SUR LA COMPLEXITÉ DES SUITES INFINIES

Jean-Paul Allouche

Résumé

On appelle *complexité* d'une suite infinie à valeurs dans un ensemble fini, la fonction qui compte le nombre de facteurs de longueur donnée de cette suite. Nous proposons un catalogue de résultats sur la complexité de certaines suites ou classes de suites.

Abstract

The *complexity* of an infinite sequence taking its values in a finite set counts the number of factors (subwords) of given length of this sequence. We give a quick view of complexity results for different sequences or classes of sequences.

1 Introduction

L'intuition suivant laquelle une suite est d'autant plus "compliquée" qu'elle a "beaucoup" de facteurs différents peut être traduite par exemple dans la définition classique de la fonction de complexité d'une suite. Soit A un alphabet (ensemble fini), et $u = (u(n))_n$ une suite à valeurs dans A. On appelle *complexité* de u, et on note p_u la fonction qui compte le nombre de facteurs de u de longueur donnée. Autrement dit p_u est définie sur les entiers par

$$p_u(n) = \#\{w \in A^n;\ \exists k,\ \ w = u(k)u(k+1)\cdots u(k+n-1)\}.$$

Cette notion permet de définir *l'entropie topologique* d'une suite comme étant :

$$h(u) = \lim_{n\to\infty} \frac{\log p_u(n)}{n \log \#A},$$

(cette limite existe toujours pour une suite infinie u). Notons que la complexité d'une suite u vérifie la double inégalité :

$$1 \le p_u(n) \le (\#A)^n,$$

et donc que son entropie vérifie :

$$0 \le h(u) \le 1.$$

Received by the editors April 1993, revised December 1993.
Communicated by M. Boffa.
AMS Mathematics Subject Classification : 68R15, 11B85.
Keywords : complexity, factors (subwords) of infinite words, substitutive sequences, automatic sequences, CD0L sequences.

Une suite “très compliquée” contiendra tous les facteurs possibles, sa complexité et son entropie seront donc maximales.

D’autres notions de complexité existent, comme celle proposée par Kolmogorov, Chaitin et Solomonoff, qui compare la donnée effective d’une suite au “plus petit programme nécessaire pour l’engendrer”, (voir par exemple [44]).

Nous nous proposons ici de donner un catalogue des propriétés de la complexité de suites classiques. Nous parlerons d’abord des suites sturmiennes qui sont les suites non ultimement périodiques de complexité minimale sur un alphabet à deux lettres, et ferons allusion à certaines de leurs généralisations. Puis nous évoquerons le cas des suites automatiques, (ou même seulement substitutives), dont le prototype est la fameuse suite de Prouhet-Thue-Morse. Puis nous insisterons sur deux familles de suites remarquables pour lesquelles nous avons obtenu des résultats que nous avons détaillés ailleurs, les suites de pliage de papier et les suites de Rudin-Shapiro. Nous finirons par quelques questions ouvertes.

2 Les suites sturmiennes

Un premier théorème ([49] et [50], [23]) stipule que *si u est une suite pour laquelle il existe un entier n tel que $p_u(n) \leq n$, alors u est ultimement périodique, (ce qui entraîne que p_u est ultimement constante).*

Autrement dit, les suites “les plus simples” non périodiques à partir d’un certain rang devraient vérifier

$$\forall n \geq 1, \quad p_u(n) = n + 1.$$

En faisant $n = 1$, on voit que ces suites sont nécessairement binaires. De telles suites existent, elles sont appelées *suites sturmiennes*, et une abondante littérature leur a été consacrée, (voir par exemple [14]). Rappelons seulement qu’on peut les engendrer en coupant le réseau $\mathbf{Z^2}$ par une droite de pente irrationnelle, ou ... en jouant au billard sur un carré : toute suite sturmienne peut s’écrire soit sous la forme $(\lfloor (n+1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor)_{n\geq 0}$, soit sous la forme $(\lceil (n+1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil)_{n\geq 0}$ pour un α irrationnel dans $[0, 1]$ et un β réel.

Plusieurs généralisations de ces suites existent, nous en citerons trois.

- On peut remplacer le réseau $\mathbf{Z^2}$ par le réseau $\mathbf{Z^n}$, ou encore jouer au billard multi-dimensionnel. C’est l’objet d’un travail d’Arnoux, Mauduit, Shiokawa et Tamura, ([10] et [11]). Les auteurs établissent en particulier que *la complexité du billard à trois dimensions est donnée par $p(n) = n^2 + n + 1$.* Ils conjecturent une jolie propriété de symétrie en d et n de l’expression de $p_d(n)$, (où p_d est la complexité du billard à d dimensions).

- On peut jouer au billard sur un triangle “rationnel”, c’est-à-dire dont les angles sont des multiples rationnels de π, comme dans un travail de Hubert ([41]). *Si l’on évite l’ensemble dénombrable d’angles d’attaque α qui donnent des résultats triviaux, la complexité de la trajectoire d’angle d’attaque α, lorsque les angles du triangle sont $\frac{a}{r}\pi, \frac{b}{r}\pi$ et $\frac{c}{r}\pi$ avec pgcd $(a, b, c, r) = 1$, est donnée par :*

$$p(n) = nr + 2r, \; n \geq n_0(\alpha).$$

Par exemple on obtient une complexité ultimement égale à $4n + 8$ pour le triangle rectangle isocèle et à $3n + 6$ pour le triangle équilatéral. Ce résultat est d’ailleurs généralisable ([41]) au cas d'un polygone à m côtés dont les angles sont des multiples rationnels de π, il faut remplacer 2 par $m-1$ dans la formule ci-dessus.

- Notons Δ l'opérateur différence première, qui à la suite u associe la suite $(u(n+1)-u(n))_n$. La suite sturmienne $(\lfloor (n+1)\alpha\rfloor - \lfloor n\alpha\rfloor)_n$ s'écrit $\Delta(\lfloor (n\alpha)\rfloor)$. Il est alors tentant de remplacer n par n^2 dans l'expression ci-dessus, quitte à remplacer Δ par Δ^2 pour que la suite obtenue ne prenne qu'un nombre fini de valeurs. Arnoux et Mauduit ont montré que *la suite* $\Delta^2(\lfloor n^2\alpha\rfloor)$*, avec* α *irrationnel dans* $[0,1]$*, ne prend que* 4 *valeurs et sa complexité vaut* $\frac{(n+1)(n+2)(n+3)}{6}$, ([12]).

Citons par ailleurs deux autres familles de suites dont la complexité a une expression ... simple :

- les suites étudiées par Rauzy, puis Arnoux et Rauzy, pour lesquelles $p_u(n) = 2n+1$, ([57], [13], voir aussi le travail de Ferenczi [40]),
- les suites étudiées par Rote, très liées aux suites sturmiennes et pour lesquelles $p_u(n) = 2n$, ([58]), et plus généralement les rotations d'angle α sur $[0,\beta]$ étudiées par Alessandri, pour lesquelles $p(n) = 2n$, pour $n \geq n_0(\alpha)$ lorsque α est convenable, ([1]).

La "philosophie" qui sous-tend les résultats de ce paragraphe est l'existence pour chacune de ces suites ou familles de suites d'une interprétation géométrique, qui explique à la fois leur basse complexité et les expressions simples que l'on obtient.

3 Les suites q-automatiques

Soit q un nombre entier supérieur ou égal à 2, nous dirons qu'une suite u est q-automatique si l'ensemble de sous-suites de u

$$\{n \to u(q^k n + a);\ k \geq 0,\ 0 \leq a \leq q^k - 1\}$$

est fini (cet ensemble est appelé *q-noyau* de la suite u). On peut aussi engendrer une telle suite par un q-automate (tag-system uniforme de module q), ou l'obtenir comme image d'un point fixe d'un morphisme uniforme de longueur q.

Rappelons que, *lorsque q est une puissance d'un nombre premier, et que u est à valeurs dans le corps fini* $\mathbf{F}_q$*, la suite u est q-automatique si et seulement si la série formelle* $\Sigma u_n X^n$ *est algébrique sur le corps de fractions rationnelles* $\mathbf{F}_q(X)$, (c'est le théorème de Christol, Kamae, Mendès France et Rauzy, [21]). Pour ces notions on pourra aussi consulter [22], [28] et [2].

Le premier résultat de complexité de ces suites est dû à Cobham ([22]) : *si une suite u est q-automatique, alors* $p_u(n) = O(n)$*, (de sorte que n est l'ordre exact de la complexité des suites automatiques non ultimement périodiques).*

Ensuite un résultat prouvé indépendamment par Brlek d'une part ([18]), et de Luca et Varricchio d'autre part ([30], voir aussi [29]), étudie le cas de la suite de Prouhet-Thue-Morse et peut s'énoncer sous la forme affaiblie suivante : *si u désigne la suite de Prouhet-Thue-Morse, alors la suite* $(p_u(n+1)-p_u(n))_n$ *est* 2*-automatique.*

Le cas de la suite de Thue-Morse sur un alphabet à plusieurs lettres a été étudié par Mouline ([52]). Puis, de façon plus générale, Tapsoba ([63]), Rauzy, Mossé

([51]) montrent que si u est une suite q-automatique vérifiant certaines conditions (techniques) supplémentaires, alors la suite $(p_u(n+1) - p_u(n))_n$ est aussi q-automatique; notons que cela entraîne que la suite p_u est q-régulière au sens de [8].

Le cas des suites de pliage de papier et des suites de Rudin-Shapiro est décrit aux paragraphes suivants.

D'autres classes de suites engendrées par des morphismes, des transducteurs, des D0L-systèmes ... ont été étudiées, citons en plus de l'article de Cobham ([22]), ceux d'Ehrenfeucht, Lee et Rozenberg, ([31] et [32]), ceux d'Ehrenfeucht et Rozenberg ([33], [34], [35], [36], [37], [38], [39]), celui de Rozenberg ([59]) et celui de Sajo ([61]). Citons aussi les articles de Bleuzen-Guernalec ([15]), de Bleuzen-Guernalec et Blanc ([16]). Citons enfin les articles de Pansiot ([53], [54] et [55]) et le travail récent de Lepistö ([43]). Nous retiendrons trois résultats frappants dans ces papiers.

- La complexité d'une suite point fixe de transduction peut être exponentielle. L'exemple donné dans [16] est la suite obtenue de la façon suivante : on note $h(n)$ l'écriture binaire renversée redoublée de l'entier n, par exemple $h(1) = 11$, $h(2) = 0011$. Puis on considère la suite

$$0\,0\,1\,h(1)\,0\,1\,h(2)\,0\,1\ \cdots h(n)\,0\,1\,h(n+1)\cdots$$

La complexité de cette suite vérifie :

$$C_1 n(\sqrt{2})^n \leq p(n) \leq C_2 n(\sqrt{2})^n,$$

où C_1 et C_2 sont des constantes strictement positives.

- Si une suite infinie est point fixe d'un morphisme, alors sa complexité vérifie :

$$C_1 f(n) \leq p(n) \leq C_2 f(n),$$

où C_1 et C_2 sont deux constantes strictement positives, et f l'une des fonctions 1, n, $n \log\log n$, $n \log n$ ou n^2, ([53] et [54]). En particulier la complexité d'une telle suite ne peut pas croître comme $n^{3/2}$, $n(\log n)^2$ ou $n \log\log\log n$.
Un exemple est donné par le point fixe commençant par 0 du morphisme

$$0 \to 0101,\ 1 \to 11,$$

dont la complexité croît comme $n \log\log n$ ([54]).

- Une suite obtenue par itération périodique de morphismes peut être de complexité supérieure à Cn^t où $t > 2$, et donc une telle suite peut ne pas être engendrée par un morphisme suivi d'un codage (voir [43]).

4 Les suites de pliage de papier

Si l'on plie une infinité de fois une feuille de papier sur elle-même, on obtient une suite infinie de “montagnes” et de “vallées” que l'on peut coder par des 0 et des 1, (voir par exemple [28]). En pliant la feuille toujours de la même façon, on obtient le “pliage régulier de papier”; en pliant parfois par dessus, parfois par dessous, on obtient un ensemble non dénombrable de suites appelées “suites de pliage de papier”. Notons que ces suites sont 2-automatiques si et seulement si les instructions de pliage (c'est-à-dire les décisions de plier vers le haut ou vers le bas) forment une suite ultimement périodique. Ce résultat a été remarqué par plusieurs auteurs, (par exemple [46], [45]).

Le résultat plutôt surprenant que nous avons prouvé dans [3] est que *toutes les suites de pliage de papier ont la même fonction de complexité, et cette fonction vérifie* $p_u(n) = 4n,\ \forall n \geq 7$. L'outil essentiel est que les suites de pliage peuvent être obtenues comme suites de Toeplitz (voir par exemple [4] pour un survol).
On peut par exemple construire le pliage régulier de papier comme suit. On part de la suite $(01)^\infty$ dans laquelle on intercale des "trous" une fois sur deux :

$$0 \quad \bullet \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \quad \bullet \quad 0 \quad \cdots,$$

on remplit ensuite les trous avec la suite $(01)^\infty$, mais seulement un trou sur deux :

$$0 \quad 0 \quad 1 \quad \bullet \quad 0 \quad 1 \quad 1 \quad \bullet \quad 0 \quad 0 \quad 1 \quad \bullet \quad 0 \quad 1 \quad 1 \quad \bullet \quad 0 \quad \cdots.$$

Il ne reste plus qu'à itérer ce procédé, obtenant successivement :

$$\begin{array}{cccccccccccccccccc}
0 & 0 & 1 & 0 & 0 & 1 & 1 & \bullet & 0 & 0 & 1 & 1 & 0 & 1 & 1 & \bullet & 0 & \cdots, \\
0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & \bullet & 0 & \cdots, \\
0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & \cdots.
\end{array}$$

Notons que, si à chaque étape on choisit d'insérer soit la suite $(01)^\infty$, soit la suite $(10)^\infty$, on obtient *toutes les suites de pliage*. Indiquons enfin que l'égalité de toutes les fonctions de complexité des suites de pliage laisse penser qu'il existe des bijections "naturelles" entre les facteurs de longueur donnée de ces suites. Dans [5] est prouvée une propriété de "synchronisation" des facteurs de ces suites : *pour chaque entier k il existe un ensemble de $4k$ positions où on est sûr de trouver une occurrence de chaque facteur de longueur k dans* **toutes** *les suites de pliage.*

5 Les suites de Rudin-Shapiro

Si u est une suite infinie à valeurs dans $\{-1, +1\}$, que peut-on dire de la quantité

$$F_N(u) = \| \sum_{n=0}^{N-1} u(n) e^{2i\pi n\theta} \|_\infty$$

lorsque N tend vers l'infini ? (le sup étant pris sur les θ réels). En majorant trivialement, puis en minorant la norme L^∞ par la norme L^2, on obtient :

$$\sqrt{N} \quad \leq \quad F_N(u) \quad \leq \quad N.$$

Par ailleurs on peut démontrer que, pour presque toute suite u, on a :

$$F_N(u) \quad \leq \quad \sqrt{N \log N}.$$

Shapiro en 1951 ([62]), puis indépendamment Rudin en 1959 ([60]) ont construit une suite u "explicite", pour laquelle on a :

$$F_N(u) \quad \leq \quad C\sqrt{N}.$$

Brillhart et Carlitz ont prouvé dans [17] que cette suite peut être définie par $u(n) = (-1)^{v(n)}$ où $v(n)$ représente le nombre de 11 (avec chevauchements éventuels) dans le développement binaire de n. En particulier u est 2-automatique.

Shallit a conjecturé en 1989 que la fonction de complexité p_u de la suite u de Rudin-Shapiro vérifie :

$$\forall n \geq 8, \ \ p_u(n) = 8n - 8.$$

Ce résultat peut en fait être déduit des travaux de Mouline [52] et de Tapsoba [63], il a été aussi prouvé indépendamment par Brlek [19].

Au lieu de compter les 11 dans le développement de n en base 2, on peut, suivant une idée de Mendès France, compter les $1 * \cdots * 1$. Notons v_d le nombre de $1 * \cdots * 1$ dans le développement binaire de n (il y a $d-1$ étoiles) : ainsi $v = v_1$, et, par exemple, v_2 compte le nombre total de $1 * 1$, c'est-à-dire le nombre total de 101 et de 111 dans n en base 2. La fonction de complexité de la suite $u_d = (-1)^{v_d}$ a aussi une expression simple (voir [9]) : *il existe un entier $n_0(d)$ tel que, pour $n \geq n_0(d)$, on a $p_{u_d}(n) = 2^{d+2}(n-1)$.*

Notons qu'il se passe pour ces suites de Rudin-Shapiro généralisées u_d un phénomène curieux. Elles sont obtenues par projection à partir de points fixes de morphismes de longueur une puissance de 2 sur un “gros” alphabet (donc 2-automatiques). La complexité d'un tel point fixe peut être calculée, puis un miracle a lieu, *cette complexité est ultimement égale à celle de la suite obtenue après projection*; en d'autres termes un facteur d'une telle suite de Rudin-Shapiro, dès qu'il est assez grand, est l'image d'un seul facteur de la suite correspondante sur le gros alphabet.

Il se trouve que *les suites u_d ont la propriété de Rudin-Shapiro, dite “du $\sqrt{N}$”, ($F_N(u_d) \leq C_d\sqrt{N}$)*, comme démontré dans [7]. Il est alors tentant de se demander quelle est la complexité d'autres suites de Rudin-Shapiro généralisées ayant aussi cette propriété. Celles introduites par Mendès France et Tenenbaum dans [47] sont fabriquées à partir des suites de pliage de papier, et forment un ensemble non dénombrable de suites ayant la propriété du $\sqrt{N}$: nous avons prouvé dans [3] que *toutes ces suites ont la même complexité, et cette complexité est égale à $8n-8$ pour $n \geq 8$.*

Signalons enfin que, contrairement au cas des suites de pliage, *on ne peut synchroniser tous les facteurs de ces suites de Rudin-Shapiro obtenues à partir des suites de pliage, mais seulement la “moitié” d'entre eux,* (voir [6]).

6 Questions

Dans ce qui précède on a vu un certain nombre de suites dont la fonction de complexité est ultimement affine (suites sturmiennes, suites de billard sur des polygones rationnels, suites d'Arnoux-Rauzy, suites de pliage de papier, suites de Rudin-Shapiro généralisées au sens de [7] et au sens de [47], suites de Rote [58]), suites de rotations sur $[0, \beta]$. Peut-on espérer donner une caractérisation de toutes les suites ayant cette propriété ? ou de toutes les suites automatiques ayant cette propriété ? Y-a-t-il un (vague) lien entre la propriété de Rudin-Shapiro et le caractère ultimemement affine de la fonction de complexité ? Signalons au passage l'intéressant article de Mignosi sur les suites à complexité majorée par une fonction linéaire, ([48]).

Peut-on se débarrasser des conditions techniques de Tapsoba pour affirmer que pour toute suite automatique u, la fonction de complexité p_u est telle que $(p_u(n+1) - p_u(n))_n$ est automatique ? et peut-on alors calculer l'automate correspondant de façon ... automatique, disons de façon algorithmique (voir à ce sujet [51]) ?

Dans quels cas la complexité d'une suite point fixe de morphisme et la complexité de sa projection sont-elles ultimement égales, comme c'est le cas pour les suites de Rudin-Shapiro généralisées u_d citées plus haut ?

Citons dans ce paragraphe la suite "auto-génératrice" de Kolakoski ([42]), pour laquelle peu de choses (en particulier concernant la complexité) sont connues (voir néanmoins [26], [27], [64], [24], [25], [56] and [20]). Cette suite est composée alternativement de blocs de 2 et de blocs de 1, et elle est égale à la suite des longueurs successives de ces blocs :

$$2\,2\,1\,1\,2\,1\,2\,2\,1\,\cdots$$

La conjecture donnée dans [27] est que la complexité de cette suite satisfait :

$$p(n) \sim cn^q, \;\; q = \frac{\log 3}{\log \frac{3}{2}}.$$

Pour conclure indiquons deux directions que nous n'avons pas abordées ici.

- La notion de fonction de récurrence d'une suite infinie. Soit u une suite u *minimale*, c'est-à-dire : pour tout entier n il existe un entier r tel que tout facteur de u de longueur r contienne tous les facteurs de u de longueur n. Soit $R_u(n)$ le plus petit tel entier r. On appelle fonction de récurrence de la suite u, la fonction $n \to R_u(n)$. Sur ce sujet le lecteur pourra lire les articles de Morse et Hedlund, ([49] et [50]), et la thèse de Mouline, ([52]).

- La notion de complexité d'un langage : c'est le nombre de mots de longueur donnée dans ce langage. Naturellement la complexité du langage des facteurs d'une suite infinie coïncide avec la complexité de cette suite.

Remerciements : l'auteur tient à remercier chaleureusement C. Mauduit pour avoir porté à sa connaissance un certain nombre de travaux, M. Bousquet-Mélou et E. Roblet pour avoir grandement contribué à améliorer une version préliminaire de cet article, et les *referees* grâce auxquels des références stratégiques ont été ajoutées.

Références

[1] P. Alessandri, Codage des rotations, Mémoire de D. E. A., Université Claude Bernard, Lyon I (1993).

[2] J.-P. Allouche, Automates finis en théorie des nombres, *Expo. Math.* **5** (1987) 239–266.

[3] J.-P. Allouche, The number of factors in a paperfolding sequence, *Bull. Austr. Math. Soc.* **46** (1992) 23–32.

[4] J.-P. Allouche et R. Bacher, Toeplitz sequences, paperfolding, towers of Hanoi and progression-free sequences of integers, *Ens. Math.* **38** (1992) 315–327.

[5] J.-P. Allouche et M. Bousquet-Mélou, Canonical positions for the factors in the paperfolding sequences, à paraître dans *Theoret. Comput. Sci.*

[6] J.-P. Allouche et M. Bousquet-Mélou, Facteurs des suites de Rudin-Shapiro généralisées, ce volume.

[7] J.-P. Allouche et P. Liardet, Generalized Rudin-Shapiro sequences, *Acta Arith.* **60** (1991) 1–27.

[8] J.-P. Allouche et J. Shallit, The ring of k-regular sequences, *Theoret. Comput. Sci.* **98** (1992) 163–187.

[9] J.-P. Allouche et J. Shallit, Complexité des suites de Rudin-Shapiro généralisées, Congrès "Thémate", Luminy 1991, à paraître.

[10] P. Arnoux, C. Mauduit, I. Shiokawa et J.I. Tamura, Complexity of sequences defined by billiards in the cube, *Bull. Soc. math. France*, à paraître.

[11] P. Arnoux, C. Mauduit, I. Shiokawa et J.I. Tamura, Rauzy's conjecture on the cubic billiards, *Tokyo J. Math.*, à paraître.

[12] P. Arnoux, C. Mauduit et G. Meigniez, Sur la complexité de certaines suites arithmétiques, en préparation.

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

[14] J. Berstel et P. Séébold, Mots de Sturm - un survol, en préparation.

[15] N. Bleuzen-Guernalec, Suites points fixes de transductions uniformes, *C.R. Acad. Sci. Paris* **300** Série I (1985) 85–88.

[16] N. Bleuzen-Guernalec et G. Blanc, Production en temps réel et complexité de structure de suites infinies, *Informatique théorique et applications, RAIRO* **23** (1989) 195–216.

[17] J. Brillhart et L. Carlitz, Note on the Shapiro polynomials, *Proc. Amer. Math. Soc.* **25** (1970) 114–118.

[18] S. Brlek, Enumeration of factors in the Thue-Morse word, *Discrete Applied Math.* **24** (1989) 83–96.

[19] S. Brlek, Communication privée.

[20] A. Carpi, Repetitions in the Kolakoski sequence, *Bull. EATCS* **50** (1993) 194–196.

[21] G. Christol, T. Kamae, M. Mendès France et G. Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. math. France* **108** (1980) 401–419.

[22] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972) 164–192.

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

[24] K. Culik II, J. Karhumäki et A. Lepistö, Alternating iteration of morphisms and the Kolakoski sequence, in G. Rozenberg et A. Salomaa, éditeurs, *Lindenmayer Systems : Impacts in Theoretical Computer Science, Computer Graphics and Developmental Biology*, Springer–Verlag (1992) 93–103.

[25] K. Culik II et J. Karhumäki, Iterative devices generating infinite words, in A. Finkel et M. Jantzen, éditeurs, STACS 1992 : 9th Annual Symposium on Theoretical Aspects of Computer Science, *Lecture Notes in Computer Science* **577** (1992) 531–543.

[26] F.M. Dekking, Regularity and irregularity of sequences generated by automata, *Séminaire de Théorie des Nombres de Bordeaux* (1979–1980), exposé n° 9, 9-01–9-10.

[27] F.M. Dekking, On the structure of self-generating sequences, *Séminaire de Théorie des Nombres de Bordeaux* (1980–1981), exposé n° 31, 31-01–31-06.

[28] M. Dekking, M. Mendès France et A.J. van der Poorten, FOLDS!, *Math. Intell.* **4** (1982) 130–138, 173–181, 190–195.

[29] A. de Luca et S. Varricchio, On the factors of the Thue-Morse word on three symbols, *Information Processing Letters* **27** (1988) 281–285.

[30] A. de Luca et S. Varricchio, Some combinatorial properties of the Thue-Morse sequence, *Theoret. Comput. Sci.* **63** (1989) 333–348.

[31] A. Ehrenfeucht, K.P. Lee et G. Rozenberg, Subword complexities of various classes of deterministic developmental languages without interaction, *Theoret. Comput. Science* **1** (1975) 59–75.

[32] A. Ehrenfeucht, K.P. Lee et G. Rozenberg, On the number of subwords of everywhere growing DT0L-languages, *Discrete Math.* **15** (1976) 223–234.

[33] A. Ehrenfeucht et G. Rozenberg, On the subword complexity of square-free D0L-languages, *Theoret. Comput. Science* **16** (1981) 25–32.

[34] A. Ehrenfeucht et G. Rozenberg, On the subword complexity of D0L-languages with a constant distribution, *Information Processing Letters* **13** (1981) 108–113.

[35] A. Ehrenfeucht et G. Rozenberg, On the subword complexity and square-freeness of formal languages, *Lecture Notes in Computer Science* **104** (1981) 1–4.

[36] A. Ehrenfeucht et G. Rozenberg, On the subword complexity of homomorphic images of languages, *Informatique théorique et applications, RAIRO* **16** (1982) 303–316.

[37] A. Ehrenfeucht et G. Rozenberg, On the subword complexity of locally catenative D0L-languages, *Information Processing Letters* **16** (1983) 7–9.

[38] A. Ehrenfeucht et G. Rozenberg, On the subword complexity of m-free D0L-languages, *Information Processing Letters* **17** (1983) 121–124.

[39] A. Ehrenfeucht et G. Rozenberg, On the size of the alphabet and the subword complexity of square-free D0L languages, *Semigroup Forum* **26** (1983) 215–223.

[40] S. Ferenczi, Les transformations de Chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité $2n+1$, Preprint (1993).

[41] P. Hubert, Dynamique symbolique des billards polygonaux, Preprint (1993).

[42] W. Kolakoski, Self-generating runs, Prob. 5304, sol. in *Amer. Math. Monthly* **73** (1966) 681–682.

[43] A. Lepistö, On the power of periodic iteration of morphisms, *Lecture Notes in Computer Science* **700** (1993) 496–506.

[44] M. Li et P. Vitányi, Kolmogorov complexity and its applications, in J. van Leeuwen éditeur, *Handbook of Theoretical Computer Science*, Volume A : Algorithms and Complexity, MIT Press (1990) 187–254.

[45] P. Liardet, communication privée.

[46] M. Mendès France et J. Shallit, Wire bending, *J. Combin. Theory, Ser. A* **50** (1989) 1–23.

[47] M. Mendès France et G. Tenenbaum, Dimension des courbes planes, papiers pliés et suites de Rudin-Shapiro, *Bull. Soc. math. France* **109** (1981) 207–215.

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

[49] M. Morse et G.A. Hedlund, Symbolic Dynamics, *Amer. J. Math.* **60** (1938) 815–866.

[50] M. Morse et G.A. Hedlund, Symbolic Dynamics II. Sturmian trajectories, *Amer. J. Math.* **62** (1940) 1–42.

[51] B. Mossé, Notions de reconnaissabilité pour les substitutions et complexité des suites automatiques, Preprint (1993).

[52] J. Mouline, Contribution à l'étude de la complexité des suites substitutives, Thèse, Université de Provence (1989).

[53] J.-J. Pansiot, Bornes inférieures sur la complexité des facteurs des mots infinis engendrés par morphismes itérés, *Lecture Notes in Computer Science* **166** (1984) 230–240.

[54] J.-J. Pansiot, Complexité des facteurs des mots infinis engendrés par morphismes itérés, *Lecture Notes in Computer Science* **172** (1984) 380–389.

[55] J.-J. Pansiot, On various classes of infinite words obtained by iterated mappings, *Lecture Notes in Computer Science* **192** (1985) 188–197.

[56] G. Păun, How much Thue is Kolakoski?, *Bull. EATCS* **49** (1993) 183–185.

[57] G. Rauzy, Suites à termes dans un alphabet fini, *Séminaire de Théorie des Nombres de Bordeaux* (1982–1983), exposé n° 25, 25-01–25-16.

[58] G. Rote, Sequences with subword complexity $2n$, *J. Number Theory*, à paraître.

[59] G. Rozenberg, On subwords of formal languages, *Lecture Notes in Computer Science* **117** (1981) 328–333.

[60] W. Rudin, Some theorems on Fourier coefficients, *Proc. Amer. Math. Soc.* **10** (1959) 855–859.

[61] A. Sajo, On subword complexity functions, *Discrete Applied Math.* **8** (1984) 209–212.

[62] H.S. Shapiro, Extremal problems for polynomials and power series, Thesis, M.I.T. (1951).

[63] T. Tapsoba, Complexité de suites automatiques, Thèse de troisième cycle, Université Aix-Marseille II (1987).

[64] W.D. Weakley, on the number of C^∞-words of each length, *J. Combin. Theory Ser. A* **51** (1989) 55–62.

Jean-Paul Allouche
C.N.R.S., U.P.R. 9016
L.M.D., Luminy, Case 930
F-13288 Marseille Cedex 9
France