Fréquences des facteurs des suites sturmiennes

Valérie Berthé
Laboratoire de Mathématiques Discrètes
CNRS-UPR 9016
Case 930, 163 avenue de Luminy
F-13288 Marseille Cedex 9
France

Résumé : Dekking a explicité les fréquences des facteurs de la suite de Fibonacci en utilisant le graphe des mots. Nous généralisons ce résultat aux suites sturmiennes en montrant, également par le graphe des mots, que les fréquences des facteurs de même longueur d'une suite sturmienne prennent au plus 3 valeurs. Nous explicitons ces valeurs et donnons, pour chacune d'elles, le nombre de facteurs ayant cette fréquence en fonction du développement en fraction continue de l'angle α de la suite sturmienne.

1 Introduction

Les suites sturmiennes, dont la plus connue est la suite de Fibonacci, point fixe de la substitution σ définie par $\sigma(a) = ab$ et $\sigma(b) = a$, ont de nombreuses caractérisations (voir, par exemple, [6] et [22]).

1. Les suites sturmiennes ont pour fonction de complexité $p(n) = n + 1$, pour tout n. Rappelons que la fonction de complexité, définie pour une suite à valeurs dans un alphabet de cardinal fini, compte le nombre de facteurs de longueur donnée de cette suite. Or une suite dont la complexité satisfait $p(n) \leq n$, pour un entier n, est ultimement périodique. Les suites sturmiennes sont donc les suites de complexité minimale parmi les suites non-ultimement périodiques (voir [9]).

2. Les suites sturmiennes sont exactement les suites équilibrées sur un alphabet à deux lettres qui sont non-ultimement périodiques (voir [9], [17], [18]). Rappelons qu'une suite équilibrée est telle que la différence entre le nombre d'occurrences d'une lettre dans deux de ses facteurs de même longueur est bornée par 1 en valeur absolue.

3. Les suites sturmiennes sont des rotations irrationnelles : ce sont exactement les suites obtenues en codant l'orbite d'un point ρ du cercle unité sous

la rotation d'angle irrationnel α, par rapport à des intervalles complémentaires du cercle unité de longueur α et $1-\alpha$ (voir [17] et [18]).

4. Notons que les rotations correspondent à des codages de trajectoires de pente initiale irrationnelle dans un billard carré, où l'on code les côtés horizontaux par a et les côtés verticaux par b (voir [19]).

5. Une seconde manière "graphique" de considérer les suites sturmiennes consiste à coder le tracé d'une demi-droite de pente irrationnelle dans le plan euclidien muni d'un repère orthonormé, de la manière suivante (voir [18]).

   Proposition 1 *Soit a une suite sturmienne. Il existe alors α irrationnel dans $]0,1[$ et ρ tels que $a = \underline{a}(\alpha,\rho)$ ou $\overline{a}(\alpha,\rho)$ et où les suites $\underline{a}(\alpha,\rho) = (\underline{a}_n)_{n\in\mathbb{N}}$ (respectivement $\overline{a}(\alpha,\rho) = (\overline{a}_n)_{n\in\mathbb{N}}$) sont définies sur $\{a,b\}$ de la manière suivante :*

$$\underline{a}_n = \begin{cases} a & si\ \lfloor (n+1)\alpha+\rho \rfloor - \lfloor n\alpha+\rho \rfloor = 0 \\ b & si\ \lfloor (n+1)\alpha+\rho \rfloor - \lfloor n\alpha+\rho \rfloor = 1, \end{cases}$$

   respectivement

$$\overline{a}_n = \begin{cases} a & si\ \lceil (n+1)\alpha+\rho \rceil - \lceil n\alpha+\rho \rceil = 0 \\ b & si\ \lceil (n+1)\alpha+\rho \rceil - \lceil n\alpha+\rho \rceil = 1. \end{cases}$$

On appelle *angle* (ou *fréquence*, ou encore *pente*) d'une suite sturmienne le réel α qui lui est ainsi associé.

La fréquence $f(B)$ d'un bloc B est définie comme la limite, si elle existe, du nombre d'apparitions de ce bloc parmi les n premières lettres de la suite, divisé par n. Notons que l'existence des fréquences de blocs pour les suites sturmiennes est assurée par la caractérisation 3.

Dekking a montré, dans [10], que les fréquences des facteurs de même longueur de la suite de Fibonacci prenaient au plus 3 valeurs; il a, de plus, explicité ces valeurs et donné, pour chacune d'elles, le nombre de facteurs ayant cette fréquence. Il étudie pour cela le graphe des mots. Le but de cet article est de montrer, également en utilisant le graphe des mots, un résultat analogue pour les suites sturmiennes. Plus précisément, nous montrons que les fréquences des facteurs de même longueur des suites sturmiennes d'angle α prennent au plus trois valeurs, valeurs que nous explicitons en fonction du développement en fraction continue de α; nous donnons, de plus, le nombre de facteurs ayant chacune de ces trois fréquences.

Théorème 1 *Considérons une suite sturmienne d'angle α. Soit $m \geq 1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs tels que $\frac{p_1}{q_1} < \alpha < \frac{p_2}{q_2}$.*

Les fréquences des facteurs de longueur m sont à valeurs dans l'ensemble :

$$p_2 - \alpha q_2,\ \alpha q_1 - p_1,\ \alpha(q_1 - q_2) + p_2 - p_1.$$

Plus précisément, soient $(\frac{p^{(n)}}{q^{(n)}})$ et $(c^{(n)})$ les suites des convergents et des quotients partiels associées à α dans son développement en fraction continue.

Supposons $kq^{(n)} + q^{(n-1)} < m < (k+1)q^{(n)} + q^{(n-1)}$, avec $n \geq 1$ et $1 \leq k \leq c^{(n+1)}$. Les fréquences des facteurs de longueur m sont à valeurs dans l'ensemble :

$$\{(-1)^n(kp^{(n)} + p^{(n-1)} - \alpha(kq^{(n)} + q^{(n-1)})), (-1)^n(\alpha q^{(n)} - p^{(n)}),$$
$$(-1)^n(-\alpha((k-1)q^{(n)} + q^{(n-1)}) + (k-1)p^{(n)} + p^{(n-1)})\}.$$

Supposons $m = kq^{(n)} + q^{(n-1)}$, avec $n \geq 1$ et $1 \leq k \leq c^{(n+1)}$. Les fréquences des facteurs de longueur m sont à valeurs dans l'ensemble :

$$\{(-1)^n(kp^{(n)} + p^{(n-1)} - \alpha(kq^{(n)} + q^{(n-1)})), (-1)^n(\alpha q^{(n)} - p^{(n)})\}.$$

De plus, il y a

- *$m - q_2 + 1$ facteurs de fréquence $p_2 - \alpha q_2$;*
- *$m - q_1 + 1$ facteurs de fréquence $\alpha q_1 - p_1$;*
- *$(q_1 + q_2) - m - 1$ facteurs de fréquence $\alpha(q_1 - q_2) + p_2 - p_1$.*

On rappelle qu'un m-point de Farey est un élément α de $[0,1]$ tel que $\alpha = \frac{p}{q}$, avec $p \geq 0$, $1 \leq q \leq m$ et $pgcd(p,q) = 1$. Les points de Farey vérifient les propriétés suivantes (voir [11]) :

Proposition 2 *1. Si $\frac{p}{q}, \frac{p''}{q''}, \frac{p'}{q'}$ sont trois m-points de Farey consécutifs, alors*

$$\frac{p''}{q''} = \frac{p+p'}{q+q'}.$$

2. Deux m-points de Farey $\frac{p}{q}$ et $\frac{p'}{q'}$ tels que $m \leq q + q' - 1$ sont consécutifs si et seulement si $p'q - q'p = 1$.

3. Soit $m \geq 2$. Deux m-points de Farey successifs n'ont pas le même dénominateur.

Notons que la connaissance des fréquences de blocs d'une suite sturmienne u permet une description précise de la mesure associée au système dynamique $(\overline{\mathcal{O}(u)}, T)$, où T est le décalage qui à la suite $(u_n)_{n\in\mathbb{N}}$ associe la suite $(u_{n+1})_{n\in\mathbb{N}}$ et où $\overline{\mathcal{O}(u)}$ est l'adhérence de l'orbite sous le décalage T de la suite u, dans $\{a,b\}^{\mathbb{N}}$ muni du produit des topologies discrètes. En effet, on définit une mesure de probabilité μ sur la famille $\mathcal{B}$ des boréliens de $\overline{\mathcal{O}(u)}$, de la manière suivante

: la mesure μ est l'unique mesure de probabilité invariante par T, telle que $\mu([w]) = f(w)$, où $[w]$ est le cylindre correspondant aux suites de $\overline{\mathcal{O}(u)}$ de préfixe w et où $f(w)$ est la fréquence d'apparition du bloc w dans la suite u. Or le système dynamique associé à une suite sturmienne est *uniquement ergodique*, c'est-à-dire qu'il existe une unique mesure T-invariante. Par conséquent, la mesure μ définie ci-dessus est l'unique mesure T-invariante associée au système dynamique $(\overline{\mathcal{O}(u)}, T)$.

Le théorème 1 peut être prouvé soit en utilisant la définition combinatoire ($p(n) = n + 1$) des suites sturmiennes, ce que nous faisons ici, soit en utilisant leur caractérisation dynamique (les suites sturmiennes sont des rotations irrationnelles) (voir [4]). En effet, on vérifie qu'à chaque facteur B de la suite on peut associer un intervalle I du cercle unité de la manière suivante : l'ensemble des entiers n tels que le facteur B apparaisse à l'indice n de la suite correspond à l'ensembles des entiers tels que $\{\alpha n + \rho\}$ appartienne à l'intervalle I, où α et ρ sont associés à la suite considérée selon la proposition 1. On déduit de l'équirépartition de la suite $(\{\alpha n + \rho\})_{n \in \mathbb{N}}$ que la fréquence $f(B)$ du facteur B est alors égale à la longueur de I. Or les $(n+1)$ intervalles I correspondant aux $(n + 1)$ blocs de longueur n sont obtenus en plaçant les points $\{-\alpha\}$, $\{-2\alpha\}$, $\cdots, \{-n\alpha\}$ sur le segment $[0, 1]$, la notation $\{x\}$ désignant la partie fractionnaire de x. Par conséquent, le théorème 1 correspond à une autre formulation du théorème des trois distances, utilisé en analyse diophantienne (voir [21], [23] ou [24]) :

Théorème des trois distances *Soit α un nombre irrationnel. Plaçons les points $\{\alpha\}$, $\{2\alpha\}$, $\cdots, \{n\alpha\}$ sur le segment $[0, 1]$. Les $(n + 1)$ segments trouvés ont au plus trois longueurs, l'une étant la somme des deux autres.*

On a bien sûr remplacé ici α par $1 - \alpha$.

Ce travail a été motivé par un article de Burrows et Sulston : ceux-ci associent, dans [8], une suite de valeurs d'entropies $(H_n)_{n \in \mathbb{N}}$ à une suite binaire u, afin de donner une mesure du désordre de la suite u et éventuellement afin de "reconnaître" les suites quasi-cristallines (c'est-à-dire les suites pouvant modéliser un réseau atomique unidimensionnel quasi-cristallin) ou, plus généralement, les suites de spectre discret. La suite (H_n) converge vers l'entropie métrique du système dynamique symbolique associé à la suite u et les termes H_n sont définis à partir de fréquences conditionnelles. Nous montrons dans [5] que cette mesure du désordre ne permet pas en fait de faire une classification entre les suites selon leurs propriétés spectrales.

Notons que les suites étudiées dans [5] sont des suites substitutives et que les techniques de calcul des fréquences employées reposent sur les substitutions sous-jacentes. Ces techniques ne peuvent donc être utilisées ici car les suites sturmiennes ne sont généralement pas substitutives (voir [7], voir aussi [3]).

2 Le graphe des mots

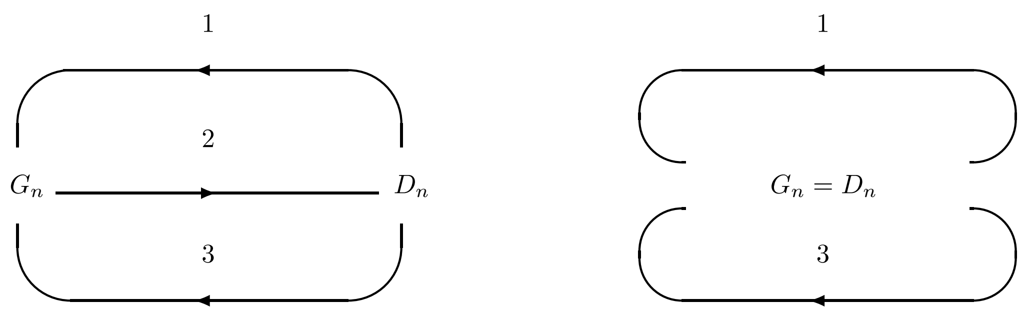
L'outil utilisé ici pour la preuve combinatoire est le graphe des mots (voir [20]) : le graphe des mots, également appelé graphe de Rauzy est un sous-graphe du graphe de de Bruijn.

Le graphe des mots de longueur n, associé à une suite, est le graphe orienté Γ_n, dont les sommets sont les facteurs de longueur n de la suite, avec une arête de U vers V si V suit U dans la suite, c'est-à-dire, plus précisément, s'il existe un mot W de longueur $n-1$ tel que

$$U = xW \text{ et } V = Wy, \text{ avec } x, y \in \{a, b\},$$

et tel que xWy soit un facteur de la suite.

Considérons une suite sturmienne. De la complexité ($p(n) = n+1$, pour tout n), on déduit l'existence d'un unique facteur D_n de longueur n biprolongeable à droite, c'est-à-dire ayant deux extensions à droite dans la suite[1]. Un tel facteur est encore appelé facteur *spécial* ou facteur *expansif.* Soit, de même, G_n l'unique facteur de longueur n biprolongeable à gauche. Une suite sturmienne présente deux types de graphes selon que $G_n = D_n$ ou que $G_n \neq D_n$:



Notons que D_{n-1} est un suffixe de D_n et que G_{n-1} est un préfixe de G_n, c'est-à-dire que l'on peut écrire $D_n = xD_{n-1}$ et $G_n = G_{n-1}y$, où x et y appartiennent à $\{a, b\}$.

Soit U un sommet de Γ_n. On note U^+ le nombre d'arêtes de Γ_n d'origine U et U^- le nombre d'arêtes d'extrémité U. Le lemme suivant permet de déduire du graphe des mots des résultats sur les fréquences.

[1]Notons qu'on entend généralement par extension d'un facteur B un facteur Bx, où x est une lettre qui suit le bloc B dans la suite. Nous appellons ici extension par abus de langage, la lettre x elle-même.

Lemme 1 *Soient U et V deux sommets reliés par une arête tels que $U^+ = 1$ et $V^- = 1$. Les facteurs U et V ont alors la même fréquence.*

En effet, écrivons $U = xW$ et $V = Wy$, où x et y sont des lettres. Comme $U^+ = 1$, le facteur U a pour unique extension droite y; de même, le facteur V a pour unique extension gauche x. Par conséquent, nous avons les égalités suivantes entre les fréquences :

$$f(U) = f(Uy) = f(xWy) = f(xV) = f(V).$$

Par branche (1) ou (3), représentées sur la figure ci-dessus, on entend tous les mots de ce chemin, D_n et G_n exclus. En revanche, D_n et G_n seront inclus dans la branche (2).

On déduit alors du lemme précédent que les mots d'une même branche ont même fréquence. On associera donc à une branche la fréquence des mots de cette branche.

3 Quelques propriétés des suites sturmiennes

Nous allons rappeler, dans ce paragraphe, quelques propriétés des suites sturmiennes.

Lemme 2 *L'ensemble des facteurs d'une suite sturmienne est stable par image miroir. On en déduit, en particulier, que G_n est l'image miroir de D_n.*

Par image miroir, on entend le retourné d'un mot. Par exemple, *abaa* est le miroir de *aaba*.

Pour montrer ce lemme (voir par exemple [2]), il suffit de considérer la caractérisation 2 des suites sturmiennes comme suites équilibrées. En effet, le cardinal d'un ensemble équilibré de facteurs de longueur n sur un alphabet à deux lettres est au plus $n + 1$ (voir [9]). Un ensemble équilibré de facteurs est tel que la différence entre le nombre d'occurrences d'une lettre dans deux de ses facteurs de même longueur est bornée par 1 en valeur absolue. Par conséquent, en adjoignant les images miroirs des $n + 1$ facteurs de longueur n d'une suite sturmienne à ces mêmes facteurs, on obtient encore un ensemble équilibré de facteurs, donc de même cardinal $n + 1$.

Lemme 3 *Les suites sturmiennes de même angle ont mêmes facteurs.*

Ce lemme est une conséquence directe de la minimalité des rotations irrationnelles (caractérisation 3) (voir [16]).

Le résultat suivant permet d'expliciter les facteurs expansifs (voir [2] ou [15]).

Lemme 4 *Le facteur expansif de longueur m d'une suite sturmienne d'angle α est le retourné du bloc $a_1 a_2 \cdots a_m$, où $(a_n)_{n\in\mathbb{N}} = \underline{a}(\alpha, 0)$.*

En effet, le bloc $a_1 a_2 \cdots a_m$ a deux prolongés à gauche suivant que l'on considère la caractérisation par partie entière inférieure ou supérieure donnée dans la proposition 1. On conclut alors grâce aux deux lemmes précédents.

On en déduit le lemme suivant ([16]).

Lemme 5 *Soit $m \geq 1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs. Les suites sturmiennes dont l'angle α vérifie $\alpha \in]\frac{p_1}{q_1}, \frac{p_2}{q_2}[$ ont même facteur expansif de longueur $m-1$.*

En effet, on a $\lfloor k\alpha \rfloor = \lfloor k\frac{p_1}{q_1} \rfloor$, pour $1 \leq k \leq m$. On déduit donc ce résultat du lemme 4.

Lemme 6 *Deux suites sturmiennes ayant le même facteur expansif de longueur $m-1$ ont les mêmes facteurs de longueur m.*

On montre ce lemme par récurrence. On vérifie qu'il est vrai pour $m = 2$. Supposons que deux suites sturmiennes ayant le même facteur expansif de longueur $m-1$ ont les mêmes facteurs de longueur m. Considérons alors deux suites sturmiennes ayant le même facteur expansif D_m de longueur m et par conséquent le même facteur biprolongeable à gauche G_m, d'après le lemme 2. En particulier, par hypothèse de récurrence, ces deux suites ont mêmes facteurs de longueur m, car elles ont le même facteur expansif de longueur $m-1$. Montrons que les facteurs de longueur m ont les mêmes extensions dans les deux suites.

Supposons que $G_{m-1} \neq D_{m-1}$. Le facteur D_m a pour extensions a et b, dans les deux suites. Les facteurs de longueur m différents de D_m ont une unique extension droite. Or le suffixe de longueur $m-1$ d'un facteur de longueur m différent de D_m est différent de D_{m-1}, car $G_{m-1} \neq D_{m-1}$; on conclut alors en notant que ce suffixe a donc une unique extension droite, qui est la même dans les deux suites, par hypothèse de récurrence.

Supposons maintenant que $G_{m-1} = D_{m-1}$. On note $D_m = xD_{m-1}$. On a, d'après le lemme 2 : $G_m = G_{m-1}x$. Notons de plus $\overline{x} = a$, si $x = b$ et $\overline{x} = b$, si $x = a$. Le facteur xD_{m-1} a pour extensions droites a et b, dans les deux suites, par définition. De même, le facteur $D_{m-1}x$ a pour extensions gauches a et b. Par conséquent, le facteur $\overline{x}D_{m-1}$ a pour unique extension droite x, dans les deux suites. Le raisonnement est le même que précédemment pour les facteurs de longueur m restants.

On déduit le lemme suivant de la représentation par parties entières.

Lemme 7 *Soit $m \geq 1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs. On considère une suite sturmienne dont l'angle α vérifie $\alpha \in]\frac{p_1}{q_1}, \frac{p_2}{q_2}[$. Supposons $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$. On a alors $D_{q_1+q_2-1} = aD_{q_1+q_2-2}$. De manière analogue, si $\frac{p_1+p_2}{q_1+q_2} < \alpha < \frac{p_2}{q_2}$, alors $D_{q_1+q_2-1} = bD_{q_1+q_2-2}$.*

Preuve Montrons que l'on a $\lfloor (q_1+q_2-1)\alpha \rfloor = p_1+p_2-1$. En effet, on a d'après la proposition 2 : $(p_1+p_2)q_1-(q_1+q_2)p_1 = 1$. On en déduit que $\frac{p_1+p_2-1}{q_1+q_2-1} \leq \frac{p_1}{q_1}$, et donc que $p_1 + p_2 - 1 \leq \alpha(q_1 + q_2 - 1)$. On montre, de même que $\frac{p_2}{q_2} \leq \frac{p_1+p_2}{q_1+q_2-1}$, ce qui implique que $(q_1 + q_2 - 1)\alpha < p_1 + p - 2$.

Supposons $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$. On a donc $\lfloor (q_1 + q_2)\alpha \rfloor \leq p_1 + p_2 - 1$, ce qui implique que $\lfloor (q_1+q_2-1)\alpha \rfloor = \lfloor (q_1+q_2)\alpha \rfloor$. On déduit de la proposition 1 et du lemme 4 que $D_{q_1+q_2-1} = aD_{q_1+q_2-2}$.

On montre, de manière analogue que si $\frac{p_1+p_2}{q_1+q_2} < \alpha < \frac{p_2}{q_2}$ alors $D_{q_1+q_2-1} = bD_{q_1+q_2-2}$.

Soit $m \geq 1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs. Ces deux m-points de Farey sont également deux $(q_1 + q_2 - 1)$-points de Farey successifs, d'après la proposition 2. Par conséquent, les suites sturmiennes dont l'angle α vérifie $\alpha \in \left]\frac{p_1}{q_1}, \frac{p_2}{q_2}\right[$ ont le même facteur expansif de longueur $q_1 + q_2 - 2$ (lemme 5) et donc les mêmes facteurs de longueur $q_1 + q_2 - 1$ (lemme 6). En particulier, le facteur $G_{q_1+q_2-2}$ a deux extensions, d'après le lemme 7, selon la position de α par rapport à $\frac{p_1+p_2}{q_1+q_2}$, ce qui implique que $G_{q_1+q_2-2} = D_{q_1+q_2-2}$, ou en d'autres termes, que $G_{q_1+q_2-2}$ est un palindrome. Plus généralement, la proposition suivante donne une caractérisation des facteurs spéciaux palindromes.

Proposition 3 *Soit $m \geq 1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs tels que $\frac{p_1}{q_1} \neq 0$ et $\frac{p_2}{q_2} \neq 1$. On considère une suite sturmienne dont l'angle α vérifie $\alpha \in \left]\frac{p_1}{q_1}, \frac{p_2}{q_2}\right[$. On a $G_m = D_m$ si et seulement si $m = q_1 + q_2 - 2$.*

Preuve La preuve de cette proposition repose sur les propriétés de bonne approximation des points de Farey.

Notons que l'hypothèse $\frac{p_1}{q_1} \neq 0$ et $\frac{p_2}{q_2} \neq 1$ implique que $G_{q_1+q_2-1}$ n'est pas un palindrome. En effet, supposons que l'on ait $G_{q_1+q_2-1} = D_{q_1+q_2-1}$. On a vu que $G_{q_1+q_2-2} = D_{q_1+q_2-2}$. On en déduit donc que $G_{q_1+q_2-1}$ est une puissance $(q_1 + q_2 - 1)$- ième d'une lettre que nous noterons x, ce qui implique, d'après le lemme 7, que $\frac{p_1}{q_1} = 0$ ou $\frac{p_2}{q_2} = 1$. Plus précisément, on vérifie que $x = a$ si $\frac{p_1}{q_1} = 0$ et que $x = b$, si $\frac{p_2}{q_2} = 1$.

Par conséquent, il suffit de montrer que l'on a $G_m \neq D_m$ pour $max(q_1, q_2) \leq m < q_1 + q_2 - 2$. Supposons donc que $G_m = D_m$, avec $max(q_1, q_2) \leq m < q_1 + q_2 - 2$. On vérifie que

$$D_m = b_1 \cdots b_m, \text{ où } (b_n) = \underline{a}(\alpha, 1 - \{\alpha(m + 2)\}),$$

selon le même raisonnement que pour le lemme 4. De l'égalité $G_m = D_m$, on déduit que la différence $\lfloor \alpha k + 1 - \{\alpha(m + 2)\} \rfloor - \lfloor \alpha k \rfloor$ est constante pour $1 \leq k \leq m + 1$. Cette différence vaut de plus 1 ou 0.

Supposons que $\lfloor \alpha k + 1 - \{\alpha(m+2)\}\rfloor - \lfloor \alpha k\rfloor = 0$, pour tout $1 \leq k \leq m+1$. On a donc

$$1 - \{\alpha(m+2)\} < 1 - \{\alpha k\}, \text{pour tout } 1 \leq k \leq m+1. \tag{1}$$

Posons $p = 1 + \lfloor \alpha(m+2)\rfloor$ et $q = m+2$. Montrons que l'on a :

$$\alpha < \frac{p_2}{q_2} < \frac{p}{q} < \frac{p - p_2}{q - q_2}. \tag{2}$$

On vérifie aisément l'inégalité $\alpha < \frac{p}{q}$. Or $\frac{p}{q}$ est un $(q_1 + q_2 - 1)$-point de Farey. Les deux m-points de Farey consécutifs $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont également, d'après la proposition 2.1, deux $(q_1 + q_2 - 1)$-points de Farey consécutifs, puisque, par hypothèse sur m, $q \leq q_1 + q_2 - 1$. On obtient donc : $\frac{p_2}{q_2} < \frac{p}{q}$, ce qui implique la dernière inégalité de (2), à savoir $\frac{p}{q} < \frac{p-p_2}{q-q_2}$. On en déduit que $p - \alpha q > p_2 - \alpha q_2$, ce qui est en contradiction avec l'équation (1), dans laquelle on a affecté à k la valeur q_2.

Dans le cas où $\lfloor \alpha k + 1 - \{\alpha(m+2)\}\rfloor - \lfloor \alpha k\rfloor = 0$, pour tout $1 \leq k \leq m+1$, on obtient également une contradiction en faisant intervenir q_1 au lieu de q_2.

Remarque Il existe de nombreuses preuves de cette proposition (voir par exemple [4], [12] ou [14]). Hubert ([14]) prouve en particulier ce résultat en exploitant également les propriétés de bonne approximation des points de Farey mais en utilisant la représentation 4 des suites sturmiennes, comme codages de trajectoires dans un billard carré.

On peut également donner une preuve de la proposition 3 en utilisant la caractérisation des mots strictement bispéciaux donnée par Mignosi et de Luca dans [12], à partir des mots standard de Rauzy (voir [19]). En effet, si on a l'égalité $D_m = G_m$, on vérifie que le mot G_m est alors *strictement bispécial*, c'est-à-dire que les quatre extensions aG_m, bG_m, $G_m a$ et $G_m b$ sont des facteurs de suites sturmiennes. Notons que sous les hypothèses de la proposition 3, le mot $G_{q_1+q_2-2}$ est alors un élément de PER, en reprenant les notations de [12], c'est-à-dire que $G_{q_1+q_2-2}$ admet deux périodes, q_1 et q_2, premières entre elles, ou en d'autres termes, que $G_{q_1+q_2-2}$ est un mot maximal pour le théorème de Fine et Wilf (voir [13]).

4 Preuve du théorème 1

On considère une suite sturmienne d'angle α. Soient $(\frac{p^{(n)}}{q^{(n)}})$ et $(c^{(n)})$ les suites des convergents et des quotients partiels associés à α dans son développement en fraction continue.

Rappelons que la suite des convergents est définie, par récurrence, de la manière suivante (voir, par exemple, [11]) :

$$\begin{cases} p^{(n+1)} = c^{(n+1)}p^{(n)} + p^{(n-1)}, \ p^{(-1)} = 1, \ p^{(0)} = c^{(0)} \\ q^{(n+1)} = c^{(n+1)}q^{(n)} + q^{(n-1)}, \ q^{(-1)} = 0, \ q^{(0)} = 1. \end{cases}$$

On a, de plus, $pgcd(p^{(n)}, q^{(n)}) = 1$ et $(\alpha - \frac{p^{(n)}}{q^{(n)}})$ du signe de $(-1)^n$. Nous allons supposer n pair pour fixer les idées, le raisonnement étant analogue si l'on suppose n impair. On vérifie que l'on a la situation suivante :

$$\frac{p^{(n)}}{q^{(n)}} < \alpha < \frac{p^{(n+1)}}{q^{(n+1)}} < \frac{kp^{(n)} + p^{(n-1)}}{kq^{(n)} + q^{(n-1)}} < \frac{p^{(n-1)}}{q^{(n-1)}},$$

pour $0 \leq k \leq c^{(n+1)}$. Or les convergents $\frac{p^{(n)}}{q^{(n)}}$ et $\frac{p^{(n-1)}}{q^{(n-1)}}$ satisfont

$$p^{(n-1)}q^{(n)} - p^{(n)}q^{(n-1)} = 1,$$

ce qui implique que

$$(kp^{(n)} + p^{(n-1)})q^{(n)} - (kq^{(n)} + q^{(n-1)})p^{(n)} = 1.$$

On déduit alors de la proposition 2.2 le lemme suivant.

Lemme 8 *Les points* $\frac{p^{(n)}}{q^{(n)}}$ *et* $\frac{kp^{(n)}+p^{(n-1)}}{kq^{(n)}+q^{(n-1)}}$ *sont deux* M*-points de Farey consécutifs, pour*
$1 \leq k \leq c^{(n+1)}$*, avec* $M = kq^{(n)} + q^{(n-1)}$.

Par conséquent, on déduit le théorème 1 du théorème suivant.

Théorème 2 *Soit* u *une suite sturmienne d'angle* α*. Soit* $m \geq 1$ *. Soient* $\frac{p_1}{q_1}$ *et* $\frac{p_2}{q_2}$ *deux* m*-points de Farey consécutifs tels que* $\frac{p_1}{q_1} < \alpha < \frac{p_2}{q_2}$.
Les fréquences des facteurs de longueur m *sont à valeurs dans l'ensemble :*

$$p_2 - \alpha q_2, \ \alpha q_1 - p_1, \ \alpha(q_1 - q_2) + p_2 - p_1.$$

Plus précisément, il y a

- $m - q_2 + 1$ *facteurs de fréquence* $p_2 - \alpha q_2$*;*
- $m - q_1 + 1$ *facteurs de fréquence* $\alpha q_1 - p_1$*;*
- $(q_1 + q_2) - m - 1$ *facteurs de fréquence* $\alpha(q_1 - q_2) + p_2 - p_1$.

La preuve de ce résultat se fait par récurrence sur m et est basée sur l'évolution du graphe des mots étudiée par Arnoux et Rauzy dans [1]. En effet, la taille des branches donne le nombre de facteurs correspondant à chacune des fréquences. De plus, les fréquences des branches reliant D_m à G_m sont données par $f(D_{m-1}a)$

et $f(D_{m-1}b)$ (si elles sont non vides), alors que la fréquence de la branche du milieu est donnée par $f(D_m)$. On utilisera donc l'hypothèse supplémentaire de récurrence suivante (sur m) : $f(D_{m-1}a) = p_2 - \alpha q_2$ et $f(D_{m-1}b) = \alpha q_1 - p_1$.

On vérifie que la propriété est vraie pour $m = 1$. En effet, la lettre a a pour fréquence $1-\alpha$ et la lettre b a pour fréquence α, d'après la caractérisation 3 des suites sturmiennes comme codages de rotations et plus précisément, d'après la propriété d'équirépartition de la suite $(\{\alpha n + \rho\})_{n\in\mathbb{N}}$, α étant irrationnel.

Supposons la proriété vraie pour $m \geq 1$. Montrons qu'elle est vraie pour $m+1$. Soient $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ deux m-points de Farey consécutifs. Soit u une suite sturmienne d'angle α tel que $\frac{p_1}{q_1} < \alpha < \frac{p_2}{q_2}$. Nous allons distinguer trois cas selon la position de m par rapport à $q_1 + q_2 - 2$: dans les deux premiers cas ($m < q_1+q_2-2$ et $m = q_1+q_2-2$), $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont également deux $(m+1)$-points de Farey consécutifs alors que dans le troisième cas ($m = q_1 + q_2 - 1$), $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ ne sont plus des $(m+1)$-points de Farey consécutifs (d'après la proposition 2). Notons que si $\frac{p_1}{q_1} = 0$ ou si $\frac{p_2}{q_2} = 1$, on ne considère que le cas $m = q_1 + q_2 - 1$. En effet, on a alors $m = \sup(q_1, q_2) = q_1 + q_2 - 1$.

- Supposons $m < q_1 + q_2 - 2$. On a donc, d'après ce qui précède, $\frac{p_1}{q_1} \neq 0$ et $\frac{p_2}{q_2} \neq 1$.

  Par hypothèse de récurrence, on a $f(D_{m-1}a) = p_2 - \alpha q_2$ et $f(D_{m-1}b) = \alpha q_1 - p_1$. Montrons que l'on a $G_{m-1} \neq D_{m-1}$. Les $(m-1)$-points de Farey consécutifs encadrant l'angle α sont, d'après la proposition 2 :

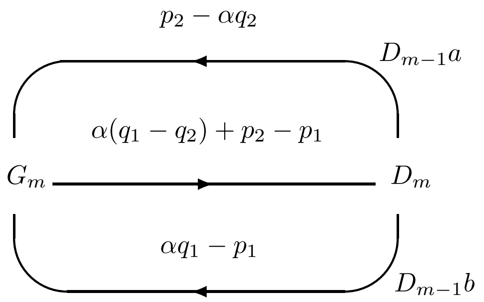
  - $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$, si $m \geq \sup(q_1, q_2) + 1$;
  - $\frac{p_1-p_2}{q_1-q_2}$ et $\frac{p_2}{q_2}$, si $m = q_1$;
  - $\frac{p_1}{q_1}$ et $\frac{p_2-p_1}{q_2-q_1}$, si $m = q_2$.

  Dans le cas où les $(m-1)$-points de Farey consécutifs encadrant l'angle α sont différents de 0 et de 1, l'inégalité $G_{m-1} \neq D_{m-1}$ résulte de la proposition 3. Dans le cas où $m = q_1$, l'égalité $\frac{p_1-p_2}{q_1-q_2} = 0$ implique que $\frac{p_1}{q_1} = \frac{1}{m}$ et que $\frac{p_2}{q_2} = \frac{1}{m-1}$. On a alors $D_{m-1} = ba^{m-2}$, d'après le lemme 7. Le facteur G_{m-1} étant l'image miroir du facteur D_{m-1}, on a donc $G_{m-1} \neq D_{m-1}$. De même, dans le cas où $m = q_2$, l'égalité $\frac{p_2-p_1}{q_2-q_1} = 1$ implique que $\frac{p_1}{q_1} = \frac{m-2}{m-1}$, $\frac{p_2}{q_2} = \frac{m-1}{m}$ et que $D_{m-1} = ab^{m-2}$. On a donc montré, dans tous les cas, que $G_{m-1} \neq D_{m-1}$.

  Par conséquent, on en déduit que

  $$f(D_m) = f(D_{m-1}) = \alpha(q_1 - q_2) + p_2 - p_1.$$

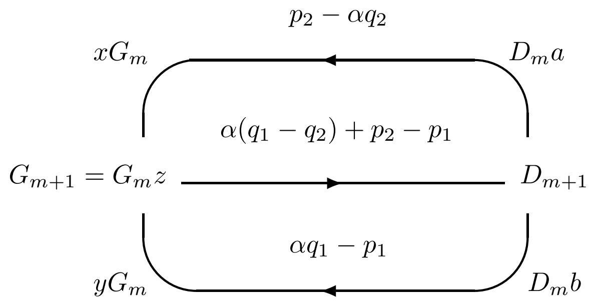
  On peut donc affecter aux trois branches du graphe des mots de longueur m les fréquences correspondantes :



Γ_m

Les mots de la branche (1) ont pour fréquence $p_2 - \alpha q_2$, les mots de la branche (2) ont pour fréquence $\alpha(q_1 - q_2) + p_2 - p_1$ et les mots de la branche (3) ont pour fréquence $\alpha q_1 - p_1$.

L'évolution du graphe des mots est la suivante :



Γ_{m+1}

On a complété le graphe par la gauche, c'est-à-dire qu'on a associé à chaque facteur distinct de G_m son unique extension à gauche et à G_m ses deux extensions à gauche.

Rappelons que l'hypothèse $m < q_1 + q_2 - 2$ implique, d'après la proposition 2, que $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont également deux $(m+1)$-points de Farey consécutifs encadrant l'angle α.

On a, de plus,

$$f(D_m x) = f(D_{m-1} x),$$

pour $x = a$ ou b. En effet, on a $D_{m-1}x \neq G_m$, car $G_{m-1} \neq D_{m-1}$. On a de même $f(D_{m+1}) = f(D_m)$, car $D_m \neq G_m$, d'après la proposition 3. Par conséquent, les mêmes fréquences sont associées aux mêmes branches.

Il reste à évaluer le nombre de facteurs ayant chacune des trois fréquences. Ces nombres sont donnés par la taille des branches du graphe. Or on constate que le nombre de facteurs de la branche (1) (respectivement de la branche (3)) augmente de 1 quand on passe de Γ_m à Γ_{m+1}, alors que le nombre de facteurs de la branche (2) diminue de 1 (voir [1]).

- Supposons maintenant que $m = q_1 + q_2 - 2$, ce qui implique que $\frac{p_1}{q_1} \neq 0$ et que $\frac{p_2}{q_2} \neq 1$. On a l'égalité $G_m = D_m$, d'après la proposition 3, mais en revanche on a $G_{m-1} \neq D_{m-1}$. En effet, si $q_1 \neq 2$ et $q_2 \neq 2$ alors $m \geq \sup(q_1, q_2)+1$, ce qui implique que $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont des $(m-1)$-points de Farey consécutifs. On conclut en appliquant la proposition 3. Supposons maintenant que q_1 ou q_2 soit égal à 2. Supposons $q_1 = 2$ pour fixer les idées, le raisonnement étant analogue si $q_2 = 2$. On a alors $q_2 > q_1 = 2$, car $\frac{p_2}{q_2} \neq 1$. Par conséquent, $\frac{p_1}{q_1}$ et $\frac{p_2-p_1}{q_2-q_1}$ sont deux $(m-1)$-points de Farey consécutifs encadrant l'angle α. Supposons de plus que $\frac{p_2-p_1}{q_2-q_1} = 1$. On a alors $\frac{p_2}{q_2} = \frac{2}{3}$, $m = 3$ et d'après le lemme 7 $D_2 = ab \neq G_2$. Supposons maintenant $\frac{p_2-p_1}{q_2-q_1} \neq 1$. On a $m - 1 = q_2 - 1 \neq q_1 + (q_2 - q_1) - 2$. On conclut ici encore en appliquant la proposition 3. On a donc montré dans tous les cas l'inégalité $G_{m-1} \neq D_{m-1}$.

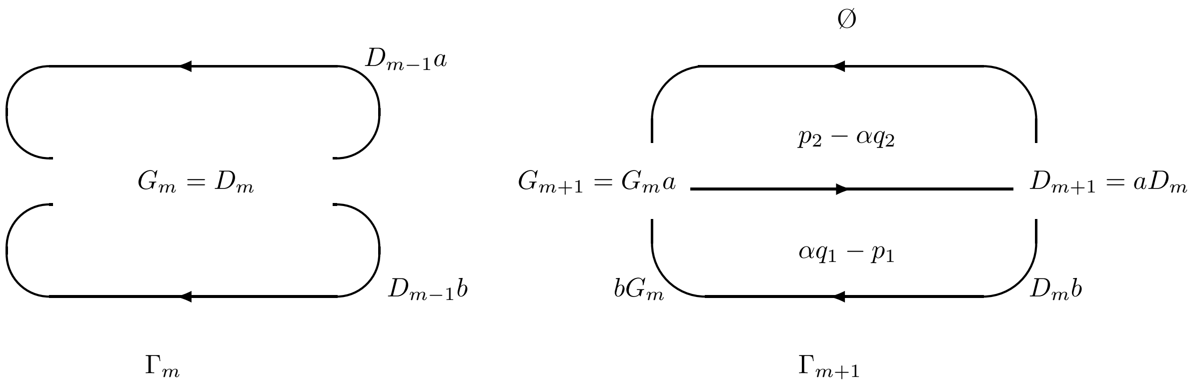
  On a donc
  $$f(D_m) = f(D_{m-1}) = \alpha(q_1 - q_2) + p_2 - p_1. \tag{3}$$

  Par hypothèse de récurrence, les mots de la branche (1) ont pour fréquence $p_2 - \alpha q_2$, les mots de la branche (2) ont pour fréquence $\alpha(q_1 - q_2) + p_2 - p_1$ et d'après l'équation (3), les mots de la branche (3) ont pour fréquence $\alpha q_1 - p_1$. La branche (1) a de plus $(q_1 - 1)$ éléments et la branche (3) comporte $(q_2 - 1)$ éléments, par hypothèse de récurrence.

  On constate que l'évolution du graphe des mots dépend de la première lettre de D_{m+1}. Par conséquent, d'après le lemme 7, il faut distinguer deux cas selon que $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$ ou que $\frac{p_1+p_2}{q_1+q_2} < \alpha < \frac{p_2}{q_2}$. Nous supposerons ici que $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$, le raisonnement étant analogue dans l'autre cas. On a donc : $D_{m+1} = aD_m$ et $G_{m+1} = G_m a$.

  L'évolution du graphe des mots est la suivante, en complétant ici encore le graphe par la gauche (voir [1]) :

  La branche (1) est vide. On vérifie que la branche (2) comporte un élément de plus que la branche (1) de Γ_m, à savoir q_1 éléments et que la branche



(3) (de Γ_{m+1}) comporte également un élément de plus que la branche correspondante de Γ_m, c'est-à-dire q_2 éléments.

On a de plus $f(D_m x) = f(D_{m-1}x)$, pour $x = a$ ou b, car $D_{m-1} \neq G_{m-1}$. Les fréquences des facteurs de longueur $m+1$ prennent donc deux valeurs données par $f(D_{m+1}) = f(D_m a)$ pour les mots de la branche du milieu de Γ_{m+1} et par $f(D_m b)$ pour les mots de la troisième branche, ce qui achève la récurrence dans ce cas, en rappelant que $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont deux $(m+1)$-points de Farey consécutifs.

- Supposons enfin que $m = q_1 + q_2 - 1$. On suppose toujours que $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$, c'est-à-dire que α est compris entre les $(m+1)$-points de Farey consécutifs $\frac{p_1}{q_1}$ et $\frac{p_1+p_2}{q_1+q_2}$. On a donc $D_m = aD_{m-1}$ et $G_m = G_{m-1}a$, d'après le lemme 7.

  Si $\frac{p_1}{q_1} \neq 0$ et $\frac{p_2}{q_2} \neq 1$, on vérifie que $\frac{p_1}{q_1}$ et $\frac{p_2}{q_2}$ sont des $(m-1)$-points de Farey successifs encadrant l'angle α, ce qui implique que $G_{m-1} = D_{m-1}$, d'après la proposition 3. Si $\frac{p_1}{q_1} = 0$, on a alors $G_{m-1} = D_{m-1} = a^{m-1}$. De même, si $\frac{p_1}{q_1} = 1$, $G_{m-1} = D_{m-1} = b^{m-1}$. On a donc, dans tous les cas, l'égalité $G_{m-1} = D_{m-1}$.

  On a $f(D_m b) = f(D_{m-1}b)$ car $D_{m-1}b = G_{m-1}b \neq G_m$. Par conséquent, on obtient que
  $$f(D_m b) = \alpha q_1 - p_1.$$
  On déduit, de plus, du lemme 1 que
  $$f(D_m) = f(G_m).$$
  Or $G_m = G_{m-1}a = D_{m-1}a$. En utilisant l'hypothèse de récurrence, on obtient donc que
  $$f(D_m) = p_2 - \alpha q_2.$$
  On en déduit que
  $$f(D_m a) = f(D_m) - f(D_m b) = -\alpha(q_1 + q_2) + p_1 + p_2.$$

Il reste à déterminer la troisième fréquence, si elle existe, et à considérer la longueur des branches, pour achever la récurrence. Pour cela, nous allons distinguer deux cas selon que $\frac{p_1}{q_1} = 0$ ou non.

- Supposons $\frac{p_1}{q_1} \neq 0$. Montrons que l'on a $G_m \neq D_m$. Si $\frac{p_2}{q_2} \neq 1$, il suffit d'appliquer la proposition 7. Si $\frac{p_2}{q_2} = 1$, on a alors $\frac{p_1}{q_1} = \frac{m-1}{m}$. Or on a supposé $\frac{p_1}{q_1} < \alpha < \frac{p_1+p_2}{q_1+q_2}$, c'est-à-dire $\frac{m-1}{m} < \alpha < \frac{m}{m+1}$. On a donc $D_m = ab^{m-1} \neq G_m$, d'après le lemme 7.
  On en déduit que
  $$f(D_{m+1}) = f(D_m) = p_2 - \alpha q_2.$$
  On constate de plus que l'évolution de la taille des branches du graphe des mots est analogue à celle du premier cas ($m < q_1 + q_2 - 2$), puisque $G_m \neq D_m$. Plus précisément, le nombre de facteurs de la branche (1) (respectivement de la branche (3)) augmente de 1 quand on passe de Γ_m à Γ_{m+1}, alors que le nombre de facteurs de la branche (2) diminue de 1. Par conséquent, parmi les $q_1 + q_2 + 1$ facteurs de longueur $q_1 + q_2$, il y a 1 (c'est-à-dire $m + 1 - (q_1 + q_2) + 1$) facteur de fréquence $-\alpha(q_1 + q_2) + p_1 + p_2$ (branche 1), $q_1 - 1$ (c'est-à-dire $2q_1 + q_2 - (m + 1) - 1$) facteurs de fréquence $p_2 - \alpha q_2$ (branche 2) et $q_2 + 1$ (c'est-à-dire $m + 1 - q_1 + 1$) facteurs de fréquence $\alpha q_1 - p_1$ (branche 3).
- Supposons que $\frac{p_1}{q_1} = 0$. On a donc $\frac{p_2}{q_2} = \frac{1}{m}$, c'est-à-dire $m = q_2$. On a de plus
  $$D_{m+1} = G_{m+1} = a^{m+1}.$$
  Par conséquent, on obtient
  $$f(D_{m+1}) = f(D_m a) = 1 - \alpha(q_2 + 1).$$
  On vérifie alors que seul D_{m+1} a pour fréquence $1 - \alpha(q_2 + 1)$ et que les q_2 facteurs de longueur $q_2 + 1$ restant ont pour fréquence α.

Remarques

- On constate que l'évolution de la forme du graphe des mots est gouvernée par la place de α par rapport aux $(q_1 + q_2)$-points de Farey $\frac{p_1}{q_1}$, $\frac{p_1+p_2}{q_1+q_2}$ et $\frac{p_2}{q_2}$. En effet, on en déduit la lettre qui prolonge $D_{q_1+q_2-2}$ dans $D_{q_1+q_2-1}$ (lemme 7), ce qui détermine, en particulier, quelle est la branche qui des branches (1) ou (3) de $\Gamma_{q_1+q_2-2}$ se transforme en branche du milieu pour $\Gamma_{q_1+q_2-1}$. En particulier, on peut démontrer le théorème 1 sans utiliser la proposition 3 en ajoutant quelques hypothèses de récurrence supplémentaires. Par souci de clarté, nous avons préféré démontrer séparément la proposition 3.

- De nouvelles fréquences n'apparaissent quand on passe des mots de longueur n aux mots de longueur $n+1$, que lorsque $G_{n-1} = D_{n-1}$, dans le cas où $\frac{p_1}{q_1} \neq 0$ et $\frac{p_2}{q_2} \neq 1$. Le processus est alors le suivant : on soustrait à la plus grande des fréquences des branches (1) et (3), la plus petite des ces fréquences, qui est une "mesure d'approximation" de l'angle α; il s'agit de l'algorithme des fractions continues.

- Arnoux et Rauzy ont également étudié dans [1] les suites de complexité $2n+1$, telles qu'il existe, pour tout n, un unique facteur de longueur n ayant trois prolongements à droite, noté D_n, et un unique facteur de longueur n ayant trois prolongements à gauche, noté G_n. Arnoux et Rauzy montrent qu'une telle suite peut se représenter géométriquement comme un échange de six intervalles sur le cercle unité. Cet échange est uniquement ergodique. Par conséquent, les fréquences de blocs existent pour une telle suite. La méthode utilisée ici se généralise facilement. En effet, le graphe des mots admet alors quatre branches, trois branches allant de D_n à G_n et une branche allant de G_n à D_n. On montre en particulier qu'il existe, pour les facteurs de longueur donnée, au plus quatre fréquences dont l'une est la somme des trois autres.

Remerciements Je remercie J.-P. Allouche, J. Berstel et P. Liardet pour leurs précieux conseils ainsi que P. Hubert pour de nombreuses discussions. Je remercie également J.-M. Dumont qui m'a indiqué la référence [21]. Enfin, je remercie le rapporteur de cet article pour ses commentaires judicieux.

Références

[1] P. ARNOUX et G. RAUZY *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. math. France **119** (1991), 199–215.

[2] J. BERSTEL *Mots de Fibonacci*, Séminaire d'Informatique Théorique, LITP, Universités Paris 6-7 (1980-81), 57–78.

[3] J. BERSTEL et P. SÉÉBOLD *Morphismes de Sturm*, Bull. Belg. Math. Soc. **1** (1994), 175–189.

[4] V. BERTHÉ *Fonctions de Carlitz et automates. Entropies conditionnelles*, Thèse, Univ. Bordeaux I (1994).

[5] V. BERTHÉ *Conditional entropy of some automatic sequences*, J. Phys. A: Math. Gen. **27** (1994) 7993–8006.

[6] T. C. BROWN *Descriptions of the characteristic sequence of an irrational*, Canad. Math. Bull. **36** (1993), 15–21.

[7] D. CRISP, W. MORAN, A. POLLINGTON et P. SHIUE *Substitution invariant cutting sequences*, Journal de Théorie des Nombres de Bordeaux **5** (1993), 123–137.

[8] B. L. BURROWS et K. W. SULSTON *Measures of disorder in non-periodic sequences*, J. Phys. A: Math. Gen. **24** (1991), 3979–3987.

[9] E. M. COVEN et G. A. HEDLUND *Sequences with minimal block growth*, Math. Systems Theory **7** (1973), 138–153.

[10] F. M. DEKKING *On the Prouhet-Thue-Morse Measure*, Acta Universitatis Carolinae, Mathematica et Physica, **33** (1992), 35–40.

[11] G. H. HARDY et E. M. WRIGHT *An introduction to the theory of numbers*, Oxford Science Publications (1979).

[12] A. DE LUCA et F. MIGNOSI *Some combinatorial properties of Sturmian words*, Theoret. Comput. Sci., à paraître.

[13] N. J. FINE et H. S. WILF *Uniquenes theorems for periodic functions*, Proc. Amer. Math. Soc. **16** (1965), 109–114.

[14] P. HUBERT *Communication privée*.

[15] F. MIGNOSI *Infinite words with linear subword complexity*, Theoret. Comput. Sci. **65** (1989), 221–242.

[16] F. MIGNOSI *On the number of factors of Sturmian words*, Theoret. Comput. Sci. **82** (1991), 71–84.

[17] M. MORSE et G. A. HEDLUND *Symbolic dynamics*, Amer. J. Math. **60** (1938), 815–866.

[18] M. MORSE et G. A. HEDLUND *Symbolic dynamics II: Sturmian trajectories*, Amer. J. Math. **62** (1940), 1–42.

[19] G. RAUZY *Mots infinis en arithmétique*, dans : M. Nivat et D. Perrin, eds., Automata on Infinite Words, Lecture Notes in Computer Science **192** (1985), 165–171.

[20] G. RAUZY *Suites à termes dans un alphabet fini*, Sém. de Théorie des Nombres de Bordeaux (1983), 25-01–25-16.

[21] N. B. SLATER *Gaps and steps for the sequence $n\theta$ mod 1*, Proc. Cambridge Phil. Soc. **63** (1967), 1115–1123.

[22] K. B. STOLARSKY *Beatty sequences, continued fractions, and certain shift operators* Canad. Math. Bull. **19** (1976), 473–482.

[23] V. T. SÓS *On the distribution mod 1 of the sequence* $n\alpha$, Ann. Univ. Sci. Budapest Eőtvős Sect. Math. **1** (1958), 127–134.

[24] S. ŚWIERCZKOWSKI *On successive settings of an arc on the circumference of a circle*, Fund. Math. **46** (1958), 187–189.