

Combinatoire des codages de rotations

par

GILLES DIDIER (Marseille)

1. Introduction. Il existe de nombreuses méthodes pour générer des suites symboliques (i.e. à valeurs sur un ensemble fini). L'une d'elles consiste à découper le cercle unité en un nombre fini d'intervalles en associant une lettre différente à chaque intervalle, puis à considérer, étant donné un point x et une transformation Δ du cercle, la suite dont le $n^{\text{ième}}$ terme est la lettre associée à l'intervalle auquel appartient le point $\Delta^n(x)$.

Nous présentons dans ce travail certaines propriétés des suites générées de cette façon dans le cas où la transformation est une rotation du cercle unité. En particulier, nous en donnons une caractérisation combinatoire complète.

Une classe particulière de ces suites est déjà bien connue. En effet, lorsqu'on considère un découpage en deux intervalles et que l'un des deux est de longueur égale à l'angle, les suites ainsi obtenues sont "sturmiennes". L'étude de ces suites est certainement un des domaines de la dynamique symbolique les plus prolifiques comme en témoigne l'abondante littérature consacrée à ce sujet (voir [5]).

L'intérêt qu'elles suscitent encore à ce jour s'explique par diverses raisons. Ce sont d'une part les suites non-ultimement périodiques les plus simples d'un point de vue combinatoire. Mais elles constituent surtout un des liens les plus établis entre arithmétique et combinatoire des mots.

En effet, dans les deux articles fondateurs de la dynamique symbolique [9, 10], M. Morse et G. A. Hedlund ont montré que les suites sturmiennes, dont on peut donner des définitions combinatoires simples en terme de fonction de complexité (fonction qui associe à tout entier naturel n le nombre de facteurs de longueur n apparaissant dans la suite) ou d'équilibre, peuvent également s'interpréter en tant que codages de rotations particuliers.

L'étude des suites sturmiennes est empreinte de cette dualité et la démonstration de leurs caractéristiques fait souvent appel à des résultats de théorie

1991 *Mathematics Subject Classification*: Primary 11B85.

des nombres. Inversement, on peut utiliser leurs propriétés combinatoires pour démontrer ou donner des démonstrations plus simples de résultats arithmétiques. Par exemple, le théorème des trois distances peut être retrouvé assez naturellement en étudiant leurs langages [2].

Les aspects combinatoires des codages de rotations associés à un découpage quelconque du cercle sont moins bien connus. On sait que leurs fonctions de complexité sont ultimement affines [1, 11] mais cette propriété ne suffit pas à les caractériser. Par exemple, dans le cas où le cercle est partitionné en deux intervalles, les fonctions de complexité sont, pour une large classe, égales à $2n$ pour tout entier naturel non nul n . Il existe cependant des suites de même fonction de complexité qui ne sont pas des codages de rotation, comme on peut s'en convaincre en lisant l'article de G. Rote consacré à ce thème [11].

Dans la même publication, l'auteur montre que, pour un découpage du cercle en deux parties égales, les langages des suites codages de rotations d'angles irrationnels sont "complémentaires symétriques" (i.e. stables par l'opération qui consiste à "renverser" les mots et "inverser" leurs lettres) et que, réciproquement, les suites satisfaisant cette condition sont des codages de rotations associés à un découpage du cercle unité en deux intervalles de longueur $1/2$.

La démonstration de ce résultat passe par l'étude des suites sturmiennes obtenues en soustrayant (modulo 2) deux à deux les termes consécutifs des codages de rotation. Dans [7], P. Hubert montre de façon tout à fait indépendante qu'un codage de rotation avec un découpage du cercle en deux intervalles est la différence terme à terme de deux suites sturmiennes.

Par une approche différente, nous explicitons dans ce travail les liens existant entre les suites sturmiennes et les codages de rotations sur un nombre quelconque d'intervalles pour en tirer une caractérisation combinatoire complète.

Dans la section 2 nous introduisons les premières notations et définitions ainsi que quelques résultats classiques sur les suites sturmiennes.

La section 3 débute avec une discussion sur le caractère récurrent des codages de rotations et se poursuit avec la présentation des outils utilisés dans nos démonstrations.

Nous distinguons dans la section 4 le cas où le cercle est découpé en deux intervalles. Nous montrons que, pour une large classe de suites codages de rotation incluant celles de complexité $2n$, cette caractérisation a une expression simple. Le cas général est ensuite traité en considérant les suites pouvant être extraites de codages de rotations. Nous terminons cette section en discutant l'effectivité de nos critères.

La section 5 est enfin consacrée à l'extension des résultats précédents au cas des partitions en un nombre fini d'intervalles.

2. Définitions et notations. Etant donné un ensemble \mathcal{E} , on note $\#\mathcal{E}$ le cardinal de \mathcal{E} . On appelle *alphabet* tout ensemble fini non-vide d'éléments, appelés lettres ou symboles, que nous noterons en gras pour éviter toute confusion.

Un *mot* sur un alphabet \mathcal{A} est une suite de lettres de \mathcal{A} indexée sur $\{0, 1, \dots, n\}$ pour un entier naturel n , si le mot est fini, et sur \mathbb{N} si le mot est infini. La *longueur* d'un mot fini w , notée $|w|$, est le nombre de lettres le composant. On désignera par :

- \mathcal{A}^n l'ensemble des mots de longueur n sur \mathcal{A} ,
- \mathcal{A}^* l'ensemble des mots finis sur \mathcal{A} ,
- $\mathcal{A}^{\mathbb{N}}$ l'ensemble des suites sur \mathcal{A} .

Le *concaténé* de deux mots $u = u_0 \dots u_{|u|-1}$ et $v = v_0 \dots v_{|v|-1}$, noté uv , est le mot $u_0 \dots u_{|u|-1}v_0 \dots v_{|v|-1}$.

Soit u un mot ou une suite sur un alphabet \mathcal{A} . Pour un entier naturel i , on dit qu'un mot w de longueur n *apparaît au rang* i de u si on a $u_i u_{i+1} \dots u_{i+n-1} = w_0 w_1 \dots w_{n-1}$. Un *facteur* de u est un mot apparaissant à un rang quelconque de u . On note alors $L_n(u)$ l'ensemble des facteurs de longueur n de u . Le *langage* de u , noté $L(u)$, est l'ensemble des facteurs finis de u (i.e. $L(u) = \bigcup_{n \in \mathbb{N}} L_n(u)$).

Plus généralement, un langage sur un alphabet \mathcal{A} est un sous-ensemble de \mathcal{A}^* . Il est dit *factoriel* si tout facteur d'un mot lui appartenant est également dans cet ensemble, et *prolongeable à droite* (resp. *prolongeable à gauche*) si chacun de ses éléments est préfixe (resp. suffixe) d'un mot lui appartenant. En particulier, le langage d'une suite est factoriel et prolongeable à droite.

Un langage sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ est dit *équilibré* si pour tout couple de mots de même longueur (w, w') lui appartenant, on a

$$||w|_{\mathbf{1}} - |w'|_{\mathbf{1}}| \leq 1,$$

où $|w|_{\mathbf{1}}$ désigne le nombre d'occurrences de la lettre $\mathbf{1}$ dans w . Une suite est dite équilibrée si son langage l'est.

Une suite u est dite *récurrente* (resp. *uniformément récurrente*) si tout facteur de u y apparaît une infinité de fois (resp. une infinité de fois à lacunes bornées).

Il est parfois possible de déterminer certaines caractéristiques d'une suite en étudiant le nombre de mots y apparaissant. On appelle *fonction de complexité* de u et on note $p(u, n)$ la fonction qui à tout entier naturel n non nul fait correspondre le nombre de facteurs de longueur n de u :

$$p(u, n) = \#L_n(u).$$

Comme son nom l'indique, la fonction de complexité d'une suite traduit assez bien l'idée que l'on se fait intuitivement de sa "complexité" (voir [3] pour un survol du sujet). En particulier, une suite est ultimement périodique (i.e.

périodique à partir d'un certain rang) — donc très “simple” — si et seulement si sa fonction de complexité est constante à partir d'un certain rang. Comme cette fonction est croissante, une suite non ultimement périodique est de complexité au moins $n + 1$. Cette borne est effectivement atteinte pour certaines suites que l'on appelle *suites sturmiennes*.

Nous rappelons deux résultats classiques sur les suites sturmiennes.

THÉORÈME 2.1 ([Coven & Hedlund [6]). *Une suite u sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ non ultimement périodique est équilibrée si et seulement si elle est sturmiennne.*

THÉORÈME 2.2 (Morse & Hedlund [9, 10]). *Une suite u sur $\{\mathbf{0}, \mathbf{1}\}$ est de complexité $p(u, n) = n + 1$ pour tout entier naturel non nul n si et seulement si il existe un nombre irrationnel $\alpha \in [0, 1[$ et un nombre réel $x \in [0, 1[$ tels que pour tout entier naturel i , on a : $u_i = \mathbf{1}$ si $x + i\alpha \bmod 1 \in I$ et $u_i = \mathbf{0}$ sinon, où I est l'intervalle $[0, \alpha[$ ou $]0, \alpha]$.*

Le langage d'une suite sturmiennne ne dépend que de l'angle de la rotation qui lui est associée (voir par exemple [4]). Deux suites sturmiennes ont le même langage si et seulement si elles codent une rotation de même angle.

Si l'on fixe le découpage, à une suite sturmiennne donnée correspond un unique point de départ.

Nous nous intéresserons aux suites produites par le même mécanisme mais en découpant tout d'abord le cercle en deux intervalles de façon quelconque.

DÉFINITION 2.1. Soient deux réels $\alpha \in [0, 1[$ et $\beta \in]0, 1[$. Une suite u à valeurs dans l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ code une rotation d'angle α avec un découpage $(\beta, 1 - \beta)$ s'il existe un réel positif x , strictement inférieur à 1, appelé *point de départ* de la rotation, tel que pour tout entier naturel k ,

$$u_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

où I est l'un des intervalles suivants : $[0, \beta]$, $[0, \beta[$, $]0, \beta]$ ou $]0, \beta[$.

La distinction des différents intervalles I possibles n'est utile que pour un ensemble dénombrable de points de départ x (de la forme $x = -n\alpha \bmod 1$ ou $x = \beta - n\alpha \bmod 1$, pour un entier naturel n).

La définition suivante est une généralisation de la précédente pour les suites sur des alphabets de cardinal supérieur à deux.

DÉFINITION 2.2. Une suite u sur un alphabet $\mathcal{A} = \{\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_g\}$ est une *codage de rotation* s'il existe deux nombres réels $\alpha \in [0, 1[$ et $x \in [0, 1[$ et une partition du cercle en $\#\mathcal{A}$ intervalles de longueurs strictement positives $\{I_0, \dots, I_g\}$, tels que, pour tout entier naturel k , on ait

$$u_k = \mathbf{a}_j \quad \text{si } x + k\alpha \bmod 1 \in I_j.$$

Notons que ces deux définitions ne concernent que les cas où à chaque lettre de la suite correspond un unique intervalle.

Nous ne nous préoccupons par la suite que du cas où l'angle α est irrationnel (si α est rationnel les suites associées sont périodiques et leur étude présente moins d'intérêt).

Sans perdre de généralité, nous supposons $\alpha < 1/2$. En effet, toute rotation d'angle α quelconque peut être considérée comme une rotation d'angle inférieur à $1/2$, quitte à remplacer α par $1 - \alpha$ et à changer le sens de rotation.

Par la suite, nous dirons parfois seulement que “ u est un codage de rotation” ou “ u code une rotation” sans en préciser les paramètres (angle, découpage et point de départ) lorsqu'il ne sera pas nécessaire de les expliciter.

3. Résultats préliminaires

3.1. Remarques sur le caractère (uniformément) récurrent des codages de rotation. Soit u une suite sur l'alphabet $\mathcal{A} = \{\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_g\}$ codant une rotation d'angle α avec la partition du cercle unité $\{I_{\mathbf{a}_0}, \dots, I_{\mathbf{a}_g}\}$. Notons \mathcal{R}_α la rotation d'angle α et, pour tout facteur $w = w_0 w_1 \dots w_{|w|-1}$ de u , K_w l'ensemble $\bigcap_{i=0}^{|w|-1} \mathcal{R}_\alpha^{-i}(I_{w_i})$. Cet ensemble est une intersection finie d'intervalles de $[0, 1[$, il est donc constitué d'une réunion finie d'intervalles.

Soit x le point de départ associé à la suite u . Un mot w apparaît au rang n de u si et seulement si $x + n\alpha \bmod 1 \in K_w$.

Si tout facteur w de u est tel que l'ensemble K_w soit de mesure non nulle, cet ensemble contient au moins un intervalle de $[0, 1[$ de longueur strictement positive et l'uniforme récurrence de u est une conséquence directe de l'équirépartition de l'ensemble $\{n\alpha \bmod 1 : n \in \mathbb{N}\}$.

On en déduit que si la suite u n'est pas uniformément récurrente alors au moins un de ses facteurs w est tel que l'ensemble K_w soit de mesure nulle. Cet ensemble ne pouvant être vide si w apparaît dans u , ceci ne peut arriver que s'il est réduit à un ensemble fini de points. Il existe alors deux entiers naturels i et j , $i < j$, strictement inférieurs à $|w|$ avec I_{w_i} et I_{w_j} tous deux fermés ou ouverts respectivement à gauche et à droite et tels que la différence entre la borne supérieure de I_{w_j} et la borne inférieure de I_{w_i} soit égale à $(j - i)\alpha \bmod 1$. De plus, si on note y la borne inférieure de I_{w_i} , comme le mot w apparaît dans u , il existe un entier naturel n tel que $x = y - n\alpha \bmod 1$.

Réciproquement, s'il existe un entier naturel n tel que le point de départ de u soit égal à $y - n\alpha \bmod 1$, la suite u n'est pas récurrente et par conséquent n'est pas uniformément récurrente.

En résumé, on obtient la proposition suivante qui reste vraie si l'on remplace “uniformément récurrent” par “récurrent” :

PROPOSITION 3.1. *Un codage de rotation n'est pas uniformément récurrent si et seulement si deux intervalles I et J de la partition, avec I fermé à gauche et J fermé à droite ou I ouvert à gauche et J ouvert à droite, sont tels que la borne inférieure de I et la borne supérieure de J appartiennent à l'orbite, sous l'action de la rotation, du point de départ associé au codage.*

On remarque que même dans ce cas, la suite u est uniformément récurrente à un préfixe de longueur finie près (i.e. il existe $k \in \mathbb{N}$ tel que $(u_{n+k})_{n \in \mathbb{N}}$ soit uniformément récurrente).

Par exemple, si u est définie pour tout entier naturel k par

$$u_k = \begin{cases} \mathbf{1} & \text{si } k\alpha \bmod 1 \in [0, \alpha], \\ \mathbf{0} & \text{sinon,} \end{cases}$$

le mot $\mathbf{11}$ n'admet qu'une seule occurrence et u n'est pas récurrente. Par contre, Tu est sturmiennne et uniformément récurrente.

3.2. Deux automates cellulaires. La caractérisation que nous proposons fait intervenir deux automates cellulaires particuliers que nous notons f_0 et f_1 , et qui, à tout mot u de $\{\mathbf{0}, \mathbf{1}\}^{\mathbb{N}}$, associent respectivement les mots $f_0(u)$ et $f_1(u)$ tels que pour tout entier naturel k , on ait

$$f_0(u)_k = \bar{f}_0(u_k u_{k+1}) \quad \text{et} \quad f_1(u)_k = \bar{f}_1(u_k u_{k+1}),$$

où \bar{f}_0 et \bar{f}_1 sont les fonctions de $\{\mathbf{0}, \mathbf{1}\}^2$ dans $\{\mathbf{0}, \mathbf{1}\}$ définies par

$$\begin{array}{cc} \bar{f}_0 & \bar{f}_1 \\ \mathbf{00} \rightarrow \mathbf{0} & \mathbf{00} \rightarrow \mathbf{0} \\ \mathbf{01} \rightarrow \mathbf{0} & \text{et } \mathbf{01} \rightarrow \mathbf{1} \\ \mathbf{10} \rightarrow \mathbf{1} & \mathbf{10} \rightarrow \mathbf{0} \\ \mathbf{11} \rightarrow \mathbf{0} & \mathbf{11} \rightarrow \mathbf{0} \end{array}$$

On remarque que si le mot $\mathbf{11}$ n'apparaît pas dans u , on a $f_0(u) = u$ et $f_1(u) = Tu$.

On peut définir des applications analogues sur l'ensemble des mots finis de longueur supérieure ou égale à 2 sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$. Par abus, nous les noterons également f_0 et f_1 .

Une propriété des automates cellulaires est que ceux-ci commutent avec le décalage.

LEMME 3.1. *Soit u une suite sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$. Si les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes alors on a*

$$L(f_0(u)) = L(f_1(u)).$$

Preuve. Les suites $f_0(u)$ et $f_1(u)$ étant sturmiennes, leurs langages sont équilibrés (théorème 2.1). Pour tout entier naturel non nul n , il existe

un unique entier $k_n^{(0)}$ (resp. $k_n^{(1)}$) tel que pour tout mot $w \in L_n(u)$ on ait $|f_0(w)|_{\mathbf{1}} \in \{k_n^{(0)}, k_n^{(0)} + 1\}$ (resp. $|f_1(w)|_{\mathbf{1}} \in \{k_n^{(1)}, k_n^{(1)} + 1\}$).

Nous allons montrer que le langage $L(f_0(u)) \cup L(f_1(u))$ est équilibré.

Pour tout mot w facteur de u , la quantité $|f_0(w)|_{\mathbf{1}} - |f_1(w)|_{\mathbf{1}}$ est, par définition, égale à la différence entre le nombre d'occurrences du facteur $\mathbf{10}$ dans w et le nombre d'occurrences du facteur $\mathbf{01}$ dans w . Cette quantité appartient à $\{-1, 0, 1\}$. En effet, entre deux occurrences successives de $\mathbf{01}$ il y a exactement une occurrence de $\mathbf{10}$ et inversement.

Plus précisément, on a

$$|f_0(w)|_{\mathbf{1}} - |f_1(w)|_{\mathbf{1}} = \begin{cases} -1 & \text{si } w_0 w_{|w|-1} = \mathbf{01}, \\ 0 & \text{si } w_0 = w_{|w|-1}, \\ 1 & \text{si } w_0 w_{|w|-1} = \mathbf{10}. \end{cases}$$

De plus, pour tout entier naturel n strictement supérieur à 1, il existe au moins deux facteurs de longueur n de u , w et w' , tels que l'on ait $w_0 w_{n-1} = \mathbf{01}$ et $w'_0 w'_{n-1} = \mathbf{10}$. Dans le cas contraire la suite u serait ultimement périodique — ce qui est incompatible avec le fait que $f_0(u)$ (ou $f_1(u)$) soit sturmienne.

Comme $w_0 w_{n-1} = \mathbf{01}$, on a $|f_0(w)|_{\mathbf{1}} = |f_1(w)|_{\mathbf{1}} - 1$, ce qui implique que $k_n^{(0)} \leq k_n^{(1)}$ (on a $k_n^{(0)} \leq |f_0(w)|_{\mathbf{1}} = |f_1(w)|_{\mathbf{1}} - 1 \leq k_n^{(1)}$). De même, comme $w'_0 w'_{n-1} = \mathbf{10}$, on a $|f_0(w')|_{\mathbf{1}} = |f_1(w')|_{\mathbf{1}} + 1$, ce qui implique que $k_n^{(0)} \geq k_n^{(1)}$ (on a $k_n^{(0)} + 1 \geq |f_0(w')|_{\mathbf{1}} = |f_1(w')|_{\mathbf{1}} + 1 \geq k_n^{(1)} + 1$). On en conclut que pour tout entier naturel non nul n , les entiers $k_n^{(0)}$ et $k_n^{(1)}$ sont égaux.

Quels que soient deux mots w et w' , facteurs de longueur n de u , on a

$$||f_0(w)|_{\mathbf{1}} - |f_1(w')|_{\mathbf{1}}| \leq 1.$$

Autrement dit, le langage $L(f_0(u)) \cup L(f_1(u))$ est équilibré.

On montre alors par récurrence que, pour tout entier naturel non nul n , $L_n(f_0(u)) = L_n(f_1(u))$:

Pour $n = 1$, on a bien $L_1(f_0(u)) = L_1(f_1(u)) = \{\mathbf{0}, \mathbf{1}\}$. Supposons qu'au rang n , on ait $L_n(f_0(u)) = L_n(f_1(u))$. Il existe donc un unique mot $w \in L_{n-1}(f_0(u)) \cup L_{n-1}(f_1(u))$ tel que $\{w\mathbf{0}, w\mathbf{1}\} \subset L_n(f_0(u)) \cup L_n(f_1(u))$.

Si l'on suppose $L_{n+1}(f_0(u)) \neq L_{n+1}(f_1(u))$ alors

$$\#(L_{n+1}(f_0(u)) \cup L_{n+1}(f_1(u))) > n + 2,$$

et il existe au moins deux mots w' et w'' appartenant à $L_n(f_0(u)) \cup L_n(f_1(u))$ avec $w' \neq w''$, tels que $\{w'\mathbf{0}, w'\mathbf{1}, w''\mathbf{0}, w''\mathbf{1}\} \subset L_{n+1}(f_0(u)) \cup L_{n+1}(f_1(u))$. Ces langages étant factoriels, le mot w est nécessairement suffixe de w' et w'' . Comme $w' \neq w''$, on a $w' = xw$ et $w'' = yw$ avec $x \neq y$ et on a donc $\{\mathbf{0}w\mathbf{0}, \mathbf{0}w\mathbf{1}, \mathbf{1}w\mathbf{0}, \mathbf{1}w\mathbf{1}\} \subset L_{n+1}(f_0(u)) \cup L_{n+1}(f_1(u))$. Ceci est en contradiction avec le fait que $L(f_0(u)) \cup L(f_1(u))$ est équilibré. On a donc nécessairement $L_{n+1}(f_0(u)) = L_{n+1}(f_1(u))$.

On peut finalement conclure que, pour tout entier naturel non nul n , on a $L_n(f_0(u)) = L_n(f_1(u))$ et, par conséquent, $L(f_0(u)) = L(f_1(u))$. ■

3.3. Suites extraites d'un codage de rotation. Les propriétés suivantes nous seront utiles dans les sections 4.2 et 5.

REMARQUE 3.1. Si u code l'orbite d'un point $x \in [0, 1[$ sous l'action d'une rotation d'angle α le cercle unité étant partitionné en un nombre fini d'intervalles, toute suite $(u_{pn+i})_{n \in \mathbb{N}}$ (où $p \in \mathbb{N}$ et $i \in \mathbb{N}$) code l'orbite du point $x + i\alpha \bmod 1$ sous l'action d'une rotation d'angle $p\alpha \bmod 1$ avec la même partition du cercle.

LEMME 3.2. *Soit un nombre irrationnel $\alpha \in [0, 1[$ et une partition du cercle unité en un nombre fini d'intervalles de longueurs strictement positives. Il existe deux entiers naturels p et q premiers entre eux tels que tout intervalle de la partition contienne un intervalle semi-ouvert de longueur $p\alpha \bmod 1$ et un intervalle semi-ouvert de longueur $q\alpha \bmod 1$.*

Preuve. Notons β la longueur du plus petit intervalle de la partition. Le nombre α étant irrationnel, la densité de $\{n\alpha \bmod 1 : n \in \mathbb{N}\}$ dans $[0, 1[$ nous assure de l'existence d'un entier p tel que $p\alpha \bmod 1 < \beta$. Ceci implique que tout intervalle de la partition contient un intervalle semi-ouvert de longueur $p\alpha \bmod 1$.

Le nombre $p\alpha \bmod 1$ étant également irrationnel, il existe un entier naturel r tel que $(rp+1)\alpha \bmod 1 < \beta$. De la même façon, en posant $q = rp+1$, tout intervalle de la partition contient également un intervalle semi-ouvert de longueur $q\alpha \bmod 1$.

La définition de q implique de plus que les entiers p et q sont premiers entre eux. ■

4. Suites sur deux lettres codant une rotation

4.1. Cas où les intervalles de la partition contiennent tous deux un intervalle semi-ouvert de longueur α

THÉORÈME 4.1. *Une suite u sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ est un codage de rotation d'angle α irrationnel associé à une partition en deux intervalles contenant tous deux un intervalle semi-ouvert de longueur α si et seulement si les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes.*

Preuve. Supposons tout d'abord que u soit un codage de rotation satisfaisant les hypothèses de la démonstration. Autrement dit, il existe un nombre irrationnel $\alpha \in [0, 1[$, un réel $x \in [0, 1[$ et un intervalle I dans $[0, 1[$, cet intervalle et son complémentaire contenant un intervalle semi-ouvert de longueur α , tels que, pour tout entier naturel k , on ait

$$u_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Par définition, pour tout entier naturel k , on a $u_k u_{k+1} = \mathbf{10}$ si et seulement si $x + k\alpha \bmod 1 \in I$ et $x + (k+1)\alpha \bmod 1 \notin I$.

Si $I = [0, \beta[$ ou $I =]0, \beta[$, ceci équivaut à $x + k\alpha \bmod 1 \in [\beta - \alpha, \beta[\cap I$. Comme I contient un intervalle semi-ouvert de longueur α , on a $[\beta - \alpha, \beta[\subset I$. On peut alors définir la suite $f_0(u)$ par

$$f_0(u)_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in [\beta - \alpha, \beta[, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

ou encore

$$f_0(u)_k = \begin{cases} \mathbf{1} & \text{si } x - (\beta - \alpha) + k\alpha \bmod 1 \in [0, \alpha[, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

De même, si $I = [0, \beta]$ ou $I =]0, \beta]$, on a $u_k u_{k+1} = \mathbf{10}$ si et seulement si $x + k\alpha \bmod 1 \in]\beta - \alpha, \beta] \cap I$. Comme I contient un intervalle semi-ouvert de longueur α , on a $] \beta - \alpha, \beta] \subset I$. La suite $f_0(u)$ est alors définie par

$$f_0(u)_k = \begin{cases} \mathbf{1} & \text{si } x - (\beta - \alpha) + k\alpha \bmod 1 \in]0, \alpha], \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Dans les deux cas, la suite $f_0(u)$ est un codage de rotation d'angle α associé à un découpage en deux intervalles semi-ouverts de longueurs respectives α et $1 - \alpha$. D'après le théorème 2.2, elle est sturmienne.

De même, pour tout entier naturel k , $u_k u_{k+1} = \mathbf{01}$ si et seulement si $x + k\alpha \bmod 1 \notin I$ et $x + (k+1)\alpha \bmod 1 \in I$.

Si $I = [0, \beta[$ ou $I = [0, \beta]$, ceci équivaut à $x + k\alpha \in [1 - \alpha, 1[\cap I^C$. Le complémentaire de I , noté I^C , contient également un intervalle semi-ouvert de longueur α et on a $[1 - \alpha, 1[\subset I^C$. On peut alors définir la suite $f_1(u)$ par

$$f_1(u)_k = \begin{cases} \mathbf{1} & \text{si } x - (1 - \alpha) + k\alpha \bmod 1 \in [0, \alpha[, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Enfin, si $I =]0, \beta]$ ou $I =]0, \beta[$, on peut de la même façon définir la suite $f_1(u)$ par

$$f_1(u)_k = \begin{cases} \mathbf{1} & \text{si } x - (1 - \alpha) + k\alpha \bmod 1 \in]0, \alpha], \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Comme la suite $f_0(u)$, $f_1(u)$ est un codage de rotation d'angle α associé à un découpage du cercle unité en deux intervalles semi-ouverts de longueurs respectives α et $1 - \alpha$, d'après le théorème 2.2, les suites $f_0(u)$ et $f_1(u)$ sont toutes deux sturmiennes.

Inversement, supposons que la suite u soit telle que les suites $f_0(u)$ et $f_1(u)$ soient sturmiennes. D'après le lemme 3.1, on a alors $L(f_0(u)) = L(f_1(u))$. D'après le théorème 2.2 et la remarque sur la correspondance entre le langage d'une suite sturmienne et l'angle irrationnel qui lui est associé, il existe un nombre irrationnel $\alpha \in [0, 1[$ et deux réels appartenant à $[0, 1[$

que l'on note x_0 et x_1 tels que, pour tout entier naturel k , on ait

$$f_0(u)_k = \begin{cases} \mathbf{1} & \text{si } x_0 + k\alpha \bmod 1 \in I_0, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

avec $I_0 = [0, \alpha[$ ou $I_0 =]0, \alpha]$, et

$$f_1(u)_k = \begin{cases} \mathbf{1} & \text{si } x_1 + k\alpha \bmod 1 \in I_1, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

avec $I_1 = [0, \alpha[$ ou $I_1 =]0, \alpha]$.

Posons $\beta = x_1 - x_0 \bmod 1$ et désignons par I'_0 l'intervalle $[\beta, \beta + \alpha \bmod 1[$ si $I_0 = [0, \alpha[$ et $] \beta, \beta + \alpha \bmod 1]$ si $I_0 =]0, \alpha]$. La suite $f_0(u)$ peut se définir par

$$f_0(u)_k = \begin{cases} \mathbf{1} & \text{si } x_1 + k\alpha \bmod 1 \in I'_0, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

On a en quelque sorte translaté le découpage associé à la suite $f_0(u)$ pour faire coïncider les points de départs.

Comme l'on ne peut avoir pour un entier naturel k simultanément $f_0(u)_k = \mathbf{1}$ et $f_1(u)_k = \mathbf{1}$, on peut supposer que les intervalles I'_0 et I_1 sont d'intersection vide. On note alors K l'intervalle situé entre I_1 et I'_0 lorsqu'on suit le sens de la rotation. On a, selon I_1 et I'_0 , $K = [\alpha, \beta]$, $[\alpha, \beta[$, $] \alpha, \beta]$ ou $] \alpha, \beta[$.

Remarquons ensuite que la donnée des suites $f_0(u)$ et $f_1(u)$ suffit pour déterminer u .

Pour cela, on considère les entiers naturels $n_0 = \min\{n \in \mathbb{N} \mid f_0(u)_n = \mathbf{1}\}$ et $n_1 = \min\{n \in \mathbb{N} \mid f_1(u)_n = \mathbf{1}\}$. La première lettre de u , u_0 est alors $\mathbf{1}$ si $n_0 < n_1$ et $\mathbf{0}$ si $n_0 > n_1$.

On peut interpréter géométriquement ce fait. On a $n_0 < n_1$ si et seulement si la trajectoire du point x_1 sous l'action de la rotation d'angle α rencontre l'intervalle I'_0 avant l'intervalle I_1 . Ceci n'est possible que si le point x_1 appartient à I'_0 ou si il appartient à K .

De plus, par définition, $I' = K \cup I'_0$ est l'un des intervalles suivants : $[\alpha, \beta + \alpha \bmod 1]$, $[\alpha, \beta + \alpha \bmod 1[$, $] \alpha, \beta + \alpha \bmod 1]$ ou $] \alpha, \beta + \alpha \bmod 1[$. On obtient finalement

$$u_0 = \begin{cases} \mathbf{1} & \text{si } x_1 \bmod 1 \in I', \\ \mathbf{0} & \text{sinon,} \end{cases}$$

ou encore, si l'on note I l'intervalle translaté de $-\alpha$ de I' (i.e. $I = [0, \beta]$, $[0, \beta[$, $]0, \beta]$ ou $]0, \beta[$ selon I') et x le réel $x_1 - \alpha \bmod 1$,

$$u_0 = \begin{cases} \mathbf{1} & \text{si } x \bmod 1 \in I, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Comme l'on peut faire le même raisonnement pour tous les décalages de u (i.e. toutes les suites de la forme $T^k u$ avec k entier naturel) et que les automates cellulaires commutent avec cette opération, pour tout entier

naturel k on a

$$f_0(T^k u) = T^k f_0(u) \quad \text{et} \quad f_1(T^k u) = T^k f_1(u),$$

et donc

$$u_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

De plus, d'après sa définition, l'intervalle I contient bien un intervalle semi-ouvert de longueur α . Comme on peut également définir son complémentaire comme union d'un intervalle semi-ouvert de longueur α et d'un autre intervalle, la partition du cercle $\{I, I^C\}$ satisfait bien les hypothèses de la proposition. ■

REMARQUES. En particulier, les suites sturmiennes satisfont les hypothèses du théorème 4.1 et, dans ce cas, les deux automates cellulaires sont respectivement l'identité et le décalage à droite qui préservent le caractère "sturmien" d'une suite.

L'hypothèse " I et son complémentaire contiennent des semi-ouverts de longueur α " ne peut pas être remplacé par " $\min\{\beta, 1 - \beta\} \geq \alpha$ ", où β est la longueur de I . En effet, pour $I =]0, \alpha[$, si le point de départ de u est de la forme $x = -n\alpha \bmod 1$, avec n entier naturel, les suites $f_0(u)$ et $f_1(u)$ ne sont pas sturmiennes.

Par contre, supposer " $\min\{\beta, 1 - \beta\} > \alpha$ " est une condition suffisante pour que l'hypothèse du théorème soit satisfaite.

Réciproquement, si u est définie pour tout entier k par

$$u_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

et si les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes, alors, en notant β la longueur de l'intervalle I , on a nécessairement $\min\{\beta, 1 - \beta\} \geq \alpha$. L'intervalle I peut éventuellement être $]0, \alpha[$ si il n'y a pas d'entier naturel n tel que $x = -n\alpha \bmod 1$.

COROLLAIRE 4.1. *Soit u une suite sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ telle que $p(u, 2) = 4$. La suite u est un codage de rotation si et seulement si les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes.*

Preuve. Si les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes, il suffit d'utiliser le théorème 4.1 pour conclure que u est un codage de rotation.

Réciproquement, si u est telle que $p(u, 2) = 4$, les mots $\mathbf{00}$ et $\mathbf{11}$ sont facteurs de cette suite. Si de plus la suite u code une rotation d'angle α , ceci implique que I et son complémentaire contiennent des intervalles fermés de longueur α (et donc des intervalles semi-ouverts de longueur α). On peut alors conclure en utilisant le théorème 4.1. ■

En particulier, ce corollaire caractérise les suites de complexité $2n$ qui sont des codages de rotations.

4.2. Cas général

LEMME 4.1. *Si u est une suite sur l'alphabet $\{0, 1\}$ définie simultanément, pour tout entier naturel k , par*

$$u_k = \begin{cases} \mathbf{1} & \text{si } x + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon} \end{cases}$$

et

$$u_k = \begin{cases} \mathbf{1} & \text{si } x' + k\alpha' \bmod 1 \in I', \\ \mathbf{0} & \text{sinon,} \end{cases}$$

où x et x' sont deux nombres réels dans $[0, 1[$, α et α' sont deux nombres irrationnels dans $[0, 1/2]$ et I et I' deux sous-intervalles de $[0, 1[$ de bornes inférieures nulles, alors on a $x = x'$, $\alpha = \alpha'$ et les intervalles I et I' sont de même longueur.

Preuve. Ce lemme est montré dans le cas où l'intervalle I (resp. I') est semi-ouvert de longueur α (resp. α') (voir les remarques suivant le théorème 2.2). Dans le cas où l'intervalle I et son complémentaire (resp. I' et son complémentaire) contiennent un intervalle semi-ouvert de longueur α (resp. α'), le théorème 4.1 implique que les suites $f_0(u)$ et $f_1(u)$ sont sturmiennes. Dans la seconde partie de la démonstration de ce théorème, nous avons exprimé l'angle, le point de départ et la longueur du découpage de la rotation associée à u en fonction des points départs et de l'angle des rotations associées aux suites $f_0(u)$ et $f_1(u)$. Le théorème 2.2 nous donne donc également l'unicité de ces paramètres dans ce cas.

Reste le cas où I ou son complémentaire ne contient pas d'intervalle semi-ouvert de longueur α et I' ou son complémentaire ne contient pas d'intervalle semi-ouvert de longueur α' . D'après le lemme 3.2, il existe deux entiers naturels p et q premiers entre eux tels que I et son complémentaire contiennent des intervalles semi-ouverts de longueur $p\alpha \bmod 1$ et $q\alpha \bmod 1$. De plus les suites $(u_{pn})_{n \in \mathbb{N}}$ et $(u_{qn})_{n \in \mathbb{N}}$ peuvent alors être définies par : pour tout entier naturel k ,

$$u_{pk} = \begin{cases} \mathbf{1} & \text{si } x + k(p\alpha \bmod 1) \bmod 1 \in I, \\ \mathbf{0} & \text{sinon} \end{cases}$$

et

$$u_{pk} = \begin{cases} \mathbf{1} & \text{si } x' + k(p\alpha' \bmod 1) \bmod 1 \in I', \\ \mathbf{0} & \text{sinon;} \end{cases}$$

$$u_{qk} = \begin{cases} \mathbf{1} & \text{si } x + k(q\alpha \bmod 1) \bmod 1 \in I, \\ \mathbf{0} & \text{sinon} \end{cases}$$

et

$$u_{qk} = \begin{cases} \mathbf{1} & \text{si } x' + k(q\alpha' \bmod 1) \bmod 1 \in I', \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Nous sommes alors ramenés au cas précédent et l'on en déduit que $p\alpha \bmod 1 = p\alpha' \bmod 1$ et $q\alpha \bmod 1 = q\alpha' \bmod 1$, avec p et q premiers entre eux.

Comme on a $\alpha \in [0, 1/2]$ et $\alpha' \in [0, 1/2]$, on en déduit que $\alpha = \alpha'$. Pour les mêmes raisons que précédemment, on a également $x = x'$ et les intervalles I et I' sont de même longueur. ■

THÉORÈME 4.2. *Soit u une suite sur l'alphabet $\{0, 1\}$. La suite u est un codage de rotation si et seulement si il existe deux entiers naturels non nuls p et q premiers entre eux tels que, pour tout entier naturel $i < p$ et tout entier naturel $j < q$, les suites $f_0((u_{pn+i})_{n \in \mathbb{N}})$, $f_1((u_{pn+i})_{n \in \mathbb{N}})$, $f_0((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1((u_{qn+j})_{n \in \mathbb{N}})$ soient sturmiennes.*

Preuve. Supposons que la suite u code une rotation d'angle α irrationnel, le cercle unité étant partitionné en deux intervalles de longueurs strictement positives.

Le lemme 3.2 nous assure de l'existence de deux entiers naturels p et q premiers entre eux tels que chaque intervalle de la partition contienne des intervalles semi-ouverts de longueur $p\alpha \bmod 1$ et de longueur $q\alpha \bmod 1$.

D'après la remarque 3.1, pour tout entier naturel i strictement inférieur à p et tout entier naturel j strictement inférieur à q , les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ sont des codages de rotations d'angles respectifs $p\alpha \bmod 1$ et $q\alpha \bmod 1$, le cercle unité tous deux associés à la partition initiale du cercle unité.

Pour tout entier naturel $i < p$ et tout entier naturel $j < q$, les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ sont des codages satisfaisant bien les hypothèses du théorème 4.1. Ceci nous permet de conclure que les suites $f_0((u_{pn+i})_{n \in \mathbb{N}})$, $f_1((u_{pn+i})_{n \in \mathbb{N}})$, $f_0((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes.

Inversement, supposons qu'il existe deux entiers naturels p et q premiers entre eux tels que, pour tout entier naturel $i < p$ et pour tout entier naturel $j < q$, les suites $f_0((u_{pn+i})_{n \in \mathbb{N}})$, $f_1((u_{pn+i})_{n \in \mathbb{N}})$, $f_0((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1((u_{qn+j})_{n \in \mathbb{N}})$ soient sturmiennes.

D'après le théorème 4.1, les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ sont de codages de rotations. De plus, comme p et q sont premiers entre eux, quitte à intervertir ces deux nombres, il existe deux entiers naturels a et b tels que $ap - bq = 1$. Notons $r = ap$ et $s = bq$. Pour tout entier naturel i et tout entier naturel j , les suites $(u_{rn+i})_{n \in \mathbb{N}}$ et $(u_{sn+j})_{n \in \mathbb{N}}$, respectivement extraites en progressions arithmétiques de $(u_{pn+(i \bmod p)})_{n \in \mathbb{N}}$ et de $(u_{qn+(j \bmod q)})_{n \in \mathbb{N}}$, sont également des codages de rotations (remarque 3.1). Autrement dit, pour tout entier naturel i , il existe un réel $x_{(r,i)} \in [0, 1[$, un nombre irrationnel $\alpha_{(r,i)} \in [0, 1/2]$ et un intervalle $I_{(r,i)}$ du cercle unité de borne inférieure nulle tels que

$$u_{rk+i} = \begin{cases} \mathbf{1} & \text{si } x_{(r,i)} + k\alpha_{(r,i)} \bmod 1 \in I_{(r,i)}, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

De même, pour tout entier naturel j , il existe un nombre réel $x_{(s,j)} \in [0, 1[$, un nombre irrationnel $\alpha_{(s,j)} \in [0, 1/2]$ et un intervalle $I_{(s,j)}$ du cercle unité

de borne inférieure nulle tels que

$$u_{sk+j} = \begin{cases} \mathbf{1} & \text{si } x_{(s,j)} + k\alpha_{(s,j)} \bmod 1 \in I_{(s,j)}, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Considérons, pour tout entier naturel i , les suites $(u_{rsn+i})_{n \in \mathbb{N}}$ qui peuvent être extraites à la fois de $(u_{rn+i})_{n \in \mathbb{N}}$ et de $(u_{sn+i})_{n \in \mathbb{N}}$ et que l'on peut donc définir de deux façons; pour tout entier naturel k , on a

$$u_{rsk+i} = \begin{cases} \mathbf{1} & \text{si } x_{(r,i)} + ks\alpha_{(r,i)} \bmod 1 \in I_{(r,i)}, \\ \mathbf{0} & \text{sinon} \end{cases}$$

et

$$u_{rsk+i} = \begin{cases} \mathbf{1} & \text{si } x_{(s,i)} + kr\alpha_{(s,i)} \bmod 1 \in I_{(s,i)}, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

On en déduit (lemme 4.1) que, pour tout entier naturel i , on a

$$x_{(r,i)} = x_{(s,i)} \quad \text{et} \quad \alpha_{(r,i)} = \frac{r}{s}\alpha_{(s,i)}$$

et on peut supposer

$$I_{(r,i)} = I_{(s,i)}.$$

Comme le décalage d'un codage rotation est également un codage de rotation dont tous les paramètres, sauf le point de départ, sont les mêmes, on a, pour tout entier naturel i ,

$$\alpha_{(r,i+r)} = \alpha_{(r,i)} \quad \text{et} \quad \alpha_{(s,i+s)} = \alpha_{(s,i)}.$$

En particulier, comme $r = s + 1$, on a pour tout entier naturel i

$$\alpha_{(r,i+r)} = \frac{r}{s}\alpha_{(s,i+1+s)}.$$

Avec la remarque précédente, on obtient

$$\alpha_{(r,i)} = \frac{r}{s}\alpha_{(s,i+1)} = \alpha_{(r,i+1)}.$$

D'où l'on déduit que pour tout couple d'entiers naturels i et j ,

$$\alpha_{(r,i)} = \alpha_{(r,j)} = \alpha_{(r)} \quad \text{et} \quad \alpha_{(s,i)} = \alpha_{(s,j)} = \alpha_{(s)} \quad \text{avec} \quad \alpha_{(r)} = \frac{r}{s}\alpha_{(s)}.$$

On procède de la même façon pour montrer que pour tout couple d'entiers naturels i et j , on a

$$I_{(r,i)} = I_{(r,j)} = I_{(s,i)} = I_{(s,j)} = I.$$

Nous avons donc démontré que, pour tout entier naturel i , les suites $(u_{rn+i})_{n \in \mathbb{N}}$ sont toutes de codage de rotations de même angle associés à un même découpage du cercle.

Il nous reste à déterminer les points de départ de ces rotations. Pour tout entier naturel i , nous avons établi

$$x_{(r,i)} = x_{(s,i)}.$$

En particulier, pour tout entier naturel i , on a

$$x_{(r,i+r)} = x_{(s,i+1+s)}.$$

Comme les suites $(u_{rn+i+r})_{n \in \mathbb{N}}$ et $(u_{sn+i+1+s})_{n \in \mathbb{N}}$ sont respectivement des décalages des suites $(u_{rn+i})_{n \in \mathbb{N}}$ et $(u_{sn+i+1})_{n \in \mathbb{N}}$, on en déduit

$$x_{(r,i)} + \alpha_{(r)} \bmod 1 = x_{(s,i+1)} + \alpha_{(s)} \bmod 1,$$

et enfin

$$x_{(r,i)} + \alpha_{(r)} \bmod 1 = x_{(r,i+1)} + \frac{s}{r} \alpha_{(r)} \bmod 1.$$

Finalement, pour tout entier naturel i , nous avons établi que

$$x_{(r,i+1)} - x_{(r,i)} \bmod 1 = \frac{1}{r} \alpha_{(r)}.$$

Ceci nous permet de conclure que la différence entre deux points de départ “successifs”, $x_{(r,i+1)} - x_{(r,i)}$, est constante et égale à $\frac{1}{r} \alpha_{(r)}$. Autrement dit, si l’on pose $\alpha = \frac{1}{r} \alpha_{(r)}$, on a, pour tout entier naturel i ,

$$x_{(r,i)} = x_{(r,0)} + i\alpha \bmod 1.$$

D’après la définition des réels $x_{(r,k)}$, on a

$$u_k = \begin{cases} \mathbf{1} & \text{si } x_{(r,k)} \in I, \\ \mathbf{0} & \text{sinon} \end{cases}$$

et l’on peut donc définir la suite u par

$$u_k = \begin{cases} \mathbf{1} & \text{si } x_{(r,0)} + k\alpha \bmod 1 \in I, \\ \mathbf{0} & \text{sinon,} \end{cases}$$

pour tout entier naturel k . Ceci termine la démonstration. ■

Remarques relatives au choix des entiers p et q . On peut, en fonction d’une suite u , donner des valeurs p et q effectives pour le critère du théorème 4.2, autrement dit, déterminer des valeurs p et q telles que, si u est un codage de rotation, les suites $f_0((u_{pn+i})_{n \in \mathbb{N}})$, $f_1((u_{pn+i})_{n \in \mathbb{N}})$, $f_0((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1((u_{qn+j})_{n \in \mathbb{N}})$ soient sturmiennes, pour tout entier naturel $i < p$ et pour tout entier naturel $j < q$.

LEMME 4.2. *Soit u un codage de rotation sur l’alphabet $\{\mathbf{0}, \mathbf{1}\}$. Si v et w sont deux suites extraites de u en progressions arithmétiques, on ne peut avoir simultanément $L_2(v) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$ et $L_2(w) = \{\mathbf{01}, \mathbf{10}, \mathbf{11}\}$.*

Preuve. Soit u codant l’orbite d’un point x sous l’action d’une rotation d’angle α (qu’il n’est pas, ici, nécessaire de supposer irrationnel), le cercle unité étant partitionné en deux sous-intervalles de $[0, 1[$, I et I^c associés respectivement à la lettre $\mathbf{1}$ et à la lettre $\mathbf{0}$.

Supposons que l’on ait v et w , deux suites extraites de u en progressions arithmétiques, telles que $L_2(v) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$ et $L_2(w) = \{\mathbf{01}, \mathbf{10}, \mathbf{11}\}$. Notons alors α_v et α_w les angles associés à ces suites. On déduit de $L_2(v) =$

$\{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$ que $\mathcal{R}_{\alpha_v}(I) \subset I^C$ et de $L_2(w) = \{\mathbf{01}, \mathbf{10}, \mathbf{11}\}$ que $\mathcal{R}_{\alpha_w}(I^C) \subset I$. Ces deux inclusions ne sont simultanément possibles que si les intervalles I et I^C sont semi-ouverts de longueur $1/2$ et les angles α_v et α_w tous deux égaux à $1/2$. On obtiendrait alors $L_2(v) = L_2(w) = \{\mathbf{01}, \mathbf{10}\}$, ce qui est en contradiction avec l'hypothèse faite sur les facteurs de longueur 2 de v et w . ■

Soit u une suite sur $\{\mathbf{0}, \mathbf{1}\}$ codant une rotation d'angle α irrationnel. Nous nous restreindrons dans cette discussion au cas où $p(u, 2) = 3$ (si $p(u, 2) = 4$, on peut utiliser le corollaire 4.1). Supposons donc, quitte à intervertir les lettres $\mathbf{0}$ et $\mathbf{1}$, que $L_2(u) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$.

On déduit du lemme précédent que, comme u est un codage de rotation d'angle irrationnel avec $L_2(u) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$, on a pour toute suite v extraite en progression arithmétique de u : soit $L_2(v) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}\}$, soit $L_2(v) = \{\mathbf{00}, \mathbf{01}, \mathbf{10}, \mathbf{11}\}$ (l'hypothèse d'irrationalité de l'angle exclut le cas où $p(v, 2) = 2$, qui implique la périodicité de v).

Soit alors p , un entier naturel non nul, tel que l'on ait, pour un entier naturel l , $u_l u_{l+p} = \mathbf{11}$. La lettre $\mathbf{1}$ revenant infiniment dans u , l'existence de cet entier est assurée.

On a, avec la remarque précédente, nécessairement $p((u_{pn+i})_{n \in \mathbb{N}}, 2) = 4$. Ceci implique que I et I^C contiennent des intervalles semi-ouverts de longueur $p\alpha \bmod 1$ (on utilise le même argument que pour le corollaire 4.1). On peut donc appliquer, pour tout entier naturel i , le théorème 4.1 aux suites $(u_{pn+i})_{n \in \mathbb{N}}$. Et, en particulier pour tout entier naturel $i < p$, les suites $f_0((u_{pn+i})_{n \in \mathbb{N}})$ et $f_1((u_{pn+i})_{n \in \mathbb{N}})$ sont sturmiennes.

Considérons alors les suites $(u_{pn})_{n \in \mathbb{N}}$ et $(u_{pn+1})_{n \in \mathbb{N}}$. Soit k le plus petit entier naturel tel que $u_{pk} = \mathbf{1}$ et r le plus petit entier naturel tel que $u_{p(k+r)+1} = \mathbf{1}$. Ici encore, la lettre $\mathbf{1}$ revenant une infinité de fois dans les suites $(u_{pn})_{n \in \mathbb{N}}$ et $(u_{pn+1})_{n \in \mathbb{N}}$, l'existence de ces deux entiers est assurée.

Si l'on pose $q = rp + 1$, on a $p((u_{qn+pk})_{n \in \mathbb{N}}, 2) = 4$ et on conclut de la même façon que, pour tout entier naturel $j < q$, les suites $f_0((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes.

5. Codages de rotations associés à une partition finie. Nous allons considérer le cas où le cercle est partitionné en un nombre quelconque d'intervalles de longueurs strictement positives. On code alors l'orbite d'un point de départ en associant une lettre différente à chaque intervalle (voir la définition 2.2). Les résultats obtenus dans la section précédente se généralisent pour caractériser les suites produites de cette façon.

Nous allons appliquer les résultats obtenus dans le cas des partitions en deux intervalles. Ceci nécessite l'introduction des fonctions indicatrices de chaque lettre de l'alphabet.

Pour tout entier naturel i strictement inférieur à $\#\mathcal{A}$, nous noterons ϕ_i l'application qui à toute suite u sur l'alphabet \mathcal{A} associe la suite $\phi_i(u)$ sur l'alphabet $\{\mathbf{0}, \mathbf{1}\}$ définie, pour tout entier naturel k , par

$$\phi_i(u)_k = \begin{cases} \mathbf{1} & \text{si } u_k = \mathbf{a}_i, \\ \mathbf{0} & \text{sinon.} \end{cases}$$

Nous verrons plus loin qu'une certaine classe de codages de rotations possède la propriété suivante.

DÉFINITION 5.1. Nous dirons qu'une suite u sur un alphabet $\mathcal{A} = \{\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_g\}$ possède la *propriété* (\dagger) si, quitte à renuméroter les lettres de \mathcal{A} , l'inclusion suivante est vérifiée :

$$L_2(u) \subset (\{\mathbf{a}_g \mathbf{a}_0, \mathbf{a}_0 \mathbf{a}_1, \dots, \mathbf{a}_{g-1} \mathbf{a}_g\} \cup \{\mathbf{a}_0 \mathbf{a}_0, \dots, \mathbf{a}_g \mathbf{a}_g\}).$$

REMARQUE 5.1. Soit u une suite sur un alphabet \mathcal{A} . Si u possède la propriété (\dagger) alors on a, pour tout entier naturel i strictement inférieur à $\#\mathcal{A}$,

$$f_0 \circ \phi_i(u) = f_1 \circ \phi_{(i+1) \bmod \#\mathcal{A}}(u).$$

Nous obtenons alors un résultat similaire à celui du lemme 3.1.

LEMME 5.1. *Soit u une suite sur un alphabet \mathcal{A} possédant la propriété (\dagger). Si, pour tout entier naturel $i \leq g$ avec $g = \#\mathcal{A} - 1$, les suites $f_0 \circ \phi_i(u)$ sont sturmiennes alors on a*

$$L(f_0 \circ \phi_0(u)) = L(f_0 \circ \phi_1(u)) = \dots = L(f_0 \circ \phi_g(u)).$$

Preuve. Il suffit d'appliquer, pour tout entier naturel i strictement inférieur à $\#\mathcal{A}$, le lemme 3.1 à chaque suite $\phi_i(u)$. La remarque précédente nous donne alors

$$L(f_0 \circ \phi_i(u)) = L(f_0 \circ \phi_{(i+1) \bmod \#\mathcal{A}}(u)). \quad \blacksquare$$

LEMME 5.2. *Une suite u sur un alphabet \mathcal{A} est un codage de rotation d'angle α irrationnel si et seulement si, pour tout entier naturel i strictement inférieur à $\#\mathcal{A}$, les suites $\phi_i(u)$ sont des codages de rotations d'angle α .*

Preuve. Si l'on suppose que la suite u code l'orbite d'un point x sous l'action d'une rotation d'angle α le cercle unité étant partitionné en $\#\mathcal{A}$ intervalles I_0, \dots, I_g , la suite $\phi_i(u)$ code l'orbite du même point sous l'action de la même transformation, le cercle unité étant alors partitionné en deux intervalles.

Réciproquement, si l'on suppose que, pour tout entier naturel $i < \#\mathcal{A}$, les suites $\phi_i(u)$ codent des rotations d'angle α , il existe $\#\mathcal{A}$ réels x_0, x_1, \dots, x_g dans $[0, 1[$ et $\#\mathcal{A}$ intervalles I'_0, I'_1, \dots, I'_g inclus dans $[0, 1[$ tels que l'on ait,

pour tout entier naturel k ,

$$u_k = \begin{cases} \mathbf{a}_0 & \text{si } x_0 + k\alpha \bmod 1 \in I'_0, \\ \mathbf{a}_1 & \text{si } x_1 + k\alpha \bmod 1 \in I'_1, \\ \vdots & \\ \mathbf{a}_g & \text{si } x_g + k\alpha \bmod 1 \in I'_g. \end{cases}$$

Ou encore en posant $x = x_0$ et pour tout entier naturel $i < \#\mathcal{A}$, $I_i = \mathcal{R}_{(x-x_i)}(I'_i)$,

$$u_k = \begin{cases} \mathbf{a}_0 & \text{si } x + k\alpha \bmod 1 \in I_0, \\ \mathbf{a}_1 & \text{si } x + k\alpha \bmod 1 \in I_1, \\ \vdots & \\ \mathbf{a}_g & \text{si } x + k\alpha \bmod 1 \in I_g. \end{cases}$$

L'ensemble $\{x + n\alpha : n \in \mathbb{N}\}$ étant dense dans $[0, 1[$, les intervalles I_0, I_1, \dots, I_g forment une partition de $[0, 1[$. ■

On remarque que si l'on n'impose pas que les angles associés aux suites $\phi_i(u)$ soient égaux, la suite u peut ne pas être un codage de rotation au sens de la définition 2.2. Les suites équilibrées sur plus de deux lettres ne satisfont pas cette définition (voir [8]). Pourtant toutes leurs images par les fonctions caractéristiques sont sturmiennes, donc des codages de rotations.

PROPOSITION 5.1. *Soit u une suite sur l'alphabet \mathcal{A} . La suite u est un codage de rotation d'angle α irrationnel associé à une partition du cercle unité en $\#\mathcal{A}$ intervalles contenant tous un intervalle semi-ouvert de longueur α si et seulement si la suite u possède la propriété (†) et les suites $f_0 \circ \phi_0(u), \dots, f_0 \circ \phi_g(u)$ sont sturmiennes.*

Preuve. Supposons tout d'abord que u soit un codage de rotation d'angle α , chaque intervalle de la partition contenant un intervalle semi-ouvert de longueur α .

Alors la suite u possède la propriété (†). En effet, si x est un réel appartenant à l'intervalle I_i , le fait que l'intervalle $I_{(i+1) \bmod \#\mathcal{A}}$ contienne un intervalle semi-ouvert de longueur α impose que le réel $x + \alpha$ appartienne soit à l'intervalle I_i , soit à l'intervalle $I_{(i+1) \bmod \#\mathcal{A}}$. De plus, cette hypothèse implique que, pour tout entier naturel $i < \#\mathcal{A}$, la suite $\phi_i(u)$ est un codage d'une rotation satisfaisant les conditions du théorème 4.1. D'où l'on déduit que les suites $f_0 \circ \phi_0(u), \dots, f_0 \circ \phi_g(u)$ sont sturmiennes.

Réciproquement, supposons que la suite u possède la propriété (†) et que les suites $f_0 \circ \phi_0(u), \dots, f_0 \circ \phi_g(u)$ soient sturmiennes. D'après le lemme 5.1, les langages de ces suites sturmiennes sont tous égaux entre eux. Elles sont donc associées à un même angle α . On a également (remarque 5.1) pour tout entier naturel $i < \#\mathcal{A}$,

$$f_0 \circ \phi_i(u) = f_1 \circ \phi_{(i+1) \bmod \#\mathcal{A}}(u).$$

D'où l'on déduit que, pour tout entier naturel $i < \#\mathcal{A}$, les suites $f_0 \circ \phi_i(u)$ et $f_1 \circ \phi_i(u)$ sont sturmiennes. Et donc, d'après le théorème 4.1, chaque suite $\phi_i(u)$ est un codage de rotation d'angle α , le cercle étant partitionné en deux intervalles contenant chacun un intervalle semi-ouvert de longueur α . Le lemme 5.2 nous permet alors de conclure que la suite u est un codage de rotation satisfaisant l'hypothèse de la proposition. ■

THÉORÈME 5.1. *Soit u une suite sur l'alphabet \mathcal{A} . La suite u est un codage de rotation d'angle α irrationnel, le cercle unité étant partitionné en $\#\mathcal{A}$ intervalles, si et seulement si il existe deux entiers naturels strictement positifs p et q premiers entre eux tels que, pour tout entier naturel $i < p$ et tout entier naturel $j < q$, les assertions suivantes soient vérifiées :*

- les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ possèdent la propriété (†),
- les suites $f_0 \circ \phi_0((u_{pn+i})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{pn+i})_{n \in \mathbb{N}})$ et $f_0 \circ \phi_0((u_{qn+j})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes.

Preuve. Si la suite u code une rotation d'angle α irrationnel, le lemme 3.1 nous assure de l'existence de deux entiers naturels p et q premiers entre eux tels que chaque intervalle de la partition associée à u contienne des intervalles semi-ouverts de longueurs $p\alpha \bmod 1$ et $q\alpha \bmod 1$. Pour tout entier naturel $i < p$ et tout entier naturel $j < q$, les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ sont donc des codages de rotations d'angles respectifs $p\alpha \bmod 1$ et $q\alpha \bmod 1$ associés la même partition que précédemment.

Ainsi on peut utiliser la proposition 5.1 pour conclure que toutes ces suites possèdent la propriété (†) et que les suites $f_0 \circ \phi_0((u_{pn+i})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{pn+i})_{n \in \mathbb{N}})$ et $f_0 \circ \phi_0((u_{qn+j})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes.

Inversement, supposons que pour tout entier naturel $i < p$ et tout entier naturel $j < q$, les suites $(u_{pn+i})_{n \in \mathbb{N}}$ et $(u_{qn+j})_{n \in \mathbb{N}}$ possèdent la propriété (†), et que les suites $f_0 \circ \phi_0((u_{pn+i})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{pn+i})_{n \in \mathbb{N}})$ et $f_0 \circ \phi_0((u_{qn+j})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{qn+j})_{n \in \mathbb{N}})$ soient sturmiennes. Comme pour la proposition 5.1, on en déduit que pour tout entier naturel $i < p$, les langages des suites $f_0 \circ \phi_0((u_{pn+i})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{pn+i})_{n \in \mathbb{N}})$ sont égaux. Ces suites sturmiennes sont associées à un même angle $\alpha_{(p,i)}$.

De même, les suites $f_0 \circ \phi_0((u_{qn+j})_{n \in \mathbb{N}}), \dots, f_0 \circ \phi_g((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes associées à un même angle $\alpha_{(q,j)}$, pour tout entier naturel $j < q$. De la remarque 5.1 appliquée aux hypothèses, on déduit que pour tout entier $l < \#\mathcal{A}$, les suites $f_0 \circ \phi_l((u_{pn+i})_{n \in \mathbb{N}}), f_1 \circ \phi_l((u_{pn+i})_{n \in \mathbb{N}}), f_0 \circ \phi_l((u_{qn+j})_{n \in \mathbb{N}})$ et $f_1 \circ \phi_l((u_{qn+j})_{n \in \mathbb{N}})$ sont sturmiennes. D'après le théorème 4.2, les suites $\phi_l(u)$ sont des codages de rotations.

On remarque ensuite que pour chaque entier $l < \#\mathcal{A}$, l'angle associé à la suite $\phi_l(u)$ est nécessairement égal à $\alpha_{(p,0)}/p$ (lemme 4.1). Pour tout entier

$l < \#\mathcal{A}$, les suites $\phi_l(u)$ codent des rotations de même angle et le lemme 5.2 termine la démonstration. ■

On peut comme pour le théorème 4.2 donner des valeurs de p et q effectives pour le critère.

Considérons, pour chaque entier $k < \#\mathcal{A}$, la suite $\phi_k(u)$. Si u est un codage de rotation, on peut, d'après $\phi_k(u)$, déterminer deux entiers p_k et q_k effectifs pour le théorème 4.2 (voir la remarque suivant ce théorème). Il existe alors nécessairement un entier $l < \#\mathcal{A}$ tel que les valeurs p_l et q_l soient effectives pour le théorème 5.1. Il suffit de considérer l'entier l associé au plus petit intervalle de la partition.

6. Conclusion. La caractérisation des codages de rotations associés à une partition en deux intervalles contenant chacun un intervalle semi-ouvert de longueur égale à l'angle permet une étude précise des langages d'une large classe de suites sur deux lettres codant des rotations (théorème 4.1 et corollaire 4.1). En particulier, cette caractérisation permet de retrouver de façon naturelle les résultats de [11] et de [1] sur la complexité des codages de ce type. La proposition 5.1 se prête au même genre d'études dans le cas où l'on considère plus de deux intervalles. Malheureusement la combinatoire s'alourdit rapidement lorsque le nombre d'intervalles augmente.

Enfin, si u est un codage de rotation, on montre facilement que pour tout entier naturel n , il existe un entier k_n tel que $u_i = u_{k_n i}$ pour tout entier $i \leq n$. On ne peut espérer que ceci soit une caractérisation des codages de rotations : ce comportement est vérifié par d'autres suites qui n'en sont notoirement pas. Par exemple, si l'on ne retient qu'un terme sur deux du mot de Morse, on retrouve ce mot infini.

Dans le même esprit, on peut remarquer qu'on peut extraire d'un codage de rotation associé à une partition en deux intervalles, des suites arbitrairement proches de suites sturmiennes (i.e. telles que leur fonction de complexité soit égale à $n + 1$ pour tout n inférieur à un entier arbitrairement grand). Cette propriété peut probablement être utilisée pour caractériser les codages de rotations associés à une partition en deux intervalles.

Références

- [1] P. Alessandri, *Codages de rotations et basses complexités*, Thèse de Doctorat, Université d'Aix-Marseille II, 1996.
- [2] P. Alessandri and V. Berthé, *Three distances theorem*, prépublication.
- [3] J.-P. Allouche, *Sur la complexité des suites infinies*, Bull. Belg. Math. Soc. Simon Stevin 1 (1994), 133–143.
- [4] P. Arnoux et G. Rauzy, *Représentation géométrique de suites de complexité $2n+1$* , Bull. Soc. Math. France 119 (1991), 199–215.

- [5] J. Berstel, *Recent results in sturmian words*, dans: Developments in Language Theory (Magdeburg, 1995), World Sci., 1996, 13–24.
- [6] E. Coven and G. A. Hedlund, *Sequences with minimal block growth*, Math. Systems Theory 7 (1973), 138–153.
- [7] P. Hubert, *Propriétés combinatoires des suites définies par le billard dans les triangles pavants*, Theoret. Comput. Sci. 164 (1996), 165–183.
- [8] —, *Suites équilibrées*, prépublication.
- [9] M. Morse and G. A. Hedlund, *Symbolic dynamics*, Amer. J. Math. 60 (1938), 815–866.
- [10] —, —, *Symbolic dynamics II. Sturmian trajectories*, ibid. 62 (1940), 1–42.
- [11] G. Rote, *Sequences with subword complexity $2n$* , J. Number Theory 46 (1994), 196–213.

Institut de Mathématiques de Luminy
CNRS-UPR 9016
Case 907, 163, avenue de Luminy
13288 Marseille Cedex 9, France
E-mail: didier@iml.univ-mrs.fr

Reçu le 24.3.1997

(3155)