

Autour du système de numération d'Ostrowski

Valérie Berthé

Résumé

Le système de numération d'Ostrowski a pour échelle de numération les dénominateurs des convergents dans le développement en fraction continue d'un réel donné. Nous nous proposons ici d'évoquer les nombreuses applications de cette numération à la combinatoire des mots. En particulier, la numération d'Ostrowski est particulièrement adaptée pour décrire d'un point de vue tant combinatoire qu'arithmétique ou ergodique les suites sturmiennes. Plus généralement, nous indiquerons comment ces résultats peuvent s'étendre à l'étude de certaines suites doubles ainsi qu'à celle de suites de complexité sous-linéaire.

1 Introduction

Une mesure classique du désordre pour une suite à valeurs dans un alphabet fini, consiste à compter le nombre $p(n)$ de facteurs distincts de longueur n . On définit ainsi une fonction p appelée fonction de complexité. Cette fonction permet de caractériser en particulier les suites périodiques (ce sont les suites pour lesquelles il existe un entier n tel que $p(n) \leq n$), ainsi que certaines suites de basse complexité, comme les suites de complexité $n + 1$. Celles-ci sont appelées suites sturmiennes et possèdent la particularité remarquable de pouvoir être étudiées tant d'un point de vue combinatoire qu'arithmétique : elles sont en effet obtenues, entre autres méthodes, comme des codages de rotations irrationnelles sur le cercle unité.

Le but de ce survol est de montrer comment le système de numération d'Ostrowski (basé sur le développement en fraction continue d'un irrationnel donné) permet de décrire naturellement les propriétés combinatoires des suites sturmiennes et de certaines de leurs généralisations unidimensionnelles et bidimensionnelles. Nous verrons qu'il ne s'agit pas tant d'appliquer des propriétés de nature arithmétique pour en déduire des résultats combinatoires, que de faire interagir arithmétique, théorie ergodique et combinatoire des mots. Cette interaction apparaît naturellement lors de l'étude de certaines suites de faible désordre (c'est-à-dire dont la fonction de complexité est sous-linéaire) : ces suites peuvent être engendrées comme limite

de compositions successives d'un nombre fini de substitutions [70]; l'ordre selon lequel ces substitutions se composent permet alors d'introduire des développements en fraction continue généralisés, une motivation étant d'obtenir des algorithmes d'approximation simultanée. Ce travail est en quelques sortes une continuation de [2], où nous étudions les applications en combinatoire des mots du théorème des trois longueurs.

Cet article est organisé comme suit. Nous introduisons le système de numération d'Ostrowski au paragraphe 2 et rappelons notamment quelques résultats classiques de discrédence, obtenus grâce à cette numération. Nous nous intéressons plus particulièrement aux suites sturmiennes au paragraphe 3 et aux relations qu'elles entretiennent naturellement avec la numération d'Ostrowski. Nous tentons d'étendre cette approche au paragraphe 4 selon deux voies de généralisation, duales d'un point de vue arithmétique : nous considérons des codages de rotations sur un tore multidimensionnel d'une part, et d'autre part des codages de l'action de plusieurs rotations sur le tore unidimensionnel. Nous évoquons en particulier la description arithmétique des codages binaires de rotations, des échanges de trois intervalles, des suites d'Arnoux-Rauzy et des suites doubles sturmiennes, description que nous illustrons sur un exemple, à travers l'étude de l'équilibre. Enfin, nous concluons ce survol par un bref rappel au paragraphe 5 des propriétés des suites de complexité sous-linéaire.

2 Approche arithmétique

Le développement en fraction continue d'un réel α permet de construire explicitement les rationnels qui approchent au mieux ce réel ; il s'agit de la propriété de **meilleure approximation** : un rationnel a/b est dit meilleure approximation de α si

$$\forall c/d \neq a/b, 0 < d \leq b, \text{ alors } |d\alpha - c| > |b\alpha - a|.$$

Les meilleures approximations d'un irrationnel α sont exactement les convergents $\frac{p_n}{q_n}$. Pour plus de détails, voir [36, 100].

Considérons maintenant le problème d'approximation non homogène associé : comment approcher modulo 1 un réel β par des points de la forme $k\alpha$? Remarquons que si α est rationnel, il existe alors des réels β tels que la quantité $|q\alpha + p - \beta|$ peut être minorée indépendamment de (p, q) . Il est donc raisonnable de penser que si α est bien approché par les rationnels (c'est-à-dire si les quotients partiels de α sont grands), il existe alors des réels β qui ne pourront pas être bien approchés par des points de la forme $q\alpha + p$, ce qui est effectivement le cas (voir par exemple [36]). Ceci indique que le développement en fraction continue de α joue encore un rôle déterminant pour le problème non homogène. Afin de construire explicitement des solutions à ce problème, il est intéressant d'introduire le système de numération suivant, dans lequel nous allons développer à la fois des entiers et des réels.

2.1 Définitions

Le **système de numération d'Ostrowski** [133] est le système de numération associé à l'échelle de numération $(q_n)_{n \in \mathbb{N}}$, où les q_n sont les dénominateurs des convergents dans le développement en fraction continue d'un réel irrationnel $0 <$

$\alpha < 1$ donné. (Notons que l'on a bien $q_0 = 1$.) On note $(a_n)_{n \in \mathbb{N}}$ la suite des quotients partiels. On a donc

$$\alpha = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}} := [0; a_1, a_2, \dots],$$

$$\frac{p_n}{q_n} = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}} = [0; a_1, \dots, a_n].$$

On a $q_0 = 1$, $q_1 = a_1$, et pour tout entier q_n , $q_{n+1} = a_{n+1}q_n + q_{n-1}$. En appliquant l'algorithme glouton (voir par exemple [77]), on obtient le résultat suivant.

Proposition 1. *Tout nombre entier naturel N s'écrit de manière unique sous la forme*

$$N = \sum_{k=1}^m d_k q_{k-1},$$

où

$$\begin{cases} 0 \leq d_1 \leq a_1 - 1, & 0 \leq d_k \leq a_k, \quad \text{pour } k \geq 2, \\ d_k = 0 \text{ si } d_{k+1} = a_{k+1}. \end{cases} \tag{1}$$

Par exemple, si $\alpha = \frac{1+\sqrt{5}}{2}$ est le nombre d'or, on retrouve la numération de Fibonacci, et la condition dite "markovienne" (1) équivaut au fait que l'on ne trouve pas deux 1 consécutifs dans la suite des coefficients de la représentation. Cette numération est alors appelée **représentation de Zeckendorf** [178].

On peut de manière analogue développer des réels. La base est donnée par la suite $(\theta_n)_{n \in \mathbb{N}}$, où $\forall n \in \mathbb{N}$, $\theta_n := q_n \alpha - p_n$. Rappelons que le signe de θ_n est égal au signe de $(-1)^n$. On obtient donc le résultat suivant (voir par exemple [53, 54, 62, 93, 115, 116, 161, 165]).

Proposition 2. *Tout réel β , avec $-\alpha \leq \beta < 1 - \alpha$, s'écrit de manière unique sous la forme suivante*

$$\beta = \sum_{k=1}^{+\infty} b_k \theta_{k-1}, \tag{2}$$

$$\text{où } 0 \leq b_1 \leq a_1 - 1, \quad 0 \leq b_k \leq a_k, \quad \text{pour } k \geq 2,$$

$$b_k = 0 \text{ si } b_{k+1} = a_{k+1},$$

$b_k \neq a_k$, pour une infinité d'entiers pairs et une infinité d'entiers impairs.

Notons que l'on a bien convergence puisque

$$\forall k, \quad b_k |q_{k-1} \alpha - p_{k-1}| \leq \frac{a_k}{q_k} \leq \frac{1}{q_{k-1}}.$$

La condition sur les coefficients d'indice pair et impair garantit l'unicité de cette écriture dans le cas où $\beta \equiv l\alpha \pmod{1}$. En effet, on a

$$\sum_{k \geq k_0} a_{2k+1} \theta_{2k} = \sum_{k \geq k_0} -\theta_{2k-1} + \theta_{2k+1} = -\theta_{2k_0-1} = \sum_{k \geq k_0+1} a_{2k} \theta_{2k-1} + (a_{2k_0-1}) + a_{2k_0-2}.$$

Notons que cette écriture peut être étendue à \mathbb{R} (voir [157]).

Ces développements nous permettent donc d'approcher arbitrairement près β modulo 1 par des points de la forme $N\alpha$: il suffit de développer β selon (2) ; les entiers N_n avec $N_n = \sum_{k=1}^n b_k q_{k-1}$ fournissent les meilleures approximations.

Supposons maintenant que l'on ne veuille approcher β que d'un côté modulo 1. Il apparaît alors naturel de développer β selon la base $(|\theta_n|)_{n \in \mathbb{N}}$. On obtient alors (voir par exemple [82, 89]) le résultat suivant.

Proposition 3. *Tout réel β , avec $0 \leq \beta < 1$, s'écrit de manière unique sous la forme :*

$$\beta = \sum_{k=1}^{+\infty} c_k |\theta_{k-1}|, \quad (3)$$

où $0 \leq c_k \leq a_k$, et $c_{k+1} = 0$ si $c_k = a_k$,

$c_k \neq a_k$ pour une infinité d'entiers pairs.

Il existe de même un système de numération associé sur les entiers : l'échelle de numération est donnée par $((-1)^n q_n)_{n \in \mathbb{N}}$ (voir par exemple [157]). Il est alors plus délicat d'exprimer les meilleures approximations, à gauche par exemple, de β par des points de la forme $N\alpha$ en raison des signes qui apparaissent : pour un algorithme explicite, voir [23].

2.2 Quelques applications

Ces développements sont particulièrement adaptés pour étudier la **répartition** de la suite $(n\alpha)_{n \in \mathbb{N}}$, c'est-à-dire la manière dont se distribuent les points $(n\alpha)_{n \in \mathbb{N}}$ modulo 1. On sait que la suite $(n\alpha)_{n \in \mathbb{N}}$ est uniformément répartie modulo 1 si α est irrationnel, c'est-à-dire que pour tout I intervalle du tore \mathbb{R}/\mathbb{Z} , la quantité

$$\frac{\text{Card}\{n \leq N, n\alpha \in I\}}{N}$$

tend vers la longueur $|I|$ de I , quand N tend vers l'infini. Plus précisément, on peut mesurer la rapidité de cette convergence vers la répartition uniforme en introduisant la notion de **discrédance** : la discrédance D_N d'une suite $(x_n)_{n \in \mathbb{N}}$ est définie comme la borne supérieure, pour I intervalle fixé du tore \mathbb{R}/\mathbb{Z} , de la quantité

$$\frac{\text{Card}\{n \leq N; x_n \in I\}}{N} - |I|.$$

Pour plus de détails, voir par exemple [111, 57].

De nombreux travaux basés sur le système de numération d'Ostrowski ont été consacrés à l'étude de la discrédance de la suite $(n\alpha)$ (voir par exemple [12, 53, 54, 57, 60, 61, 62, 63, 80, 115, 116, 138, 155, 161, 164], voir aussi [13] pour le cas multidimensionnel). Plus précisément les problèmes de minoration ou de majoration de quantités du type $\text{Card}\{n \leq N, n\alpha \in I\} - N|I|$, pour I intervalle du tore \mathbb{R}/\mathbb{Z} , ont été largement étudiés [60, 61, 62, 99]. Nous y reviendrons au paragraphe 4.5.

L'étude des bornes inférieures ou supérieures des expressions

$$c_\alpha(N) := \sum_{1 \leq k \leq N} (\{k\alpha\} - 1/2)$$

a également donné lieu à de nombreux travaux (voir par exemple [83, 84, 114, 115, 116, 133, 159, 160] et le survol [154]). Le cas non homogène (on considère $c_{\alpha,\beta}(N) := \sum_{1 \leq k \leq N} (\{k\alpha + \beta\} - 1/2)$) a été en traité dans [135, 136] où il est montré que, de manière surprenante, l'on peut perdre l'oscillation qui se produit entre valeurs positives et négatives dans le cas homogène pour des quotients partiels bornés. En effet, si les quotients partiels de α sont bornés par A , il existe une constante C_A telle que pour une infinité d'entiers N , l'on ait $c_\alpha(N) < C_A \log N$ et pour une infinité d'entiers N , l'on ait $c_\alpha(N) > -C_A \log N$. Néanmoins dans cas non homogène, on n'a plus forcément une telle oscillation. Par exemple, $c_{\sqrt{2},1/2}(N) > 0$, pour tout N [135, 136]. Voir le paragraphe 4.5 pour une application combinatoire de ces estimations.

Hardy et Littlewood étudient également dans [83, 84] les prolongements analytiques des séries de Dirichlet $\sum_{k \geq 1} \frac{\{k\alpha\}}{k^s}$, afin de compter le nombre de points d'un réseau dans un triangle (pour plus de détails, voir aussi [14, 15, 172]).

La numération d'Ostrowski est également particulièrement adaptée à l'étude diophantienne de formes linéaires non homogènes, et en particulier, à la minimisation de quantités du type $\|q\alpha + \beta\|$, ou $q\|q\alpha + \beta\|$, où $\|x\|$ désigne la distance de x à l'entier le plus proche entier (voir par exemple [25, 35, 50, 53, 54, 160, 162]). Pour un rappel des différents algorithmes de fractions continues non homogènes correspondant à ce type de problèmes, voir [107].

On obtient de nombreuses propriétés métriques et ergodiques concernant les fractions continues en introduisant la transformation de Gauss $T :]0, 1[\rightarrow]0, 1[$, $T(x) \mapsto \{1/x\}$, si $x \neq 0$, et $T(0) = 0$. On peut de même décrire les développements (2) et (3) comme un produit croisé de la transformation T des fractions continues. Ito donne ainsi dans [89] les transformations définies sur $]0, 1[\times]0, 1[$ qui produisent respectivement les développements (2) et (3); ces transformation produisent des quotients partiels (a_n, b_n) correspondant respectivement au développement en fraction continue de α et au développement d'Ostrowski de β . Ces transformations sont analogues à la transformation suivante :

$$\tilde{T} :]0, 1[^2 \rightarrow]0, 1[^2, \quad \tilde{T}(\alpha, \beta) = (\{\frac{1}{\alpha}\}, \{\frac{\beta}{\alpha}\}). \quad (4)$$

On obtient ainsi une réalisation de leur extension naturelle et une expression des mesures invariantes (voir aussi [93, 165]). Ito et Nakada en déduisent dans [92, 93] des résultats de distribution, comme par exemple le résultat suivant de convergence vers la constante de Lévy : pour presque tout (α, β) ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log N_n = \frac{\pi^2}{12 \log 2},$$

où N_n est la suites des approximations de β par les points $k\alpha$. Une caractérisation de l'extension quadratique $\mathbb{Q}(\alpha)$ (pour α quadratique) est de plus donnée dans [82] en fonction de ces développements (voir aussi [53, 54] et [94]). Ce résultat est étendu dans [90] à des nombres cubiques en utilisant l'algorithme de Jacobi-Perron modifié. Voir également [98, 97] pour une construction de mesures non singulières, non atomiques et quasi-ergodiques associées à ces systèmes de numération.

Sidorov et Vershik étudient dans [157] les propriétés "adiques" et ergodiques des systèmes de numération d'Ostrowski. En particulier, ils établissent une réalisation adique des rotations irrationnelles sur le cercle unité en explicitant un isomorphisme

métrique entre la rotation d'angle α et la transformation adique correspondant respectivement aux développements (2) et (3). Ils en déduisent une nouvelle preuve du théorème d'Erdős sur la singularité (par rapport à la mesure de Lebesgue sur $[0, 1 + \alpha]$) de la distribution de la variable aléatoire $\sum_{n \geq 1} x_n \alpha^n$, où $x_n = 0, 1$ avec une probabilité $1/2$. Liardet décrit de même dans [119] les propriétés de l'odomètre construit sur le système de numération d'Ostrowski (il est continu, donc surjectif et minimal) en relation avec l'étude dynamique des suites α -multiplicatives. Voir aussi les travaux de Coquet [43, 44, 45, 46, 47] et le texte de Liardet qui leur est consacré [118].

Pour une approche algébrique et géométrique de ces questions, voir [6, 8]. On sait qu'il est possible de coder le flot géodésique sur la surface modulaire (le quotient du plan hyperbolique par le groupe modulaire) par les fractions continues (voir par exemple [153, 5]). Arnoux et Fisher introduisent alors le flot scénographique (qui se projette sur le flot géodésique) qui permet d'étendre l'approche modulaire à ce contexte non homogène, c'est-à-dire affine (on travaille ici avec $SA(2, \mathbb{Z}) \backslash SL(2, \mathbb{R})$ au lieu de $SL(2, \mathbb{Z}) \backslash SL(2, \mathbb{R})$) : en d'autres termes, le flot scénographique est au système de numération d'Ostrowski ce que le flot géodésique est aux fractions continues. Le produit croisé (4) des fractions continues associé à la numération d'Ostrowski est ainsi explicité comme une transformation de premier retour sur une section du flot.

Enfin, l'algorithme d'Ostrowski intervient également en informatique théorique pour le problème des arrondis dans le calcul des fonctions élémentaires afin d'établir des tables de logarithmes, de sinus etc... (pour plus de détails, voir [112, 113]).

3 Approche combinatoire

Le but de ce paragraphe est de montrer d'une part comment les fractions continues interviennent de manière naturelle pour exprimer de nombreuses propriétés du langage des facteurs d'une suite sturmiennne. Nous verrons d'autre part que si l'on veut affiner ces propriétés, c'est-à-dire considérer des propriétés métriques ou combinatoires d'une suite sturmiennne donnée, le système de numération d'Ostrowski est particulièrement adapté.

3.1 Suites sturmiennes et fonction de complexité

La fonction de complexité est une notion très utile pour l'étude des suites et des systèmes dynamiques symboliques. Soit $u = (u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans l'alphabet fini \mathcal{A} . On appelle **fonction de complexité** de u , et l'on note p , la fonction (définie sur les entiers) qui compte le nombre de facteurs de u de longueur donnée :

$$p(n) = \text{Card}\{w; w \text{ est facteur de } u \text{ et } |w| = n\}.$$

Pour plus d'informations sur la fonction de complexité, voir par exemple les survols [3, 72].

La fonction de complexité permet de caractériser les suites périodiques. On a en effet le résultat classique suivant (voir [130, 48] et pour une preuve, voir par exemple [16]) : une suite indexée par \mathbb{N} est ultimement périodique si et seulement s'il existe un entier $n \in \mathbb{N}$ non nul tel que $p(n) \leq n$. Il apparaît alors naturel de

s'intéresser aux suites de complexité $n + 1$, c'est-à-dire telles que $p(n) = n + 1$, pour tout n . Les suites de complexité $n + 1$ (indexées par \mathbb{N}) sont appelées **suites sturmiennes**. Les suites sturmiennes sont donc les suites de complexité minimale parmi les suites non ultimement périodiques. L'exemple le plus classique de suite sturmienne est la suite de Fibonacci, point fixe de la substitution σ définie sur $\{0, 1\}$ par $\sigma(0) = 01$ et $\sigma(1) = 0$. (Une **substitution** est un endomorphisme pour la concaténation du monoïde libre \mathcal{A}^* , formé des mots finis sur \mathcal{A} ; on peut étendre la définition d'une substitution à l'ensemble des suites $\mathcal{A}^{\mathbb{N}}$.) Les suites sturmiennes sont donc définies de manière purement combinatoire, mais ce qui est remarquable, c'est qu'elles peuvent être également représentées de manière géométrique (voir [131]) : les suites sturmiennes 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 longueurs α et $1 - \alpha$.

Dans tout ce qui suit R_α désigne la rotation définie sur le tore $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ de dimension 1, par $R_\alpha x = x + \alpha$ modulo 1. S'il n'y a pas d'ambiguïté, nous omettrons de préciser que nous travaillons modulo 1.

Théorème 1 (Morse-Hedlund [131]). *Soit u une suite sturmienne à valeurs dans $\{0, 1\}$. Il existe alors α irrationnel dans $]0, 1[$ et $\rho \in \mathbb{R}$ tels que l'on ait :*
soit

$$\forall k \in \mathbb{N}, (u_k = 0 \iff R_\alpha^k(\rho) = \rho + k\alpha \in [0, 1 - \alpha]),$$

soit

$$\forall k \in \mathbb{N}, (u_k = 0 \iff R_\alpha^k(\rho) = \rho + k\alpha \in]0, 1 - \alpha]).$$

*On appelle **angle** d'une suite sturmienne le réel α qui lui est ainsi associé et **partition**, la partition de \mathbb{T}^1 selon laquelle on code, c'est-à-dire soit $\{[0, 1 - \alpha[, [1 - \alpha, 1[\}$, soit $\{]0, 1 - \alpha],]1 - \alpha, 1\}$.*

Notons que les suites sturmiennes ont de nombreuses autres caractérisations, tant géométriques que combinatoires (voir, par exemple, les survols [16, 124, 29]). On peut par exemple les considérer comme des codages de droites discrètes (voir par exemple [123]) ou comme des codages de trajectoires dans un billard carré.

Les techniques les plus généralement utilisées pour étudier ces problèmes sont d'une part l'étude des graphes des mots et l'induction (nous reviendrons sur ces notions aux paragraphes 3.3 et 3.4 respectivement).

3.2 Suites sturmiennes et fractions continues

Un des intérêts de la représentation géométrique des suites sturmiennes comme codages de rotations est qu'elle fournit une description simple des facteurs en termes d'intervalles du cercle unité. En effet, on peut associer bijectivement les facteurs de taille donnée et les intervalles d'une certaine partition du cercle unité de la manière suivante (voir par exemple [102, 128]). Soit u une suite sturmienne d'angle α et de partition $\{I_0, I_1\}$; il existe donc ρ tel que

$$\forall k \in \mathbb{N}, (u_k = 0 \iff \rho + k\alpha \in I_0).$$

Les facteurs de longueur n de la suite u sont en bijection avec les $n + 1$ intervalles de la partition du cercle unité par les points $-k\alpha$, $0 \leq k \leq n$, par l'application qui à un facteur $w_1 \dots w_n$ associe l'unique intervalle du cercle unité $I_{w_1 \dots w_n}$ défini par

$$I_{w_1 \dots w_n} = \bigcap_{1 \leq i \leq n} R_\alpha^{-i+1} I_{w_i}.$$

(Notons que la connexité des ensembles $I_{w_1 \dots w_n}$ vient de $|I_0|, |I_1| \leq \sup(1 - \alpha, \alpha)$.) On a alors

$$\rho + k\alpha \in I_{w_1 \dots w_n} \text{ si et seulement si } w_1 \dots w_n \text{ apparaît à l'indice } k,$$

c'est-à-dire que

$$w_1 = u_k, w_2 = u_{k+1}, \dots, w_n = u_{k+m-1}.$$

Notons que l'on déduit de cette propriété que deux suites sturmiennes ont le même **langage** (c'est-à-dire même ensemble de facteurs) si et seulement si elles ont le même angle. En d'autres termes, pour toute suite sturmienne u , le système dynamique $(\overline{\mathcal{O}(u)}, T)$ est minimal, T désignant le **décalage unilatéral** : $T((u_n)_{n \in \mathbb{N}}) = (u_{n+1})_{n \in \mathbb{N}}$, et $\overline{\mathcal{O}(u)}$ désignant la clôture de l'orbite de u sous l'action de T dans $\{0, 1\}^{\mathbb{N}}$ muni du produit des topologies discrètes. Rappelons qu'un système dynamique $(\overline{\mathcal{O}(u)}, T)$ engendré par une suite u est dit **minimal** si l'adhérence de l'orbite de toute suite est égale au système tout entier, ce qui revient à dire que toutes les suites ont même ensemble de facteurs. Pour plus de détails, voir [137]. Il n'est pas difficile de voir par un argument de compacité [137] que la minimalité du système est équivalente à l'uniforme récurrence de la suite u (une suite est dite **uniformément récurrente** si les lacunes entre deux occurrences d'un même facteur sont bornées). En effet, soit w un facteur donné de la suite u et $[w]$ l'ensemble des suites de $\overline{\mathcal{O}(u)}$ qui commencent par w ; on peut écrire par minimalité

$$\overline{\mathcal{O}(u)} = \bigcup_{n \geq 0} T^{-n}[w],$$

et par compacité

$$\overline{\mathcal{O}(u)} = \bigcup_{1 \leq i \leq p} T^{-n_i}[w],$$

ce qui implique que pour tout $k \geq 0$, $T^k u \in \bigcup_{1 \leq i \leq p} T^{-n_i}[w]$, et donc w apparaît à lacunes bornées dans u . Ceci implique en particulier qu'une suite sturmienne est uniformément récurrente.

De nombreuses propriétés des langages sturmiens se déduisent aisément de cette correspondance entre facteurs et intervalles, et s'énoncent naturellement en fonction

du développement en fraction continue de l'angle. Nous considérons ainsi au paragraphe 3.3 les fréquences des facteurs de longueur donnée.

Nous allons voir de plus sur deux questions précises (puissances de facteurs et engendrement par substitutions) que si les fractions continues permettent de décrire des propriétés concernant les suites sturmiennes de même angle, dès que l'on veut obtenir quelque chose de précis concernant une suite sturmienne donnée, le système de numération d'Ostrowski apporte les informations nécessaires grâce au développement du point dont on code l'orbite.

3.3 Fréquences et graphe des mots

On déduit de la représentation par intervalles, par uniforme répartition de la suite $(k\alpha)_{k \in \mathbb{N}}$ (rappelons que α est irrationnel), que la fréquence d'apparition d'un facteur $w_1 \cdots w_n$ est égale à la longueur de l'intervalle $I_{w_1 \dots w_n}$. Cette longueur s'exprime en fonction du développement en fraction continue de l'angle α . Rappelons que la **fréquence** $f(w)$ du facteur w est définie comme la limite (si elle existe) du quotient du nombre d'occurrences de ce facteur dans les N premiers termes de la suite, par N . Il s'agit donc d'une notion métrique qui permet de produire une mesure de probabilité borélienne invariante par le décalage sur le système dynamique engendré par la suite.

Notons de plus que les longueurs des intervalles délimités par les points $-k\alpha$, $0 \leq k \leq n$, prennent trois valeurs au plus : il s'agit du **théorème des trois longueurs** (voir [161, 167, 168] ou le survol [2]). Ces longueurs s'expriment en fonction des convergents intermédiaires dans le développement en fraction continue de l'angle α , c'est-à-dire en fonction des approximations de α par des points de Farey. En effet, ces trois longueurs sont exactement les distances des points situés de part et d'autre de 0 (quand on les considère dans \mathbb{R}/\mathbb{Z}) et leur somme, d'où la recherche de la minoration de quantités de la forme $|q\alpha + p|$. Or ces longueurs correspondent aux fréquences des facteurs de longueur n . Ce résultat peut s'obtenir à l'aide de la notion de graphe des mots de Rauzy [18]. On obtient ainsi une preuve purement combinatoire du théorème des trois longueurs.

Le **graphe des mots** Γ_n des facteurs de longueur n d'une suite u à valeurs dans l'alphabet fini \mathcal{A} est un graphe orienté (voir, par exemple, [142]), qui est un sous-graphe du graphe des mots de de Bruijn (voir [52]). Le graphe Γ_n a pour sommets les facteurs de longueur n de la suite, avec une arête de U vers V , s'il existe un mot W de longueur $n-1$ tel que $U = xW$ et $V = Wy$, avec $x, y \in \mathcal{A}$, et tel que xWy soit un facteur de la suite.

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

dit **bispécial** ; par exemple, quand $G_n = D_n$, ce facteur est bispécial.) Une étude

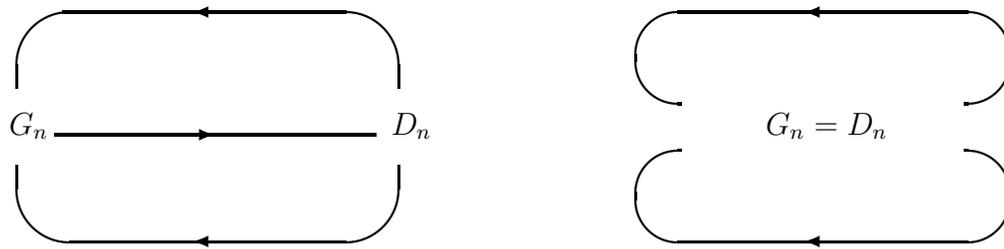


FIG. 1 – Graphes des mots pour les suites sturmiennes

précise des fréquences permet de déduire des résultats ergodiques concernant les nombres de recouvrement. Soit (X, T, μ) un système dynamique métrique. On peut associer à un tel système des invariants pour la notion d'isomorphisme métrique ou topologique, appelés **nombres de recouvrement**, permettant également de définir la notion de **rang** (voir par exemple le survol [71] ainsi que [70, 38]). Rappelons qu'une tour de Rokhlin est une union d'ensembles disjoints de la forme $B, TB, \dots, T^{h-1}B$. Les nombres de recouvrement (voir [39, 67, 69, 71, 101]) mesurent la tour de Rokhlin de plus grande taille et de hauteur arbitrairement grande h , contenue dans X , dont la base B possède certaines propriétés de régularité, comme le fait d'être un intervalle ou un ensemble de petit diamètre, si $X = \mathbb{T}^1$, ou bien d'être un cylindre, si X est un système dynamique symbolique. Combinatoirement, cela revient à considérer la proportion maximale d'une suite que l'on peut recouvrir avec des occurrences sans chevauchement de facteurs arbitrairement grands, ou en termes géométriques, la proportion maximale du cercle unité que l'on peut recouvrir avec un nombre arbitrairement grand d'images successives disjointes sous l'action de la rotation R_α d'un intervalle (voir [38, 39]).

3.4 Induction et substitutions

L'étude de l'évolution du graphe des mots est une méthode puissante qui permet de décrire avec précision les suites sturmiennes. Arnoux et Rauzy [11] ont ainsi prouvé la "*S*-adicité" des systèmes dynamiques sturmiens : l'ensemble des facteurs d'une suite sturmienne est égal à l'ensemble des facteurs d'une suite engendrée comme limite de compositions successives de deux substitutions. On a plus précisément le résultat suivant.

Théorème 2. *Soit α un irrationnel de développement en fraction continue $\alpha = [0; a_1 + 1, a_2, \dots]$. Soient σ_0 et σ_1 définies par $\sigma_0(0) = 01$, $\sigma_0(1) = 1$, et $\sigma_1(1) = 10$, $\sigma_1(0) = 0$. L'ensemble des facteurs des suites sturmiennes d'angle α est égal à celui de la suite (sturmienne)*

$$\lim_{n \rightarrow \infty} \sigma_0^{a_1} \sigma_1^{a_2} \sigma_0^{a_3} \sigma_1^{a_4} \dots \sigma_0^{a_{2n-1}} \sigma_1^{a_{2n}}(0).$$

Ce type d'engendrement est appelé ***S*-adique** selon la terminologie "adique" due à Vershik (voir par exemple [173]). Nous reviendrons sur cette notion au paragraphe 5. Une autre manière de voir ce développement consiste à considérer des règles caténatives, comme les règles standard de Rauzy [145]. Une vaste littérature existe sur ce sujet (voir par exemple le survol [29] et [152]) motivée entre autres choses par la question de la caractérisation des suites sturmiennes points fixes de substitution [49]. Le cas des suites sturmiennes caractéristiques (correspondant au problème d'approximation homogène de 0 par les points $k\alpha$) a été résolu dans [49] (pour d'autres preuves, voir les références de [124]). Pour le cas général non homogène (traité par Yasutomi dans [176] et Komatsu dans [109]), nous avons non seulement besoin de l'information donnée par le développement en fraction continue de l'angle, mais aussi du développement d'Ostrowski du point ρ dont l'orbite est codée. On peut ainsi exprimer algorithmiquement une suite sturmienne donnée, en "itérant" des substitutions élémentaires gouvernées par le développement d'Ostrowski (voir [7] et aussi [95, 176, 103, 104, 105, 106, 107, 108, 109, 110, 132, 87]).

L'argument essentiel de la preuve de [7] (donnant le développement *S*-adique explicite d'une suite sturmienne donnée) repose sur un procédé classique d'**induction**. L'idée en est la suivante. Considérons une rotation R d'angle α sur \mathbb{R}/\mathbb{Z} . Nous allons induire sur le plus grand des deux intervalles $[0, 1 - \alpha[$ et $[1 - \alpha, 1[$ (notons qu'il ne s'agit pas exactement de l'induction de Rauzy [140]). Supposons par exemple $\alpha < 1/2$. Soit donc $I = [0, 1 - \alpha[$. Soit $\rho \in I$. Soit u la suite sturmienne d'angle α définie par

$$\forall k \in \mathbb{N}, (u_k = 0 \iff R^k(\rho) = \rho + k\alpha \in I).$$

Soit $R^{(I)}$ la **rotation induite** sur l'intervalle I , c'est-à-dire **l'application de premier retour** sur I :

$$R^{(I)}(x) : I \rightarrow I, x \mapsto R^{n(x)}(x),$$

avec

$$n(x) := \min\{n \in \mathbb{N}^*, R^{n(x)}(x) \in I\}.$$

(Notons que l'induite d'une rotation est encore une rotation si l'on induit sur un intervalle dit **admissible**, c'est-à-dire de longueur égale à un multiple de α modulo 1.) On voit aisément que

$$\begin{cases} R^{(I)}(x) = x + \alpha & \text{si } x \in [0, 1 - 2\alpha[, \\ R^{(I)}(x) = x + 2\alpha - 1, & \text{si } x \in [1 - 2\alpha, 1 - \alpha[. \end{cases}$$

Soit v la suite sturmienne d'angle $\frac{\alpha}{1-\alpha}$ codant l'orbite de ρ selon la rotation d'angle α modulo $1 - \alpha$:

$$\forall k \in \mathbb{N}, (v_k = 0 \iff \rho + k\alpha \in [0, 1 - 2\alpha[\text{ modulo } (1 - \alpha)).$$

Soit τ_0 la substitution définie sur $\{0, 1\}$ par $\tau_0(0) = 0$, $\tau_0(1) = 01$. On a alors $u = \tau_0(v)$. En effet, supposons que pour un indice k , on ait $v_k = 0$. Alors $\rho + k\alpha \in [0, 1 - 2\alpha[$ et donc $u_k = 0$. Si $v_k = 1$, alors $\rho + k\alpha \in [1 - 2\alpha, 1 - \alpha[$, et $\rho + (k + 1)\alpha \in [1 - \alpha, 1[$, c'est-à-dire $u_k = 0$ et $u_{k+1} = 1$.

Notons que l'angle $\alpha < 1/2$ est transformé en $\frac{\alpha}{1-\alpha}$: on reconnaît (sur $[0, 1/2]$) la version additive de l'algorithme de Gauss qui produit les convergents intermédiaires dans le développement en fraction continue. L'idée consiste maintenant à itérer

ce procédé, afin d'engendrer u par composition successive d'un nombre fini de substitutions. On voit pour cela que la position précise de ρ par rapport aux intervalles successifs d'induction est capitale. Or les bornes des intervalles d'induction sont des points de la forme $-k\alpha$, $k \in \mathbb{N}$. On a donc besoin de savoir comment approcher et situer ρ par rapport aux points de la forme $-k\alpha$, d'où l'intervention naturelle du système de numération d'Ostrowski.

Ce type de problèmes d'engendrement par substitutions peut également se traiter de manière purement combinatoire (voir par exemple [6, 8, 87]) en utilisant le fait que l'une des lettres apparaît toujours de manière isolée dans une suite sturmienne et que l'on peut ainsi recoder de manière naturelle une suite sturmienne en une autre suite, qui est encore sturmienne (c'est une conséquence directe de l'équilibre ; pour une définition, voir paragraphe 4.5).

3.5 Puissances de facteurs et de préfixes

Considérons les puissances des facteurs des suites sturmiennes. Une suite sturmienne a des puissances bornées si et seulement si le développement en fraction continue de l'angle a des quotients partiels bornés [127]. Vandeth donne dans [171] une évaluation explicite de l'**exposant critique**, c'est-à-dire de la puissance fractionnaire maximale qui peut apparaître dans une suite sturmienne d'angle à quotients partiels bornés. Voir aussi [17, 51]. L'exposant critique est directement lié aux valeurs limites de la fonction de récurrence quotient (c'est-à-dire $\limsup \frac{R(n)}{n}$), dont le spectre des valeurs a été étudié par Cassaigne [32]. Dans un même ordre d'idées, Justin et Pirillo calculent dans [96] la longueur $L(m)$ du plus long facteur de u ayant pour période m . Holton et Zamboni introduisent dans [87] l'**exposant critique initial** comme la puissance fractionnaire maximale des préfixes. Soit $m(\alpha)$ le minimum des exposants critiques initiaux de toutes les suites sturmiennes de même angle α ; on a alors $2 \leq m(\alpha) \leq 1 + \frac{1+\sqrt{5}}{2}$. De manière surprenante, pour tout α , il existe des suites sturmiennes d'angle α d'exposant critique initial inférieur ou égal à celui de la suite de Fibonacci.

Holton et Zamboni caractérisent de plus les points dont le codage sturmien de l'orbite donne les cas d'égalité ; cette caractérisation s'énonce à partir du développement d'Ostrowski de ces points. Enfin pour les suites de même angle que la suite de Fibonacci, l'exposant critique initial est supérieur ou égal à trois, sauf sur l'orbite de la suite de Fibonacci, où il vaut $1 + \frac{1+\sqrt{5}}{2}$; de plus, une infinité non dénombrable de suites de la clôture de l'orbite de la suite de Fibonacci ont un exposant critique initial égal à 3.

4 Quelques généralisations des suites sturmiennes

Il existe plusieurs types d'objets combinatoires unidimensionnels et multidimensionnels généralisant les suites sturmiennes de manière naturelle, et permettant de produire des algorithmes d'approximation simultanée. On peut par exemple coder l'orbite d'une rotation irrationnelle R_α sur \mathbb{R}/\mathbb{Z} non plus selon une partition sturmienne (c'est-à-dire de la forme $\{[0, 1 - \alpha[, [1 - \alpha, 1[\}$) mais selon une partition en deux intervalles de longueur quelconque, voire en un nombre fini d'intervalles. C'est ce

que l'on appelle un **codage de rotation**. Sont intimement liés à ce type de suites les échanges de trois intervalles ; la connexion est rendue explicite par induction. En effet, l'induction d'un codage binaire sur un intervalle adéquat est un échange de trois intervalles ; inversement, on peut "exclure" d'un échange de trois intervalles un codage binaire de rotation en "rajoutant" un intervalle (pour plus de détails, voir par exemple [23]).

Une seconde voie de généralisation consiste non plus à coder des orbites d'une rotation irrationnelle mais à introduire un second angle β . On a alors deux approches duales : on code soit une rotation d'angle (α, β) sur le tore bidimensionnel \mathbb{T}^2 , soit une \mathbb{Z}^2 -action par deux rotations irrationnelles R_α et R_β sur le tore unidimensionnel \mathbb{T}^1 . Dans le premier cas, on obtient une suite unidimensionnelle, dans le second, une suite double. Nous allons considérer dans ce paragraphe ces différentes généralisations dans le but, soit de mettre en valeur le lien qu'elles entretiennent encore naturellement avec le système de numération d'Ostrowski, soit d'évoquer les développements en fraction continue généralisés qui permettent de les décrire, via un système de numération d'Ostrowski généralisé. Nous illustrons cette approche par l'étude de l'équilibre pour ces différentes suites.

4.1 Codages binaires de rotations

On appelle **codage binaire de rotation** un codage selon une partition en deux intervalles (semi-ouverts) du cercle unité de l'orbite d'un point sous l'action d'une rotation d'angle irrationnel. En particulier, pour une telle suite $u \in \{0, 1\}^{\mathbb{N}}$, il existe $\rho \in \mathbb{R}/\mathbb{Z}$, $\alpha \notin \mathbb{Q}$, un intervalle I de longueur $\beta \in]0, 1[$ tels que

$$\forall k \in \mathbb{N}, (u_k = 0 \iff R_\alpha^k(\rho) = \rho + k\alpha \in I).$$

Si l'on considère la correspondance bijective entre intervalles et facteurs obtenue dans le cas sturmien, cela nous amène à considérer une partition du cercle unité par des points de l'ensemble $\mathcal{P}_n = \{-k\alpha + \beta, -k\alpha, 0 \leq k \leq n-1\}$. Notons que les ensembles ainsi associés aux facteurs ne sont pas toujours des intervalles mais que la condition $\sup(\beta, 1-\beta) \leq \sup(\alpha, 1-\alpha)$ garantit la connexité (voir par exemple [1, 21]). Dans tout ce qui suit, nous supposons cette hypothèse vérifiée quand nous considérerons des codages binaires de rotations. Si l'on suppose de plus $\beta \notin \mathbb{Z}\alpha + \mathbb{Z}$, on déduit alors de $\text{Card}(\mathcal{P}_n) = 2n$ que la complexité $p(n)$ est égale à $2n$, pour tout n . (Pour plus de détails sur les codages binaires de rotations, voir par exemple [1, 2, 148, 56] et [144, 146] pour le lien avec la discrétion des suites $(k\alpha)_{k \in \mathbb{N}}$).

Les longueurs des intervalles de la partition par les points \mathcal{P}_n correspondant aux facteurs de longueur n s'expriment alors en fonction de la numération d'Ostrowski. Notons que ces longueurs prennent au plus 5 valeurs, ceci est un cas particulier d'un théorème dû à Geelen et Simpson [79]. Un calcul explicite des fréquences est donné dans [23] : l'étude des fréquences des codages binaires permet d'exprimer certains nombres de recouvrement pour les systèmes dynamiques associés ainsi que d'étendre ces résultats à des échanges de trois intervalles. On peut en effet associer classiquement par induction à un codage binaire de rotation irrationnelle un échange de trois intervalles. On en déduit alors dans [23] l'expression de nombres de recouvrement pour des échanges de trois intervalles. On obtient ainsi que tout

échange ergodique de trois intervalles est de spectre simple et ne peut être isomorphe en mesure au système dynamique associé à la suite de Thue-Morse. On obtient de plus des exemples d'échanges de trois intervalles induits par la même rotation, mais non conjugués topologiquement.

Didier introduit dans [55] un développement en fraction continue décrivant un développement S -adique pour le langage des codages binaires de rotations. Un second algorithme est également proposé dans [24], inspiré de l'algorithme étudié dans [74, 75, 76] décrivant les échanges de trois intervalles. Notons que cet algorithme est produit combinatoirement par la donnée des règles caténatives de production des facteurs bispéciaux, ou géométriquement, par observation de l'évolution des intervalles correspondant aux facteurs bispéciaux. Il reste à comparer ces divers algorithmes. En particulier, s'il est clair que l'algorithme de [74, 75, 76] n'est pas obtenu à partir d'un produit croisé de type (4) de la transformation de Gauss analogue aux produits croisés décrits dans [89, 92, 93], il est naturel de se demander si l'on peut le décrire comme une section du flot scénographique décrit dans [8], ce qui permettrait d'unifier d'un point de vue arithmétique suites sturmiennes, codages binaires et échanges de trois intervalles. Un indice dans cette direction est que ces divers algorithmes satisfont une condition de type Lagrange : les nombres quadratiques ont des développements en fraction continue qui sont ultimement périodiques.

4.2 Échanges de trois intervalles

Les échanges d'intervalles forment une classe riche de transformations aux propriétés ergodiques et arithmétiques particulièrement intéressantes. Leur étude arithmétique a été inspirée par Rauzy, qui produit dans [139] un algorithme d'approximation simultanée, grâce à un procédé d'induction (l'induction de Rauzy [140]), aux nombreuses conséquences ergodiques.

Ferenczi, Holton et Zamboni développent dans [74, 75, 76] un algorithme multidimensionnel qui permet d'engendrer les orbites des points de discontinuité et d'explorer les propriétés ergodiques et spectrales des échanges d'intervalles (voir aussi [156] et [120, 121, 122] pour un développement S -adique). En effet, cet algorithme permet entre autre chose de caractériser combinatoirement les suites de complexité $2n + 1$ qui sont des codages naturels des orbites des points de discontinuité sous l'action d'un échange de trois intervalles, d'exprimer une condition nécessaire et suffisante (en fonction des quotients partiels fournis par l'algorithme) pour qu'un réel soit une valeur propre d'un échange de trois intervalles, ou pour que la transformation soit de mélange faible. Enfin, l'algorithme permet de produire de manière surprenante des exemples d'échanges de trois intervalles avec des valeurs propres irrationnelles qui sont isomorphes en mesure à une rotation et qui sont donc de spectre discret.

Cet algorithme satisfait un théorème de Lagrange, c'est-à-dire que l'algorithme est ultimement périodique si et seulement si les deux paramètres développés appartiennent à la même extension quadratique. Plus généralement Boshernitzan et Carroll ont montré dans [28] que si tous les paramètres numériques d'un échange d'intervalles appartiennent à un même corps quadratique de nombres, alors la suite des transformations induites successives (à condition d'induire toujours selon la même règle) est ultimement périodique.

4.3 Suites d'Arnoux-Rauzy

Les suites d'Arnoux-Rauzy généralisent de manière naturelle la propriété combinatoire des suites sturmiennes d'avoir un seul facteur spécial à droite et un seul facteur spécial à gauche de longueur donnée. En leur imposant de plus d'être récurrentes, on obtient ainsi, sur un alphabet à trois lettres, une famille de suites uniformément récurrentes de complexité $2n + 1$. Notons que contrairement au cas sturmien, la complexité ne suffit plus à les caractériser.

La plus remarquable d'entre elles correspond au point fixe commençant par 0 de la substitution de Rauzy : $\sigma(0) = 01$, $\sigma(1) = 02$, $\sigma(2) = 0$; la représentation géométrique du système dynamique engendré par cette suite (également appelée **suite de Tribonacci**) est connue sous le nom de **fractal de Rauzy** [141, 142, 91, 125, 126, 42] : la dynamique est donnée par un échange de morceaux fractals pavant \mathbb{R}^2 , qui peut être alors en fait décrit comme une rotation sur le tore \mathbb{T}^2 . L'étude de cette suite a été motivée par des problèmes de discrédance multidimensionnels; en effet, la construction du fractal de Rauzy est fondée sur le lien qui existe entre le point fixe de la substitution de Rauzy et la répartition dans \mathbb{R}^2 modulo \mathbb{Z}^2 de la suite $(N\eta)_{N \in \mathbb{N}}$, où le vecteur $\eta = (\eta_1, \eta_2)$ est tel que $(1, \eta_1, \eta_2)$ est une base du module des entiers algébriques de $\mathbb{Q}(\zeta)$, ζ étant l'unique racine réelle du polynôme $X^3 + X^2 + X - 1$. Pour une étude détaillée du fractal de Rauzy et du système de numération en base Tribonacci, voir par exemple [125, 126, 42].

Si l'on cherche à décrire ces suites arithmétiquement, il est alors naturel de se demander quel type de représentation géométrique elles admettent. Arnoux et Rauzy ont montré dans [11] qu'elles correspondent à des échanges de six intervalles sur le cercle unité, en construisant un développement S -adique de leur langage à l'aide de l'étude de l'évolution des graphes des mots (similaire à celle des suites sturmiennes). Il semble raisonnable de conjecturer que les propriétés de la suite de Tribonacci sont en un sens génériques pour les suites d'Arnoux-Rauzy, c'est-à-dire plus précisément, qu'elles engendrent des systèmes dynamiques qui sont des codages naturels de rotations sur le tore \mathbb{T}^2 . Or il a été récemment prouvé qu'il n'en est rien [34]. Nous reviendrons sur cette question au paragraphe 4.5.

Néanmoins il est possible, par une approche purement combinatoire, de préciser l'arithmétique de ces suites, à travers un algorithme multidimensionnel de fractions continues défini sur un ensemble de mesure nulle (voir [147, 175, 177, 40]). On peut alors en déduire comme dans le cas sturmien les valeurs des fréquences des facteurs de longueur donnée [175], un calcul explicite des nombres de recouvrement [40], l'expression de la fonction de récurrence [41], une étude de leurs propriétés d'équilibre [34] ou un décompte de tous les facteurs de taille donnée des suites d'Arnoux-Rauzy [129]. Notons que ces suites possèdent une combinatoire proche de celle des suites sturmiennes qui permet de généraliser les propriétés de palindromie [59] et les liens entre palindromes et théorème de Fine et Wilf [37]. Dans le cas où ces suites sont de plus substitutives, Arnoux et Ito explicitent dans [9] la représentation géométrique du système dynamique engendré, généralisant l'approche de Ito et Kamura [91] du fractal de Rauzy, construisant ainsi de manière combinatoire des partitions de Markov pour des automorphismes hyperboliques sur le tore. Voir aussi [90] pour une approche diophantienne. Enfin, voir [78] pour un système d'Ostrowski généralisé décrivant les suites de Kronecker $(n(\alpha, \beta))_{n \in \mathbb{N}}$ dans \mathbb{T}^2 .

4.4 Suites sturmiennes doubles

La question de la définition de la fonction de complexité se pose dès que l'on passe en dimension supérieure. Considérons des suites doubles, c'est-à-dire des fonctions de \mathbb{Z}^2 à valeurs dans un ensemble fini. Une notion naturelle de complexité consiste à compter les facteurs rectangulaires distincts de taille donnée : on définit ainsi la notion de **fonction de complexité rectangulaire**. Cette notion peut sembler peu satisfaisante puisque nous privilégions ainsi une base du réseau \mathbb{Z}^2 . Néanmoins elle est relativement bien adaptée aux suites que nous allons considérer dans ce paragraphe. Pour une définition plus générale de la complexité, voir par exemple [150]. Les questions suivantes sont alors naturelles : quelles fonctions de complexité existent ? Peut-on obtenir une caractérisation géométrique d'une suite à partir de sa fonction de complexité ? Que se passe-t-il en particulier dans le cas périodique ?

Une suite double est dite **périodique** si elle est invariante par translation, c'est-à-dire si elle admet un vecteur de périodicité non nul et à coordonnées entières. Notons que le fait que le réseau des vecteurs de périodicité soit de rang 2 est caractérisé par une complexité rectangulaire bornée. Néanmoins, il n'existe pas de caractérisation par la fonction de complexité rectangulaire des suites périodiques. On peut en effet construire des suites doubles admettant un vecteur périodique non nul, et de complexité arbitrairement grande : soit $(u_n)_{n \in \mathbb{Z}}$ une suite unidimensionnelle de complexité 2^n ; la suite double $(u_{m+n})_{(m,n) \in \mathbb{Z}^2}$ a pour complexité 2^{m+n-1} . Réciproquement, Nivat a conjecturé que si la fonction de complexité rectangulaire $P(m, n)$ d'une suite double est telle qu'il existe (m_0, n_0) tel que $P(m_0, n_0) \leq m_0 n_0$, alors la suite double est périodique.

Notons que les travaux d'Epifanio, Mignosi et Koskas [66] laissent entrevoir une voie d'accès à cette conjecture à travers une version bidimensionnelle du théorème de Fine et Wilf. Sanders et Tijdeman ont prouvé la conjecture dans [151] pour des facteurs de taille $(2, n)$ ou $(n, 2)$: s'il existe n tel que $P(2, n) \leq 2n$ ou $P(n, 2) \leq 2n$, alors la suite est périodique. Une conjecture plus générale est donnée dans [149, 150, 151], en étendant la notion de complexité. En revanche, dès que l'on passe en dimension supérieure ou dès que l'on cherche à énoncer une conjecture utilisant des motifs autres que des rectangles, on peut produire des contre-exemples.

Cassaigne étudie et caractérise dans [33] les suites de complexité $mn + 1$, cas limite, selon la conjecture, entre le cas périodique et le cas non périodique. Ces exemples sont non uniformément récurrents et correspondent en un sens, aux cas dégénérés des suites de complexité $n + 1$ sur \mathbb{Z} , comme la suite

$$\dots 0000001000000 \dots$$

Une suite double est dite **uniformément récurrente** si pour tout entier n , il existe N tel que tout facteur carré de taille N contienne tout facteur carré de taille n . Comment alors construire des suites de basse complexité uniformément récurrentes ?

Une idée naturelle afin de construire une suite qui soit à la fois de basse complexité et non périodique, consiste alors à disposer en ligne des suites sturmiennes [20]. Supposons donc que l'on dispose en ligne des suites sturmiennes de même angle, afin de réduire le "désordre". On peut ainsi définir des suites de la manière suivante : soit $(\alpha, \beta) \in \mathbb{R}^2$, tel que $1, \alpha, \beta$ sont rationnellement indépendants et $\sup(\beta, 1 - \beta) \leq$

$\sup(\alpha, 1 - \alpha)$, et soit $\rho \in \mathbb{R}$; soit U la suite double définie sur $\{0, 1\}$ par

$$\forall(k, k') \in \mathbb{Z}^2, (U(k, k') = 0 \Leftrightarrow k\alpha + k'\beta + \rho \in [0, \alpha]).$$

Une telle suite double est appelée **suite sturmiennne double** d'angle (α, β) . On obtient ainsi des suites uniformément récurrentes de complexité $mn + n$ (le comptage des facteurs de taille donnée se fait aisément comme dans le cas sturmien en associant bijectivement facteurs et intervalles du cercle unité). Réciproquement toute suite double uniformément récurrente et de complexité $mn + n$ admet une telle description géométrique [20]. Notons que l'on obtient en colonne des suites qui sont des codages binaires de rotations de complexité $2n$.

Voyons pourquoi l'on peut considérer cette famille de suites doubles comme une généralisation à deux dimensions des suites sturmiennes. Ce sont les suites de plus basse complexité rectangulaire connues parmi les suites doubles uniformément récurrentes et non périodiques. Ces suites présentent de plus des propriétés de symétrie généralisant la notion de palindromie unidimensionnelle [22] : on obtient ainsi une caractérisation de ces suites, inspirée par la caractérisation des suites sturmiennes par palindromes donnée dans [58]. Enfin, elles sont obtenues par une projection lettre-à-lettre de suites doubles (à valeurs dans un alphabet à trois lettres) codant l'approximation d'un plan : on peut approcher un plan de normale irrationnelle par des faces carrées orientées selon les trois plans de coordonnées ; cette approximation est appelée **plan discret** ; si l'on projette cette surface sur le plan $x + y + z = 0$, selon la direction $(1, 1, 1)$, on obtient un pavage du plan par trois sortes de losanges, à savoir les projections des trois types de faces ; on peut coder cette projection sur \mathbb{Z}^2 en associant aux losanges le type de la face projetée ; on obtient ainsi une suite définie sur un alphabet à trois lettres ; on montre alors que ces suites codent une action de \mathbb{Z}^2 sur le cercle unité [21], ce qui permet de montrer qu'elles sont de complexité $mn + m + n$.

Ces dernières suites (que nous appellerons suites sturmiennes doubles sur trois lettres) sont de plus engendrées par des substitutions bidimensionnelles dont l'itération est gouvernée par l'algorithme de Jacobi-Perron [10]. La preuve repose encore (comme dans le cas sturmien) sur un processus d'induction/exduction. Notons qu'il ne semble pas que l'algorithme de Jacobi-Perron joue ici un rôle privilégié. En effet, l'introduction d'autres algorithmes de fractions continues bidimensionnelles permet de définir de nouvelles substitutions, que l'on peut itérer sans crainte de chevauchements. Néanmoins, pour vérifier que ces substitutions engendrent bien une suite définie sur tout \mathbb{Z}^2 , des problèmes techniques d'ordre combinatoire apparaissent, que l'on ne sait résoudre actuellement que dans le cas de l'algorithme de Jacobi-Perron. L'étude des fréquences des facteurs rectangulaires conduit, de plus, à des considérations diophantiennes de minimisation de formes linéaires. Ceci semble donc corroborer le fait que l'algorithme de Jacobi-Perron ne résout pas toutes les questions concernant ces suites, contrairement au cas unidimensionnel totalement décrit par l'algorithme des fractions continues. Notons que l'on peut définir à travers cette approche combinatoire de l'algorithme de Jacobi-Perron un développement de type Ostrowski pour les réels permettant d'engendrer par compositions de substitutions une suite sturmiennne double sur trois lettres donnée.

4.5 Équilibre et ensembles à restes bornés

Le but de ce paragraphe est d'illustrer sur un exemple précis (la question de l'équilibre) comment arithmétique, théorie ergodique et combinatoire interagissent. Les suites sturmiennes sont exactement les suites équilibrées indicées par \mathbb{N} à valeurs dans un alphabet à deux lettres et non ultimement périodiques [48]. Une suite à valeurs dans $\{0, 1\}$ est dite **équilibrée** si pour tout couple de facteurs (v, w) de même longueur, on a

$$||v|_0 - |w|_0| \leq 1,$$

où $|v|_0$ désigne le nombre d'occurrences de la lettre 0 dans le mot v . Supposons que l'on tente de généraliser cette condition d'équilibre. On peut soit considérer un alphabet de taille supérieure, soit imposer aux différences des nombres d'occurrence des lettres d'être bornées non plus par 1 mais par une constante uniforme. Dans le premier cas (équilibre sur un alphabet fini), on constate que les suites équilibrées sur un alphabet fini de taille quelconque et non périodiques dérivent des suites sturmiennes en remplaçant les lettres cycliquement par d'autres lettres (voir [88]). On obtient alors des suites pour lesquelles deux lettres au moins ont même fréquence. Cette étude est reliée au problème du recouvrement de \mathbb{N}^* par deux ensembles disjoints de la forme $\{[\alpha_1 n + \beta_1], n \in \mathbb{N}\}$ et $\{[\alpha_2 n + \beta_2], n \in \mathbb{N}\}$ (voir [81]). Plus généralement si l'on veut recouvrir \mathbb{N}^* par n ensembles disjoints $\{[\alpha_i n + \beta_i], n \in \mathbb{N}\}$, avec $n \geq 3$, $\alpha_i > 1$, et $\alpha_i \neq \alpha_j$, si $i \neq j$, alors les nombres α_i doivent être nécessairement rationnels [81]. Fraenkel a conjecturé que

$$\{\alpha_1, \dots, \alpha_n\} = \left\{ \frac{2^n - 1}{2^k}, 0 \leq k < n \right\}.$$

En termes combinatoires, cela revient à considérer des suites équilibrées à fréquences rationnelles (voir [4, 170]). Pour plus de références sur la conjecture de Fraenkel et ses relations avec les suites équilibrées, voir le survol [169].

Considérons maintenant les suites C -équilibrées, avec C constante strictement positive. Une suite u à valeurs dans l'alphabet \mathcal{A} est dite C -équilibrée si pour tout couple de facteurs (v, w) de même longueur et si pour toute lettre $a \in \mathcal{A}$, on a

$$||v|_a - |w|_a| \leq C,$$

où $|v|_a$ désigne le nombre d'occurrences de la lettre a dans le mot v . Il n'est pas difficile de voir que la condition de C -équilibre implique non seulement l'existence des fréquences $f(a)$ pour les lettres a de \mathcal{A} , mais aussi qu'il existe une constante C' telle que :

$$\forall a \in \mathcal{A}, \forall N \in \mathbb{N}, |\text{Card}\{n \leq N, u_n = a\} - Nf(a)| \leq C'. \quad (5)$$

Cette propriété combinatoire se traduit en termes arithmétiques de la manière suivante pour un codage binaire de rotation, par exemple. Avec les notations précédentes, il existe un intervalle I de \mathbb{T}^1 tel que la suite u satisfait :

$$\forall k \in \mathbb{N}, (u_k = 0 \iff \rho + k\alpha \in I).$$

La condition (5) se traduit par :

$$\forall N \in \mathbb{N}, |\text{Card}\{n \leq N, \rho + n\alpha \in I\} - N|I|| \leq C'.$$

Un tel intervalle I est appelé **ensemble à restes bornés**. Or d'après un résultat de Kesten [99], les intervalles à restes bornés (par rapport à la rotation d'angle α) sont les intervalles de longueur égale à un multiple entier α , modulo 1. (Notons que la preuve de Kesten utilise la numération d'Ostrowski.)

Un codage binaire de rotation (avec $1, \alpha, |I|$ rationnellement indépendants) n'est donc jamais C -équilibré, quelle que soit la constante C . Ceci est donc vrai en particulier pour les suites sturmiennes doubles (dont les suites en colonne sont des codages binaires de rotations), en considérant le C -équilibre sur les facteurs rectangulaires de même taille. Notons de plus que les suites doubles 1-équilibrées sont totalement périodiques [19] (une suite double est dite totalement périodique si le réseau de ses vecteurs de périodicité est de rang 2). Enfin, on peut donner une mesure quantitative de leur déséquilibre en introduisant la notion suivante. On définit une **fonction d'équilibre** $E(m, n)$ associée à une suite u indicée par \mathbb{Z}^2 à valeurs dans l'alphabet \mathcal{A} de la manière suivante :

$$E(m, n) := \max_{a \in \mathcal{A}} \max_{v, w \text{ de taille } (m, n)} ||v|_a - |w|_a|.$$

On obtient alors (voir [19]) pour une suite sturmienne double de paramètres (α, β) que $E(m, n) = o(\sup(m, n))$ et que si α ou β est à quotients partiels bornés, alors $E(m, n) = O(\log(\sup(m, n)))$. La preuve repose sur l'estimation de sommes du type $\sum_{j=1}^N (\{j\alpha + \rho\} - 1/2)$ (voir par exemple [30] pour le cas homogène et [135, 136], sinon), estimations produites grâce à l'introduction du système d'Ostrowski.

On peut également se poser la question de l'équilibre pour les suites d'Arnoux-Rauzy. Un exemple de suite d'Arnoux-Rauzy est donné dans [74] qui n'est jamais C -équilibrée quelle que soit la constante C . Cet exemple a été construit grâce à l'algorithme de fraction continue de [147, 177] : il correspond à des quotients partiels qui tendent très rapidement vers l'infini, et n'est pas linéairement récurrent au sens de [65, 64] (nous reviendrons sur cette notion au paragraphe suivant). On en déduit qu'il existe des suites d'Arnoux-Rauzy qui ne sont pas codages naturels de rotation sur le tore bidimensionnel \mathbb{T}^2 , contrairement à ce qui a été longtemps cru. L'idée de la preuve de ce résultat est similaire à ce que nous avons utilisé pour les codages binaires. La négation de la condition de C -équilibre sur une lettre implique que le cylindre correspondant n'est pas à restes bornés. Or d'après [143, 68], si la transformation induite d'une rotation sur le tore par rapport à un ensemble A (appelé admissible) est encore une rotation, alors A est à restes bornés : à une dimension, les intervalles admissibles sont les intervalles de longueur multiple de α modulo 1. Or l'induite d'un codage naturel de rotation sur un cylindre est encore un codage naturel de rotation, d'où le résultat. Voir également [117] pour une étude des ensembles à restes bornés pour une rotation irrationnelle sur le tore.

5 Arithmétique des suites de complexité sous-linéaire

Nous allons voir dans ce paragraphe que nous pouvons étendre cette approche au cas des suites de complexité sous-linéaire, c'est-à-dire telles qu'il existe $C \in \mathbb{N}$, tel que pour tout n , $p(n) \leq Cn$. De nombreuses propriétés tant combinatoires, ergodiques, qu'arithmétiques peuvent se déduire de cette simple indication sur l'ordre de croissance

de la fonction de complexité. Mais ce qui est particulièrement intéressant ici, c'est que l'on peut les engendrer à l'aide d'un nombre fini de substitutions. L'intérêt de ce mode d'engendrement réside dans le fait qu'on peut lui associer naturellement dans de nombreux cas un développement en fraction continue généralisé.

Le théorème suivant est un premier pas dans cette direction [31].

Théorème 3. *Si la complexité $p(n)$ d'une suite est sous-linéaire, alors $p(n+1) - p(n)$ est borné.*

Notons que ce résultat est faux pour la complexité quadratique : un exemple de suite de complexité quadratique dont les différences secondes $p(n+2) + p(n) - 2p(n+1)$ ne sont pas bornées est donné dans [70].

5.1 La conjecture S -adique

Ferenczi déduit du théorème précédent le résultat suivant [70] : les systèmes minimaux de complexité sous-linéaire sont engendrés par un nombre fini de substitutions. Une suite engendrée par l'itération d'un nombre fini de substitutions est appelée **suite S -adique**. On a plus précisément.

Théorème 4. *Soit u une suite uniformément récurrente de complexité sous-linéaire définie sur l'alphabet \mathcal{A} ; il existe alors un nombre fini c de substitutions σ_i , $1 \leq i \leq c$, définies sur un alphabet \mathcal{B} , une projection φ de \mathcal{B} dans \mathcal{A} et une suite infinie $(i_n)_{n \geq 1}$ à valeurs dans $\{1, \dots, c\}$ tels que d'une part*

$$\lim_{n \rightarrow +\infty} \inf_{b \in \mathcal{B}} |\sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_n}(b)| = +\infty$$

et d'autre part il existe une lettre b de \mathcal{B} pour laquelle tout facteur de la suite u est facteur de $\varphi(\sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_n}(b))$, pour un certain entier n .

Le nombre c de ces substitutions peut être de plus majoré explicitement dans le cas où : $\forall n, p(n+1) - p(n) \leq 2$.

On a vu que l'on connaît explicitement un développement S -adique pour les systèmes dynamiques engendrés par les suites sturmiennes, par les suites d'Arnoux-Rauzy [11, 147, 177], pour les systèmes engendrés par certains codages binaires de rotations [55] et pour les systèmes engendrés par certains échanges de trois intervalles [139, 121, 122, 74, 75, 76].

La réciproque du théorème 4 est fautive. Il suffit en effet de considérer une substitution de complexité quadratique pour construire un contre-exemple, comme le point fixe commençant par 0 de la substitution $0 \mapsto 01, 1 \mapsto 12, 2 \mapsto 2$ [134]. Il reste à définir une notion plus forte de S -adité qui permettrait de caractériser les suites de complexité sous-linéaire ; c'est ce que l'on entend par **conjecture S -adique**. Considérons donc une suite u engendrée par l'itération d'un nombre fini de substitutions ; quelles restrictions faut-il imposer aux substitutions et à leur composition pour que la suite u soit de complexité sous-linéaire ?

Durand donne une condition suffisante dans [64] pour qu'une suite S -adique soit de complexité sous-linéaire. Cette condition est trop forte puisque les suites obtenues sont non seulement de complexité sous-linéaire mais aussi linéairement récurrentes au sens de [65]. Soit u une suite récurrente donnée et w un facteur. On appelle **mot de retour** sur w tout mot v tel que vw est facteur de la suite u , w

est préfixe de vw et w a exactement deux occurrences dans vw . Une suite est dite **linéairement récurrente** s'il existe une constante $C > 0$ telle que pour tout facteur w , la longueur de tout mot de retour v de w satisfait $|v| \leq C|w|$. Notons qu'une telle suite est de complexité sous-linéaire [65]. Durand montre dans [64] qu'une suite est linéairement récurrente si et seulement si elle est obtenue par un développement S -adique primitif à quotients partiels bornés (chaque substitution revient avec des plages bornées). En particulier, une suite sturmienne est linéairement récurrente si et seulement si les coefficients dans le développement de son angle en fraction continue sont à quotients partiels bornés, ce qui montre bien que les suites S -adiques ne sont pas nécessairement linéairement récurrentes.

5.2 Quelques propriétés ergodiques

La complexité sous-linéaire implique de plus les propriétés ergodiques suivantes dans le cas des suites minimales. Soit u une suite primitive de complexité sous-linéaire. Une première conséquence est que le décalage unilatère sur $\overline{\mathcal{O}}(u)$ peut être rendu bijectif sauf sur un nombre fini d'orbites (voir par exemple [73]). La sous-linéarité est de plus invariante par isomorphisme topologique. Boshernitzan a montré que l'on obtient une majoration explicite (en fonction de l'ordre de croissance de la fonction de complexité) du nombre n de mesures ergodiques [26, 27]. En particulier les conditions suivantes impliquent l'unique ergodicité [26, 27] : $\liminf_{n \rightarrow +\infty} \frac{p(n)}{n} < 2$, ou $\limsup_{n \rightarrow +\infty} \frac{p(n)}{n} < 3$. Enfin, Ferenczi a obtenu une majoration explicite du rang dans [70], ainsi que l'absence de mélange fort.

Remerciements Je remercie J.-P. Allouche et N. Mercier pour une relecture attentive de ce travail, ainsi que P. Liardet pour de nombreuses références bibliographiques.

Références

- [1] P. ALESSANDRI *Codages de rotations et basses complexités*, Université Aix-Marseille II, Thèse, 1996.
- [2] P. ALESSANDRI, V. BERTHÉ *Three distance theorems and combinatorics on words*, Enseign. Math. **44** (1998), 103–132.
- [3] J.-P. ALLOUCHE *Sur la complexité des suites infinies*, Bull. Belg. Math. Soc. **1** (1994), 133–143.
- [4] E. ALTMAN, B. GAUJAL, A. HORDIJK *Balanced sequences and optimal routing*, to appear in J. ACM.
- [5] P. ARNOUX *Le codage du flot géodésique sur la surface modulaire*, Enseign. Math. **40** (1994), 29–48.
- [6] P. ARNOUX *Recoding Sturmian sequences on a subshift of finite type. Chaos from order, a worked out example*, Proceedings Fiesta Summer School 98, Chili.
- [7] P. ARNOUX, S. FERENCZI, P. HUBERT *Trajectories of rotations*, Acta Arith. **87** (1999), 209–217.

- [8] P. ARNOUX, A. FISHER *The scenery flow for geometric structures on the torus : the linear setting*, prépublication IML, 2000–08.
- [9] P. ARNOUX, S. ITO *Pisot substitutions and Rauzy fractals*, Bull. Belg. Math. Soc. Simon Stevin, à paraître.
- [10] P. ARNOUX, V. BERTHÉ, S. ITO *Discrete planes, \mathbb{Z}^2 -action, Jacobi Perron algorithm and substitutions*, prépublication 2001.
- [11] P. ARNOUX, G. RAUZY *Représentation géométrique de suites de complexité $2n + 1$* , Bull. Soc. math. France **119** (1991), 199–215.
- [12] C. BAXA, J. SCHOISSENGEIER *Minimum and maximum order of magnitude of the discrepancy of $(n\alpha)$* , Acta Arith. **68** (1994), 281–290.
- [13] J. BECK *Probabilistic Diophantine approximation I. Kronecker sequences*, Ann. of Math. (2) **140** (1994), 449–502.
- [14] H. BEHNKE *Über die Verteilung von Irrationalitäten*, Abh. Math. Sem. Hamburg **1** (1922), 252–267.
- [15] H. BEHNKE *Zur Theorie der Diophantischen Approximationen*, Abh. Math. Sem. Hamburg **3** (1924), 261–318.
- [16] J. BERSTEL *Recent results in Sturmian words*, Developments in Language Theory II (Dassow, Rozenberg, Salomaa, eds.) World Scientific 1996, 13–24.
- [17] J. BERSTEL *On the index of Sturmian words*, Jewels are forever, Springer, Berlin (1999), 287–294.
- [18] V. BERTHÉ *Fréquences des facteurs des suites sturmiennes*, Theoret. Comput. Sci. **165** (1996), 295–309.
- [19] V. BERTHÉ, R. TIJDEMAN *Balance properties of multi-dimensional words*, Theoret. Comput. Sci., à paraître.
- [20] V. BERTHÉ, L. VUILLON *Suites doubles de basse complexité*, J. Théor. Nombres Bordeaux **12** (2000), 179–208.
- [21] V. BERTHÉ, L. VUILLON *Tilings and rotations on the torus : a two-dimensional generalization of Sturmian sequences*, Discrete Math. **223** (2000), 27–53.
- [22] V. BERTHÉ, L. VUILLON *Palindromes and two-dimensional Sturmian sequences*, J. Autom. Lang. Comp., à paraître.
- [23] V. BERTHÉ, N. CHEKHOVA, S. FERENCZI *Covering numbers : arithmetics and dynamics for rotations and interval exchanges*, J. Anal. Math. **79** (1999), 1–31.
- [24] V. BERTHÉ, C. HOLTON, L. ZAMBONI *The structure of binary codings of rotations*, travail en cours.
- [25] J. M. BORWEIN, P. B. BORWEIN *On the generating function of the integer part : $[n\alpha + \gamma]$* , J. Number Theory **43** (1993), 293–318.

- [26] M. BOSHERNITZAN *A unique ergodicity of minimal symbolic flows with linear block growth*, J. Anal. Math. **44** (1984), 77–96.
- [27] M. BOSHERNITZAN *A condition for unique ergodicity of minimal symbolic flows*, Ergodic Theory Dynam. Sys. **12** (1992), 425–428.
- [28] M. BOSHERNITZAN, C. R. CARROLL *An extension of Lagrange's theorem to interval exchange transformations over quadratic fields*, J. Anal. Math. **72** (1997), 21–44.
- [29] T. C. BROWN *Descriptions of the characteristic sequence of an irrational*, Canad. Math. Bull. **36** (1993), 15–21.
- [30] T. C. BROWN, P. J.-S. SHIUE *Sums of fractional parts of integer multiples of an irrational*, J. Number Theory **50** (1995), 181–192.
- [31] J. CASSAIGNE *Special factors of sequences with linear subword complexity*, Developments in Language Theory II (DLT'95), Magdeburg (Allemagne), World Scientific (1996), 25–34.
- [32] J. CASSAIGNE *Limit values of the recurrence quotient of Sturmian sequences*, Theoret. Comput. Sci. **218** (1999), 3–12.
- [33] J. CASSAIGNE *Two dimensional sequences with complexity $mn + 1$* , J. Autom. Lang. Comb. **4** (1999), 153–170.
- [34] J. CASSAIGNE, S. FERENCZI, L. ZAMBONI *Imbalances in Arnoux-Rauzy sequences*, Ann. Inst. Fourier **50** (2000), 1265–1276.
- [35] J. W. S. CASSELS *Über $\lim_{x \rightarrow +\infty} x|\theta x + \alpha - y|$* , Math. Annalen **127** (1954), 288–304.
- [36] J. W. S. CASSELS *An introduction to Diophantine approximation*, Cambridge, Cambridge University Press, 1957.
- [37] M. G. CASTELLI, F. MIGNOSI, A. RESTIVO *Fine and Wilf's theorem for three periods and a generalization of Sturmian words*, Theoret. Comput. Sci. **218** (1999), 83–94.
- [38] N. CHEKHOVA *Nombres de recouvrement*, Thèse, Université Aix-Marseille II, 1997.
- [39] N. CHEKHOVA *Covering numbers of rotations*, Theoret. Comput. Sci. **230** (2000), 97–116.
- [40] N. CHEKHOVA *Algorithme d'approximation et propriétés ergodiques des suites d'Arnoux-Rauzy*, prépublication 1999.
- [41] N. CHEKHOVA *Fonctions de récurrence des suites d'Arnoux-Rauzy et réponse à une question d'Hedlund et Morse*, prépublication 1999.
- [42] N. CHEKHOVA, P. HUBERT, A. MESSAOUDI *Propriétés combinatoires, ergodiques et arithmétiques de la substitution de Tribonacci*, J. Théor. Nombres Bordeaux, à paraître.
- [43] J. COQUET *Sur certaines suites uniformément équiréparties modulo 1*, Acta Arith. **36** (1980), 157–162.

- [44] J. COQUET *Sur certaines suites uniformément équiréparties modulo 1 (II)*, Bull. Soc. Roy. Sci. Liège **48** (1979), 426–431.
- [45] J. COQUET *Répartition de la somme des chiffres associée à une fraction continue*, Bull. Soc. Roy. Sci. Liège **51** (1982), 161–165.
- [46] J. COQUET, P. TOFFIN *Représentations des entiers naturels et indépendance statistique*, Bull. Sci. Math. (2) **105** (1981), 289–298.
- [47] J. COQUET, G. RHIN, P. TOFFIN *Représentations des entiers naturels et indépendance statistique 2*, Ann. Inst. Fourier (Grenoble) **31** (1981), 1–15.
- [48] E. M. COVEN, G. A. HEDLUND *Sequences with minimal block growth*, Math. Systems Theory **7** (1973), 138–153.
- [49] D. CRISP, W. MORAN, A. POLLINGTON, P. SHIUE *Substitution invariant cutting sequences*, J. Théor. Nombres Bordeaux **5** (1993), 123–137.
- [50] T. W. CUSICK, A. M. ROCKETT, P. SZÜSZ *On inhomogeneous Diophantine approximation*, J. Number Theory **48** (1994), 259–283.
- [51] D. DAMANIK, R. KILLIP, D. LENZ *Uniform spectral properties of one-dimensional quasicrystals, III. α -continuity*, Comm. Math. Phys. **212** (2000), 191–204.
- [52] N. G. de BRUIJN *A combinatorial problem*, Rominklijke Netherland Academic Van Wetenschepen Proc. **49** Part. 20 (1946), 758–764.
- [53] R. DESCOMBES *Sur la répartition des sommets d'une ligne polygonale régulière non fermée*, Ann. Sci. École Norm. Sup. **75** (1956), 284–355.
- [54] R. DESCOMBES *Sur un problème d'approximation diophantienne I, II*, C. R. Acad. Sci. Paris **242** (1956), 1669–1772.
- [55] G. DIDIER *Codages de rotations et fractions continues*, J. Number Theory **71** (1998), 275–306.
- [56] G. DIDIER *Combinatoire des codages de rotations*, Acta Arith. **85** (1998), 157–177.
- [57] M. DRMOTA, R. F. TICHY *Sequences, discrepancies and applications*, Lecture Notes Math. **1651**, Springer Verlag (1996–1997).
- [58] X. DROUBAY, G. PIRILLO *Palindromes and Sturmian words*, Theoret. Comput. Sci. **223** (1999), 73–85.
- [59] X. DROUBAY, J. JUSTIN, G. PIRILLO *Episturmian words and some constructions of de Luca and Rauzy*, Theoret. Comput. Sci., à paraître.
- [60] Y. DUPAIN *Intervalles à restes majorés pour la suite $\{n\alpha\}$* , Acta Math. Acad. Sci. Hung. **29** (1977), 289–303.
- [61] Y. DUPAIN *Intervalles à restes majorés pour la suite $\{n\alpha\}$ II*, Bull. Soc. Math. France **106** (1978), 153–159.
- [62] Y. DUPAIN, V. T. SÓS *On the one-sided boundedness of the discrepancy-function of the sequence $n\alpha$* , Acta Arith. **27** (1980), 363–374.

- [63] Y. DUPAIN, V. T. SÓS *On the discrepancy of $(n\alpha)$ sequences*, Coll. Math. Soc. János Bolyai **34**, Topics in classical number theory, Vol. I, II (Budapest, 1981), 355–387, North-Holland, Amsterdam-New York, 1984. q
- [64] F. DURAND *Linearly recurrent subshifts have a finite number of non-periodic subshift factors*, Ergodic Theory Dynam. Systems **20** (2000), 1061–1078.
- [65] F. DURAND, B. HOST, C. SKAU *Substitutions, Bratelli diagrams and dimension groups*, Ergodic Theory Dynam. Sys. **19** (1999), 953–993.
- [66] C. EPIFANIO, P. MIGNOSI, M. KOSKAS *On a conjecture on bidimensional words*, prépublication, 1999.
- [67] S. FERENCZI *Systèmes localement de rang un*, Ann. Inst. Henri Poincaré **20** (1984), 35–51.
- [68] S. FERENCZI *Bounded remainder sets*, Acta Arith. **61** (1992), 319–326.
- [69] S. FERENCZI *Tiling and local rank properties of the Morse sequence*, Theoret. Comput. Sci. **129** (1994), 369–383.
- [70] S. FERENCZI *Rank and symbolic complexity*, Ergodic Theory Dynam. Systems **16** (1996), 663–682.
- [71] S. FERENCZI *Systems of finite rank*, Colloq. Math. **73** (1997), 35–65.
- [72] S. FERENCZI *Complexity of sequences and dynamical systems*, Discrete Math. **206** (1999), 145–154.
- [73] S. FERENCZI *Substitutions and symbolic dynamical systems*, Introduction to finite automata and substitution dynamical systems, prépublication IML, 2000–43.
- [74] S. FERENCZI, C. HOLTON, L. ZAMBONI *The structure of three-interval exchange transformations I : an arithmetic study*, Ann. Inst. Fourier, à paraître.
- [75] S. FERENCZI, C. HOLTON, L. ZAMBONI *The structure of three-interval exchange transformations II : a combinatorial study*, travail en cours.
- [76] S. FERENCZI, C. HOLTON, L. ZAMBONI *The structure of three-interval exchange transformations III : ergodic and spectral properties*, travail en cours.
- [77] A. S. FRAENKEL *Systems of numeration*, Amer. Math. Monthly **92** (1985), 105–114.
- [78] T. FUJITA, S. ITO, S. NINOMYA *Symbolical and geometrical characterizations of Kronecker sequences by using the accelerated Brun algorithm*, J. Math. Sci. Univ. Tokyo **7** (2000), 163–193.
- [79] A. S. GEELLEN, R. J. SIMPSON *A two dimensional Steinhaus theorem*, Australas. J. Combin. **8** (1993), 169–197.
- [80] A. GILLET *Sur la répartition modulo 1*, Thèse, Université Aix-Marseille II, 1968.
- [81] R. L. GRAHAM *Covering the positive integers by q disjoint sets of the form $\{[n\alpha + \beta]\} : n = 1, 2, \dots\}$* , J. Combin. Theory Ser. A **15** (1973), 354–358.

- [82] Y. HARA, S. ITO *On real quadratic fields and periodic expansions*, Tokyo J. Math. **12** (1989), 357–370.
- [83] G. H. HARDY, J. LITTLEWOOD *The lattice points of a right-angled triangle*, Proc. London Math. Soc. **20** (1922), 15–36.
- [84] G. H. HARDY, J. LITTLEWOOD *Some problems of Diophantine approximation; the lattice points of a right-angled triangle*, Abh. Math. Sem. Univ. Hamburg **1** (1922), 212–249.
- [85] G. H. HARDY, E. M. WRIGHT *An introduction to the theory of numbers*, Oxford Science Publications (1979).
- [86] E. HECKE *Über analytische Funktionen und die Verteilung von Zahlen mod. Eins*, Abh. Math. Sem. Hamburg **1** (1922), 54–76.
- [87] C. HOLTON, L. ZAMBONI *Initial powers of Sturmian words*, travail en cours.
- [88] P. HUBERT *Well balanced sequences*, Theoret. Comput. Sci. **242** (2000), 91–108.
- [89] S. ITO *Some skew product transformations associated with continued fractions and their invariant measures*, Tokyo J. Math. **9** (1986), 115–133.
- [90] S. ITO *On periodic expansions of cubic numbers and Rauzy fractals*, Dynamical systems, From Crystal to Chaos, Proceedings of the conference in honor of Gérard Rauzy on his 60th birthday, (J.-M. Gambaudo, P. Hubert, P. Tisseur, S. Vaienti, ed.) pp. 144–164, World Scientific (2000).
- [91] S. ITO, M. KIMURA *On Rauzy fractal*, Japan J. Indust. Appl. Math **8** (1991), 461–486.
- [92] S. ITO, H. NAKADA *On natural extensions of transformations related to Diophantine approximations*, Tokyo J. Math. **9** (1986), 115–133.
- [93] S. ITO, H. NAKADA *Approximations of real numbers by the sequence $n\alpha$ and their metrical theory*, Acta Math. Hung. **52** (1988), 91–100.
- [94] S. ITO, H. TACHII *A Diophantine algorithm and a reduction theory of ternary forms*, Tokyo J. Math. **16** (1993), 261–289.
- [95] S. ITO, S.-i. YASUTOMI *On continued fractions, substitutions, and characteristic sequences $[nx + y] - [(n - 1)x + y]$* , Japan J. Math. **16** (1990), 287–306.
- [96] J. JUSTIN, G. PIRILLO *Fractional powers in Sturmian words*, Theoret. Comput. Sci., à paraître.
- [97] M. KEANE *Irrational rotations and quasi-ergodic measures*, Publications des séminaires de Mathématiques de l’université de Rennes (1970), 17–26.
- [98] M. KEANE *Sur les mesures quasi-ergodiques des translations irrationnelles*, C. R. Acad. Sci. Paris **272** (1971), 54–55.
- [99] H. KESTEN *On a conjecture of Erdős and Szűsz related to uniform distribution mod 1*, Acta Arith. **12** (1966), 193–212.

- [100] A. Y. KHINTCHINE *Continued fractions*, P. Noordhoff, Ltd., Groningen 1963.
- [101] J. L. KING *Joining-rank and the structure of finite-rank mixing transformations*, J. Anal. Math. **51** (1988), 182–227.
- [102] D. E. KNUTH *Sequences with precisely $k + 1$ k -blocks*, solution du problème E2307, Amer. Math. Monthly **79** (1972), 773–774.
- [103] T. KOMATSU *On the characteristic word of the inhomogeneous Beatty sequence*, Bull. Austral. Math. Soc. **51** (1995), 337–351.
- [104] T. KOMATSU *The fractional part of $n\theta + \Phi$ and Beatty sequences*, J. Théor. Nombres Bordeaux **7** (1995), 387–406.
- [105] T. KOMATSU *A certain power series and the inhomogeneous continued fraction expansions*, J. Number Theory **59** (1996), 291–312.
- [106] T. KOMATSU *A certain power series associated with a Beatty sequence*, Acta Arith. **76** (1996), 109–129.
- [107] T. KOMATSU *On inhomogeneous continued fraction expansions and inhomogeneous Diophantine approximation*, J. Number Theory **62** (1997), 192–212.
- [108] T. KOMATSU *On inhomogeneous Diophantine approximation and the Nishioka-Shiokawa-Tamura algorithm*, Acta Arith. **86** (1998), 305–324.
- [109] T. KOMATSU *Substitution invariant inhomogeneous Beatty sequences*, Tokyo J. Math. **22** (1999), 235–243.
- [110] T. KOMATSU, A. J. van der POORTEN *Substitution invariant Beatty sequences*, Japan. J. Math. (N.S.) **22** (1996), 349–354.
- [111] L. KUIPERS, H. NIEDERREITER *Uniform distribution of sequences*, Wiley-Interscience Publ., 1974.
- [112] V. LEFÈVRE *An algorithm that computes a lower bound on the distance between a segment and \mathbb{Z}^2* , Proceedings of SCAN-98, 199–208, Budapest, Hungary, 1998, Kluwer Academic Publishers.
- [113] V. LEFÈVRE, J.-M. MULLER, A. TISSERAND *Towards correctly rounded transcendentals*, IEEE Transactions on Computers **47** (1998), 1235–1243.
- [114] M. LERCH *Question 1547*, L'intermédiaire des mathématiciens **11** (1904), 145–146.
- [115] J. LESCA *Sur la répartition modulo 1 des suites $(n\alpha)$* , Séminaire Delange-Pisot-Poitou (1966-67), Théorie des Nombres, Fasc. 1, Exp. 2, 9 pp.
- [116] J. LESCA *Sur la répartition modulo 1 de la suite $n\alpha$* , Acta Arith. **20** (1972), 345–352.
- [117] P. LIARDET *Regularities of distribution*, Compositio Mathematica **61** (1987), 267–293.
- [118] P. LIARDET *Propriétés harmoniques de la numération suivant Jean Coquet*, Colloque de Théorie Analytique des Nombres “Jean Coquet” (Marseille, 1985), 1–35, Publ. Math. Orsay, 88-02, Univ. Paris XI, Orsay, 1988.

- [119] P. LIARDET *Dynamical properties of the Ostrowski α -expansion*, travail en cours.
- [120] L.-M. LOPEZ, P. NARBEL *DOL-systems and surface automorphisms*, Mathematical foundations of computer science, 1998 (Brno), 522–532, Lecture Notes in Comput. Sci. **1450**, Springer, Berlin, 1998.
- [121] L.-M. LOPEZ, P. NARBEL *Substitutions from Rauzy induction*, Developments in Language Theory 99 (Aachen) (W. Thomas, ed.), 1999, pp. 224–233.
- [122] L.-M. LOPEZ, P. NARBEL *Substitutions and interval exchange transformations of rotation class*, Theoret. Comput. Sci, à paraître.
- [123] M. LOTHAIRE *Mots. Mélanges offerts à M.-P. Schützenberger*, J. Berstel, Tracé de droites, fractions continues et morphismes itérés, 298–309, Langue, Raisonnement, Calcul, Hermès, 1990.
- [124] M. LOTHAIRE *Algebraic Combinatorics on Words*, Chapitre 2 : Sturmian words, par J. Berstel et P. Séébold, à paraître.
- [125] A. MESSAOUDI *Propriétés arithmétiques et dynamiques du fractal de Rauzy*, J. Th. Nombres de Bordeaux **10** (1998), 135–162.
- [126] A. MESSAOUDI *Frontière du fractal de Rauzy et système de numération complexe*, Acta Arith. **95** (2000), 195–224.
- [127] F. MIGNOSI *Infinite words with linear subword complexity*, Theoret. Comput. Sci. **65** (1989), 221–242.
- [128] F. MIGNOSI *On the number of factors of Sturmian words*, Theoret. Comput. Sci. **82** (1991), 71–84.
- [129] F. MIGNOSI, L. Q. ZAMBONI *On the number of Arnoux-Rauzy words*, prépublication 2000.
- [130] M. MORSE, G. A. HEDLUND *Symbolic dynamics*, Amer. J. Math. **60** (1938), 815–866.
- [131] M. MORSE, G. A. HEDLUND *Symbolic dynamics II : Sturmian trajectories*, Amer. J. Math. **62** (1940), 1–42.
- [132] K. NISHIOKA, J. TAMURA, I. SHIOKAWA *Arithmetical properties of a certain power series*, J. Number Theory **42** (1992), 61–87.
- [133] A. OSTROWSKI *Bemerkungen zur Theorie der Diophantischen Approximationen I, II*, Abh. Math. Sem. Hamburg I (1922), 77–98 et 250–251.
- [134] J.-J. PANSIOT *Complexité des facteurs des mots infinis engendrés par morphismes itérés*, Lecture Notes Comput. Sci. **172** (1984), 380–389.
- [135] C. PINNER *On sums of fractional parts $\{n\alpha + \gamma\}$* , J. Number Theory **65** (1997), 48–73.
- [136] C. PINNER *On the one-sided boundedness of sums of fractional parts $(\{n\alpha + \gamma\} - 1/2)$* , J. Number Th. **81** (2000), 170–204.

- [137] M. QUEFFÉLEC *Substitution dynamical systems. Spectral analysis*, Lecture Notes in Mathematics **1294**, Springer-Verlag, 1987.
- [138] L. RAMSCHAW *On the discrepancy of the sequence formed by the multiples of an irrational number*, J. Number Theory **13** (1981), 138–175.
- [139] G. RAUZY *Une généralisation du développement en fraction continue*, Sémin. Delange-Pisot-Poitou, exp. 15 (1976–1977), 1–16.
- [140] G. RAUZY *Échanges d'intervalles et transformations induites*, Acta Arith. **34** (1979), 315–328.
- [141] G. RAUZY *Nombres algébriques et substitutions*, Bull. Soc. math. France **110** (1982), 147–178.
- [142] G. RAUZY *Suites à termes dans un alphabet fini*, Séminaire de Théorie des Nombres de Bordeaux, 1982-1983, exposé 25, 16 pp.
- [143] G. RAUZY *Ensembles à restes bornés*, Séminaire de Théorie des Nombres de Bordeaux, 1983-1984, exposé 24, 12 pp.
- [144] G. RAUZY *Des mots en arithmétique*, Avignon conference on language theory and algorithmic complexity (Avignon, 1983), 103–113, Publ. Dép. Math. Nouvelle Sér. B, 84-6, Univ. Claude-Bernard, Lyon, 1984.
- [145] G. RAUZY *Mots infinis en arithmétique*, in Automata on infinite words (M. Nivat, D. Perrin, eds.) Lecture Notes in Comput. Sci. **192** (1985), 167–171.
- [146] G. RAUZY *Sequences defined by iterated morphisms*, Sequences (Naples/Positano, 1988), 275–286, Springer, New York, 1990.
- [147] R. N. RISLEY, L. Q. ZAMBONI *A generalization of Sturmian sequences; combinatorial structure and transcendence*, Acta Arith. **95** (2000), 167–184.
- [148] G. ROTE *Sequences with subword complexity $2n$* , J. Number Theory **46** (1994), 196–213.
- [149] J. W. SANDER, R. TIJDEMAN *Low complexity functions and convex sets in \mathbb{Z}^k* , Mathem. Zeitschrift **223** (2000), 205–218.
- [150] J. W. SANDER, R. TIJDEMAN *The complexity of functions on lattices*, Theoret. Comput Sci. **246** (2000), 195–225.
- [151] J. W. SANDER, R. TIJDEMAN *The rectangle complexity of functions on two-dimensional lattices*, Theoret. Comput Sci., à paraître.
- [152] C. SERIES *The geometry of Markoff numbers*, Math. Intelligencer **7** (1985), 20–29.
- [153] C. SERIES *The modular surface and continued fractions*, J. London Math. Soc. (2) **31** (1985), 69–80.
- [154] J. SHALLIT *Real numbers with bounded partial quotients : a survey*, Enseign. Math. **38** (1992), 151–187.
- [155] J. SCHOISSENGEIER *On the discrepancy of $n\alpha$* , Acta Arith. **44** (1984), 241–279.

- [156] F. SCHWEIGER *Multidimensional continued fractions*, Oxford Science Publications, Oxford University Press 2000.
- [157] N. A. SIDOROV, A. M. VERSHIK *Arithmetic expansions associated with rotations of the circle and continued fractions*, St. Petersburg Math. Journ. **5** (1994), 1121–1136.
- [158] N. B. SLATER *Gaps and steps for the sequence $n\theta \bmod 1$* , Proc. Cambridge Philos. Soc. **63** (1967), 1115–1123.
- [159] V. T. SÓS *On the theory of diophantine approximations I (On a problem of A. Ostrowski)*, Acta Math. Acad. Sci. Hungar. **8** (1957), 461–472.
- [160] V. T. SÓS *On the theory of Diophantine approximations II (Inhomogeneous problems)*, Acta Math. Acad. Sci. Hungar. **9** (1957), 229–241.
- [161] V. T. SÓS *On the distribution mod 1 of the sequence $n\alpha$* , Ann. Univ. Sci. Budapest, Eötvös Sect. Math. **1** (1958), 127–134.
- [162] V. T. SÓS *On a problem of S. Hartman about normal forms*, Colloq. Math. **7** (1959/1960), 155–160.
- [163] V. T. SÓS *On a problem in the theory of simultaneous approximation*, Ann. Univ. Sci. Budapest, Eötvös Sect. Math. **3–4** (1960/1961), 291–294.
- [164] V. T. SÓS *On the discrepancy of the sequence $n\alpha$* , Coll. Math. Soc. J. Bolyai **13** (1976), 359–367.
- [165] M. STEWART *Irregularities of uniform distribution*, Acta Math. Acad. Scient. Hung. **37**(1981), 185–221.
- [166] K. STOLARSKY *Beatty sequences, continued fractions, and certain shift operators*, Canad. Math. Bull. **19** (1976), 473–482.
- [167] J. SURÁNYI *Über die Anordnung der Vielfachen einer reellen Zahl mod 1*, Ann. Univ. Sci. Budapest, Eötvös Sect. Math. **1** (1958), 107–111.
- [168] S. ŚWIERCZKOWSKI *On successive settings of an arc on the circumference of a circle*, Fundamenta Math. **46** (1958), 187–189.
- [169] R. TIJDEMAN *Exact covers of balanced sequences and Fraenkel’s conjecture*, Algebraic number theory and Diophantine analysis (Graz, 1998), 467–483, de Gruyter, Berlin, 2000.
- [170] R. TIJDEMAN *Fraenkel’s conjecture for six sequences*, Discrete Math. **222** (2000), 223–234.
- [171] D. VANDETH *Sturmian words and words with a critical exponent*, Theoret. Comput. Sci. **242** (2000), 283–300.
- [172] I. VARDI *Lattice points in a triangle*, prépublication 99.
- [173] A. M. VERSHIK *Arithmetic isomorphisms of hyperbolic automorphisms of the torus and sofic shifts*, Functional analysis and its applications **26** (1992), 170–173.

- [174] L. VUILLON *Combinatoire des motifs d'une suite sturmiennne bidimensionnelle*, Theoret. Comput. Sci. **209** (1998), 261–285.
- [175] N. WOZNY, L. Q. ZAMBONI *Frequencies of factors in Arnoux-Rauzy sequences*, Acta Arith. **96** (2001), à paraître.
- [176] S.-I. YASUTOMI *On Sturmian sequences which are invariant under some substitutions*, Number theory and its applications (Kyoto, 1997), 347–373, Dev. Math., 2, Kluwer Acad. Publ., Dordrecht, 1999.
- [177] L. Q. ZAMBONI *Une généralisation du théorème de Lagrange sur le développement en fraction continue*, C. R. Acad. Sci. Paris **327**, Série I (1998), 527–530.
- [178] E. ZECKENDORF *Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas*, Bull. Soc. Roy. Sci. Liège **41** (1972), 179–182.

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