BULLETIN DE LA S. M. F.

SÉBASTIEN FERENCZI

Les transformations de Chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité $2n+1$

Bulletin de la S. M. F., tome 123, nº 2 (1995), p. 271-292

<http://www.numdam.org/item?id=BSMF_1995__123_2_271_0>

Bull. Soc. math. France,
123, 1995, p. 271–292.

LES TRANSFORMATIONS DE CHACON : COMBINATOIRE, STRUCTURE GÉOMÉTRIQUE, LIEN AVEC LES SYSTÈMES DE COMPLEXITÉ $2n+1$

PAR

SÉBASTIEN FERENCZI

RÉSUMÉ. — Nous montrons que la transformation de Chacon est un système de complexité $(2n-1)$; nous décrivons le graphe des mots associé, et en déduisons une forme primitive de la substitution, et une représentation géométrique de la transformation comme exduction d'une rotation triadique ; nous calculons enfin la complexité de systèmes plus généraux, substitutifs ou non, qui sont les systèmes faiblement mélangeants les plus simples connus.

ABSTRACT. — We show that Chacon's map is a system of complexity $(2n-1)$; we describe the associated graph of words, and use it to give a primitive form of the substitution, and a geometric representation of the transformation, as an exduction of a triadic rotation ; then we compute the complexity of some more general systems, substitutive or not, which are the simplest known weakly mixing systems.

1. La transformation de Chacon classique

La transformation de Chacon classique se définit (*cf.* [2]) de la manière suivante : on construit des blocs de zéros et de uns par la règle

$$B_0 = 0, \quad B_{n+1} = B_n B_n 1 B_n,$$

où la multiplication dénote la juxtaposition des blocs. La transformation T sera le décalage $(Tx)_n = x_{n+1}$ sur l'ensemble $\mathcal{X}$ des suites bi-infinies (x_n) de zéros et de uns vérifiant : pour tout $s < t$, il existe n tel que le mot $x_s ... x_t$ soit un sous-mot de B_n. On appelle *langage associé à la*

(*) Texte reçu le 17 novembre 1993, révisé le 14 mars 1994.
Sébastien FERENCZI, Laboratoire de Mathématiques Discrètes, CNRS-UPR 9016, Case 930, 163 avenue de Luminy, 13288 Marseille CEDEX 9, France.
Email : ferenczi@lmd.univ-mrs.fr.
Classification AMS : 28D.

transformation T l'ensemble de tous les mots $x_s \cdots x_t$ qui sont contenus dans au moins un B_n. La suite infinie u commençant par B_n pour tout $n \geq 0$ est dite *suite associée à*, ou *suite définissant* la transformation de Chacon T.

La transformation de Chacon est connue pour ses nombreuses propriétés ergodiques, voir par exemple [4]; nous allons voir que sa combinatoire est intéressante également, car, à une petite modification près, qui est de toute manière un isomorphisme topologique, elle rentre dans la classe déjà connue des *systèmes de complexité* $2n+1$ (*cf.* [1]); de plus, nous pouvons faire une étude complète de sa combinatoire, ce qui est intéressant pour l'étude générale de cette classe, et en dégager une structure géométrique : la transformation de Chacon apparaît comme une *tour* au-dessus d'une rotation triadique — l'existence de telles structures dans la classe des systèmes de complexité $2n+1$ n'était pas connue jusqu'ici. Nous généraliserons enfin ces propriétés, notamment l'algorithme de calcul de la complexité, à des transformations voisines, y compris à des cas où le découpage est variable et donc le système n'est pas substitutif; les complexités de ces systèmes oscillent en général entre $(1+\epsilon)n$ et $2n$ (des complexités de systèmes non substitutifs n'étaient connues jusqu'ici que pour des échanges d'intervalles, et par des méthodes différentes car géométriques); certains d'entre eux sont de bons candidats pour réaliser la complexité minimale d'une transformation faiblement mélangeante.

2. Le langage associé

Soit h_n la longueur du bloc B_n, donnée par la formule de récurrence

$$h_{n+1} = 3h_n + 1$$

à partir de $h_0 = 1$, ce qui implique $h_n = \frac{1}{2}(3^{n+1} - 1)$. Soit ℓ un entier compris entre h_n et h_{n+1}; un mot de longueur ℓ du langage est repéré par son abscisse k de départ dans le bloc B_n, avec $0 \leq k \leq h_n - 1$, ou bien peut commencer par une lettre qui n'est pas dans B_n; cette lettre est alors nécessairement un 1, et — dans la tradition ergodique — on l'appelle un *spacer*, ou, en français, *séparateur*; dans ce dernier cas, nous noterons $k = -1$. Un mot W de longueur ℓ est entièrement connu si on connaît k, le nombre et la configuration des séparateurs situés entre les blocs B_n qui apparaissent dans W; suivant la valeur de ℓ, le mot W coupe de un à quatre blocs B_n, et on vérifie qu'entre eux toutes les configurations de séparateurs peuvent exister, compte tenu des contraintes suivantes : entre deux blocs consécutifs, il y a zéro ou un séparateur; entre quatre blocs consécutifs, il y a au moins un séparateur et il ne peut y en avoir trois.

Nous pouvons dresser la liste exhaustive de tous les mots que l'on peut voir pour des valeurs données de k et ℓ (ces mots n' existent pas pour toutes les valeurs de k et de ℓ, et n'ont pas de raison d'être tous différents).

On a les noms suivants, si l'on note par des traits pleins le n-bloc et par des 1 les séparateurs (le premier indice indiqué entre parenthèses est l'abscisse de départ du nom dans son bloc B_n, le deuxième est son abscisse d'arrivée dans un bloc B_n ultérieur ; l'abscisse de départ est bien sûr omise quand le nom commence par un séparateur) :

$$
\begin{array}{rl}
A(k,\ell): & -(k)--(k+\ell-1)-\ , \\
B(k,\ell): & -(k)-1-(k+\ell-2)-\ , \\
A'(k,\ell): & -(k)-\text{——}-(k+\ell-1)-\ , \\
B'(k,\ell): & -(k)-1\text{——}-(k+\ell-2)-\ , \\
C'(k,\ell): & -(k)-\text{——}1-(k+\ell-2)-\ , \\
D'(k,\ell): & -(k)-1\text{——}1-(k+\ell-3)-\ , \\
A''(k,\ell): & -(k)-1\text{——}\ \text{——}-(k+\ell-2)-\ , \\
B''(k,\ell): & -(k)-\text{——}1\text{——}-(k+\ell-2)-\ , \\
C''(k,\ell): & -(k)-\text{——}\ \text{——}1-(k+\ell-2)-\ , \\
D''(k,\ell): & -(k)-1\text{——}\ \text{——}1-(k+\ell-3)-\ , \\
E''(k,\ell): & -(k)-\text{——}1\text{——}1-(k+\ell-3)-\ , \\
F''(k,\ell): & -(k)-1\text{——}1\text{——}-(k+\ell-3)-\ , \\
Z(-1,\ell): & 1\text{——}(\ell-2)\,, \\
A(-1,\ell): & 1\text{——}-(\ell-2)-\ , \\
B(-1,\ell): & 1\text{——}1-(\ell-3)-\ , \\
A'(-1,\ell): & 1\text{——}\ \text{——}-(\ell-3)-\ , \\
B'(-1,\ell): & 1\text{——}1\text{——}-(\ell-3)-\ , \\
C'(-1,\ell): & 1\text{——}\ \text{——}1-(\ell-3)-\ , \\
C''(-1,\ell): & 1\text{——}\ \text{——}\ \text{——}1-(\ell-3)-.
\end{array}
$$

3. Calcul de la complexité

Toutes les transformations que l'on considèrera sont données avec une suite infinie u et un langage associés ; ce ne sont pas des invariants, topologiques ou autres, du système considéré. Avec les restrictions ci-dessus, nous appellerons *fonction de complexité* , ou simplement *complexité* d'un système dynamique symbolique (X,T), ou d'une transformation T, donnés avec un langage, la fonction qui à n associe le nombre de mots de longueur n du langage associé.

Si nous regardons maintenant quels mots peuvent exister suivant les valeurs de k et ℓ, nous obtenons le tableau suivant, que nous appellerons *le tableau d'ordre n*, pour $h_n + 1 \leq \ell \leq h_{n+1} + 1$ et $-1 \leq k \leq h_n - 1$:

$\ell =$	$k=-1$	$k=0$					$k=h_n-1$
h_n+1	Z	AB					AB
h_n+2	AB	AB					$BA'C'$
h_n+3	AB	AB					$A'B'C'D'$
			$k+1-\ell=2h_n-1$	$k+\ell-1=2h_n$	$k+\ell-1=2h_n+1$		
	AB	AB	AB	$BA'C'$	$A'-D'$		$A'-D'$
$2h_n+1$	AB	$BA'C'$					$A'-D'$
$2h_n+2$	$BA'C'$	$A'-D'$					$B'C'D'C''$
$2h_n+3$	$A'-C'$	$A'-D'$					$D'A''-E''$
			$k+\ell-1=3h_n-1$	$k+\ell-1=3h_n$	$k+\ell-1=3h_n+1$	$k+\ell-1=3h_n+2$	
	$A'-C'$	$A'-D'$	$A'-D'$	$B'C'D'C''$	$D'A''-E''$	$A''-F''$	$A''-F''$
$3h_n+1$	$A'-C'$	$B'C'D'C''$					$A''-F''$
$3h_n+2$	$B'C'C''$	$D'A''-E''$					$A''-F''$

Sur la ligne ℓ, nous montrons tous les mots de longueur ℓ, avec dans la case k ceux qui commencent à l'abscisse k dans le bloc B_n.

On constate que sur la ligne ℓ il y a $(2\ell - 1)$ mots possibles; nous allons montrer qu'il y a exactement $(2\ell - 1)$ mots différents. En effet, il y a trois mots de longueur 2, 00, 01 et 10.

On fait l'hypothèse de récurrence suivante : *sur la première ligne du tableau d'ordre n il y a $2h_n + 1$ mots différents* (c'est-à-dire que tous les mots écrits sur la première ligne du tableau sont effectivement différents). Pour passer d'une ligne à la ligne suivante, les mots situés dans la case k se prolongent en mots situés dans la case k juste en-dessous; ces mots de la ligne $\ell + 1$ sont différents s'ils ont des préfixes différents de longueur ℓ. Le seul problème possible correspond aux cases où un mot de longueur ℓ a deux successeurs de longueur $\ell + 1$; on vérifie que ce cas se produit dans deux cases par ligne et qu'on est dans la situation suivante : l'un des successeurs se termine par un séparateur, donc un 1, et l'autre par la première lettre de B_n qui est toujours un 0, donc ces deux successeurs sont différents (et différents de tous les autres mots de la ligne puisqu'ils ont des préfixes différents). Donc, si tous les mots écrits sur une ligne sont différents, il y en a exactement deux de plus sur la ligne suivante, et sur celle-ci à son tour tous les mots écrits sont différents.

Nous avons donc montré, puisque les tableaux d'ordre n se recouvrent, que l'hypothèse de récurrence pour n entraîne celle pour $n + 1$, et que :

PROPOSITION 1. — *Le nombre de mots de longueur ℓ est égal à $(2\ell - 1)$ pour tous les $\ell \geq 2$. On a en outre une description complète du langage : chaque mot correspond à une référence et une seule dans le tableau d'ordre n.*