

La construction qui précède fournit des points génériques pour nombre de processus, en particulier pour les systèmes dynamiques intrinsèquement ergodiques d'entropie topologique  $h$  munis d'un code  $X$  vérifiant  $\rho_{X^*} < \rho_X$  et  $h = -\text{Log } \rho_{X^*}$ : nous obtenons un point générique pour la mesure maximale.

C'est en particulier le cas pour les systèmes sofiques puisqu'ils peuvent être définis par de tels codes [3].

Lorsque  $X$  est un code synchronisant associé à un système synchronisant [3] nous obtenons un point générique pour la mesure de synchronisation.

Exemple. Soit  $S$  le  $\theta$ -shift ([7]); si le développement en base  $\theta$  du nombre  $\theta > 1$  est

$$\theta = a_0 + \frac{a_1}{\theta} + \dots + \frac{a_n}{\theta^n} + \dots$$

alors le système  $S$  est associé au code:

$$X = \{0 \dots (a_0 - 1); a_0 0, a_0 1, \dots, a_0(a_1 - 1); a_0 a_1 0 \dots a_0 a_1(a_2 - 1) \dots\}$$

avec la convention suivante: si  $a_n = 0$  l'ensemble des mots de longueur  $n + 1$  du code est vide.

Le code  $X$  vérifie  $\rho_{X^*} < \rho_X$  et  $\text{Log } \theta = -\text{Log } \rho_{X^*}$  [2]; l'entropie topologique du  $\theta$ -shift est  $\log \theta$  et il est intrinsèquement ergodique [5]. La mesure de Champernowne n'est autre que la mesure maximale et le nombre

$$x = \sum_{n \geq 0} \frac{e_n}{\theta^n}$$

où  $(e_n)_{n \geq 0}$  est définie comme à la proposition I est normal en base  $\theta$  et admet la suite  $(e_n)_{n \geq 0}$  pour  $\theta$ -développement. Ito et Shiokawa ont montré ([6]) qu'on trouve d'autres nombres normaux en alignant les mots admissibles de longueur  $n$  à la queue leu leu mais la suite  $(\omega_n)$  obtenue ainsi se trouve dans  $A^{\mathbb{N}} \setminus S$  bien qu'elle possède la bonne répartition; Rényi a montré que le nombre de mots de longueur  $n$  du  $\theta$ -shift est inférieur à  $\theta^{n-1} / \theta - 1$ ; comme un mot admissible se met sous la forme  $ma_0 a_1 \dots a_k$  où  $m$  est un message de  $X$ , le résultat d'Ito et Shiokawa est un case particulier de la proposition II:

PROPOSITION II. Soit  $X$  un code préfixe vérifiant  $\rho_{X^*} < \rho_X$  sur un alphabet  $A$ ; soit  $S$  le système dynamique associé. Supposons que

(1) tout mot  $u$  du langage associé à  $S$  se met sous la forme  $u = mb$  où  $m$  appartient à  $X^*$ , où  $b$  est le début d'un mot de  $X$ .

(2) Il existe une constante  $C$  telle que le nombre  $w_n$  de mots de longueur  $n$  du langage associé à  $S$  vérifie  $w_n < C/(\rho_{X^*})^n$ .

Alors la suite obtenue en alignant les mots de longueur 1, puis 2, ... admissibles pour  $S$  est un point générique pour la mesure de Champernowne sur  $S$ ; de plus la limite  $w_n/(\rho_{X^*})^n$  existe.

Notons qu'en général la suite obtenue n'est pas dans  $S$ , alors que celle de la proposition I s'y trouve toujours.

Nous dirons qu'un système dynamique symbolique  $S$  sur un alphabet  $A$  est synchronisant s'il existe un mot  $u$  du langage  $L$  associé à  $S$  tel que

$$\forall u, v \in A \text{ avec } wu \text{ et } vu \in L, \forall t \in A^*, \quad wut \in L \Leftrightarrow vut \in L$$

( $wu$  et  $vu$  ont le même avenir); de tels systèmes sont codés par exemple par l'ensemble  $Y$  des mots  $xu$  de  $L$  tels que  $uxu$  appartienne à  $L$  et qu'aucun facteur gauche de  $xu$  ne possède ces deux propriétés (voir [3] où les codes utilisés sont un tout petit peu différents mais possèdent exactement les mêmes propriétés).

**PROPOSITION III.** Soit  $S$  un système dynamique intrinsèquement ergodique et synchronisant dont la mesure maximale charge tous les mots; appelons  $\log \theta$  l'entropie topologique de  $S$  et supposons qu'il existe un nombre  $C$  tel que le nombre  $w_n$  de mots de longueur  $n$  du langage associé à  $S$  soit majoré par  $C\theta^n$ ; alors la suite obtenue en alignant les mots de longueur un, puis deux, puis trois et ainsi de suite du langage dans un ordre quelconque est un point générique pour la mesure maximale sur  $S$ . De plus  $w_n/\theta^n$  admet une limite lorsque  $n$  tend vers l'infini.

*Remarque I.* Les conditions ci-dessus sont équivalentes trois conditions suivantes:

$S$  est intrinsèquement ergodique d'entropie  $\log \theta$

$w_n$  est majoré par  $C\theta^n$

le code  $Y$  défini ci-dessus vérifie  $\text{Log } \theta = -\text{Log } f_{Y^*}$  et  $f_{Y^*} < f_Y$ .

*Remarque II.* En réalité nous montrerons les conclusions de la proposition III pour les systèmes synchronisants vérifiant  $w_n < C\theta^n$ ,  $\theta$  désigne l'entropie du système, pour lesquels le code  $Y$  associé au mot synchronisant  $u$  vérifie

$$\text{Log } \theta = \rho_{Y^*} \quad \text{et} \quad \rho_{Y^*} < \rho_Y$$

et où le nombre  $p_n$  de mots qui ne contiennent pas plus d'une fois le mot  $u$  vérifie

$$\lim P_n/\theta^n = 0$$

Nous dirons qu'un langage  $L$  possède la propriété de spécification s'il existe un entier  $k$  tel que pour tout couple  $(a, b)$  de mots de  $L$ , il existe un mot  $u$  de longueur inférieure ou égale à  $k$  tel que  $aub$  soit encore dans  $L$ ; les langages associés aux systèmes de Markov ou même aux systèmes sofiques admettent la spécification; lorsque le langage associé à un système dynamique admet la spécification le système est intrinsèquement ergodique, synchronisant et la mesure maximale charge tous les mots d'où le corollaire qui suit; lorsque le langage associé à un système mélangeant admet la propriété de spécification nous disons que le système lui même

admet la spécification; s'il n'est pas mélangeant nous disons qu'il admet la semi-spécification.

**COROLLAIRE.** *Soit  $S$  un système admettant la propriété de semi-spécification. Alors la suite obtenue en alignant les mots de longueur un puis deux, . . . , et ainsi de suite, du langage associé à  $S$  est générique pour la mesure maximale.*

Lorsque  $S$  est l'ensemble des suites sur  $\{0, 1, 2, \dots, 9\}$  et  $X$  le code  $\{0, 1, 2, \dots, 9\}$ , les hypothèses des propositions I et II sont toutes vraies et les deux types de suites obtenues sont les mêmes.

*Etude de la tour associée à un code*

Soit  $X$  un code préfixe, soit  $\Omega \subset X^{\mathbb{N}} \times \mathbb{N}$  l'ensemble des couples  $((x_n)_{n \in \mathbb{N}}, i)$  tels que  $1 \leq i \leq |x_0|$ ; définissons l'application  $T$  de  $\Omega$  dans lui-même par:

$$T((x_n)_{n \geq 0}, i) = \begin{cases} (x_n)_{n \geq 0}, i + 1 & \text{si } i < |x_0| \\ (x_{n+1})_{n \geq 0}, 1 & \text{si } i = |x_0| \end{cases}$$

$X^{\mathbb{N}}$  et  $\Omega$  sont supposés munis de leurs tribus boréliennes respectives.

Le système dynamique  $(\Omega, T)$  est dit tour unilatérale associée à  $X$ ;  $X^{\mathbb{N}}$  s'identifie à la base  $(X^{\mathbb{N}}, 1)$  de  $\Omega$ .

Hansel et Blanchard [3] ont démontré que lorsque  $\bar{\mu}$  décrit l'ensemble des mesures de probabilités invariantes sur  $(\Omega, T)$  alors

$$\sup h(\bar{\mu}) = -\text{Log } \rho_{X^*}$$

et les conditions 1 et 2 sont équivalentes:

(1) le code préfixe  $X$  vérifie

$$\sum_{x \in X} \rho_{X^*} = \sum_{n > 0} b_n \rho_{X^*}^n = 1$$

et

$$\sum_{x \in X} |x| \rho_{X^*} = \sum_{n > 0} n b_n \rho_{X^*}^n < \infty$$

( $b_n$  désigne toujours le nombre de mots de  $X$  de longueur  $n$ ).

On dit alors que le code  $X$  est récurrent positif.

(2) Il existe une mesure invariante  $\bar{\mu}_X$  sur  $\Omega$  telle que:

$$h(\bar{\mu}_X) = -\text{Log } \rho_{X^*}.$$

La mesure  $\mu$  est alors unique et c'est l'unique mesure induisant sur  $X^{\mathbb{N}}$  la mesure de Bernoulli définie par

$$\mu([x]) = (\rho_{X^*})^{|x|}$$

pour tout  $x$  appartenant à  $X$ .

Remarquons que si  $\rho_{X^*} < \rho_X$ , alors  $X$  est récurrent positif.

Soit  $f: \Omega \rightarrow A^{\mathbb{N}}$  l'application projection définie par:

$$fT = Tf$$

$f((x_n)_{n \geq 0}, i) = a_i$  où  $a_i$  désigne la  $i$ ème lettre du mot  $x_0$ .

C'est une application continue et toujours selon Hansel et Blanchard pour toute mesure  $\bar{\mu}$  ergodique sur  $(\Omega, T)$

$$h_T(\bar{\mu}f^{-1}) = h_T(\bar{\mu}).$$

PROPOSITION I bis. Soit  $X$  un code préfixe vérifiant  $\rho_{X^*} < \rho_X$ . Soit  $(\Omega, T)$  la tour unilatérale associée. Soit  $(e_n)_{n \geq 0}$  la suite définie à la proposition I.

Alors  $(e_n)_{n \geq 0}$  est l'image par  $f$  d'un point  $(\varepsilon_n)_{n \geq 0}$  de  $\Omega$  et le point  $(\varepsilon_n)_{n \geq 0}$  est générique pour l'unique mesure  $\bar{\mu}_X$  d'entropie  $-\text{Log } \rho_{X^*}$  sur la tour; la mesure de Champernowne  $\mu_X$  est l'image par  $f$  de la mesure  $\bar{\mu}_X$ .

Notons que  $\bar{\mu}_X$  est définie dès que  $X$  est récurrent positif alors que nous ne définissons les points génériques de Champernowne que lorsque

$$\rho_{X^*} < \rho_X.$$

Il peut exister plusieurs codes de rayons différents induisant le même système dynamique: par exemple si  $A = \{0, 1\}$  les codes

$$X = \{0, 1\} \quad \text{avec } \rho_{X^*} = \frac{1}{2} < \rho_X = \infty$$

et  $Y = \{u_1, \dots, u_n \in A^+; \forall k \in \mathbb{N}, u_1, \dots, u_k \text{ contient plus de 0 que de 1 et } u_1, \dots, u_n \text{ contient autant de 0 que de 1}\}$  avec

$$\rho_{Y^*} = \frac{1}{\sqrt{2}} = \rho_Y$$

induisent tous deux le système  $(A^{\mathbb{N}}, T)$  mais  $Y$  ne vérifie pas  $\rho_{Y^*} < \rho_Y$  et n'est même pas récurrent positif.

Nous conjecturons que dans certains cas les codes récurrents positifs associés à un même système  $S$  (ou tout au moins ceux qui vérifient  $\rho_{X^*} < \rho_X$ ) ont tous le même rayon  $\rho_{X^*}$  et donnent lieu à une unique mesure de Champernowne.

Par ailleurs, étant donné un système dynamique codé  $S$ , que peut on dire de l'ensemble  $B = \{\rho_{X^*}; X \text{ est associé au code } X\}$ .  $B$  est-il un intervalle de  $\mathbb{R}$ ?

La preuve des propositions I et I bis repose sur le lemme de comptage suivant:

LEMME I. Soit  $X$  un code préfixe vérifiant

$$\rho_{X^*} < \rho_X$$

et soit

$$\lambda = \frac{1}{\frac{b_1}{\theta} + \frac{2b_2}{\theta^2} + \dots + \frac{nb_n}{\theta^n} + \dots}.$$

Si le P.G.C.D. des longueurs des mots du code est 1, il existe  $d < 1/\rho_{X^*}$  et  $B$  appartenant à  $\mathbb{R}$  tels que:

$$c_n = \frac{\lambda}{(\rho_{X^*})^n} + v_n \quad \text{où } |v_n| < Bd^n.$$

Si le P.G.C.D. des longueurs des mots du code est  $q \neq 1$ , il existe  $d < 1/\rho_{X^*}$  et  $B$  appartenant à  $\mathbb{R}$  tels que:

$$c_n = 0 \quad \text{si } n \not\equiv 0 \pmod q$$

$$c_n = \frac{q\lambda}{(\rho_{X^*})^n} + v_n \quad \text{où } |v_n| < Bd^n \text{ si } n \equiv 0 \pmod q.$$

Exemples. Soit  $A = \{0, 1\}$ , soit  $X = \{0, 10\}$ ; le rayon  $\rho_X$  est infini et  $\rho_{X^*} = 2/(1 + \sqrt{5})$ ; dans ce cas:

$$c_n = \frac{5 + \sqrt{5}}{2} \left(\frac{1 + \sqrt{5}}{2}\right)^n + \frac{5 - \sqrt{5}}{2} \left(\frac{1 - \sqrt{5}}{2}\right)^n.$$

Soit toujours  $A = \{0, 1\}$ , soit  $X = \{00, 1111\}$ ;  $q$  est égal à 2,  $\rho_X$  est infini et  $\rho_{X^*} = \sqrt{2}/(1 + \sqrt{5})$

$$c_n = \frac{5 - \sqrt{5}}{10} \left(\sqrt{\frac{1 + \sqrt{5}}{2}}\right)^n + \frac{5 + \sqrt{5}}{10} \left(\sqrt{\frac{1 - \sqrt{5}}{2}}\right)^n \quad \text{lorsque } n \text{ est pair,}$$

$$c_n = 0 \quad \text{si } n \text{ est impair.}$$

III. Démonstrations

Preuve du lemme I. Il est bien connu que lorsque  $\rho_{X^*}$  est strictement inférieur à  $\rho_X$  alors [3]:

$$\sum_{n>0} b_n (\rho_{X^*})^n = 1$$

et qu'inversement si la seule solution réelle positive de l'équation  $\sum_{n \geq 0} b_n z^n = 1$  est strictement inférieure à  $\rho_X$  elle est égale à  $\rho_{X^*}$ .

Supposons donc  $\rho_{X^*} < \rho_X$ . Comme les  $b_n$  sont positifs ou nuls,  $\rho_{X^*}$  est racine simple de l'équation  $\sum_{n>0} b_n z^n - 1 = 0$  car la dérivée  $\sum_{n>0} n b_n z^{n-1}$  de  $\sum_{n>0} b_n z^n - 1$  ne s'annule pas en  $\rho_{X^*}$ .

S'il existe d'autres solutions à cette équation elles sont de module supérieur ou égal à  $\rho_{X^*}$  car si  $|\alpha| \leq \rho_{X^*}$  alors:

$$\sum_{n>0} b_n \alpha^n = \sum_{n>0} b_n (\rho_{X^*})^n \Rightarrow \forall n > 0 \quad b_n \alpha^n = b_n (\rho_{X^*})^n$$

et donc

ou bien le P.G.C.D. des longueurs des mots de  $X$  est 1 et  $\alpha = \rho_{X^*}$

ou bien le P.G.C.D. des longueurs des mots de  $X$  est  $q$  et pour tout entier  $m$

$$\alpha = e^{2i\pi m/q} \rho_{X^*}$$

est solution de l'équation, simple pour la même raison que ci-dessus.

Remarquons que  $c_k$  est égal au coefficient du terme de degré  $k$  du développement de la série formelle

$$\frac{1}{1 - \sum_{n>0} b_n Y^n} = 1 + \left(\sum_{n>0} b_n Y^n\right) + \left(\sum_{n>0} b_n Y^n\right)^2 + \dots$$

Il existe un disque  $D$  de centre 0, de rayon strictement supérieur à  $\rho_{X^*}$  dans lequel la fonction  $1/(1 - \sum_{n \geq 0} b_n Y^n)$  est méromorphe; dans le cas où le P.G.C.D. des longueurs des mots du code est 1:

$$\frac{1}{1 - \sum_{n>0} b_n Y^n} = \frac{1}{1 - (Y/\rho_{X^*})} + H(Y)$$

où

$$\lambda = \left(\frac{b_1}{\theta} + \frac{2b_2}{\theta^2} + \frac{3b_3}{\theta^3} + \dots\right)^{-1}$$

et faute d'autres pôles la fonction  $H(Y)$  est analytique dans  $D$ . Ainsi:

$$1 - \sum_{n>0} b_n Y^n = \sum_{n \geq 0} \frac{\lambda}{(\rho_{X^*})^n} + v_n$$

où  $v_n$  vérifie la condition demandée.

Le cas d'un P.G.C.D.  $q$  différent de 1 se traite de même où se ramène au cas précédent en considérant le code comme un code sur  $A^q$ , le P.G.C.D. de la longueur des mots étant alors 1.

*Preuve de la proposition I.* L'idée de la construction du point générique est puisée dans Champernowne [4]. Nous poserons dans la suite - c'est plus parlant -

$$\theta = \frac{1}{\rho_{X^*}}.$$

Soit  $B_n$  le bloc formé par les  $c_n$  mots de  $C_n$  mis bout à bout; la longueur de  $B_n$  est  $nc_n$ ; calculons la longueur  $l_n$  du bloc formé par la juxtaposition des blocs  $B_1 B_2 \dots B_k$ :

$$l_n = \sum_{k=1}^n kc_k = \sum_{k=1}^n (\lambda k \theta^k + kv_k).$$

Posons:

$$u_n = \sum_{k=1}^n kv_k$$

$$|u_n| \leq B \left| \sum_{k=1}^n kd^k \right| \leq \frac{Bd(nd^{n+1} - (n+1)d^n + 1)}{(d-1)^2}$$

et il existe une constante  $B'$  telle que

$$|u_n| < B'nd^n.$$

Ainsi:

$$l_n = \frac{\lambda \theta}{(\theta - 1)^2} (n\theta^{n+1} - (n+1)\theta^n + 1) + u_n \left. \vphantom{l_n}} \right\} \tag{1}$$

avec

$$u_n < B'nd^n.$$

Considérons la suite  $e = (e_n)_{n \geq 0}$  définie dans l'énoncé de la proposition. Lorsqu'il s'y rencontre un message  $m$  de longueur  $p$ , trois cas se présentent:

I: Le message se situe à cheval sur deux messages de longueur  $k$  ou  $k+1$  alignés consécutivement dans le bloc  $B_k$  ou à cheval sur  $B_k B_{k+1}$ , cas rare qui n'intervient pas dans les calculs de fréquences.

II: Ce message se situe à l'intérieur d'un des messages de longueur  $k$ ,  $m_1$ , alignés dans  $B_k$  et

$$m_1 = amb$$

où  $a$  et  $b$  sont encore des messages.

III: Ce message se situe à l'intérieur des messages de longueur  $k$  alignés dans le bloc  $B_k$ ,  $m_1$ , tel que

$$m_1 = amb$$

où  $a$  ou  $b$  n'est pas un message.

LEMME II. *La fréquence avec laquelle apparaît un message donné  $m$  de longueur  $|m| = p$  en situation II dans la suite  $(e_n)_{n \geq 0}$  est égale à  $\lambda / \theta^{|m|}$ .*

Le nombre de fois où ce message  $m$  apparaît dans le bloc  $B_k$  des messages de longueur  $k$  en situation II est:

$$\begin{aligned} f_k &= \sum_{i=0}^{k-p} c_i c_{k-p-i} = \sum_{i=0}^{k-p} (\lambda \theta^i + v_i)(\lambda \theta^{k-p-i} + v_{k-p-i}) \\ &= \sum_{i=0}^{k-p} \lambda^2 \theta^{k-p} + 2 \sum_{i=0}^{k-p} v_i \lambda \theta^{k-p-i} + \sum_{i=0}^{k-p} v_i v_{k-p-i} \\ &= x_k + y_k + z_k. \end{aligned}$$

Le terme  $x_n$  vaut  $(k-p)\lambda^2\theta^{k-p}$ .

Le terme  $y_k$  est majoré comme suit:

$$\begin{aligned} |y_k| &< 2 \sum_{i=0}^{k-p} B \lambda d^i \theta^{k-p-i} = 2 B \lambda \theta^{k-p} \left( \sum_{i=0}^{k-p} \frac{d^i}{\theta^i} \right) \\ &= \frac{2 B \lambda \theta^{k-p}}{\left(\frac{d}{\theta}\right) - 1} \left( \left(\frac{d}{\theta}\right)^{k-p-1} - 1 \right) < C \theta^{k-p} \end{aligned}$$

où  $C$  ne dépend ni de  $k$  ni de  $p$ .

Le terme  $z_k$  se majore ainsi:

$$|z_k| \leq \sum_{i=0}^{k-p} B d^i d^{k-p-i} = \sum_{i=0}^{k-p} B_i d^{k-p} < D d^{k-p}$$

où  $D$  ne dépend ni de  $k$  ni de  $p$ .

Le nombre de fois où  $m$  apparaît en situation II dans le bloc  $B_1 B_2 \cdots B_n$  est:

$$g_n = \sum_{k=1}^n f_k = \sum_{k=1}^n (x_n + y_n + z_n)$$

et

$$\begin{aligned} \sum_{k=1}^n x_n &= \sum_{n \geq k \geq p} (k-p)\lambda^2\theta^{k-p} \\ &= \lambda^2 \frac{\theta}{(\theta-1)^2} ((n-p)\theta^{n+1-p} - (n-p+1)\theta^{n-p} + 1) \\ \left| \sum_{k=1}^n y_n \right| &\leq \sum_{k=1}^n C \theta^{k-p} \leq C \frac{\theta^{n+1} - 1}{\theta - 1} \\ \left| \sum_{k=1}^n z_n \right| &\leq \sum_{k=1}^n D d^{k-p} \leq D \frac{d^{k-p+1} - d^{k-1}}{d - 1} \end{aligned}$$

la limite  $g_n/l_n$  est donc égale à  $\lambda / \theta^p$ : un message de longueur  $m$  se trouve en position II dans la suite  $c_n$  avec la fréquence  $\lambda / \theta^{|m|}$ .

Supposons maintenant le P.G.C.D. des longueurs des mots de  $X$  égal à 1.

LEMME III. Désignons par  $\delta_n$  la mesure de Dirac en un point  $u = (u_n)_{n \geq 0}$  de  $S$ . Il existe une seule mesure  $\mu$  adhérente à la suite de mesures

$$\frac{1}{l_n} \sum_{h=1}^{l_n} \delta_{T^h((e_n)_{n \geq 0})}$$

et cette mesure est mélangeante, donc ergodique.

Soit  $\mu$  une mesure adhérente à la suite de mesures en question.

Posons pour tout mot  $\omega$ ,  $\mu(\omega) = \mu([\omega]^0)$ .

Etant donné un message  $m$ , soit  $K(m)$  l'ensemble des messages  $m_1$  tels que:

Il existe  $u$  et  $v$  appartenant à  $A^*$  mais pas à  $X^*$  tels que  $m_1 = umv$ .

$u$  ne peut se mettre sous la forme  $u = m_2u_1$  où  $m_2$  est un message et  $v$  ne peut se mettre sous la forme  $v = v_1m_3$ , où  $m_3$  est un message.

Le message  $m$  apparaît dans la suite  $(e_n)_{n \geq 0}$  chaque fois qu'un message de  $K(m)$  apparaît en situation II (nous négligeons toujours le cas I) et

$$\mu(m) = \sum_{m_1 \in K(m)} a(m_1) \frac{\lambda}{\theta^{|m_1|}}$$

où  $m_1$  contient  $a(m_1)$  fois  $m$ .

Etant donné un mot  $\omega$ , soit  $H(\omega)$  l'ensemble des messages  $m_4$  tels que  $m_4 = u\omega v$  ou  $u$  et  $v$  ne sont pas des messages et où  $u$  ne se met pas sous la forme

$$u = m_5u_1$$

ni  $v$  sous la forme

$$v = v_1m_6$$

où  $m_5$  et  $m_6$  désignent encore des messages.

Alors

$$\mu(\omega) = \sum_{m \in H(\omega)} b(m)\mu(m)$$

où  $b(m)$  est le nombre d'occurrences de  $\omega$  dans  $m$ .

Ceci montre que la suite  $(1/l_n) \sum_{h=1}^{l_n} \delta_{T^h((e_n)_{n \geq 0})}$  n'admet qu'une seule mesure adhérente, donc une limite  $\mu = \mu_X$ .

Supposons maintenant le P.G.C.D. des longueurs des mots de  $X$  égal à 1 et montrons que la mesure  $\mu_X$  est mélangeante.

Soient  $m$  et  $m'$  deux messages, soit  $\varepsilon > 0$ .

Soit  $K_\varepsilon(m)$  un sous-ensemble fini de  $K(m)$  tel que

$$\left| \mu \left( \bigcup_{m_1 \in K_\varepsilon(m)} \mu(m_1) - \mu(m) \right) \right| < \varepsilon$$

et soit  $K_\varepsilon(m')$  défini de même.

Soit  $s$  un entier fixé, que nous supposerons supérieur à 3,

$$\sup \{ |m_1|; m_1 \in K_\varepsilon(m) \cup K_\varepsilon(m') \}.$$

Soit  $f_s(s)$  la fréquence avec laquelle apparaît dans la suite  $(e_n)_{n \geq 0}$  l'ensemble des blocs de la forme  $mam'$  où  $a$  est un mot de longueur  $s$ .

Soit  $g_\varepsilon(s)$  la fréquence avec laquelle apparaissent en situation II l'ensemble des messages de la forme  $m_1 b m_2$  où  $m_1 \in K_\varepsilon(m)$  et  $m_2 \in K_\varepsilon(m')$ , et où  $b$  est un message; nous écrivons:

$$m_1 b m_2 = c_1 m d_1 b d_2 m' c_2.$$

Alors:

$$|g_\varepsilon(s) - f_\varepsilon(s)| < \varepsilon.$$

Evaluons  $g_\varepsilon(s)$ :

$$g_\varepsilon = \sum \frac{\lambda a(m_1) a(m_2)}{\theta^{|m_1|} \theta^{|m_2|} \theta^{|b|}}$$

où la somme est prise sur l'ensemble des triplets  $m_1, m_2, b$  où  $m_1$  appartient à  $K_\varepsilon(m)$ ,  $m_2 \in K_\varepsilon(m')$ ,  $b \in F(X^*)$  et  $|b| = s - |d_1| - |d_2|$ .

Comme le nombre de messages de longueur  $n$  est pour tout  $n$  (le P.G.C.D. des longueurs des mots de  $X$  est 1)  $\lambda \theta^n + v_n$ ,

$$g_\varepsilon(s) = \sum_{\substack{m_1 \in K_\varepsilon(m) \\ m_2 \in K_\varepsilon(m')}} \frac{\lambda^2 (\theta^{|b|} + v_b)}{\theta^{|m_1|} \theta^{|m_2|} \theta^{|b|}} \cdot a(m_1) a(m_2)$$

et comme  $b$  tend vers l'infini avec  $s$ :

$$\begin{aligned} \lim_{s \rightarrow \infty} g_\varepsilon(s) &= \sum_{\substack{m_1 \in K_\varepsilon(m) \\ m_2 \in K_\varepsilon(m')}} \frac{\lambda^2}{\theta^{|m_1|} \theta^{|m_2|}} \cdot a(m_1) a(m_2) \\ &= \mu \left( \bigcup_{m_1 \in K_\varepsilon(m)} m_1 \right) \cdot \mu \left( \bigcup_{m_2 \in K_\varepsilon(m')} m_2 \right). \end{aligned}$$

Si  $\varepsilon$  tend vers 0: les cas  $m_1 b m_2$  ou  $b$  n'est pas un message se raréfient et

$$\lim_{s \rightarrow \infty} \mu(T^{-s} m \cap m') = \mu(m) \mu(m').$$

Nous passerions comme plus haut au cas de mots quelconques par l'intermédiaire des ensembles  $H(\omega)$ , et la mesure  $\mu_X = \mu$  est donc fortement mélangeante, donc ergodique.

Montrons que  $\mu_X$  est la seule mesure associée à la suite  $e$ .

Supposons qu'une autre mesure  $\nu$  soit adhérente à la suite de mesures

$$\frac{1}{n} \sum_{h < n} \delta_{T^h(e)};$$

comme  $l_{n+1}/l_n$  est borné par  $2\theta$  la mesure  $\nu$  serait absolument continue par rapport à la mesure ergodique  $\mu_X$ : c'est impossible et

$$\mu_X = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{h < n} \delta_{T^h(e)}.$$

Il est clair que la mesure ainsi construite charge tous les messages, donc tous les mots facteurs de mots de  $X^*$ .

Déterminons l'entropie de la mesure  $\mu_X$

Soit  $E_n = \{U[m]^0; m \in C_n\}$ .

Comme il y a environ  $\lambda\theta^n$  messages de longueur  $n$ , comme la mesure de  $[m]^0$  est minorée par  $\lambda/\theta^n$ , si  $n$  est assez grand:

$$\mu(E_n) > \frac{\lambda^2}{2}.$$

Soit  $h$  l'entropie de la mesure  $\mu_X$ .

Le théorème de Shannon-Breiman-MacMillan [8] dit que étant donné  $\delta > 0$  et  $\varepsilon > 0$  il existe  $n(\varepsilon, \delta)$  tel que les mots  $(a_1 \cdots a_n)$  de  $A^n$  se partagent en deux classes:

Le groupe  $M$  des 'mauvais atomes' qui vérifie:

$$\sum_{a_1, \dots, a_n \in M} \mu_X([a_1 \cdots a_n]^0) < \delta.$$

Le groupe  $B$  des 'bons atomes': si  $a_1, \dots, a_n$  appartient à  $B$ , alors:

$$e^{-n(h+\varepsilon)} < \mu([a_1 \cdots a_n]^0) < e^{-n(h-\varepsilon)}.$$

Comme sur l'ensemble  $E_n$  de mesure minorée par  $\lambda^2/2$

$$\frac{\lambda}{\theta^n} < \mu([a_1 \cdots a_n]) < e^{-n(h-\varepsilon)}$$

il vient immédiatement

$$h \leq \text{Log } \theta.$$

Supposons  $h < \text{Log } \theta$  et montrons que c'est impossible.

Soit:

$$\varepsilon = \frac{1}{2}(\text{Log } \theta - h).$$

Si  $\delta$  est pris assez petit et  $n$  assez grand la moitié au moins des atomes de  $E_n$  se trouve dans le groupe  $B$ . Soit  $D_n$  la réunion des cylindres en question:

$$\begin{aligned} e^{-n(h+\varepsilon)} < \mu_X([a_1 \cdots a_n]^0) & \quad \text{si } [a_1 \cdots a_n]^0 \in D_n \\ e^{-n(\text{Log } \theta - \varepsilon)} < \mu_X([a_1 \cdots a_n]^0) & \quad \text{si } [a_1 \cdots a_n]^0 \in D_n. \end{aligned}$$

$D_n$  est réunion de  $\lambda\theta^n/4$  atomes au moins et donc

$$\begin{aligned} \mu_X(D_n) & > \frac{\lambda\theta^n}{4} \cdot e^{-n(\text{Log } \theta - \varepsilon)} \\ & > \frac{\lambda}{4} \frac{\theta^n}{\theta'^n} \quad \text{où } \text{Log } \theta' = \text{Log } \theta - \varepsilon. \end{aligned}$$

Comme  $\theta'$  est strictement inférieur à  $\theta$ ,  $\lim_{n \rightarrow \infty} \mu_X(D_n) = \infty$  ce qui n'est pas: ainsi  $h = \text{Log } \theta$ . □

Indiquons comment l'on peut montrer que la partition  $\{[a]^0; a \in A\}$  est un générateur de Bernoulli (nous ne détaillerons pas la preuve que l'on peut déduire plus aisément de la remarque sur la tour étant donné que la tour associée à un code est un système de Markov dénombrable et que son générateur canonique est faiblement Bernoulli, propriété qui passe au facteur).

Soit

$$Z_{n,q} = \sum_a \sum_b |\mu([a_1, \dots, a_n]^0) \cap T^{-q}([b_1 \cdots b_n]^0) - \mu([a_1 \cdots a_n]^0) \cdot \mu([b_1, \dots, b_n]^0)|$$

où  $a = (a_1, \dots, a_n)$  et  $b = (b_1, \dots, b_n)$  décrivent  $A^n$ .

Soit  $K(a_1 \cdots a_n)$  l'ensemble des messages dont  $a_1 \cdots a_n$  est facteur et qui ne contiennent aucun autre message dont  $a_1, \dots, a_n$  est facteur:  $m = ua_1 \cdots a_nv$ ; nous écrivons

$$Z_{n,q} = \sum_a \sum_b \left| \sum_{\substack{m \in K(a_1 \cdots a_n) \\ m' \in K(b_1 \cdots b_n)}} \mu([m]^0 \cap [m']^{q-|u'|+|u|}) - \mu([m]^0) \cdot \mu([m']^0) \right|$$

Comme d'après la démonstration qui précède

$$\mu([m]^0) = \lambda / \theta^m$$

et

$$\mu([m] \cap [m']^{q-|u'|+|u|}) = \frac{\lambda}{\theta^{|m|+|m'|}} \left( 1 + \frac{v_q - |u'| + |u|}{\lambda \theta^{q-|u'|+|u|}} \right)$$

considérant que

$$\sum_{x \in X} |x| \rho_X^{|x|} = \sum_{x \in X} \frac{|x|}{\theta^{|x|}} < \infty, \quad \text{si } \sum_{\substack{x \in X \\ |x| > h}} \frac{x}{\theta^{|x|}} < \varepsilon$$

on peut 'négliger' les cas où  $m = ua_1 \cdots a_nv$  avec  $|u|$  ou  $|v|$  supérieur à  $h$  et on en déduit aisément que si  $q$  est grand  $Z_{n,q}$  est inférieur, pour tout  $n$ , à tout nombre fixé d'avance.

*Cas où le P.G.C.D. des longueurs des mots de  $X$  est supérieur à 1*

Soit  $X$  un code préfixe sur un alphabet  $A$ , soit  $q \neq 1$  le P.G.C.D. des longueurs des mots de  $X$ ; soit  $Y$  le code préfixe sur  $A^q$ , dont les mots sont les mots de  $X$ , considéré comme mots su  $A^q$ . Le P.G.C.D. des longueurs des mots de  $Y$  est alors 1 et

$$\begin{aligned} \rho_Y &= (\rho_X)^q \\ \rho_{Y^*} &= (\rho_{X^*})^q. \end{aligned}$$

Si nous alignons les messages de longueur 1, puis 2, puis 3  $\cdots$  de  $Y^*$  pour former une suite  $(f_n)_{n \geq 0}$  de  $(A^q)^{\mathbb{N}}$  - ce qui revient à aligner les messages de longueur  $q$ , puis  $2q$ , puis  $3q \cdots$  de  $X^*$  pour former une suite  $(e_n)_{n \geq 0}$  de  $A^{\mathbb{N}}$ , alors la suite  $(f_n)_{n \geq 0}$  est générique pour une mesure  $\mu_c^q$  sur  $(A^q)^{\mathbb{N}}$  fortement mélangeante, ergodique et d'entropie  $-q \text{Log } \rho_{X^*}$  par rapport au shift sur  $(A^q)^{\mathbb{N}}$ , et il s'ensuit que la suite  $(e_n)_{n \geq 0}$  est générique pour la mesure  $T$ -invariante sur  $A^{\mathbb{N}}$

$$\mu_c = \frac{1}{q} (\mu_c^q + T^{-1} \mu_c^q + \cdots + T^{-q+1} \mu_c^q)$$

qui est une mesure d'entropie  $-\text{Log } \rho_{X^*}$  et ergodique, et charge tous les mots; le cas du générateur de Bernoulli se traiterait comme précédemment, en remarquant que si la mesure  $\mu_X$  est fortement mélangeante:

$$\lim_{q \rightarrow \infty} \sum_{m \in K(a_1 \cdots a_n)} \sum_{m' \in K(b_1 \cdots b_n)} \mu([m]^0) \mu(T^{q-|u'|+|u|}[m']^0) = \mu([m]^0) \cdot \mu([m']^0).$$

*Preuve de la proposition I bis.* Les  $c_n$  messages de  $c_n$  s'écrivent, si nous les décomposons en mots de  $X$ :

$$(x_{1,n}^1, \dots, x_{1,n}^{k_1}, x_{2,n}^1, \dots, x_{2,n}^{k_2}, \dots, x_{c_n,n}^1, \dots, x_{c_n,n}^{k_{c_n}})$$

où pour tout  $i, j$ ,  $x_{i,n}^j$  est un mot de  $X$  et pour tout  $i$ :

$$\sum_{j=1}^{k_i} |x_{i,n}^j| = n.$$

La suite  $(\varepsilon_n)_{n \geq 0}$  définie par:

$$\begin{aligned}
 (\varepsilon_0, \varepsilon_1, \dots) = & (x_{1,1}^1, 1)(x_{1,1}^1, 2) \cdots (x_{1,1}^1, |x_{1,1}^1|)(x_{2,1}, 1)(x_{2,1}, 2) \cdots \\
 & \cdots (x_{c_{n-1}, n-1}^k, n-1)(x_{c_{n,1}}^1, 1)(x_{c_{n,1}}^1, 2) \cdots (x_{c_{n,1}}^1, |x_{c_{n,1}}^1|)(x_{c_{n,1}}^2, 1) \cdots \\
 & \cdots (x_{c_{n,1}}^{k_1}, |x_{c_{n,1}}^{k_1}|)(x_{c_{n,2}}^1, 1) \cdots
 \end{aligned}$$

est l'antécédent par  $f$  de la suite  $(e_n)_{n \geq 0}$  définie à la proposition I et nous montrerions comme précédemment que le point  $(\varepsilon_n)_{n \geq 0}$  est générique pour une mesure  $\nu$  dont l'image par  $f$  est la mesure  $\mu_X$ ; l'entropie de  $\nu$  est donc minorée par  $-\text{Log } \rho_{X^*}$  qu'elle ne peut dépasser:  $\nu$  est donc l'unique mesure  $\bar{\mu}$  d'entropie  $-\text{Log } \rho_{X^*}$  sur la tour et  $\mu_X = f(\bar{\mu})$ .

Nous aurions pu, au lieu de calculer directement l'entropie de la mesure  $\mu_X$ , montrer que la restriction de  $\nu$  à la base de la tour ([3]) est à un facteur près la mesure de Bernoulli d'entropie  $-\text{Log } \rho_{X^*}$  ce qui implique ([3]) que  $\nu = \bar{\mu}$  et comme  $\mu_X = f(\bar{\mu}_X)$ :

$$h(\mu_X) = h(\bar{\mu}_X) = -\text{Log } \rho_{X^*}.$$

*Preuve de la proposition II*; (on supposera  $X$  de P.G.C.D. un, ce qui ne change pas le fond du problème); Soit  $X$  un code préfixe vérifiant les hypothèses de II; le nombre de mots de longueur  $n$  du langage  $L$  associé à  $S$  se met sous la forme

$$w_n = c_n + c_{n-1}v_1 + c_{n-2}v_2 + \cdots + c_{n-k}v_k + \cdots + c_0v_n$$

où  $c_h$  est le nombre de messages de longueur  $h$  et  $v_h$  le nombre de facteurs gauches de longueur  $h$  des mots de  $X$  (un mot de longueur  $n$  s'écrit  $ma$  où  $m$  appartient à  $X$  et où  $a$  est le début d'un mot de  $X$ , de façon unique car  $X$  est préfixe); nous poserons  $v_0 = 1$  pour avoir de plus jolies formules.

Posons  $\theta = 1/\rho_{X^*}$ ; d'après le lemme I,  $c_n = \lambda\theta^n + d_n$  où  $d_n$  est majoré par  $Bd^n$  avec  $d < \theta$ ;  $v_n$  est majoré comme  $w_n$  par  $C\theta^n$ ; ainsi

$$\begin{aligned}
 w_n = & \lambda(v_0\theta^n + v_1\theta^{n-1} + \cdots + v_k\theta^{n-k} + \cdots + v_n) + d_nv_0 + d_{n-1}v_1 + \cdots + d_0v_n. \\
 \frac{w_n}{\theta^n} = & \lambda\left(v_0 + \frac{v_1}{\theta} + \cdots + \frac{v_n}{\theta^n}\right) + \frac{1}{\theta^n}(v_0d_n + v_1d_{n-1} + \cdots + v_nd_0)
 \end{aligned}$$

Le deuxième terme est majoré par

$$\lambda B\left(\frac{d^n}{\theta^n} + \frac{d^{n-1}}{\theta^{n-1}} + \cdots + \frac{d}{\theta} + 1\right)$$

et donc la série de terme général  $v_k/\theta^k$  converge puisque  $w_n/\theta^n$  est borné; ainsi lorsque  $n$  tend vers l'infini,  $(d_nv_0 + \cdots + d_0v_n)$  tend vers 0 (c'est le terme général du produit des séries absolument convergentes  $v_k/\theta^k$  et  $d_k/\theta^k$ ). Alors  $w_n/\theta^n$  admet aussi une limite  $D$  lorsque  $n$  tend vers l'infini et

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \left| \frac{w_n}{\theta^n} - \frac{c_nv_0 + c_{n-1}v_1 + \cdots + c_{n-\lfloor \sqrt{n} \rfloor}v_{\lfloor \sqrt{n} \rfloor}}{\theta^n} \right| & \leq \lim_{n \rightarrow \infty} \frac{c_{n-\lfloor \sqrt{n} \rfloor-1}v_{\lfloor \sqrt{n} \rfloor+1} + \cdots + c_0v_n}{\theta^n} \\
 & = 0
 \end{aligned}$$

Montrons que la mesure de Champernowne sur  $S$  est la seule mesure associée à la suite  $g^1g^2\cdots$  obtenue par concaténation des mots de longueur croissante qui sont admissibles pour  $S$ ; nous donnerons seulement la trame de la démonstration car

les justifications des passages à la limite et le choix des morceaux de suite à négliger se font comme pour la proposition I; considérons le bloc des mots de longueur  $n$ ; sa longueur totale est un équivalent de  $nD^n$ ; la fréquence d'apparition des mots composés d'un message de longueur inférieure à  $n - \sqrt{n}$  et d'un autre mot est négligeable d'après la dernière inégalité obtenue; seuls comptent les mots formés d'un message de longueur supérieure à  $n - \sqrt{n}$  et d'un bout de mot du code de longueur inférieure à  $\sqrt{n}$ ; dans un tel mot le bout de longueur inférieure à  $\sqrt{n}$  est négligeable; le bloc des mots de longueur  $n$  se ramène donc à l'ensemble des messages de longueur  $k$ , répétés  $v_{n-k}$  fois où  $k$  parcourt l'ensemble des entiers compris entre  $n - \sqrt{n}$  et  $n$ ; comme  $v_{n-k}$  n'est pas trop grand (la série de terme général  $v_n/\theta^h$  converge) les  $v_k$  répétitions et l'ordre choisi pour aligner les mots n'influent pas sur la répartition qui est la même que celle rencontrée à la proposition I. (ici encore l'ergodicité de la mesure de Champernowne intervient).

*Justification de la remarque I.* D'après [3] s'il existe une mesure d'entropie  $\log \theta$  qui charge tous les mots du langage  $L$  associé à  $S$ , elle est l'image d'une mesure sur la tour associée au code  $Y$  et  $\rho_{Y^*} = 1/\theta$ ; si  $\rho_{Y^*} = \rho_Y$ , enlevons un mot  $su$  à  $Y$ ; le rayon  $\rho_Y$  ne change pas, le rayon  $\rho_{Y^*}$  diminue car le code  $Z = Y - \{su\}$  définit un système dynamique strictement inclus dans  $S$  puisque le mot  $usu$  n'y figure pas, et donc d'entropie strictement inférieure à celle de  $S$  à cause de l'ergodicité intrinsèque; on a donc  $\rho_{Z^*} > \rho_Z$  ce qui est absurde.

*Preuve de la proposition III.* Soit  $v$  un mot de longueur  $n$  du langage  $L$  associé à  $S$ ; s'il contient au moins une fois le mot  $u$  il s'écrit  $v = autub$  où  $tu$  est un message ou  $aub$  ( $aumb$  où  $m$  est le mot vide) d'après [3]; comme  $u$  est synchronisant il y a 'indépendance' entre  $a$ ,  $b$  et  $tu$ : si  $au$ ,  $utu$  et  $ub$  sont dans  $L$   $autub$  y est aussi ([3]) et le nombre  $w_n$  de mots de longueur  $n$  de  $L$  s'écrit  $w_n = v'_0c_n + v'_1c_{n-1} + \dots + v'_nc_n + p_n$  où  $p_n$  désigne le nombre de mots de longueur  $n$  qui ne contiennent pas plus d'une fois le mot  $u$ ; si nous montrons que  $p_n/\theta^n$  tend vers 0 lorsque  $n$  tend vers l'infini, alors nous terminerons la démonstration comme à la proposition II; si cela était faux il existerait  $b$  avec  $p_n > b\theta^n$  pour une infinité de  $n$ ; soient  $f$  et  $f+2g$  de tels entiers, et soit  $E_g$  l'ensemble des mots de longueur  $f$  qui se prolongent en un mot  $arb$  de  $L$  où  $a$  et  $b$  sont de longueur  $g$ , soit  $a_f\theta^f$  le cardinal de  $E_g$ ; un mot ne contenant pas deux fois le mot  $u$  est produit de trois mots du même type de longueurs respectives  $g$ ,  $f$ ,  $g$  et

$$b\theta^{f+2g} < P_{f+2g} < C\theta^g \cdot a_f\theta^f \cdot C\theta^g \quad \text{et donc } q_f > b/C^2;$$

comme les  $E_g$  sont emboîtés les uns dans les autres il apparaît que l'ensemble des suites qui ne contiennent pas deux fois le mot  $u$  est un fermé invariant d'entropie  $\log \theta$  ce qui est impossible cause de l'ergodicité intrinsèque.

Je remercie Roland Perrier qui m'a aimablement signalé une erreur dans une erreur dans une preuves.

BIBLIOGRAPHIE

[1] R. Adler & P. Shields. Irreducible Markov Shifts. *Ann. of Math. Stat.* **43** (1972).  
 [2] A. Bertrand-Mathis. Développements en base  $\theta$ , répartition modulo un de la suite  $(x\theta^n)_{n \geq 0}$ , langages codés et  $\theta$ -shifts. *Bull. de la S.M.F.* (1986).

- [3] F. Blanchard & G. Hansel. Systèmes codés. *Theoretical comp. Sci.* (1986).
- [4] D. Champernowne. The construction of decimals normal in the scale of ten. *J. London Math. Soc.* **8** (1950), 254–260.
- [5] F. Hofbauer.  $\beta$ -shifts have unique maximal measure. *Monatshefte* **85** (1978), 189–198.
- [6] S. Ito & I. Shiokawa. A construction of  $\beta$  normal sequences. *J. Math. Soc. Japan*, vol. **27**, 1 (1975).
- [7] W. Parry. On the  $\beta$ -expansion of real numbers. *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.
- [8] M. Smorodinsky. *Ergodic Theory, Entropy*. Lecture Notes in Mathematics, n° 247 (1971).
- [9] M. Smorodinsky. Normal numbers for Markov Chains. To appear.
- [10] M. Smorodinsky.  $\beta$ -automorphism are Bernoulli shifts. *Acta Math. Acad. Sci. Hungaricae* **T24**, (3–4) (1973), 273–8.